Random Network Consideration: Theory and Experiment *

Ian Chadd†

November 5, 2018

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, etc. However, it is not straightforward whether, given choice data, a) preferences among alternatives can be revealed, or b) the network of consideration spillovers can be revealed. Using a novel laboratory experiment, I test a deterministic Network Choice model proposed in previous work and find a plethora of violations thereof, even at the individual level. I then propose a stochastic model, Random Network Choice, and analyze its properties regarding the formation of consideration sets. When applied to the laboratory data, I find considerable consistency with the general Random Network Choice model. Armed with a model of network choice consistent with my experimental data, I consider one application in the realm of advertising to show that such a generalization of so-called "positive spillovers" in attention is necessary to avoid misleading welfare analysis.

JEL Codes: D83, D90
Keywords: Limited consideration, random attention, product network
Most Recent Version: `http://econweb.umd.edu/~chadd/files/jmp.pdf`

*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.

†Department of Economics, University of Maryland, Email: chadd@econ.umd.edu

1 Introduction

In decision environments with a large number of alternatives, decision makers (DMs hereinafter) may structure search according to a network of connections between these alternatives. For example, a shopper on Amazon.com utilizes a list of "suggested items" to navigate between available goods. The network of connections between options need not be exogeneously provided by some firm, however. Consider a DM who is considering donating to some charity from the set {Animal Shelter in DC, Animal Shelter in NYC, Homeless Shelter in NYC}. Then a DM who initially considers donating to the Animal Shelter in DC may subsequently consider donating to the Animal Shelter in NYC because of the shared attribute of being an Animal Shelter. Similarly, the same DM may then consider the Homeless Shelter in NYC because of the shared geographic location with the Animal Shelter in NYC. The DM will eventually consider both the Animal Shelter in DC and the Homeless Shelter in NYC, even though the two charities share no common attributes. If attention "spills over" between options in this manner in some decision making environment, it would be important for firms to be able to properly elicit the network from the choices of DMs and attention data, if observable.

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[1] to a competing restaurant by round 4%. 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 <u>categories</u>: 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 modelling 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 <u>network of connections</u> between options, with this category-specific treatment as a special case.

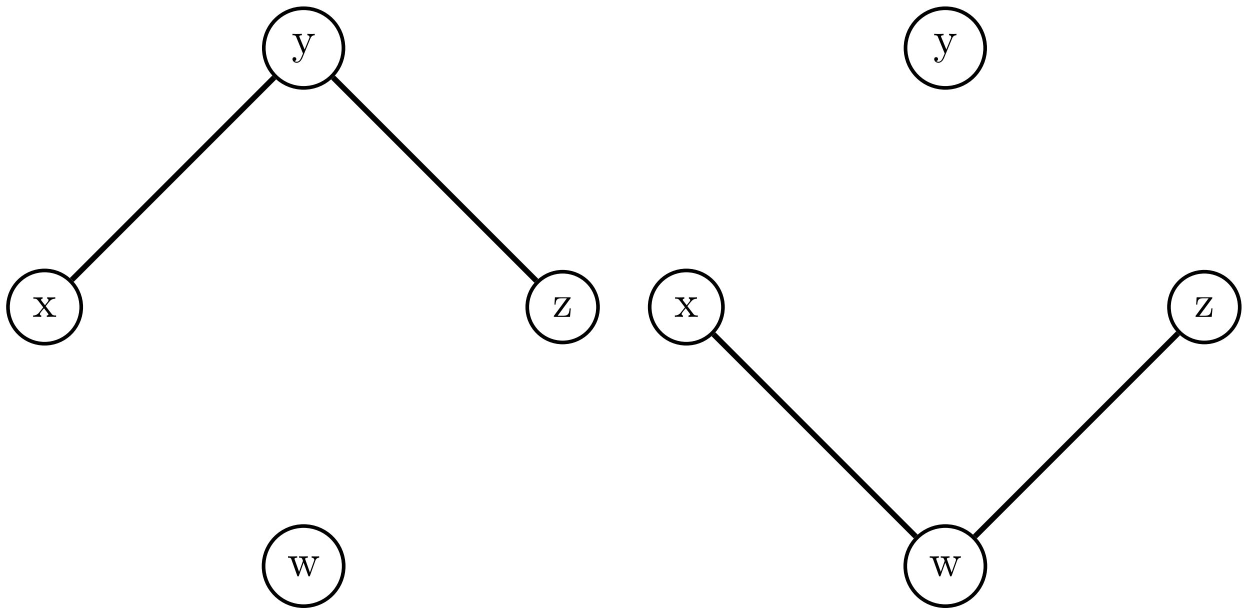
Beyond the importance that a model of such network consideration has for firms, a precise model of network consideration is important for welfare analysis. Indeed, a common refrain among these marketing

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

studies of attention spillovers is that these positive externalities lead to an under-allocation of advertising relative to the social optimum; Shapiro (2018) presents a supply-side model to make this case. However, if consideration is modelled as following a more general network structure, this is not necessarily true. I show this much in the Section 7.

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 (2017) and which I'll call "Network Choice" (NC hereinafter), is leveraged to structure the parameterization of the laboratory experiment. NC also serves as a deterministic baseline model against which to test the elicited attention 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 Choice (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 general random product 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 and better fits the experimental data. To preview how this model works, consider the following example:

Figure 1: Example random graph on four options



(a) $f(g_1) = \frac{1}{3}$ (b) $f(g_2) = \frac{2}{3}$

In the above example, the options $\{w, x, y, z\}$ are connected to one another in a random graph structure. The random graph is represented as a distribution over the set of all possible graphs on these four options. In Figure 1, the graph g_1 occurs with probability $\frac{1}{3}$, and the graph g_2 occurs with probability $\frac{2}{3}$. All other possible graphs occur with probability 0. Consider a DM who starts at option x and considers options according to RNC. Then with probability $\frac{1}{3}$, the network g_1 is in effect, in which case attention spills over from option x to option y, then from option y to option z. Let $\Gamma_x(T \mid S)$ be the probability that set T is considered when S is available and the starting point is x. Since there are no other networks that connect the set $\{x, y, z\}$, $\Gamma_x(\{x, y, z\} \mid \{w, x, y, z\}) = f(g_1) = \frac{1}{3}$. Notice that, similarly, $\Gamma_x(\{w, x, z\} \mid \{w, x, y, z\}) = f(g_2) = \frac{2}{3}$. The RNC model works as follows: given a starting point x and an available set S, the DM forms stochastic consideration sets according to some distribution over possible networks. From each consideration set, the DM chooses the option that is maximal according to some partial order $\succ$. 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 some dataset that is consistent with RNC, the properties of RNC are such that it may admit an infinite number of representations. In some settings, this may not be a desirable property. For this reason, I then consider a special case of RNC, which I dub the "Pseudo-Markovian RNC" model (PM-RNC hereinafter) due to its proximity to "Markov networks," common in the network analysis body of literature.[2] Under PM-RNC, I show that the PM-RNC representation of some set of stochastic choice data must be

[2]See Frank and Strauss (1986) for a discussion of Markov networks.

unique (up to permutations of preferences between alternatives for which preferences cannot be revealed with the given dataset). I present an additional necessary consideration set mapping property for PM-RNC, Binary Separability, and take this to the experimental data. I find mixed evidence of consistency with Binary Separability, suggesting that there is likely a family of RNC special cases between the most general RNC model and PM-RNC that i) adds structure beyond RNC in the direction of PM-RNC, but ii) is similarly consistent with my experimental data. An exploration of such classes of models is beyond the scope of the current work and would make for a fruitful next step in the study of network consideration.

This paper proceeds as follows. Related literature, both theoretical and experimental studies, are 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 and PM-RNC models which are tested in Section 6. These results are discussed in light of an application to advertising 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.[3] 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 (2017) 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. (2018) 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

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

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. (2015) 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 (2017) 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 and PM-RNC are both decision theoretic models and involve no strategic interaction between multiple agents.

2.2 Theory

The proposed RNC model contained herein is closely related to several models of path-dependent attention and choice. Masatlioglu and Suleymanov (2017) present a model of Network Choice (NC hereinafter) where attention spills over between options in a given deterministic network. In the realm of stochastic path-dependent models of limited consideration, Suleymanov (2018) presents a Path-Dependent Consideration model that is also similar to RNC, in that consideration follows a path of connections between available options according to some stochastic process. Suleymanov (2018) assigns probabilities in this model to <u>paths</u> with some initial starting point, where RNC assigns probabilities to more general <u>networks</u>. Further, Suleymanov (2018) builds on earlier work contained in Masatlioglu and Nakajima (2013) where new elements are added to the consideration set only when they dominate everything that has already been included in the consideration set. RNC does not share this feature, instead allowing consideration sets to evolve stochastically, independent of the preference relation. The same approach is used in NC, though for a deterministic setting.

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. This results in choice probabilities that violate a regularity condition of Luce (1959), where adding an element to the feasible set should not increase the choice frequency of a given element previously available. In more recent work, Cattaneo et al. (2017) present a "Random Attention Model" (RAM hereinafter), that actually relies on violations of the Luce regularity condition to reveal preference. They apply a monotonicity condition on attention rules of the following form:

$$\text{For any } a \in S - T,\, \Gamma(T \mid S) \leq \Gamma(T \mid S - a).$$

where $\Gamma(T \mid S)$ is the probability that the set T is considered when S is available. In Section 5, I show that RNC satisfies a starting-point contingent version of this monotonicity condition. This allows me to directly connect the revealed preference approach in Cattaneo et al. (2017) to that in RNC.[4]

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).

3 Experiment

In order to test the deterministic NC model, we construct a laboratory environment with several goals. First, the environment must mimic a choice setting where distinct options are linked to one another via a product network. Second, the environment must induce the subject to behave as if they were in the real world analogue to the laboratory environment - that is, choice must be properly incentivized. Finally, we err on the side of creating an overly 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 more complicated considerations or fewer restrictions.

[4] Manzini and Mariotti (2014), Suleymanov (2018), and Cattaneo et al. (2017) are not the only examples of random attention models. See Cattaneo et al. (2017) for a full review of random attention models and their connection to RAM, of which RNC is a starting-point contingent special case.

Option X

Shape	Pattern	Size	Number
□	═	SMALL	5

Table 1: Option Example

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. On average, subjects earned \$23.63 for approximately 90 minutes of time spent in the lab.[5]

It is helpful to consider the experimental environment from the perspective of a given subject. The subject faces 31 distinctive extended problems, each defined by a starting point x and a set of available options, S, just as in the theory. 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. For each extended decision problem, the subject's payoff is simply the value of the option they have chosen, converted to cash. While subjects make decisions in each of 31 extended decision problems, they are only paid for one, which is chosen randomly at the end of experiment. Subjects do not know 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.

Each option 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, denominated in Experimental Currency Units (ECUs). Each attributes can take on one of 5 values, from 1 ECU to 5 ECU, resulting in 625 distinct possible options, with values ranging from 5 to 25 ECU. A full table of attribute values can be found in Appendix D. For clarity, consider the following option described by these four attributes:

Option X in Table 1 is described by 4 attributes: Square, Two-Bar Pattern, Small, and 5. These each pay off 2 ECU, 3 ECU, 2 ECU, and 5 ECU, respectively. Then the value of Option X is 13 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 information for the starting point and no other available option. This information includes an option identification label, unique at the extended decision problem level (i.e. "Option 5" is displayed at the top of the screen when information for

[5]A single pilot session was conducted with 16 subjects in order to test the experimental software and receive feedback regarding clarity of the instructions used. A few minor changes were made to the design and instructions following this pilot experiment, including, but not limited to, the use of extended decision problem unique option labels. These subjects are not included in any of the analysis contained herein.

Option 5 is presented), attribute information for the displayed option, as well as two lists of information (explained below). In addition, the interface displays information for the subject's provisional choice at all times (explained in detail in Section 3.2).

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 may simply click on the option label in one of these lists and then click a box labelled "View the Selected Option" pertaining to that list. At that time, all relevant information on the screen will update to reflect information for the option to which the subject has navigated.

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 available option also had the Shape attribute "Square" and the Number attribute "5," it would be included in this list of linked options for Option X. 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 are simply agnostic of characteristics of said options, subjects may ask themselves why providing the network structure is necessary in the first place. This may lead to the perception of some deception on the part of the subject or general confusion. 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. Other experiments, both in psychology and experimental economics, use similar designs. 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. In all of the following, whenever statistical tests are conducted separately based on this variation, I use "Baseline" to refer to the original context-less 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, 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, one of the 31 extended decision problems was chosen at random (with each extended decision problem chosen with equal likelihood), and subjects were paid for that single choice only. Once these extended decision problems were completed, they were asked a set of demographic questions on 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.

3.2 Choice Process Data

The experimental design elicits choice process data a la Caplin et al. (2011) in the following manner. Choices in each extended decision problem are treated as provisional, in that choosing a new option does not end the current period. This simply updates the subject's provisional choice, allowing the subject to make a number of switches between provisionally chosen options within a single period. At all times, information regarding the subject's provisional choice was displayed in the upper-right portion of the experimental interface, including the option label (e.g "Option 12") for the provisional choice and attribute information. This information was provided as a reference for the subject to avoid concerns of imperfect recall during the option evaluation process.[6]

[6]It should be noted that, while the appearance of this information differs between the design used herein and that used in Caplin et al. (2011), the two share the feature of always displaying the subject's provisional choice. In Caplin et al. (2011),

At the end of each period, a "decision time" was chosen randomly from a uniform distribution from 2 to 75 seconds. The provisional choice held by the subject at the realized decision time was then treated as the final choice for the extended decision problem and subjects were paid the value of the option held at that time.

In each period, while subjects were initially shown the information for the starting point, they did not initially have any option provisionally chosen. They must then choose some option to serve as their initial provisional choice (usually the starting point itself). For this reason, the lower end of the decision time support was 2 seconds, giving the subject time to choose an initial provisional choice and thus minimizing the number of observations for which the subject might be paid nothing for a given extended decision problem.

3.3 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.

3.3.1 NC: Upward Monotonicity

The first property of consideration sets in NC that I utilize to create extended decision problems is 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:

B1. 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

the provisional choice is indicated by a selected row in a list of continuously displayed available options. In the experiment contained herein, this information is contained in a portion of the interface that simply updates when the provisional choice is altered by the subject.

"reachable" from the starting point x and are also in T.[7] Then when the DM is confronted with (x, S) for $T \subseteq S$, it should be clear that 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). Moreover, since $T \subseteq S$, these options are also still available and therefore should 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 $\delta_i = (x, A_i)$ be one of these five extended decision problems. Each A_i was then chosen such that $A_1 \subseteq A_2 \subseteq ... \subseteq A_5$. The starting point x was chosen such that $x \in A_1$. A violation of Upward Monotonicity would then look like $\Gamma_x(A_i) \nsubseteq \Gamma_x(A_{i+1})$. From these five extended decision problems, we then have 10 separate tests of Upward Monotonicity per subject, or 1070 in total across 107 subjects.

3.3.2 Symmetry

The next property of NC to be tested concerns the "undirectedness" of how products are connected in NC. By the definition of "reachable," it should be clear that if y is reachable from x, x is also reachable from y. This is simply the result of consideration spilling over in either direction of a connection between options, regardless of the origin. This has a clear implication for how consideration sets should compare across the same available set, but given distinct starting points, which is captured in the NC Symmetry property:

B2. Symmetry: If $y \in \Gamma_x(S)$, $\Gamma_x(S) = \Gamma_y(S)$ for all S.

To test the Symmetry Axioms of NC, I repeat the available sets in each of δ_i above, letting γ_i denote one such extended decision problem. These Symmetry extended decision problems use a distinct starting point $y \neq x$, with $y \in A_1$. A test of these Symmetry Axioms would then involve a comparison between consideration sets in δ_i and γ_i. Then, in total, these ten extended decision problems create a maximum of five tests of Symmetry for each subject. However, notice that the Symmetry property only applies if $y \in \Gamma_x(S)$, which may not be born out in the data for a given subject. The actual total number of tests of Symmetry will then be endogenously determined by consideration behavior of subjects.

3.3.3 Path Connectedness

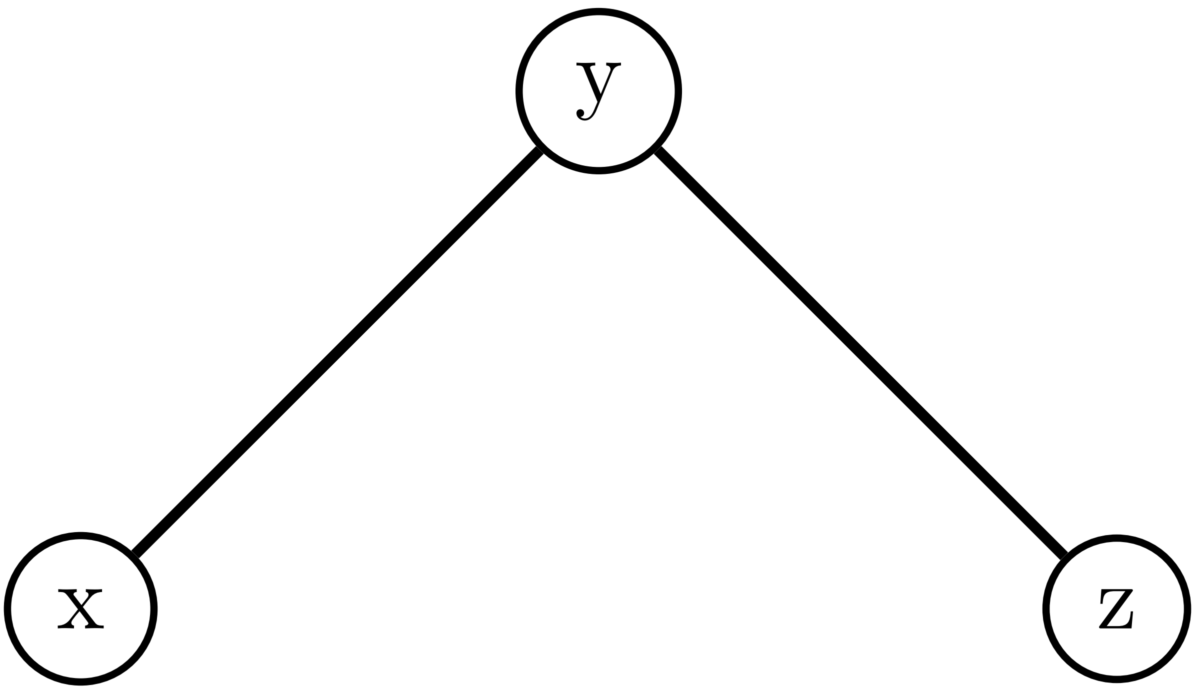
The final property of NC to be tested concerns the impact of an option that uniquely provides a connection between two other options. This property, Path Connectedness, essentially states that the revelation that some option y is required to make z reachable from x should also reveal i) that y is reachable from x in the absence of z and ii) that z is reachable from y in the absence of x. Formally, Masatlioglu and Suleymanov (2017) write this property as follows:

[7] Masatlioglu and Suleymanov (2017) say that y is "reachable" from x if there exists a path from x to y in the network of connected options. I follow in their footsteps and use the same terminology here.

B3. 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 Figure 2, 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. Path Connectedness essentially identifies the fact that this tells us two things. 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.

Figure 2: Example graph with three options



This property utilizes four separate extended decision problems: (x, S), $(x, S \setminus y)$, $(x, S \setminus z)$, and $(y, S \setminus x)$. Further, note that the hypothesis of the property, similar to that of the Symmetry property, is going to be endogenously determined by subject consideration data: it may be the case that we present both (x, S) and $(x, S \setminus y)$, to the subject and that their behavior does not satisfy the hypothesis of this statement. In order to increase the probability that there is a sufficient number of observations where the hypothesis is satisfied, two separate options for y and z in the above are presented to each subject, holding x and S fixed. There are thus four possible tests of Path Connectedness for each subject, constructed using seven extended decision problems. The actual number of tests conducted are a function of the consideration set data.

In total, these three properties of NC lead to the creation of 17 extended decision problems to be used in the laboratory experiment. The remaining 14 (out of 31) were constructed to test axioms on choice data of NC, along with the choice axioms of a related model contained in Suleymanov (2018). The results of these tests are not included, as the focus of the main body of the paper is an exploration of consideration set formation in NC.

4 Results: NC

The results of tests of consistency with NC are presented below. Given data on both choices and consideration sets, one can check for consistency in two separate ways: simultaneous and sequential. Under a simultaneous test, one would test whether the resultant choices were consistent with some NC representation. This is, in general, the approach taken when consideration sets are unobserved and consistency with some decision theoretic model can only be tested using choice data. In this experiment, consideration sets are elicited, so one can take a sequential approach, paying more attention to the consideration set formation process. In a sequential test consistency with NC is separated into two questions. First, in each extended decision problem, does the subject choose a money-maximizing element of the consideration set? Second, is the formation of random consideration set mappings consistent with an NC representation (i.e. do consideration sets satisfy Upward Monotonicity, Symmetry, and Path Connectedness)? Subsection 4.1 presents general experimental results and demographic information. Subsection 4.2 answers the first question on optimality of choice. Finally, Subsection 4.3 reports tests of the NC properties.

4.1 General Results

In all of the below, statistical tests were conducted on aggregate data, pooled across the Baseline and Context displays, except where explicitly mentioned. Tests of differences between the two displays that were omitted in the main text can be found in Appendix C. Upon completion of the experimental task, subjects filled out a brief demographic questionnaire which asked questions on Age, self-reported SAT and ACT scores (if any), self-reported GPA, and Gender. Results are presented in Table 2.

Table 2: Demographic Information

	Age	SAT	ACT	GPA	Female
mean	20.68224	1810.833	30.11429	3.363364	.4485981
sd	1.551616	324.1159	3.946491	.438533	.4996913
min	18	1100	20	2	0
max	27	2360	36	4	1
count	107	84	35	107	107

In order for data on self-reported GPA, SAT, and ACT to be used in subsequent analysis, responses were normalized as in Cohen et al. (1999), Filiz-Ozbay et al. (2016), and Chadd et al. (2018): Let j be the variable under consideration with $j \in$ {GPA, SAT, ACT}, μ_i^j be the value of variable j for subject i, μ_{max}^j be the maximum value of j in the subject population, and μ_{min}^j be the minimum value of j in the subject population. Then let $\hat{\mu}_i^j$, the normalized value of variable j for subject i, be defined as follows:

$$\hat{\mu}_i^j = \frac{\mu_i^j - \mu_{min}^j}{\mu_{max}^j - \mu_{min}^j}$$

such that $\hat{\mu}_i^j$ can be interpreted as the measure of j for subject i, normalized by the distribution of j in the subject population. Some subjects were missing one or more measures for $j \in$ {GPA, SAT, ACT}, since these measures were self-reported (and some subjects could not recall their scores on one or more of these measures). Additionally, some responses were outside the range of feasible scores (for example, an SAT score of 20). All subjects could reliably self-report a feasible GPA from the range of 2 to 4, so $\hat{\mu}_i^{GPA}$ will be used for any subsequent analysis involving cognitive ability. Normalized scores are reported in Table 3 along with an additional measure for Cognitive Score. For some subject i Cognitive Score is taken to be $\hat{\mu}_i^{SAT}$ if the subject reported a feasible SAT score and $\hat{\mu}_i^{ACT}$ if the subjects did not report a feasible SAT score and reported a feasible ACT score. Cognitive Score is lower than $\hat{\mu}_i^{GPA}$ and has higher variance (likely due to more imperfect recall of SAT/ACT scores relative to GPA).

Table 3: Cognitive Scores

$\hat{\mu}_{SAT}$	0.564
	(0.257)
$\hat{\mu}_{ACT}$	0.632
	(0.247)
$\hat{\mu}_{GPA}$	0.682
	(0.219)
Cognitive Score	0.579
	(0.256)
Observations	107

4.2 Choice Optimality

For the purposes of testing “choice optimality,” I say that a subject chose a “highest-valued option” in a given extended decision problem when the option that was last chosen by the subject before the end of the period (i.e. before 75 seconds elapsed) led to 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. (2018), for example. Here, we only say that a "mistake" was made when the subject ended up choosing a lower-valued option than one that was actively considered in the current period. Note: we view an option as having been "considered" if the subject navigated to its information at some point in the period. Since we are testing for choice optimality in the context of NC, which says that a DM will choose the optimal option in the DM's consideration set, it is natural to define "mistakes" as described above.

Subjects chose the highest-valued option in 85.675% of extended decision problems (Wilcox $p < 0.001$ for $H_0 : \mu = 1$). Given that the overall mistake rate is non-trivial, I further investigate the determinants of choice optimality through several logistic regression specifications. In Table 4, the dependent variable is Correct, a binary variable that takes the value 1 if the subject chose the consideration set optimal option in the extended decision problem and 0 otherwise. Context is a binary variable used to indicate whether the observation came from the Context display. Period goes from 1 to 31, indicating the period in which the extended decision problem was completed. CS_N^ϵ is the residual generated from an OLS regression of CS_N onto Period and N. Both the size of the available set (N) and the period in which the extended decision problem is conducted affect the size of the consideration set.[8] 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, the size of the available set. Female and GPA are defined as above (i.e. GPA is normalized according to the POMP procedure described in Subsection 4.1). In both model specifications, marginal effects from a logistic regression are reported, along with robust standard errors clustered at the subject level.

From Table 4, we can see that the prevalence of sub-optimal choice can partly be attributed to learning: each additional period increases the probability that the choice will be consideration set optimal by 0.1 percentage points, resulting in higher rates of sub-optimal choice in earlier periods. Further evidence of this can be seen by looking at the final period only, where 94.33% of observations are consideration set optimal. 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 0.953 percentage points. Somewhat surprisingly, the size of the consideration set itself matters - an additional option considered (holding N and Period fixed) decreases the probability of consideration set optimal choice by about 1.3 percentage points. Note that in neither specification do Context, Female, or GPA matter. This brings us to our first two results on consideration set optimality of choice:

[8] An interested reader can find these results in the Appendix C. 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.

Table 4: Determinants of Optimal Choice

	Model 1	Model 2
Context	-0.0156	-0.0164
	(0.031)	(0.031)
Period	0.00175**	0.00174**
	(0.001)	(0.001)
CS_N^ϵ	-0.0126***	-0.0129***
	(0.003)	(0.003)
N	-0.00953***	-0.00952***
	(0.001)	(0.001)
Female		-0.0359
		(0.027)
$\hat{\mu}_{GPA}$		0.0358
		(0.080)
Observations	3276	3276

Standard errors in parentheses

Marginal effects from logit regression specifications

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Result 1. *Consideration set choice optimality in the aggregate is broadly consistent with NC:*

- *85.675% of choices are consideration set optimal*

Result 2. *There is a non-trivial number of observations where choices are not consideration set optimal in a manner not predicted by NC:*

- *larger consideration sets decrease the likelihood of consideration set optimality*
- *larger available sets decrease the likelihood of consideration set optimality*

4.3 Property Tests

In Subsections 4.3.1 - 4.3.3, we consider subject-level data to test the deterministic consideration set properties of NC.

4.3.1 Upward Monotonicity

The five extended decision problems constructed to test Upward Monotonicity, when repeated twice (in order to test Symmetry - results in Subsection 4.3.2) comprise 20 tests of Upward Monotonicity per subject. In the aggregate, 79.8% of these observations were inconsistent with Upward Monotonicity, as seen in Table 5. Moreover, an analysis of the CDF of the proportion of Upward Monotonicity violations per subject, displayed in Figure 3, 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.

Table 5: Aggregate Test of NC Upward Monotonicity

	Monotonicity Violation
Mean	0.798
Std Error	0.000
N	2140

$p < 0.001$ for aggregate test of $H_0 : \mu = 0$

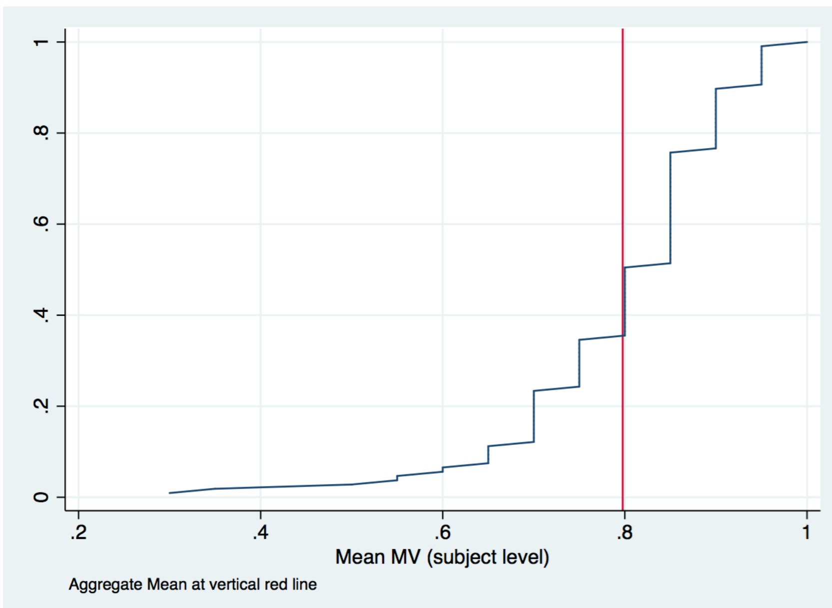


Figure 3: Cumulative of Mean Monotonicity Violations by Subject

Given the prevalence of violations of Upward Monotonicity, we ask what characteristics of extended decision problems and the choice environment affect the likelihood of observing a violation. First, note that Upward Monotonicity does not take into account the "distance" between the sizes of the relevant available sets. At first glace, this appears as if it should not matter. If set $\Gamma_x(T)$ is considered when T is available, then this indicates that option x is connected to elements in T through some combination of paths. When $S \supseteq T$ is available, these same paths are still present, so at least all of the elements in T should again be considered. However, if there is some probability that the DM switches from one path or sub-network to another when encountering a new extended decision problem with the same starting point, then the size of $S \setminus T$ may matter.[9] As $S \setminus T$ gets larger, the number of sub-networks on S relative to T also gets considerably

[9]Note that, in NC, the probability that the DM "switches" consideration sub-networks across identical or nested extended decision problems is implicitly 0. We mention this switching behavior as an empirical possibility in the environments NC is modelling, not as behavior that is consistent with the MS model itself.

larger, leading to an increased likelihood that the DM follows some sub-network that is distinct from the one they follow when T is available. We may thus expect that the likelihood of observing a violation of Upward Montonicity to be increasing in $|S \setminus T|$. We see that this is the case in Figure 4 below. For extended decision problems (x, A_i), for $A_i \subseteq A_j$, we define the Distance between A_i and A_j to be equal to $|A_j| - |A_i|$. In Figure 4, we see 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).

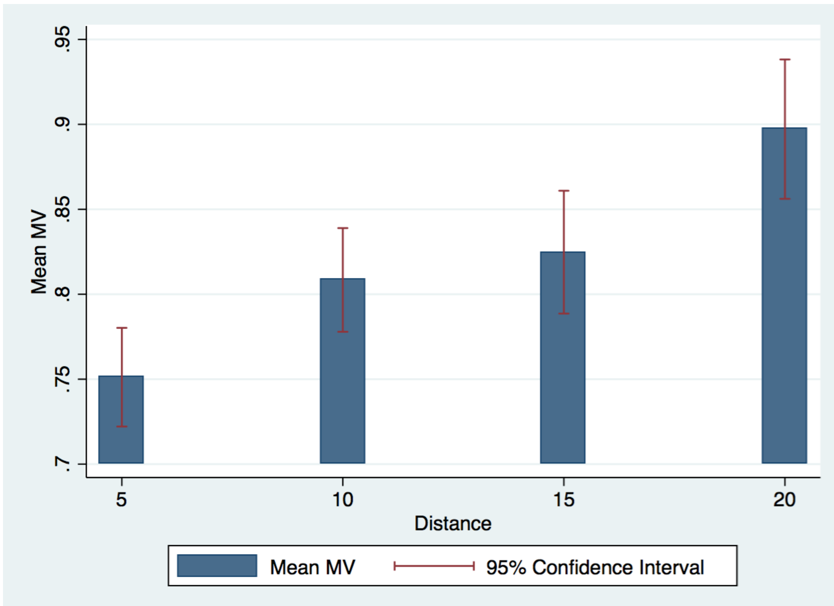


Figure 4: MV by $|A_j| - |A_i|$, $A_i \subseteq A_j$

To further investigate the determinants of Upward Monotonicity violations in our data, Table 6 reports the results of several logistic regressions. In each model, Context is a dummy variable used to indicate whether the observation came from the Context display, and Female, GPA, and Age are as they were defined previously. Reported coefficients are marginal effects from logistic regressions and standard errors are robust and clustered at the subject level. From Table 6, it initially appears that Distance increases the likelihood of Upward Monotonicity violations by approximately 0.93 percentage points, per the reported marginal effect in Model 1. However, the entirety of this effect in the aggregate is driven by the tests of Upward Monotonicity involving $A_1 \subseteq A_2$. These available sets are of size 5 and 10, respectively. Thus, it appears as if Upward Monotonicity violations are ubiquitous regardless of Distance, provided that the available sets involved are sufficiently large.

Table 6: Determinants of Monotonicity Violations

	Model 1	Model 2	Model 3	Model 4
Distance	0.00927***	0.00241	0.00241	0.00241
	(0.00153)	(0.00178)	(0.00178)	(0.00178)
A_1 to A_2		-0.234***	-0.234***	-0.234***
		(0.0259)	(0.0257)	(0.0263)
Context			0.0138	0.0181
			(0.0232)	(0.0223)
Female				0.00377
				(0.0242)
GPA				0.0679*
				(0.0354)
Age				-0.00599
				(0.00586)
Observations	2140	2140	2140	2140

Standard errors in parentheses
Marginal effects from logistic regressions reported
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.3.2 Symmetry

Recall that extended decision problems were constructed such that the five nested available sets, $A_1 \subseteq A_2 ... \subseteq A_5$, were each used in two extended decision problems: (x, A_i) and (y, A_i), $x, y \in A_1$. 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:

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

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

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 we exclude tests of condition 2 unless condition 1 is satisfied only trivially (i.e. $y \notin \Gamma_x(A_i)$). We will thus only include a maximum of five tests of symmetry per subject.

Out of a possible maximum of 535 tests of deterministic symmetry, there were 401 observations where the hypothesis of this axiom was satisfied. In the aggregate, 84.29% of these observations violated symmetry (Wilcox sign rank $p < 0.001$ for $H_0 : \mu = 0$). Table 7 presents aggregate summary statistics. Hypothesis is a dummy variable indicating whether the observation satisfied at least one hypothesis contained in conditions 1 and 2, Symmetric is a dummy variable indicating whether the implication in the relevant condition is satisfied (conditional on the hypothesis being satisfied), and Violation is simply equal to 1 - Symmetric.

Table 7: NC Symmetry Summary Statistics

	hypothesis	symmetric	violation
Mean	.7495327	.1571072	.8428928
SD	.4336877	.3643564	.3643564
N	535	401	401

At the subject level, we also examine Symmetry violation counts and rates. 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). We can further examine the distribution of Symmetry violation rates in Figure 5. Notably, more than half of subjects violated MS Symmetry in 100% of their valid tests.

Table 8: NC Symmetry Subject Level Summary Statistics

	Mean	SD	Min	Max	N
Hypothesis N	3.748	(1.237)	0.000	5.000	107.000
Symmetric N	0.589	(0.672)	0.000	3.000	107.000
Violation N	3.159	(1.175)	0.000	5.000	107.000

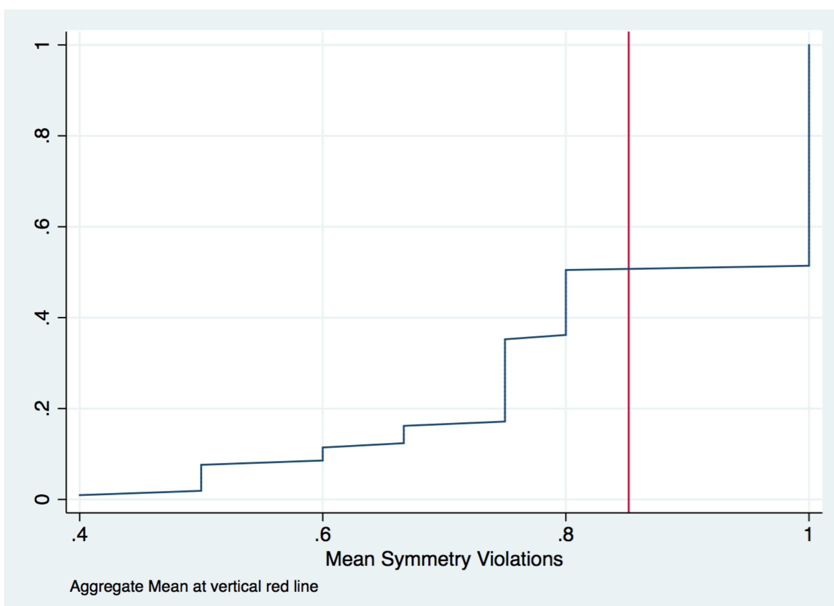


Figure 5: Cumulative Mean Symmetry Violations per Subject

We conjecture that one reason Symmetry may fail in this context is that little information is available regarding an individual option prior to its consideration. The deterministic MS Symmetry axiom requires much of the DM: conditional on arriving at some node in the network, the DM follows the same pattern of search for a given available set. If a given subject then follows different paths of consideration starting at, say, option y in (x, A_i) and (y, A_i), then their consideration set will not satisfy symmetry. In the Baseline environment, no information is available to the subject regarding an individual option prior to its consideration. Thus, if we make some information available to the subject prior to an object being considered, this may increase the likelihood of symmetry consideration paths across extended decision problems. We test this hypothesis by comparing the rate of violations of Symmetry in Table 9. In the Context environment, 81.9% of observations violated symmetry compared to 87.4% in the Baseline, which runs counter to this hypothesis. However, this difference is not statistically significant (Wilcox $p > 0.10$).

Table 9: Symmetry Violations by Context

	Baseline	Context
Mean	0.819	0.874
Std Error	0.026	0.025
N	227	174

Wilcox $p > 0.10$ for $H_0 : \mu_{Baseline} = \mu_{Context}$

While a simple analysis of aggregate Symmetry violations is helpful, it is illuminating to consider subject-level mean violations of symmetry. Recall that we may be including a maximum of five tests of Symmetry for a given subject. Table 8 presents summary statistics for subject-level data on the number of tests per subject (Hypothesis) and violations of MS Symmetry at the subject level. Further, from Figure 5 we can see that nearly 50% of all subjects violated Symmetry in each observation where the hypothesis was satisfied. When we decompose this cumulative distribution function by informational environment, we can see that a larger proportion of subjects have a 100% Symmetry violation rate in the Context environment than in the Baseline (42.11% vs 58.33%; Mann-Whitney $p < 0.10$).

Therefore, on the whole, subjects tend to exhibit behavior more consistent with Symmetry in the Baseline than when Context is provided, though behavior in both is largely inconsistent with Symmetry.

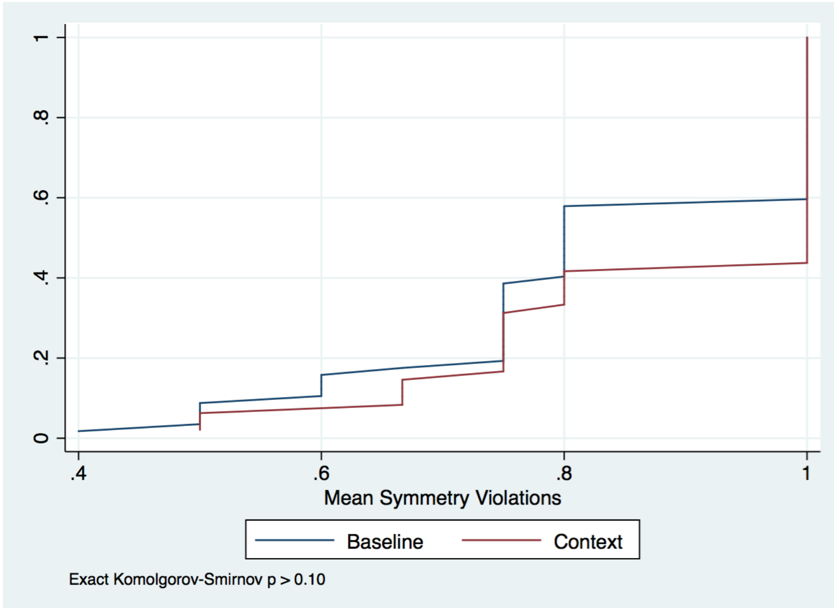


Figure 6: Cumulative of Per Subject Mean Symmetry Violations by Context

4.3.3 Path Connectedness

With 107 subjects and four possible tests of Path Connectedness per subject, we have a maximum of 428 tests. In the aggregate, only about 23% of possible tests were such that the hypothesis of the MS Path Connectedness axiom was satisfied, making for 99 total tests used, across 67 subjects. Of these 99 tests, roughly 45.5% were consistent with with Path Connectedness, as reported in Table 10.

Table 10: Aggregate Test of NC Path Connectedness

	Path Connectedness
Mean	0.455
Std Error	0.050
N	99

Wilcoxon signed-rank $p < 0.001$ for aggregate test of $H_0 : \mu = 1$

Thus, taken together, the experimental results pertaining to consideration set data largely reject the NC deterministic model at the subject level. This leads me to the next result:

Result 3. *Consideration set probabilities are largely inconsistent with the deterministic NC model in the experimental data:*

- *Nearly 80% of observations violate Upward Monotonicity*
- *Approximately 84% of observations violate Symmetry*
- *45.5% of observations violate Path Connectedness*

5 Random Network Choice

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. First, I propose the most general Random Network Choice (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. The RNC model has a feature that leads to an infinite number of representations for a given set of choice data that is consistent with this model. Guided by the notion that eliciting a unique network structure from choice or consideration set data may be desirable in many applications, I present a special case, Pseudo-Markovian Random Network Choice (PM-RNC hereinafter) which does have a unique representation.

I should note that the following section is meant only to provide suitable modelling alternatives to the NC model that may be consistent with the experimental dataset contained herein. While the NC model of Masatlioglu and Suleymanov (2017) is an axiomatic characterization of choice in the deterministic setting, such an axiomatic characterization of RNC and PM-RNC is beyond the scope of the current work. Instead, I focus on necessary conditions of consideration set mappings to be tested against the experimental data.

5.1 Random Network Consideration

Let X be a finite set of alternatives with 2^X as the set of all subsets of X. I consider random networks on X. To that end, let the random variable G_{ij} indicate whether items i and j are connected ($G_{ij} = 1$) or not ($G_{ij} = 0$). Let G be the incidence matrix consisting of these random variables and $\mathcal{G}$ be the collection of all such incidence matrices.

To consider subsets $S \in 2^X$, I must restrict attention to only those nodes that are in S. We denote G^S as the incidence matrix G with attention restricted to S; in other words, G^S is simply G with G_{ij} replaced with 0 for all $\{i, j\} \nsubseteq S$. Let F be an element of $\Delta(\mathcal{G})$, with $f(g|S)$ as the conditional probability that $G = g$ when the set of available options is S. We say that a distribution is **S-Independent** if $f(g|S) = f(g|S') = f(g)$ for all $S, S' \in 2^X$ and for all $g \in \mathcal{G}$. For the remainder of the analysis, I restrict attention to only **S-Independent** distributions, though I mention this generalization as a possible path to pursue in future theoretical explanations of deviations from the RNC model.

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 $(x_0, x_1, ...x_n)$ with $x_0 = i$ and $x_n = j$ where $y_{x_k,x_{k+1}} = 1$ 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*** *if*

1. *t and t' are connected in g for all $t, t' \in T$ with $t \neq t'$*

2. *$\nexists k \notin T$ and $t \in T$ such that k and t are connected under g*

Given an extended problem (S, x) which consists of an alternative set S and starting point $x \in S$, the DM forms a consideration set stochastically. Let $\Gamma_x(T \mid S)$ be a random consideration set mapping that gives the probability that the consideration set is T, given that the starting point is x under alternative set S.

A random network consideration set mapping, $G_x(T \mid S)$ is defined in the following manner. For ease of notation, let $\mathcal{G}_T^S = \{g \in \mathcal{G} \mid g^S \text{ is T-Connected}\}$.

Definition 2. *Given a distribution F on $\Delta(\mathcal{G})$, a* ***random network consideration set mapping*** *is defined as follows:*

$$G_x(T \mid S) = \begin{cases} 1 & \textit{if } \{x\} = T \subseteq S \\ \sum_{g \in \mathcal{G}_T^S} f(g) & \textit{if } \{x\} \subseteq T \subseteq S \\ 0 & \textit{otherwise} \end{cases} \tag{3}$$

The extreme cases, where $\{x\} = T \subseteq S$, where $x \notin T$, or where $T \nsubseteq S$ are trivial. For the non-trivial case, a random network consideration set mapping can be thought of 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 due to some element $k \in X \setminus S$. 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 considered is simply the sum of the probabilities of each network occurring.

5.2 Necessary Properties

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 stay T-Connected for the same set under $S \cup \{a\}$. The new element a may be connected to some $t \in T$ or it may not. 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:

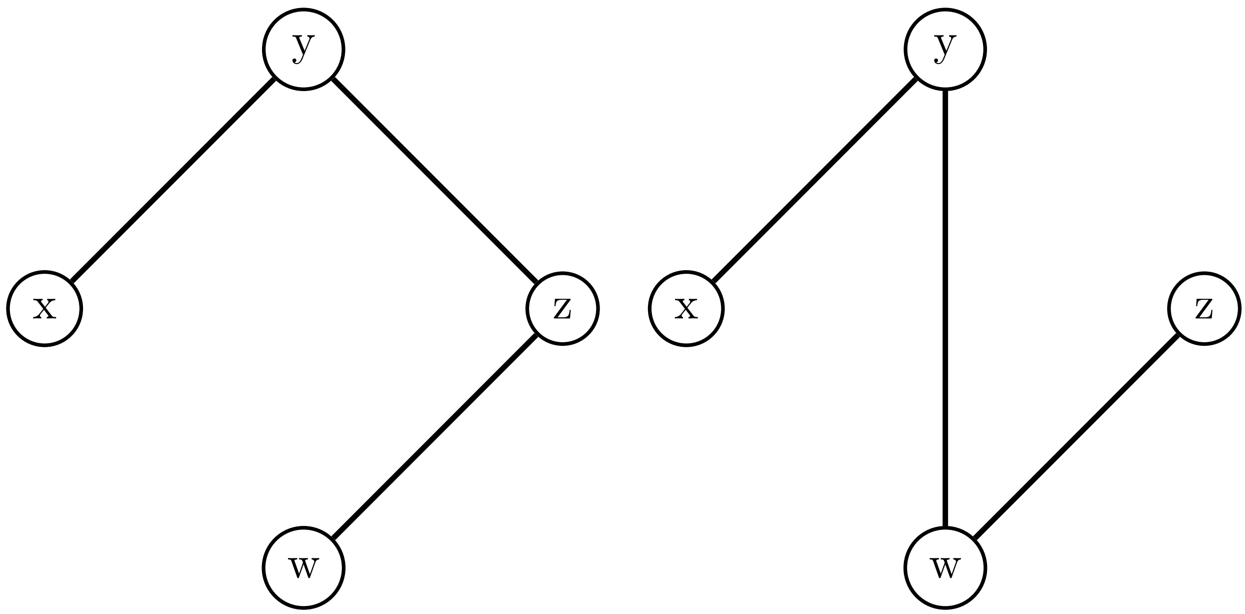
Lemma 1. *For any g such that g^S is T-Connected for some $T \subseteq$, $g^{S'}$ is T'-Connected for some unique T' such that $T \subseteq T'$ and $T' \subseteq S'$ for all $S \subseteq S'$. Equivalently, $\mathcal{G}_T^S \subseteq \bigcup_{T' \subseteq S' : T \subseteq T'} \mathcal{G}_{T'}^{S'}$ for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

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

A1. RNC Upward Monotonicity For each $x \in T \subseteq S$, $\Gamma_x(T \mid S) \leq \sum_{T' \subseteq S' : T \subseteq T'} \Gamma_x(T' \mid S')$

That random network consideration set mappings should satisfy this condition should be obvious. If a network with attention restricted to S leads to set T being considered (i.e. it is part of the sum that makes up $\Gamma_x(T \mid S)$), then by Lemma 1, that same network is T'-Connected for some $T' \subseteq S'$ with $S \subseteq S'$. Then that same network will appear as part of the sum that makes up some (unique) $\Gamma(T' \mid S')$. In short, if a network is included on the left-hand side of A1, it will show up on the right-hand side as well. To see why this expression does not hold with equality, consider the following example network:

Figure 7: An RNC Monotonicity Example



(a) y_1 (b) y_2

Then when $S = \{x, y, z\}$ is available, the set $T = \{x, y, z\}$ is considered with probability $f(g_1)$ when the starting point is x. However, when $S' = \{w, x, y, z\}$ is available, the probability that T' is considered is $f(y_1) + f(y_2)$. Under y_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 the RNC Upward Monotonicity axiom, but not the left hand side when $T = \{x, y, z\}$. Notice that, in this example, had we considered $T = \{x, y\}$, the expression would have held with equality.

This property is clearly a stochastic generalization of the Upward Monotonicity property of Masatlioglu and Suleymanov (2017). When we restrict attention to $\Gamma_x(T \mid S) \in \{0, 1\}$, A1 is equivalent to B1. 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 what the effect of changing the starting point might have on the probability of a given subset being considered. First, it should be clear that $\Gamma_x(T \mid S)$ to $\Gamma_y(T \mid S)$ for some $y \notin T$ is essentially a trivial comparison. If $x \in T$, $y \notin T$ will imply that T cannot be considered from y, and so $\Gamma_x(T \mid S) \geq \Gamma_y(T \mid S) = 0$. But this is not informative, since $\Gamma_x(T \mid S) \geq 0$ by definition. However, when switching from x to y while both are in T, we reveal a fundamental characteristic of random network consideration:

A2. RNC Symmetry: $\Gamma_x(T|S) = \Gamma_y(T|S)$ 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 $t \in T$. Thus, in the non-trivial case in Definition 2, which is satisfied by both x and y (since $\{x, y\} \subseteq T$), we are summing over the same set of networks, resulting in the same consideration probabilities for each $t \in T$ as the starting point.

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 A2, 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 $\Gamma_x(T \mid S) > 0$. This then implies that there exists some network that connects x to z when S is available. 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 $\Gamma_x(T' \mid S \setminus \{y\}) > 0$, 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. This information illuminates direct relationships between x and y and between y and z and it leads us to our final axiom used to characterize random network consideration set mappings:

A3. RNC Path Connectedness If $\exists T \subseteq S$ with $z \in T$ such that $\Gamma_x(T \mid S) > 0$ and $\nexists T' \subseteq S \setminus \{y\}$ such that $z \in T'$ and $\Gamma_x(T' \mid S \setminus \{y\}) > 0$, then the following must hold:

(a) $\exists T'' \subseteq S \setminus \{z\}$ with $y \in T''$, such that $\Gamma_x(T'' \mid S \setminus \{z\}) > 0$

(b) $\exists T''' \subseteq S \setminus \{x\}$ with $z \in T'''$, such that $\Gamma_y(T''' \mid S \setminus \{x\}) > 0$

The hypothesis, as mentioned previously, reveals that y is required to establish a connection between z and x. 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 A1 - A3 are stochastic generalizations of consideration set properties of NC. The following Proposition captures this notion, with the proof contained in the Appendix.

Proposition 1. *If $\Gamma_x(T \mid S)$ is 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 A1 - A3, then $\Gamma_x(S)$ satisfies B1 - B3 where $\Gamma_x(S) = T$ for $\Gamma_x(T \mid S) = 1$.*

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, with the proof contained in the Appendix.

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

5.3 Choice Rule

The DM is also endowed with an antisymmetric and transitive preference relation, $\succ$. Given the consideration set $T \subseteq S$, after the realization of the random network process, the DM chooses the $\succ -$maximal element of T. We thus define a Random Network Choice (or, abusing abbreviations, RNC) as follows:

Definition 3. *A choice rule π is a* ***random network choice (RNC)*** *if there exists a set of networks $\mathcal{G}$ on X, a probability distribution over these networks F, and an antisymmetric, transitive preference relation $\succ$ on X, such that:*

$$\pi_y(x \mid S) = \sum_{\{x,y\} \subseteq T \subseteq S} \mathbb{1}\{x \textit{ is } \succ\textit{-best in } T\} G_y(T \mid S) \tag{4}$$

where G is a random network consideration set mapping.

5.4 Revealed Preference

In general, it may be possible for there to be multiple RNC representations of a given π. Suppressing notation for X and $\mathcal{G}$, we denote a given RNC representation using only the distribution over networks and the preference relation, $(F, \succ)$. Given some choice rule π, we denote the set of possible RNC representations as $(F^\pi, \succ^\pi) = \{(F^1, \succ^1), ..., (F^N, \succ^N)\}$. Following Masatlioglu et al. (2012) and Lleras et al. (2017) and erring on the side of being conservative in assessing revealed preference, we say that "x is revealed preferred to y" (denoted $x \succ y$) if $x \succ^i y$ for all $\succ^i$ such that $(F^i, \succ^i) \in (F^\pi, \succ^\pi)$.

With the possibility of multiple RNC representations for a given π, under what conditions can we guarantee that $x \succ y$ for each representation? It turns out that a very simple condition captures all aspects of revealed preference.

Lemma 2. *For any RNC π, x is revealed preferred to y if $\exists S \subseteq X$ such that:*

$$\pi_y(x \mid S) > 0 \tag{5}$$

Given this method to reconstruct $\succ$ for an RNC π, we then ask whether this condition is sufficient to reveal preferences completely. We first define the following binary relations to assist in exploring this idea:

Definition 4. *Let P be a binary relation X^2 such that xPy if $\exists S \subseteq X$ such that:*

$$\pi_y(x \mid S) > 0 \tag{6}$$

Further, let P_R be the transitive closure of P on X^2.

We utilize this binary relation to obtain the following helpful result:

Lemma 3. *For some RNC π, x is revealed preferred to y if, and only if, xP_Ry.*

5.4.1 Connection to RAM

As mentioned previously, the random consideration set mappings of RNC satisfy a starting-point contingent version of the monotonicity condition laid out in Cattaneo et al. (2017). We call this condition "Starting Point Monotonicity" and it is as follows:

$$\text{Starting-Point Monotonicity: } \Gamma_x(T \mid S) \leq \Gamma_x(T \mid S \setminus \{a\}) \tag{7}$$

for $a \notin S$. It should be clear that, provided Γ_x has an RNC representation, it will satisfy this Starting-Point Monotonicity condition. If an element a is removed from Cattaneo et al. (2017) show, albeit with no starting-point contingent attention, that if some random choice rule has a RAM representation, preferences can be revealed in the following manner. Under RAM, x is revealed preferred to y if, and only if, the following holds:

$$\exists S \text{ such that } \pi_z(x \mid S) > \pi_z(x \mid S \setminus \{y\}) \tag{8}$$

One might surmise that since RNC satisfies Starting-Point Monotonicity, that preferences could also be revealed using condition 8. In this case, we may be missing some revealed preferences by only considering P_R as defined above. Lemma 4 shows that this worry is unfounded: under RNC, if x and y satisfy condition 8, $(x, y) \in P$ as defined above.

Lemma 4. *Let π be an RNC and let x and y be such that there exists some set $S \supseteq \{x, y, z\}$ such that the following holds:*

$$\pi_z(x \mid S) > \pi_z(x \mid S \setminus \{y\}) \tag{9}$$

Then $(x, y) \in P$ and x is revealed preferred to y.

5.5 Pseudo-Markovian RNC

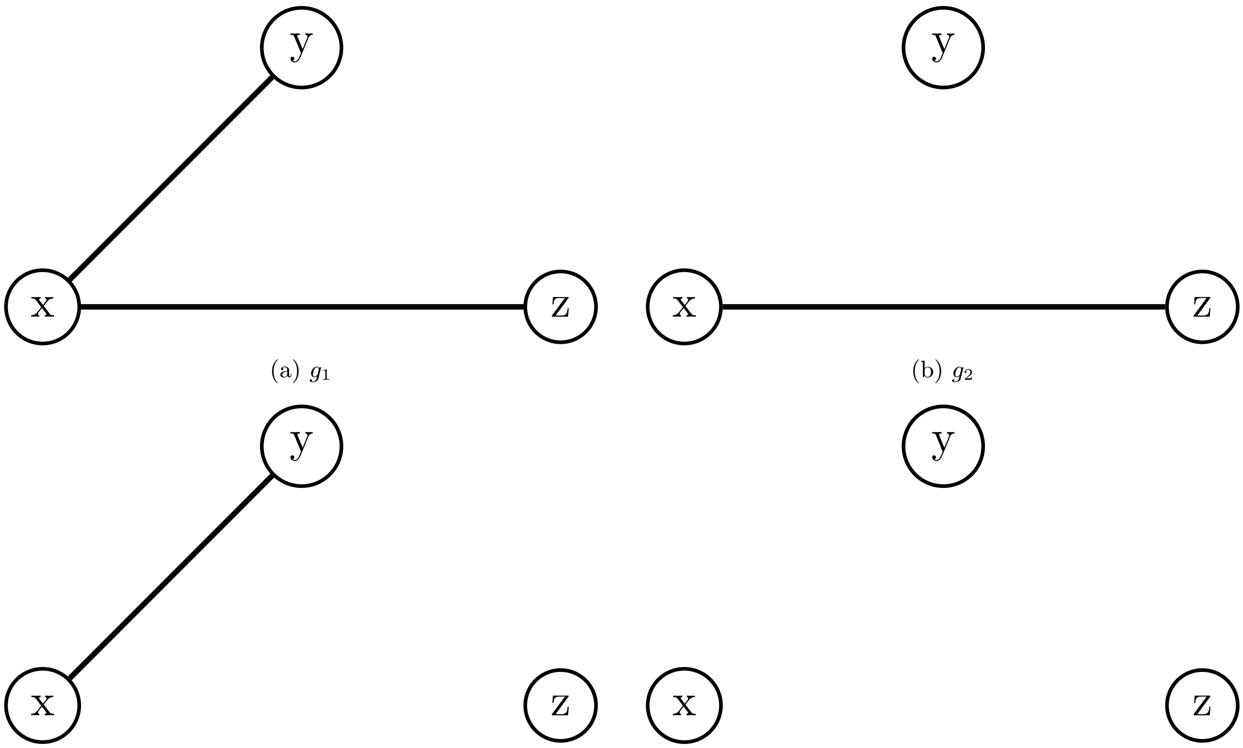
As mentioned in Section 5.4, it is possible that, for a given set of consideration set or choice data consistent with RNC, there may be multiple $(F, \succ)$ representations thereof. Consider the following example. The choice probabilities in Table 11 come from an RNC where all possible networks on three options having the same probability of $\frac{1}{8}$ and $\succ$ is such that $x \succ y \succ z$. Given this choice data, one could work in the opposite direction, uniquely identifying probabilities of a subset of networks that are consistent with an RNC representation. As it turns out, in this example, one cannot uniquely identify probabilities for a subset of these networks given choice probabilities alone.

	Available Set S			
	$\{x, y\}$	$\{x, z\}$	$\{y, z\}$	$\{x, y, z\}$
$\pi_z(x \mid S)$	-	$\frac{1}{2}$	0	$\frac{5}{8}$
$\pi_z(y \mid S)$	-	0	$\frac{1}{2}$	$\frac{1}{8}$
$\pi_z(z \mid S)$	-	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$
$\pi_y(x \mid S)$	$\frac{1}{2}$	-	0	$\frac{5}{8}$
$\pi_y(y \mid S)$	$\frac{1}{2}$	-	1	$\frac{3}{8}$

Table 11: Choice Data for $f(g) = \frac{1}{8}$

To see this starkly, first consider the RNC representation of this choice data that actually generated these choice probabilities (i.e. F such that $f(g) = \frac{1}{8}$ for all g). Now, consider those networks under which z is isolated or only connected to x, as displayed in Figure 8.

Figure 8: Networks under which z is isolated or only connected to x



(a) g_1 (b) g_2 (c) g_3 (d) g_4

It is easy to see that, for any $\epsilon \in [0, \frac{1}{8}]$, a newly constructed pair $(F', \succ)$ will also represent the data in Table 11, where F' is constructed as follows:

$$f'(g) = \begin{cases} \frac{1}{8} & \text{if } g \notin \{g_1, g_2, g_3, g_4\} \\ \frac{1}{8} + \epsilon & \text{if } g \in \{g_1, g_3\} \\ \frac{1}{8} - \epsilon & \text{if } g \in \{g_2, g_4\} \end{cases} \tag{10}$$

with g_1, g_2, g_3, and g_4 as in the figure above. In words, F' is simply F adjusted by ϵ for some of the networks. Thus, even for this very simple case, the most general RNC model can lead to an infinite number of plausible representations of some given π. However, this may not be a desirable property in some empirical environments. In this section, I consider a stochastic special case of RNC, Pseudo-Markovian RNC (PM-RNC, hereafter) that does end up exhibiting a unique representation for some π.

In this special case, we consider only RNC representations of a particular form, where we add restrictions on F in the following manner:

Definition 5. *An RNC $(F, \succ)$ is a **Pseudo-Markovian RNC (PM-RNC)** if there exists a matrix μ with entries $\mu_{ij} \in [0,1]$ where for each network $g \in \mathcal{G}$, the probability of g occurring, $f(g)$ can be written as follows:*

$$f(g) = \prod_{(i,j)\in X^2} [\mathbb{1}\{g_{ij}=1\}\mu_{ij} + \mathbb{1}\{g_{ij}=0\}(1-\mu_{ij})] \tag{11}$$

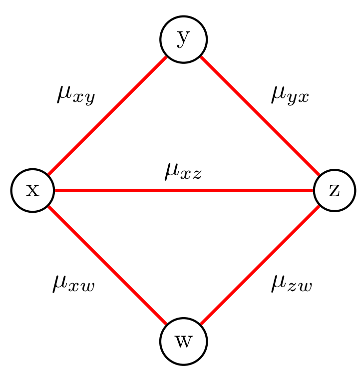
For PM-RNC, we use the notation $M_x(T \mid S)$ to refer to the RNC G_x under this particular formulation. Further, it suffices to represent the entire distribution over networks as a weighted network, with the weight on the connection between options i and j as μ_{ij}. We denote a PM-RNC representation as $(\mu, \succ)$.

A benefit of considering the special case of PM-RNC is that the distribution over networks in each PM-RNC representation is unique:

Lemma 5. *Let π be a PM-RNC with representations $(\mu^\pi, \succ^\pi)$. Then $\mu^i = \bar{\mu}$ for all $(\mu^i, \succ^i)$ representations of π (i.e. μ is unique).*

We can build up intuition for this form of a starting point contingent random consideration set mapping by working through the following example:[10]

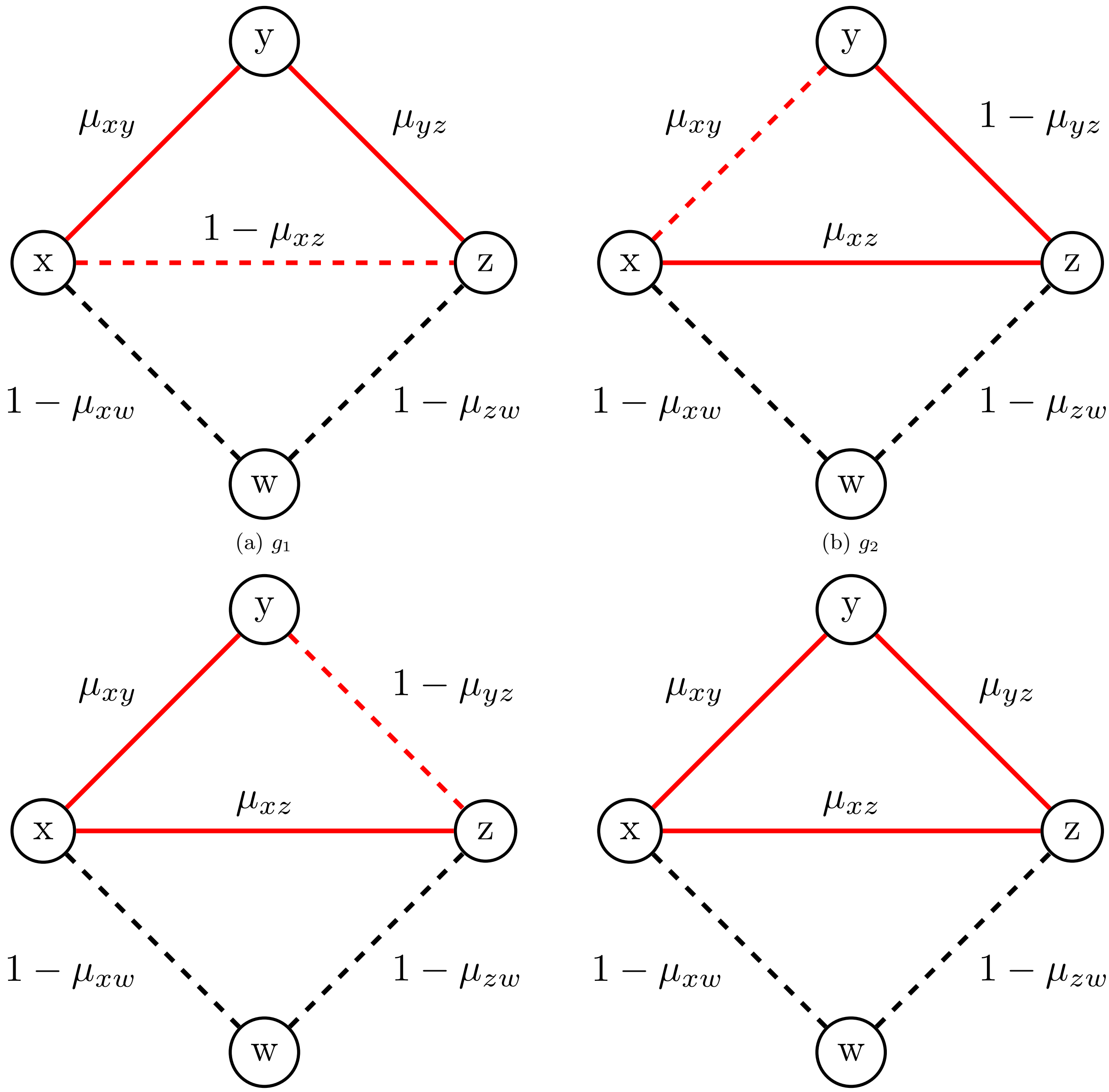
Figure 9: Network g with Four Nodes



Suppose that we are interested in computing $G_x(T \mid S)$ under the network in Figure 9. Consider $G_x(\{x,y,z\} \mid \{w,x,y,z\})$. The DM starts at option x, considering node x with probability 1. Consideration of the set $\{x,y,z\}$ from the set $\{w,x,y,z\}$ can then follow any of the T-Connected networks under S, shown in Figure 10.

[10]Note that, in all figures, $\mu_{ij} = 0$ is represented as the absence of a connection between nodes i and j.

Figure 10: Networks that generate Consideration Set $\{x, y, z\}$



(a) g_1 (b) g_2 (c) g_3 (d) g_4

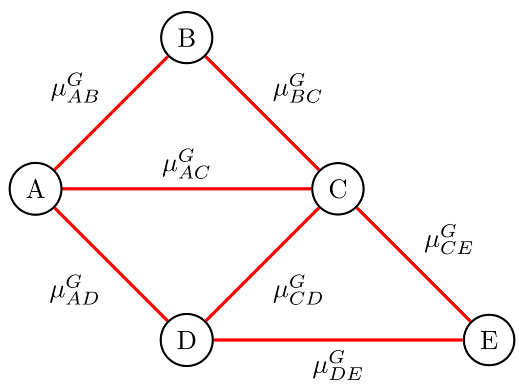
In each case, we consider the probability that consideration spills over from node x to nodes y and z, but not to node w (hence its inclusion in each of the networks in Figure 10). Taking g_1 as an example, we then construct $f(g_1)$ by multiplying together μ_{xy} and μ_{yz} (consideration spills over from node x to y, then from y to z), then $(1 - \mu_{xz})$ (consideration does not spillover from x to z or vice versa), and finally by $(1 - \mu_{xw})$ and $(1 - \mu_{zw})$ (consideration does not spillover from x or z to w).

We can then calculate the probability of the consideration set being $\{x, y, z\}$ as follows:

$$
\begin{aligned}
G_x(\{x,y,z\} \mid \{x,y,z,w\}) = & \quad f(g_1)+f(g_2)+f(g_3)+f(g_4) \\
= & \quad \mu_{xy}\mu_{yz}\cdot(1-\mu_{xz})\cdot(1-\mu_{xw})(1-\mu_{zw}) \\
& +\mu_{xz}\mu_{yz}\cdot(1-\mu_{xy})\cdot(1-\mu_{xw})(1-\mu_{zw}) \\
& +\mu_{xy}\mu_{xz}\cdot(1-\mu_{yz})\cdot(1-\mu_{xw})(1-\mu_{zw}) \\
& +\mu_{xy}\mu_{xz}\mu_{yz}\cdot(1-\mu_{xw})(1-\mu_{zw})
\end{aligned} \tag{12}
$$

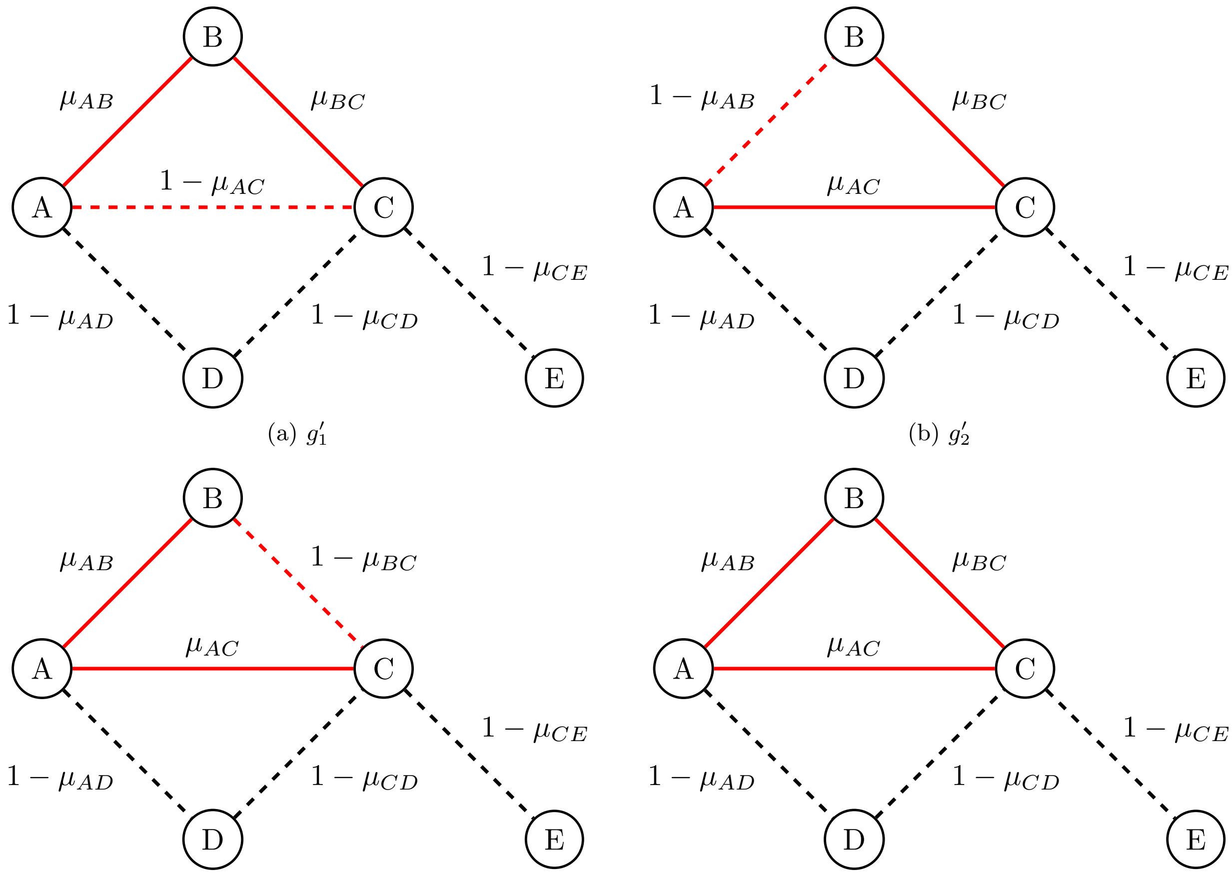
We can thus use the above procedure to calculate $G_x(T \mid \{x,y,z,w\})$ for any T. But what happens when we add an element to the grand set of alternatives? Now consider the following example:

Figure 11: Network g' with Five Nodes



When inspecting the same consideration set $T = \{x,y,z\}$ under $\{x,y,z,w,v\}$, we have to include the possibility that consideration now spills over to node v from node z. In this case, we use the following constrained sub-networks of G':

Figure 12: Networks that generate Consideration Set $\{A, B, C\}$



(a) g_1' (b) g_2' (c) g_3' (d) g_4'

Solid lines indicate connections

Looking just at nodes $\{x, y, z\}$ and connection weights between them, we can see that this is identical to the networks in Figure 10. This is straightforward, as nothing has changed in connection weights within the set $\{x, y, z\}$. Additionally, in each network displayed in Figure 12, we have to account for the possibility that consideration spills over to node w, just as we did for the networks in Figure 10. However, we now have to consider the possibility that node v is considered once z is considered. For this reason, we can write $G_x(\{x, y, z\} \mid \{x, y, z, w, v\})$ as follows:

$$\begin{aligned} G_x(\{x, y, z\} \mid \{x, y, z, w, v\}) &= f(g_1') + f(g_2') + f(g_3') + f(g_4') \\ &= (1 - \mu_{zv})[f(g_1) + f(g_2) + f(g_3) + f(g_4)] \\ &= (1 - \mu_{zv}) G_x(\{x, y, z\} \mid \{x, y, z, w\}) \end{aligned} \tag{13}$$

Then as we inspect the same potential consideration set T in larger and larger sets, the probability of set T being considered aggregates by including the probability that none of the added alternatives are considered, accounting for all links between the added alternatives and the set T. Notice that in the above, we don't consider any connection weights between alternatives w and v in any of the networks used to construct $G_x(\{x, y, z\} \mid \{x, y, z, w, v\})$. This is intuitive: since we aren't considering cases where alternative w is considered, we never need to account for the possibility that consideration spills over from w to v (even though there is a positive probability of this happening under the network g', represented by μ_{wv}).

From the above examples, one can intuit an additional property of PM-RNC consideration set mappings beyond those of the general RNC case. The networks enumerated in Figures 10 and 12 should illuminate the fact that RNC Symmetry holds under PM-RNC. However, a comparison between the examples in Figures 9 and 11 suggests a more strict form of monotonicity than that required in RNC Upward Monotonicity. This characteristic is given below:

A1. PM-RNC Binary Separability $\Gamma_x(T \mid S \cup \{S'\}) = \Gamma_x(T \mid S) \prod_{z \in S' \setminus S} \prod_{t \in T} \Gamma_z(\{z\} \mid \{t, z\})$,

Proposition 3. *If an RNC G_x has a PM-RNC representation, it satisfies PM-RNC Binary Separability.*

6 Results: Random Network Choice

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 subject, 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 $\Gamma_x(T)$ to the sum of probabilities over supersets of T for some presented superset of S, offering a direct test of the RNC Monotonicity property. Table 12 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 on $T \subseteq S$, but they do not occur with positive probability. Then RNC Monotonicity will, by default, be satisfied trivially. While these observations are technically consistent with RNC Monotonicity, they are excluded in the column labeled "NT", for "Non-Trivial" in Table 12. In the aggregate, 12.5% 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 considerably lower under RNC than under NC at 39.7% (Wilcoxon signed-rank $p < 0.01$ for $H_0 : \mu = 0.8$).

Table 12: Aggregate Test of RNC Monotonicity

	All	NT
Mean	0.125	0.397
Std Error	0.000	0.000
N	9974	3132

Wilcoxon signed-rank $p < 0.01$ for aggregate test of $H_0 : \mu = 0$ for both All and NT
Mann-Whitney $p < 0.01$ for $H_0 : \mu_{All} = \mu_{NT}$
NT results exclude observations where $\Gamma_x(T \mid S) = 0$

Considering the presence of RNC Monotonicity violations even when "trivial" observations are included, I further investigate the determinants of these violations. For a given (x, S) and T pairing, one can imagine several measures as generalizations of the "Distance" measure used in Section 4.3.1. First, conditional on S, larger consideration sets T leave less room for supersets to be included in the right-hand side of the RNC Monotonicity expression. Larger sets T more closely approach full consideration of the set S, leaving less room for non-trivial observations of supersets of T under $S' \supseteq S$. Thus, we may expect that violations are more likely to occur as $S \setminus T$ increases in size across observations. Second, when comparing to some set $S' \supseteq S$, the size of $S' \setminus T$ relative to T may have a similar effect. These factors are considered in Table 13. Additional options in $S \setminus T$ actually increase the likelihood of a violation occurring by 0.299 percentage points each, though this effect is only marginally significant. Consistent with the hypothesis presented above, each additional option in $S' \setminus T$ decreases the likelihood of a violation occurring by 2.57 percentage points each. This result is further confirmed in Figure 13, where average RNC monotonicity violations are plotted by quartile of $|S' \setminus T|$. This leads to an interesting comparative result relative to NC. Recall that in Section 4.3.1, it is shown that monotonicity violations were ubiquitous once the distance between S and S' became sufficiently large. Here, the opposite is true: holding the size of T constant, as options are added to S', RNC Monotonicity violations become less common.

While none of these results on the determinants of RNC Monotonicity Violations are implied by the RNC model directly - clearly the RNC model implies no violations of RNC Monotonicity at all - I view the analysis above as reasonable starting points for further generalizations of network consideration in the stochastic case. It is possible that generalizing RNC further to include adaptive consideration behavior, directed random network structures, and/or other features may directly imply the results above.

Table 13: Determinants of RNC Monotonicity Violations

	Model 1
$\|S \setminus T\|$	0.00299*
	(0.00175)
$\|S' \setminus T\|$	-0.0257***
	(0.00185)
Observations	9974

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

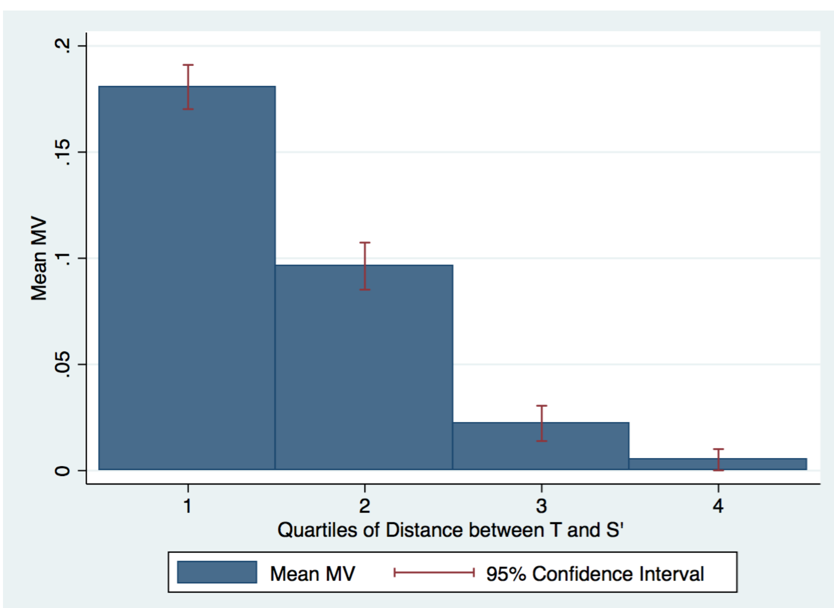


Figure 13: RNC Monotonicity Violations by $|S' \setminus T|$ Quartile

6.2 RNC Symmetry

While the subject-level data reveals a number of violations of the deterministic Symmetry property in NC, it may be possible that aggregate consideration set probabilities are consistent with RNC. This 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 14. 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 15. There are thus broad, early indications that consideration set formation is consistent with RNC Symmetry.

While aggregate results support RNC Symmetry according to these conditional logistic regression models, it is possible that additional information provided to the subject in the Context environment might have an effect on RNC Symmetry, especially in light of the slight, though statistically insignificant, difference

Table 14: RNC Symmetry: Aggregate

	All
Choose	
Starting Point	-0.0862
	(0.0848)
Observations	38082

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

Table 15: 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	13448	11088	7670

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

in symmetry violations at the individual level between the Baseline and Context environments. The same regression specification used in Table 15 and is conducted separately for each environment in Tables 16 and 17. This reveals a significant, though mixed, effect of the starting point in certain treatment - N combinations.

Table 16: RNC Symmetry Regressions: Baseline

	N = 5	N = 10	N = 15	N = 20	N = 25
Choose					
Starting Point	-0.319	-0.358	-0.407*	-0.437	-0.443
	(0.232)	(0.229)	(0.236)	(0.274)	(0.302)
Observations	400	3000	8036	6006	4602

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

Table 17: RNC Symmetry Regressions: Context

	N = 5	N = 10	N = 15	N = 20	N = 25
Choose					
Starting Point	-0.178	0.527*	0.288	0.614*	1.147***
	(0.244)	(0.291)	(0.312)	(0.345)	(0.434)
Observations	376	2100	5412	5082	3068

Standard errors in parentheses

Odds ratios from conditional logit regression specifications

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Each consideration set observed in the data individually can be examined directly by looking at $\Gamma_x(T \mid S)$ 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.[11] 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 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.[12] Of these 69 tests, only 5 resulted in a statistically significant rejection of $H_0 : \Gamma_x(T \mid S) = \Gamma_y(T \mid S)$, only 7.23%. We therefore find robust support for RNC Symmetry regardless of the test method (conditional logit vs. rank-sum at the consideration set - available set pair level).

6.3 RNC Path Connectedness

Finally, RNC Path Connectedness is tested by systematically considering subsets of the experimental data according to their satisfication of hypotheses of RNC Path Connectedness for each potential case. Recall that, in the construction of the extended decision problems used to test Path Connectedness in NC, four different cases resulted from varying the option used in each extended decision problem. Table 18 presents the results of aggregate tests of RNC Path Connectedness separately for each case.

Recall that under the RNC Path Connectedness, the hypothesis of this property would be endogenously determined by consideration set data in the experiment:

[11] Consideration set - available set pair has 214 observations, 107 for each starting point. If the consideration set only occurs once, the sign-rank test will result in $p > 0.10$.

[12] 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.

$$\exists T \subseteq S \text{ with } z \in T \text{ such that } \Gamma_x(T \mid S) > 0 \text{ and } \nexists T' \subseteq S \setminus \{y\} \text{ such that } z \in T' \text{ and } \Gamma_x(T' \mid S \setminus \{y\}) > 0,$$

In order to structure tests of RNC Path Connectedness, I then consider, for each case, the largest number of observations such that this hypothesis is satisfied when aggregating over subjects to construct each $\Gamma_x(T \mid S)$. Then N in Table 18 can be interpreted as the number of subjects (out of 107) who satisfy the particular hypothesis of the case under consideration. Cases 2 and 4 are clearly less stringent tests of RNC Path Connectedness; all 107 subjects satisfy the hypothesis of these cases.

The implication portion of the RNC Path Connectedness property is framed around "the existence" of some set that includes z (y) and is considered with positive probability. This is equivalent to a positive frequency of $z(y)$ being observed for the observations considered in each case. Then for each case, the fact that $Prob(Y)$ and $Prob(Z)$ are positive in Table 18 indicates that the dataset, as a whole, is consistent with this property. This should not be a surprise, even considering the fact that there are non-trivial violations of other properties of RNC consideration sets documented in this section. RNC Path Connectedness, like Path Connectedness, is a relatively weak requirement to impose on consideration set probabilities.

Table 18: RNC Path Connectedness

	Case 1		Case 2		Case 3		Case 4	
	Prob(Y)	Prob(Z)	Prob(Y)	Prob(Z)	Prob(Y)	Prob(Z)	Prob(Y)	Prob(Z)
Mean	0.327	0.327	0.925	0.561	0.227	0.091	0.879	0.178
Std Error	0.068	0.068	0.026	0.048	0.045	0.031	0.032	0.037
N	49	49	107	107	88	88	107	107

Result 4. *Aggregate consideration set frequencies are largely consistent with RNC:*

- *Only 12.5% of all observations violate RNC 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*

6.4 PM-RNC Binary Separability

In addition to RNC Monotonicity, RNC Symmetry, and RNC Path Connectedness, consideration set data consistent with PM-RNC necessarily must satisfy the additional Binary Separability property. Results of tests thereof using this experimental data are presented in this section.

Notice that Binary Separability will necessarily result in consideration set probabilities such that $\Gamma_x(T \mid S \cup S')$ is weakly less than $\Gamma_x(T \mid S)$, since PM-RNC simply takes the latter and multiplies it by the product of a number of probabilities between 0 and 1, inclusive. A clear violation of Binary Separability would thus consist of an observation of $\Gamma_x(T \mid S \cup S') > \Gamma_x(T \mid S)$. Aggregate tests of violations of this type, which I term "First-Order Binary Separability Violations," are presented in Table 19.

Table 19: Aggregate Test of First Order PM-RNC Binary Separability Violations

	All	NT
Mean	0.122	0.282
Std Error	0.003	0.007
N	9974	4308

Wilcoxon signed-rank $p < 0.01$ for aggregate test of $H_0 : \mu = 0$ in each case
NT: $\Gamma_x(T \mid S) = \Gamma_x(T \mid S') = 0$ excluded

In the aggregate, 12.2% of observations constitute first-order Binary Separability violations, with the proportion of such violations jumping to 28.2% when we isolate attention to only "Non-Trivial" observations as defined above.

In order to provide a more finely-tuned test of Binary Separability, the expression provided in the PM-RNC Binary Separability property is constructed for each observation that does not constitute such a first-order Binary Separability Violation. Recall this expression:

$$\Gamma_x(T \mid S \cup S') = \Gamma_x(T \mid S) \prod_{z \in S' \setminus S} \prod_{t \in T} \Gamma_z(\{z\} \mid \{t, z\})$$

In the above, the construction of $\Gamma_x(T \mid S \cup S')$ and $\Gamma_x(T \mid S)$ is straightforward and conducted as in previous analyses in this section. However, the construction of product on the right-hand side, $\prod_{z \in S' \setminus S} \prod_{t \in T} \Gamma_z(\{z\} \mid \{t, z\})$, is not so clear, given that subjects are never presented with an extended decision problem of the form $(z, \{t, z\})$ for any pair of options. In the PM-RNC model, $\Gamma_z(\{z\} \mid \{t, z\})$ is clearly equal to $1 - \mu_{zt}$, or the probability that consideration does not spill over from z to t in the binary comparison. Thus, in order to construct this nested product, I utilize an estimated $1 - \mu_{zt}$ as a proxy for $\Gamma_z(\{z\} \mid \{t, z\})$. For each pair of options (i, j) presented to subjects in the experiment, an observed μ_{ij} is estimated by calculating the frequency with which subjects navigate from i to j, or from j to i, conditional on a link being provided in the experimental interface. Note that this frequency is across all observed links between i and j, regardless of the direction followed, so that the resulting estimated weighted network of μ_{ij} is undirected. Only those estimated μ_{ij} which were statistically significantly greater than 0 at the $\alpha = 0.10$ level were taken as non-zero frequencies.[13]

Mean direct Binary Separability violations are reported in Table 20. Results in Table 20 are presented only for those observations that did not constitute a first-order Binary Separability violation. Across all

[13]Each μ_{ij} was tested using a Wilcoxon signed-rank test.

Table 20: Aggregate Test of Binary Separability Violations

	All	NT
Mean	0.349	0.989
Std Error	0.005	0.002
N	8759	3093

Wilcoxon signed-rank $p < 0.01$ for aggregate test of $H_0 : \mu = 0$ in each case
NT: $\Gamma_x(T \mid S) = \Gamma_x(T \mid S') = 0$ excluded

such observations, 34.9% violate Binary Separability directly along with a staggering 98.9% of Non-Trivial observations (as defined above). The latter result is not particularly surprising, since a "violation" as defined in this section does not allow for additional noise in consideration probabilities beyond that directly implied by the strong form of PM-RNC: unless the realized left-hand side of the Binary Separability expression was exactly equal to the estimated right-hand side, the observation was coded as a Binary Separability violation. A more thorough investigation of Binary Separability would thus necessitate analyzing the size of errors in estimation.

Table 21 reports the summary statistics of two measures of errors in the estimation of these Binary Separability expressions, compared to the distribution of positive $\Gamma_x(T \mid S \cup S')$ for reference. In an abuse of notation for the sake of brevity, in Table 21 and in this discussion, let Γ refer to $\Gamma_x(T \mid S \cup S')$ and $\hat{\Gamma}$ refer to the estimate of $\Gamma_x(T \mid S) \prod_{z \in S' \setminus S} \prod_{t \in T} \Gamma_z(\{z\} \mid \{t, z\})$. Then $\frac{\Gamma - \hat{\Gamma}}{\Gamma}$ gives the normalized difference between the two expressions, conditional on positive Γ, and $\hat{\Gamma} \mid \Gamma = 0$ gives the estimated right-hand side expression, conditional on Γ being equal to zero. Both of these measures of Binary Separability errors are presented only for those observations where a violation is observed.

Table 21: Binary Separability Error Size Summary Statistics

	Γ	$\frac{\Gamma - \hat{\Gamma}}{\Gamma}$	$\hat{\Gamma} \mid \Gamma = 0$
Mean	.011	.031	.005
SD	.008	.759	.004
N	1427	178	2881
p1	.009	-3.591	.000
p25	.009	-.143	.002
p50	.009	.253	.005
p75	.009	.471	.007
p99	.047	.938	.016

Notes:
Γ s.t. $\Gamma_x(T \mid S \cup S') > 0$

First compare the first two columns for Γ and $\frac{\Gamma - \hat{\Gamma}}{\Gamma}$. Note that the interquartile range deflation rate of Γ (given by $\frac{\Gamma - \hat{\Gamma}}{\Gamma}$ for p50) is -0.143 to 0.471 implying that there is considerable spread in the PM-RNC Binary Separability estimate of Γ. This implies that PM-RNC Binary Separability is likely too strong an assumption for the given experimental data. However, the distribution of $\hat{\Gamma} \mid \Gamma = 0$ appears to closely resemble that of Γ suggesting that $\hat{\Gamma}$ may yet still present a reasonable estimate of Γ for out-of-sample consideration sets; it's possible that because the experimental data includes a considerable number of observations of $\Gamma = 0$, that PM-RNC Binary Separability is failed because of data limitations. There are thus only mixed, and generally negative, results concerning the fit of PM-RNC Binary Separability to the experimental data set. I view further exploration of PM-RNC Binary Separability using a large field dataset to be a worthwhile avenue for future research on this topic.

7 Discussion

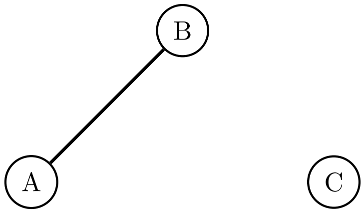
The NC and RNC models can be applied to many settings, but in this section I focus on one particular application of note: empirical studies of the effects of advertising. As mentioned previously, Shapiro (2018),Sahni (2016), and Lewis and Nguyen (2015) report evidence of so-called "positive spillovers in advertising," where attention spills over from an advertised good to similar competitor goods. A common refrain in studies such as these is that the positive externalities resulting from these attention spillovers lead to an under-allocation of advertising in the competitive equilibrium. Indeed, Shapiro (2018) presents a supply-side model and accompanying estimates which support this claim.

However, an implicit assumption in the models commonly assumed in the above studies is that these positive attention spillovers affect competing goods at the category level. When this framework is rooted in an RNC model setup, it can be easily seen that a more general treatment of these spillovers as coming from a generalized random network structure will result in more ambiguous welfare implications.

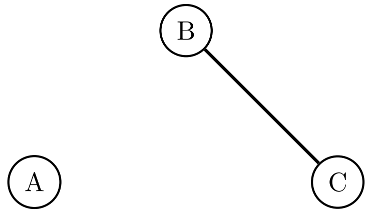
7.1 A Simple Advertising Model

Consider a simple RNC Advertising Game with firms A, B, and C, each producing a single good (which I will also denote A, B, and C, respectively), and each choosing whether advertise or not, such that the strategy for each firm is $\sigma_i \in \{0, 1\}$. Advertising ($\sigma_i = 1$) involves a fixed cost of $c > 0$, which is identical across firms. There is a continuum of DMs of mass 1, each of whom exhibits limited stochastic attention according to the RNC model. The distribution over possible networks on $X = \{A, B, C\}$ is given in Figure 14, with a restriction on α such that $\alpha \in [0, \frac{1}{2}]$.

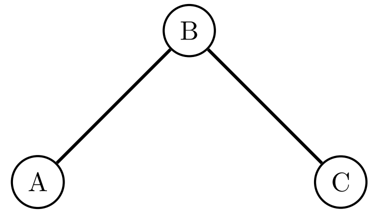
Figure 14: Example Distribution on $X = \{A, B, C\}$



(a) $f(g_1) = \alpha$



(b) $f(g_2) = \alpha$



(c) $f(g_3) = 1 - 2\alpha$

The DM's choice of any firm leads to the same level of utility, so I assume that a DM who considers more than one firm is equally likely to choose any of the firms they consider.[14] Further, since RNC is starting-point dependent, I assume that DMs set their starting points based on the firms that are advertising. If more than one firm advertises, the mass of DMs is divided equally among all those firms who advertise. The sequence of the game then works as follows:

1. Firms simultaneously choose whether to advertise or not
2. DMs observe who advertises and sets a firm as their starting point with equal probability among all those who advertise
   (a) If no firm advertises, DMs consider no firm and choose nothing (resulting in 0 profit for all firms)
3. DMs form stochastic consideration set mappings according to their starting points and the distribution over networks on the set of firms
4. DMs choose each firm that they consider with equal probability
5. Firms realize profit, which is simply the mass of DMs who choose their firm minus the fixed cost of advertising (if the firm chose to advertise)

For a given network g, let $N_i(g)$ represent the set of firms connected to firm i under g by some path (i.e. firm i's "neighbors"). Given a strategy profile σ, let the set of firms who advertise under σ be equal to σ_a. Then for a given strategy profile and distribution over networks on $\{A, B, C\}$, firm i's expected profit is equal to the following:

$$\pi_i = \sum_{j=1}^{3} f(g_j) \frac{|\sigma_a \cap N_i(g_j)|}{|\sigma_a| \cdot |N_i(g_j)|} - c \cdot \sigma_i \tag{14}$$

[14] Note that this random choice procedure is indeed a departure from the choice procedure used in RNC. This is done for tractability in this particular application and is of no particular consequence to the main point under discussion.

Given this setup, it is straightforward to show that there exists a non-empty subset of the $\alpha - c$ parameter space that lead to i) $\sigma^* = (1, 0, 1)$ as a Nash Equilibrium and ii) $\sigma' = (0, 1, 0)$ not as a Nash Equilibrium, where profit is strictly higher under σ'. This is captured in the following proposition, with the proof included in the Appendix:

Proposition 4. *In the RNC Advertising Game with parameters α, β, and c, such that (α, β, c) is in the unit cube, there exist non-empty subsets of the parameter space where:*

1. $\sigma^* = (1, 0, 1)$ *is supported as a Nash Equilibrium*

2. $\sigma' = (0, 1, 0)$ *is not supported as a Nash Equilibrium*

3. *Aggregate profit under σ' is higher than under σ^**

This proposition then tells us that the welfare properties of advertising in such an environment will depend on the structure of distribution in the RNC limited consideration on the part of the consumers. In particular, the "category" spillover approach implicitly assumed in previous work in the marketing body of literature is nested in the RNC Advertising Game where $\alpha = 0$. When this is true, advertising will never be undertaken by multiple firms simultaneously in equilibrium. Indeed, for sufficiently high costs ($c > \frac{1}{3}$), no firm will advertise. This is a stark example of the free-riding effect documented in Shapiro (2018), but is only implied by a narrow interpretation of attention spillovers as spillovers occurring equally across an entire category of goods. In the more general case described here, positive attention spillovers actually result in an over-allocation of advertising relative to an alternative allocation with the same levels of revenue.

8 Conclusion

In this work, I present a the results of an experiment designed to test for the validity of the deterministic Network Choice (NC) model of Masatlioglu and Suleymanov (2017). Overall inconsistency of the data with NC led to a proposed stochastic generalization thereof in RNC and PM-RNC. These models have applications in the realms of choice architecture, web platform search optimization, and advertising. In the latter case, I view a significant contribution of this work as illuminating a possible model for positive spillovers in advertising behavior, which highlights the need to examine product network structures empirically, especially for the purposes of welfare analysis.

It should be noted that RNC is only one version of a stochastic generalization of NC. Indeed, Cattaneo et al. (2017) present another version in which the network is deterministic and starting points are stochastically determined. While the experimental data presented herein was predominantly consistent with RNC, there remain non-trivial violations thereof to be investigated in either further generalizations of RNC or other attempts at modelling stochastic network consideration. This is likely a fruitful avenue for future research.

Additionally, the mixed results of the PM-RNC model, combined with the RNC results, suggest that there is a nested model of network consideration with more structure than RNC (which potentially leads to a unique representation), but less than that of PM-RNC. Future theoretical research can likely shed light on the structure and validity of models between the two.

References

A. Caplin and M. Dean. Search, choice, and revealed preference. Theoretical Economics, 6(1):19–48, 2011.

A. Caplin and M. Dean. Revealed preference, rational inattention, and costly information acquisition. American Economic Review, 105(7):2183–2203, 2015.

A. Caplin, M. Dean, and D. Martin. Search and satisficing. American Economic Review, 101(7):2899–2922, 2011.

M. D. Cattaneo, X. Ma, Y. Masatlioglu, and E. Suleymanov. A random attention model. arXiv preprint arXiv:1712.03448, 2017.

I. Chadd, E. Filiz-Ozbay, and E. Y. Ozbay. The relevance of irrelevant information. 2018.

G. Charness, F. Feri, M. A. Meléndez-Jiménez, and M. Sutter. Experimental games on networks: Underpinnings of behavior and equilibrium selection. Econometrica, 82(5):1615–1670, 2014.

S. Choi, E. Gallo, and S. Kariv. Networks in the laboratory. 2015.

P. Cohen, J. Cohen, L. S. Aiken, and S. G. West. The problem of units and the circumstance for POMP. Multivariate Behavioral Research, 34(3):315–346, 1999.

M. Dean and N. Neligh. Experimental tests of rational inattention. Technical report, 2017.

M. Dean, Ö. Kıbrıs, and Y. Masatlioglu. Limited attention and status quo bias. Journal of Economic Theory, 169:93–127, 2017.

E. Filiz-Ozbay, J. C. Ham, J. H. Kagel, and E. Y. Ozbay. The role of cognitive ability and personality traits for men and women in gift exchange outcomes. Experimental Economics, https://doi.org/10.1007/s10683-016-9503-2, 2016.

O. Frank and D. Strauss. Markov graphs. Journal of the american Statistical association, 81(395):832–842, 1986.

X. Gabaix, D. Laibson, G. Moloche, and S. Weinberg. Costly information acquisition: Experimental analysis of a boundedly rational model. American Economic Review, 96(4):1043–1068, 2006.

S. Geng. Decision time, consideration time, and status quo bias. Economic Inquiry, 54(1):433–449, 2016.

R. Lewis and D. Nguyen. Display advertisings competitive spillovers to consumer search. Quantitative Marketing and Economics, 13(2):93–115, 2015.

J. S. Lleras, Y. Masatlioglu, D. Nakajima, and E. Y. Ozbay. When more is less: Limited consideration. Journal of Economic Theory, 170:70–85, 2017.

R. D. Luce. Individual choice behavior. 1959.

P. Manzini and M. Mariotti. Stochastic choice and consideration sets. Econometrica, 82(3):1153–1176, 2014.

Y. Masatlioglu and D. Nakajima. Choice by iterative search. Theoretical Economics, 8(3):701–728, 2013.

Y. Masatlioglu and E. A. Ok. A canonical model of choice with initial endowments. Review of Economic Studies, 81(2):851–883, 2013.

Y. Masatlioglu and E. Suleymanov. Decision making within a product network. Technical report, Working Paper, University of Michigan, 2017.

Y. Masatlioglu, E. A. Ok, et al. Rational choice with status quo bias. Journal of Economic Theory, 121(1): 1–29, 2005.

Y. Masatlioglu, D. Nakajima, and E. Y. Ozbay. Revealed attention. American Economic Review, 102(5): 2183–2205, 2012.

N. Neligh. Vying for dominance: An experiment in dynamic network formation. Technical report, Working Paper, 2017.

J. L. Orquin and S. M. Loose. Attention and choice: A review on eye movements in decision making. Acta psychologica, 144(1):190–206, 2013.

E. Reutskaja, R. Nagel, C. F. Camerer, and A. Rangel. Search dynamics in consumer choice under time pressure: An eye-tracking study. American Economic Review, 101(2):900–926, 2011.

N. S. Sahni. Advertising spillovers: Evidence from online field experiments and implications for returns on advertising. Journal of Marketing Research, 53(4):459–478, 2016.

B. T. Shapiro. Positive spillovers and free riding in advertising of prescription pharmaceuticals: The case of antidepressants. Journal of Political Economy, 126(1):381–437, 2018.

C. A. Sims. Implications of rational inattention. Journal of Monetary Economics, 50(3):665–690, 2003.

C. A. Sims. Rational inattention: Beyond the linear-quadratic case. American Economic Review, 96(2): 158–163, 2006.

E. Suleymanov. Stochastic attention and search. 2018.

A Proofs for Section 5 (Random Network Choice)

Lemma 1. *For any g such that g^S is T-Connected for some $T \subseteq$, $g^{S'}$ 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 $g^S = g - \sum_{\{i,j\} \subseteq X \setminus S} y_{ij}$. 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 $g^{S'} = g^S + \sum_{i \in S'} \sum_{j \in S' \setminus S} y_{ij}$. 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, consider $T'' \subseteq S'$, but $T'' \neq T$. Note that $T \subset T''$, by construction, so either i) $\exists k \in T''$ such that $k \notin T'$ 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 k is connected to some $t \in T$ (by $T \subseteq T''$) and $T' \cup \{k\} \supseteq T'$, 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. □

Proposition 1. *If $\Gamma_x(T \mid S)$ is 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 A1 - A3, then $\Gamma_x(S)$ satisfies B1 - B3 where $\Gamma_x(S) = T$ for $\Gamma_x(T \mid S) = 1$.*

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 A1 - A3. Let $\Gamma_x(S) = T$ for $\Gamma_x(T \mid S) = 1$

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

- **B2.** 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 $\Gamma_x(T \mid S) = 1 = \Gamma_y(T' \mid S)$. Since $y \in \Gamma_x(S)$, $y \in T$ and $T = T'$, since by A2 $y \in T$ implies $\Gamma_x(T \mid S) = \Gamma_y(T \mid S) = 1$. Then $\Gamma_x(S) = \Gamma_y(S)$ and $\Gamma_x(S)$ satisfies **??**.

- **B3.** Let $z \in \Gamma_x(S)$ and $z \notin \Gamma_x(S \setminus \{y\})$. Then $z \in T$ for T such that $\Gamma_x(T \mid S) = 1$ and $z \notin T'$ for $\Gamma_x(T' \mid S \setminus \{y\}) = 1$. Note that T' is the <u>only</u> subset of $S \setminus \{y\}$ for which $\Gamma_x(T' \mid S \setminus \{y\}) > 0$, by definition of $\Gamma_x(T \mid S)$ and $\Gamma_x \in \{0, 1\}$. Since $\Gamma_x(T \mid S)$ satisfies A3, $\nexists T''$ with $y \in T''$ such that $\Gamma_x(T'' \mid S \setminus \{z\}) > 0$ and $\nexists T'''$ with $z \in T'''$ such that $\Gamma_y(T'' \mid S \setminus \{x\}) > 0$. Then, by definition of $\Gamma_x(S)$, $y \notin \Gamma_x(S \setminus \{z\})$ and $z \notin \Gamma_y(S \setminus \{x\})$. Therefore, $\Gamma_x(S)$ satisfies B3.

$\square$

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

Proof. Suppose a random consideration set mapping Γ has a random network consideration set mapping G:

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 $y \in \mathcal{G}$ 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. Then 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 Let T be such that $z \in T$ with $G_x(T \mid S) > 0$. Then there exists some $g^S \in \mathcal{G}_T^S$ where $f(g^S) > 0$. Since $\nexists T' \subseteq S \setminus \{y\}$ with $z \in T'$ such that $G_x(T' \mid S \setminus \{y\}) > 0$, 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 $G_x(T' \mid S \setminus \{y\}) > 0$ for $T' = \{j \mid j \text{ is connected to some node on this } x\text{-}z \text{ path in } g^S\}$ since $f(g^S) > 0$. Since there exists an x-y-z path in y^S, we can consider each sub-path independently.

Consider the x-y sub-path. When z is removed from S, this path survives, and if we let $T'' = \{j \in S \setminus \{z\} \mid j \text{ is connected to some node on the } x\text{-}z \text{ path in } g^S\}$, then $G_x(T'' \mid S \setminus \{z\}) > 0$, since $f(g^S) > 0$.

By similar logic, $G_y(T''' \mid S \backslash \{x\}) > 0$ for $T''' = \{j \in S \backslash \{x\} \mid j \text{ is connected to some node on the } y\text{-}z \text{ path in } g^S\}$.

$\square$

Lemma 2. *For any RNC π, x is revealed preferred to y if $\exists S \subseteq X$ such that:*

$$\pi_y(x \mid S) > 0 \tag{15}$$

Proof. For some π with an RNC representation, suppose that there exists $S \subseteq X$ such that $\pi_y(x \mid S)$. Choose some RNC representation of π, call it $(\hat{F}, \hat{\succ})$. By definition of π with an RNC representation, this implies that $\exists T \subseteq S$ with $\{x, y\} \subseteq T$ where $G_y^F(T \mid S) > 0$. Further, by definition, x is $\hat{\succ}$−best in T. Since $y \in T$, $x \hat{\succ} y$ according to this representation.

To show that $x \succ y$ for all RNC representations, suppose not. Then $\exists (F', \succ') \neq (\hat{F}, \hat{\succ})$ that also represents π, but for which $y \succ' x$. However, $\pi_y(x \mid S) > 0$ implies that $x \succ' y$, by the above logic, a contradiction. $\square$

Lemma 3. *For some RNC π, x is revealed preferred to y if, and only if, xP_Ry.*

Proof. $\rightarrow$:

The if part of this proof is trivial. Since xP_Ry, either xPy, indicating that x is revealed preferred to y directly by Lemma 2 or $\exists k$ such that xPk and kPy. In the latter case, xPk and kPy in <u>all</u> representations of π, again by Lemma 2, which, by transitivity of $\succ$ implies that xP_Ry in all RNC representations of π. Thus, x is revealed preferred to y.

$\leftarrow$:

Suppose that x is revealed preferred to y, but not xP_Ry. Since P_R is the transitive closure of P, it can be written as $P_R = \bigcup_{i=1,2,3,...} P^i$, where $P^1 = P$ and $P^{i+1} = P \circ P^i$. Then, if xP_Ry, $(x,y) \in P^i$ for some i and there then exists some finite sequence $\{k_0, k_1, ..., k_n\}$ where $xPk_0P...Pk_nPy$. Since $(x,y) \notin P_R$, there exists no such finite sequence.

By Lemma 2, all representations of π will be such that $P \subseteq \succ$. Select one, calling it $(F, \succ)$. From this, we construct an additional RNC $(F, \succ')$, where $\succ'$ is such that:

1. $P \subseteq \succ'$
2. $(y,x) \in \succ'$
3. $\succ'$ is transitive

Note that by the last requirement, $P_R \subseteq \succ'$ (all transitive supersets of P will include P_R). Since $(x,y) \notin P_R$, the construction of $\succ'$ to include P and (y,x) is valid.

We claim that $(F, \succ')$ also represents π. To show this, let $\pi^{\succ'}$ be RNC choice probabilities under $\succ'$ and consider $\pi_w(z \mid S)$ and $\pi_w^{\succ'}(z \mid S)$ for some arbitrary $w, z \in S$.

Case 1: $(z,w) \in P$ or $(w,z) \in P$

Suppose, to the contrary, that $\pi_w(z \mid S) > \pi_w^{\succ'}(z \mid S)$, without loss of generality. Then $\exists$ some $T \subseteq S$ with $\{w, z\} \subseteq T$ and z as $\succ$-best in T, but z is not $\succ'$ $-$best in T, with some $g^S \in \mathcal{G}_T^S$ such that $f(g^S) > 0$. Let $t^{\succ'}$ be the $\succ'$-best element in T. Note that $t^{\succ'} \neq w$, since $(z,w) \in P \subseteq \succ'$. Let $T' \subseteq T$ be the set of all nodes on some $w - t^{\succ'}$ path in g^S. Now consider $\pi_{t^{\succ'}}(z \mid T')$. Since $f(g^S) > 0$, $t^{\succ'}$ and z are connected under g^S, and $\pi_{t^{\succ'}}(z \mid T') > 0$, since z is $\succ$-best in $T \supseteq T'$. Then $(z, t^{\succ'}) \in P \subseteq \succ'$, a contradiction.

For $(w,z) \in P$, follow the above logic, reversing the roles of w and z.

Case 2: (z,w) and $(w,z) \notin P$

Since neither (z, w) nor $(w, z) \in P$, by definition $\pi_w(z \mid S) = \pi_z(w \mid S) = 0$ for all $S \subseteq X$. Suppose under $\succ$, $z \succ w$. Then for $\pi_w(z \mid S) = 0$ to be true, either i) $f(g^S) = 0$ for all g^S such that z and w are connected or ii) for all g^S such that $f(g^S) > 0$ and z is connected to w, all $w - z$ paths in g^S include some node k such that $k \succ z$ and $k \succ w$. Clearly, if F such that (i) holds, $\pi_w^{\succ'}(z \mid S) = \pi_z^{\succ'}(w \mid S) = 0$.

Then if (ii) holds, for each g^S such that $f(g^S) > 0$ and z is connected to w, and for each a $w-z$ path that includes some node k such that $k \succ z$, consider independently the $w-k$ and $k-z$ sub-paths. Denote the sets of nodes for each of these sub-paths $T_{w-k}^{g^S}$ and $T_{k-z}^{g^S}$, respectively. Then $\pi_w(k \mid T_{w-k}^{g^S}) > 0$ and $\pi_z(k \mid T_{k-z}^{g^S}) > 0$, since $f(g^S) > 0$. It follow then, that $\{(k, w), (k, z)\} \subseteq P \subseteq \succ'$ and $\pi_w^{\succ'}(z \mid S) = \pi_z^{\succ'}(w \mid S) = 0$ for all $S \subseteq X$.

Therefore, $\pi^{\succ'} = \pi$ and $(F, \succ')$ also represents π. Since both $(F, \succ)$ and $(F, \succ')$ represent π, but $(y, x) \in \succ'$, x is not revealed preferred to y, a contradiction. $\square$

Lemma 4. *Let π be an RNC and let x and y be such that there exists some set $S \supseteq \{x, y, z\}$ such that the following holds:*

$$\pi_z(x \mid S) > \pi_z(x \mid S \setminus \{y\}) \tag{16}$$

Then $(x, y) \in P$ and x is revealed preferred to y.

Proof. Let x and y be such that there exists some set S and $z \in S$ such that $\pi_z(x \mid S) > \pi_z(x \mid S \setminus \{y\})$. Then when S is available, there exists some network $g^S \in \mathcal{G}_T^S$ for some $T \supseteq \{x, z\}$ such that x is $\succ$ $-$best in T and $f(g^S) > 0$. Suppose, to the contrary, that $y \notin T$. Then if g^S is T-Connected, $g^{S \setminus \{y\}}$ is also T-Connected. Then $\pi_z(x \mid S) = \pi_z(x \mid S \setminus \{y\})$, since this will hold for all T such that $y \notin T$ and x is $\succ$-best in T, a contradiction. Then $y \in T$.

Note that if g^S is T-Connected, g^T is T-Connected. Then $\pi_y(x \mid T) > 0$, since $f(g^S) = f(g^T) > 0$, $\{x, y\} \subseteq T$, x is $\succ$-best in T, and g^T is T-Connected. Therefore, $(x, y) \in P$ and x is revealed preferred to y. $\square$

Lemma 5. *Let π be a PM-RNC with representations $(\mu^\pi, \succ^\pi)$. Then $\mu^i = \bar{\mu}$ for all $(\mu^i, \succ^i)$ representations of π (i.e. μ is unique).*

Proof. This lemma is fairly straightforward and is a function of the restrictions imposed by the particular consideration structure.

Note that for this lemma to hold, we must show that for any $(succ_i, \mu_{i)}$ and $(\succ_j, \mu^j)$ that both represent π, $\mu_{kl}^i = \mu_{kl}^j$ $k, l \in X$ such that $k \neq l$.

Suppose $k \succ l$ (or k is revealed preferred to l). Then the following must be true $\pi_l(k \mid \{k, l\}) = \mu_{kl}^i$ since μ^i represents π. Similarly, $\pi_l(k \mid \{k, l\}) = \mu_{kl}^j$. It follows that $\mu_{kl}^i = \mu_{kl}^j$ for all k, l such that k is revealed preferred to l.

But what if k cannot be revealed preferred to l? Assume to the contrary that $\mu^i_{kl} > 0$ for some μ^i that represents π. By definition,

$$\pi_l(k \mid \{k,l\}) = \begin{cases} \mu^i_{kl}, & \text{for } k \succ^i l \\ 0, & \text{for } l \succ^i k \end{cases} \tag{17}$$

Since $\mu^i_{kl} > 0$ and $\not(k \succ l)$, $l \succ^i k$. By definition,

$$\pi_k(l \mid \{k,l\}) = \begin{cases} \mu^i_{kl}, & \text{for } l \succ^i k \\ 0, & \text{for } k \succ^i l \end{cases} \tag{18}$$

Since $\mu^i_{kl} > 0$ and $l \succ^i k$, $\pi_k(l \mid \{k,l\}) > 0$, which implies that $l \succ k$, a contradiction.

Thus, for any two elements k and l where we cannot reveal preference, $\mu^i_{kl} = 0$ for any μ^i that represents π.

□

Proposition 3. *If an RNC G_x has a PM-RNC representation, it satisfies PM-RNC Binary Separability.*

Proof. The proof is written for $S' = \{z\}$, but the aggregate of the logic to larger S' is trivial.

Let G_x be an RNC with a PM-RNC representation, which is denoted as the matrix of consideration weights μ. For any network g, the probability that it occurs can be written as follows:

$$f(g) = \prod_{(i,j)\in X^2} [\mathbb{1}\{g_{ij} = 1\}\mu_{ij} + \mathbb{1}\{g_{ij} = 0\}(1 - \mu_{ij})]$$

Then for any starting point x in set S, the probability that set $T \subseteq S$ is considered is given by the following, for any non-trivial probability:

$$\begin{aligned} \Gamma_x(T \mid S) &= \sum_{g\in\mathcal{G}^S_T} f(g) \\ &= \sum_{g\in\mathcal{G}^S_T} \prod_{(i,j)\in X^2} [\mathbb{1}\{g_{ij} = 1\}\mu_{ij} + \mathbb{1}\{g_{ij} = 0\}(1 - \mu_{ij})] \end{aligned}$$

Note that by Lemma 1, if $g^{S\cup\{z\}}$ is T-Connected, then g^S is also T-Connected. In other words, $\Gamma_x(T \mid S\cup\{z\})$ can be constructed by beginning with $\mathcal{G}^S_T$ and subtracting out those networks for which g^S is not T-Connected.

Consider a partition $\mathcal{P}(\mathcal{G}^S_T)$ of $\mathcal{G}^S_T$ into subsets where g and g' are included in the same subset if $g_{ij} = g'_{ij}$ for all $\{i,j\} \neq \{t,z\}$ for some $z \notin S$ and $t \in T$. Let P be an arbitrary one of these subsets. We define the following:

$$f(P) \equiv \sum_{g \in P} f(g) \tag{19}$$

where $f(P)$ is taken to be probability over all networks in P. Recalling that on each $g \in P$ is such that g^S is T-Connected, restrict attention only to those in P such that $g^{S\cup\{z\}}$ is T-Connected. There is clearly a single $g \in P$ such that $g^{S\cup\{z\}}$ is T-Connected, the network g such that $g_{tz} = 0$ for all $t \in T$ (otherwise, $g^{S\cup\{z\}}$ would not be T-Connected). Let this unique $g \in P$ be denoted $g^*(P)$. Since $g^*_{tz}(P) = 0$ for all $t \in T$, it follows that $f(g^*(P)) = \prod_{t\in T}(1-\mu_{tz})f(P)$.

Then $\Gamma_x(T \mid S \cup \{z\})$ can be constructed as follows:

$$\begin{aligned}
\Gamma_x(T \mid S \cup \{z\}) &= \sum_{g \in \mathcal{G}_T^{S\cup\{z\}}} f(g) \\
&= \sum_{P \in \mathcal{P}(\mathcal{G}_T^S)} f(g^*(P)) \\
&= \sum_{P \in \mathcal{P}(\mathcal{G}_T^S)} \prod_{t\in T}(1-\mu_{tz}) f(P) \\
&= \sum_{P \in \mathcal{P}(\mathcal{G}_T^S)} \prod_{t\in T}(1-\mu_{tz}) \sum_{g\in P} f(g) \\
&= \prod_{t\in T}(1-\mu_{tz})\Gamma_x(T \mid S)
\end{aligned}$$

The result then directly follows an observation that $(1-\mu_{tz})$ is simply equal to $\Gamma_z(\{z\} \mid \{t, z\})$ in the PM-RNC model. □

B Proofs for Section 7 (Discussion)

Proposition 4. *In the RNC Advertising Game with parameters α, β, and c, such that (α, β, c) is in the unit cube, there exist non-empty subsets of the parameter space where:*

1. *$\sigma^* = (1, 0, 1)$ is supported as a Nash Equilibrium*
2. *$\sigma' = (0, 1, 0)$ is not supported as a Nash Equilibrium*
3. *Aggregate profit under σ' is higher than under σ^**

Proof. 1. $\sigma^* = (1, 0, 1)$ **is supported as a Nash Equilibrium**:

First, consider each firm's incentives under $\sigma^* = (1, 0, 1)$. Firm A's expected profit under σ^* is as follows:

$$\pi_A(\sigma^*) = \frac{3\alpha}{4} + \frac{1-2\alpha}{3} - c$$

In order for Firm A to not have an incentive to unilaterally deviate to $\sigma_A = 0$, the following condition must then hold:

$$\begin{gathered}\frac{3\alpha}{4} + \frac{1-2\alpha}{3} - c > \frac{1-2\alpha}{3} \\ \frac{3\alpha}{4} > c \end{gathered} \tag{20}$$

The condition for Firm C is identical.

Conditional on Firm's A and C choosing to advertise, Firm B has no incentive to advertise if the following holds:

$$\begin{gathered}\frac{2\alpha}{3} + \frac{1-2\alpha}{3} - c < \frac{\alpha}{2} + \frac{1-2\alpha}{3} \\ \frac{\alpha}{6} < c \end{gathered} \tag{21}$$

Clearly there exist positive α and c such that conditions 20 and **??** hold.

2. $\sigma' = (0, 1, 0)$ **is not supported as a Nash Equilibrium**:

   For Firm A, there exists an incentive to unilaterally deviate from $\sigma'_A = 1$ if the following holds:

$$\begin{gathered}\alpha + \frac{1-2\alpha}{3} - c > \frac{\alpha}{2} + \frac{1-2\alpha}{3} \\ \frac{\alpha}{2} > c \end{gathered} \tag{22}$$

   The condition is identical for Firm C.

   Clearly, for any feasible and strictly positive α, 20 - 22 are satisfied for any $c \in (\frac{\alpha}{2}, \frac{3\alpha}{4})$.

3. **Aggregate profit under σ' is higher than under σ^***

This should be clear from the definition of the profit function. Under σ^*, aggregate profit is equal to $1 - 2c$, since two firms are advertising, whereas under σ' it is simply $1 - c$.

$\square$

C Additional Results

C.1 Results by Baseline/Context

Table 22: Correct Rate by Treatment

	NC	C
Mean	0.851	0.863
Std Error	0.009	0.009
N	1733	1555

Wilcox $p > 0.10$ for $H_0 : \mu_C = \mu_{NC}$.

Table 23: Monotonicity Violations by Context

	Baseline	Context
Mean	0.791	0.805
Std Error	0.012	0.013
N	1140	1000

Wilcox $p > 0.10$ for $H_0 : \mu_{Baseline} = \mu_{Context}$

Table 24: Path Connectedness by Context

	Baseline	Context
Mean	0.456	0.452
Std Error	0.067	0.078
N	57	42

Mann-Whitney $p > 0.10$ for $H_0 : \mu_{Baseline} = \mu_{Context}$

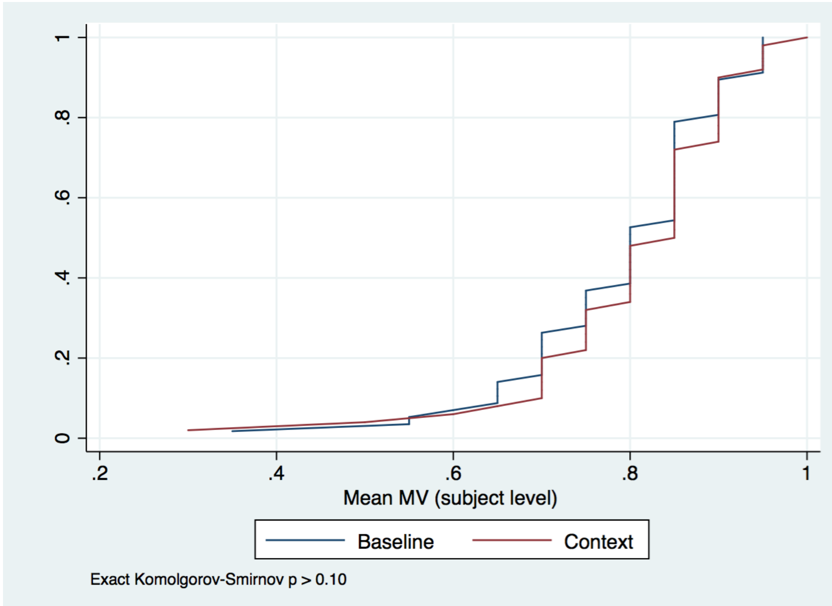


Figure 15: Cumulative Distribution of Mean MV by Subject and Treatment

C.2 Omitted Results

Table 25: Determinants of Consideration Set Size

	Model 1	Model 2	Model 3
N	0.217***	0.218***	0.204***
	(0.011)	(0.011)	(0.012)
Period	0.0271***	0.0264***	0.0189**
	(0.008)	(0.008)	(0.009)
Female		0.146	0.529
		(0.466)	(0.512)
$\hat{\mu}_{GPA}$		2.131	
		(1.418)	
Cognitive Score			-0.450
			(1.137)
Constant	4.241***	2.687**	4.520***
	(0.234)	(1.106)	(0.859)
Observations	25811	25811	21525

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

D Instructions

D.1 Baseline

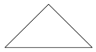
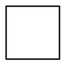
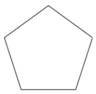
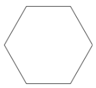
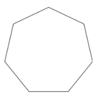
Instructions

Thank you for participating in the experiment today. At this time, please be sure that your cell phone is turned off and stored away. At no point during this experiment should you use your cell phone or any other electronic device. Also, please refrain from communicating with any other subject in the lab today. Failure to follow these rules will result in your expulsion from the lab and you will forfeit any cash earnings you may have otherwise received.

This is an experiment in decision making. You will be paid a $7 guaranteed show-up fee in addition to earnings based on your decisions in the experiment.

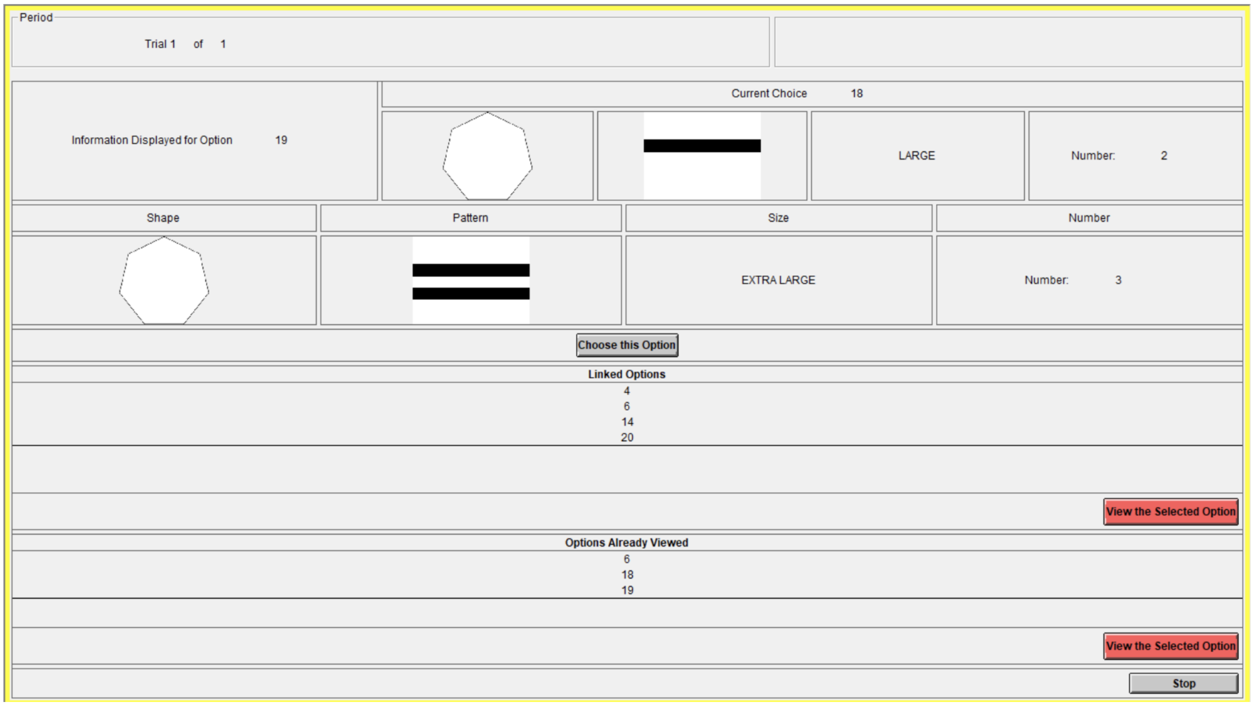
Decision Environment

In each of 31 periods, you will be faced with a number of options from which you can select one. Each option has 4 attributes: Shape, Pattern, Size, and Number. The value of an individual attribute is given in Experimental Currency Units (ECUs) in the following table:

Shape	Value	Pattern	Value	Size	Value	Number	Value
	1 ECU		1 ECU	EXTRA SMALL	1 ECU	1	1 ECU
	2 ECU		2 ECU	SMALL	2 ECU	2	2 ECU
	3 ECU		3 ECU	MEDIUM	3 ECU	3	3 ECU
	4 ECU		4 ECU	LARGE	4 ECU	4	4 ECU
	5 ECU		5 ECU	EXTRA LARGE	5 ECU	5	5 ECU

The value of a given option is the sum of the value of its attributes as per the table above.

In each period, you will be shown a version of the following screen:



The screen is composed of the following parts (left to right, top to bottom):

Option Label: this is the option for which information is currently displayed. In the example, Option 19 is shown along with information on the Attributes of Option 19. Option Labels have been chosen randomly for each Period and do not reflect the value of the option. Moreover, two options with the same option label may have different values in different periods.

Current Choice: this is the option that you are currently holding as your choice. This will be explained in detail below. In addition to the label for your Current Choice, you are shown information about the Attributes for your Current Choice for your reference.

Attribute Information: these are the attributes for the option currently displayed. The value of each option is the sum of its attributes as according to the table above. For example, the value of Option 19 in the example above is 16 ECU (5 for Heptagon + 3 for Two-Bar Pattern + 5 for Extra Large + 3 for Number 3 = 16).

Choose this Option Button: you can click this button to change your Current Choice to the option that is currently displayed. If your Current Choice is the option that is currently displayed, this portion of the screen with display The option currently displayed is your Current Choice.

Linked Options: this is a clickable list of options that are Linked to the currently displayed option. An option is Linked to the currently displayed option if it shares 2 or more attributes with the currently displayed option. For example, if there was another Option in the current Period with a Two-Bar Pattern and the

Number 3, it would be shown in the list of Linked Options for Option 19 in the screenshot above. Note that this list may be quite long, in which case you will see a scroll bar next to the list of Linked Options.

Options Already Viewed: this is a clickable list of options that have already been viewed by you in the given period. You can click on any option in this list and click View Selected Option to view information for that option again. Again, if the list of Options Already Viewed gets sufficiently long, you will be shown a scroll bar next to the list. **Note**: you can only view information for options other than the one currently displayed by either clicking on it in the Linked Options menu or the Options Already Viewed menu.

Stop: if you would like to Stop looking at information for the available options and would not like to change your Current Choice, you can click the Stop button.

Period Duration

In each period, you will have up to 75 seconds to evaluate all of the available options and make choices. At any time, you can click Stop and you will not be shown any more information on any of the options for the given period. Note: in order to move on to the next period, you must wait for the entire 75 seconds to pass in the current period. Thus, if you Stop after, say, 45 seconds in the current period, you will still have to wait the remaining 30 seconds for the period to end in order to move on to the next period.

Choices

At the end of each Period, a random time between 2 and 75 seconds will be chosen and your Current Choice held at that time will be implemented as your chosen option for that Period. Each time between 2 and 75 seconds is equally likely to be chosen.When evaluating options during the 75 seconds, you will not know at what time your Current Choice will be implemented as your chosen option for that Period. At the beginning of each Period, you will start off being shown information for one particular Option but will have no Current Choice. If you do not have a Current Choice at the time chosen randomly by the computer program to implement your choice, you will be paid nothing for the current Period. Thus, it is in your best interest to choose any Option as soon as possible. You can then replace it with a better option when/if you find an option that has a higher value.

At the end of each period, you will be told i) at what time your Current Choice was implemented, ii) which Option you held at that time, and iii) what the value (in ECU) of that option was.

For clarification purposes, consider the following example: a subject is in Period 4 with a time limit of 75 seconds and they immediately choose the first option shown to them, Option 1, which has a value that they have determined to be 10 ECU. After 30 seconds, the subject changes their Current Choice to Option 2 with a value of 12 ECU and continues evaluating the available options. After 10 more seconds (40 in total since the start of the period), the subject selects a new Current Choice of Option 3 with a value of 14 ECU. The subject makes no further choices and the 75 seconds runs out. The subject's choices are shown in the

Time	0 Seconds	30 Seconds	40 Seconds
Option	Option 1	Option 2	Option 3
Value	10 ECU	12 ECU	14 ECU

following table:

If the time chosen randomly by the system at the end of the Period is anywhere between 2 and 30 seconds, the subject will be paid 10 ECU (for holding Option 1). If it is anywhere between 30 and 40 seconds, they will be paid 12 ECU (for holding Option 2). If it is 40 seconds or higher, they will be paid 14 ECU.

Earnings

You will be paid a guaranteed show-up fee of $7 in addition to your earnings for your decision. The value of the option that is treated as your final choice for a period (i.e. the option held by you at the time chosen by the system) is the sum of the value of its Attributes as given in the table above. Experimental Currency Units (ECUs) will be converted to cash (USD) at a rate of 1 ECU = $1 USD.

Though you will make decisions in each of 31 periods, you will only be paid for 1 of these periods. Which period will be paid will be chosen at random at the end of the experiment, with each period being equally likely to be chosen. Thus, it is in your best interest to behave in each period as if it is the period for which you will be paid.

D.2 Environment with Context

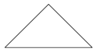
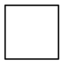
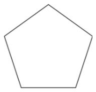
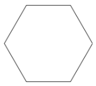
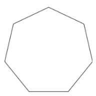
Instructions

Thank you for participating in the experiment today. At this time, please be sure that your cell phone is turned off and stored away. At no point during this experiment should you use your cell phone or any other electronic device. Also, please refrain from communicating with any other subject in the lab today. Failure to follow these rules will result in your expulsion from the lab and you will forfeit any cash earnings you may have otherwise received.

This is an experiment in decision making. You will be paid a $7 guaranteed show-up fee in addition to earnings based on your decisions in the experiment.

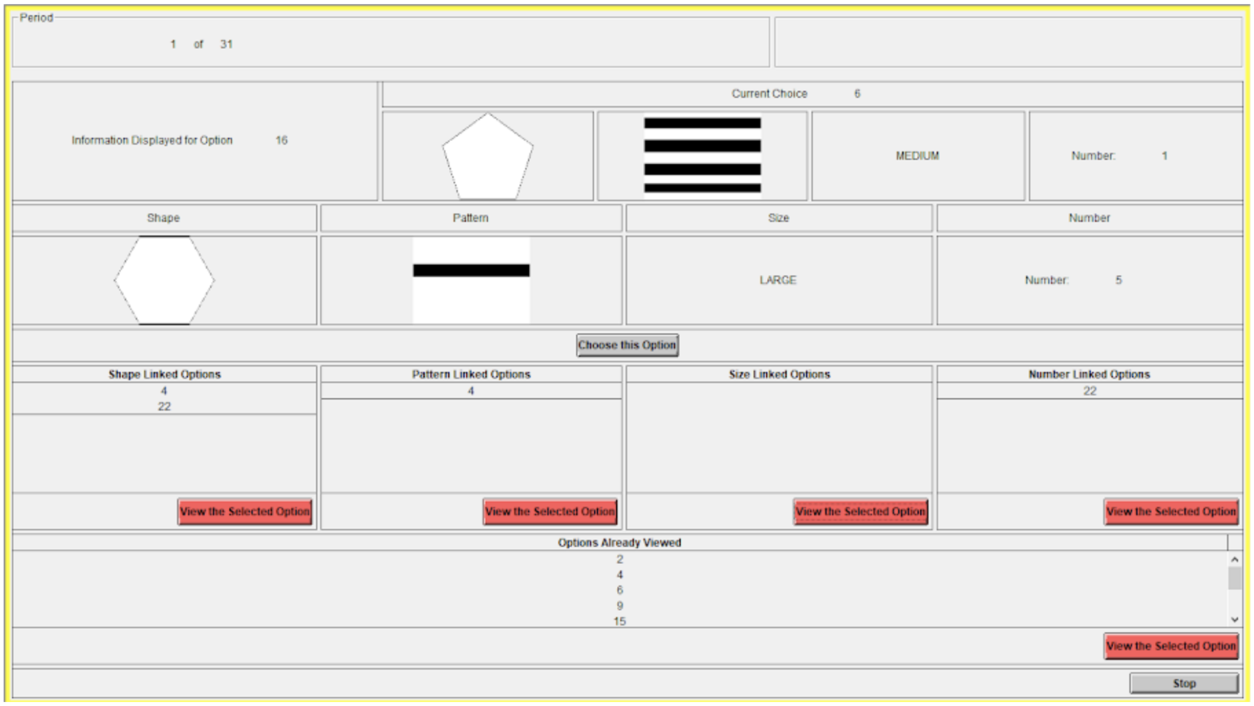
Decision Environment

In each of 31 periods, you will be faced with a number of options from which you can select one. Each option has 4 attributes: Shape, Pattern, Size, and Number. The value of an individual attribute is given in Experimental Currency Units (ECUs) in the following table:

Shape	**Value**	**Pattern**	**Value**	**Size**	**Value**	**Number**	**Value**
	1 ECU		1 ECU	EXTRA SMALL	1 ECU	1	1 ECU
	2 ECU		2 ECU	SMALL	2 ECU	2	2 ECU
	3 ECU		3 ECU	MEDIUM	3 ECU	3	3 ECU
	4 ECU		4 ECU	LARGE	4 ECU	4	4 ECU
	5 ECU		5 ECU	EXTRA LARGE	5 ECU	5	5 ECU

The value of a given option is the sum of the value of its attributes as per the table above.

In each period, you will be shown a version of the following screen:



The screen is composed of the following parts (left to right, top to bottom):

Option Label: this is the option for which information is currently displayed. In the example, Option 16 is shown along with information on the Attributes of Option 16. Option Labels have been chosen randomly for each Period and do not reflect the value of the option. Moreover, two options with the same option label may have different values in different periods.

Current Choice: this is the option that you are currently holding as your choice. This will be explained in detail below. In addition to the label for your Current Choice, you are shown information about the Attributes for your Current Choice for your reference.

Attribute Information: these are the attributes for the option currently displayed. The value of each option is the sum of its attributes as according to the table above. For example, the value of Option 16 in the example above is 15 ECU (4 for Hexagon + 2 for One-Bar Pattern + 4 for Large + 5 for Number 5 = 15).

Choose this Option Button: you can click this button to change your Current Choice to the option that is currently displayed. If your Current Choice is the option that is currently displayed, this portion of the screen with display The option currently displayed is your Current Choice.

Linked Options: there are four clickable lists of options that are Linked to the currently displayed option. An option is Linked to the currently displayed option if it shares 2 or more attributes with the currently displayed option. For example, if there was another Option in the current Period with a One-Bar Pattern and the Number 5, it would be shown in a list of Linked Options for Option 16 in the screenshot above. Note that a list may be quite long, in which case you will see a scroll bar next to the list of Linked Options. The full list of Linked options is separated into four different fields, one for each Attribute: Shape, Pattern, Size, and Number. An option will be displayed in the relevant field if it meets two criteria: i) the option shares at least two Attributes with the currently displayed option and ii) it shares the Attribute for the relevant field with the currently displayed option.

For example, Option 4 also has the Hexagon Shape Attribute and the One-Bar Pattern Attribute. Since Option 16 (the currently displayed option) has both of these Attributes, Option 4 is linked to Option 16. Since it has the same Shape as Option 16, it will be listed in the Shape Linked Options list. Since it has the same Pattern as Option 16, it will also be listed in the Pattern Linked Options list. Consider, for example, another option, call it Option 12 (not displayed in the above screenshot). It has the Attributes: Square, One-Bar, Small, 4. Notice that it shares the Pattern Attribute with Option 16 (the currently displayed option): both have the Pattern One-Bar. But it does not share any other Attributes with Option 16. Therefore, it will not show up in any of the link lists when Option 16 is the currently displayed option. Especially note that it will not show up in the Pattern Linked Options list for Option 16, even though they have the same Pattern Attribute, because it does not share two or more Attributes with Option 16.

Options Already Viewed: this is a clickable list of options that have already been viewed by you in the given period. You can click on any option in this list and click View Selected Option to view information for that option again. Again, if the list of Options Already Viewed gets sufficiently long, you will be shown a scroll bar next to the list. **Note**: you can only view information for options other than the one currently displayed by either clicking on it in one of the Linked Options menus or the Options Already Viewed menu and clicking the View the Selected Option button for that list. Whenever you click on a new option from one of these lists and click View the Selection Option, all of the information on the screen (Option Label, Attribute Information, Linked Options, Options Already Viewed) will update to display information for the option to which you are navigating. The Current Choice information in the upper right of the screen will only ever change if you change your Current Choice (by choosing a new option using the Choose this Option button).

Stop: if you would like to Stop looking at information for the available options and would not like to change your Current Choice, you can click the Stop button.

Period Duration In each period, you will have up to 75 seconds to evaluate all of the available options and make choices. At any time, you can click Stop and you will not be shown any more information on any of the options for the given period. Note: in order to move on to the next period, you must wait for the entire 75 seconds to pass in the current period. Thus, if you Stop after, say, 45 seconds in the current period,

Time	0 Seconds	30 Seconds	40 Seconds
Option	Option 1	Option 2	Option 3
Value	10 ECU	12 ECU	14 ECU

you will still have to wait the remaining 30 seconds for the period to end in order to move on to the next period.

Choices

At the end of each Period, a random time between 2 and 75 seconds will be chosen and your Current Choice held at that time will be implemented as your chosen option for that Period. Each time between 2 and 75 seconds is equally likely to be chosen.When evaluating options during the 75 seconds, you will not know at what time your Current Choice will be implemented as your chosen option for that Period. At the beginning of each Period, you will start off being shown information for one particular Option but will have no Current Choice. If you do not have a Current Choice at the time chosen randomly by the computer program to implement your choice, you will be paid nothing for the current Period. Thus, it is in your best interest to choose any Option as soon as possible. You can then replace it with a better option when/if you find an option that has a higher value.

At the end of each period, you will be told i) at what time your Current Choice was implemented, ii) which Option you held at that time, and iii) what the value (in ECU) of that option was.

For clarification purposes, consider the following example: a subject is in Period 4 with a time limit of 75 seconds and they immediately choose the first option shown to them, Option 1, which has a value that they have determined to be 10 ECU. After 30 seconds, the subject changes their Current Choice to Option 2 with a value of 12 ECU and continues evaluating the available options. After 10 more seconds (40 in total since the start of the period), the subject selects a new Current Choice of Option 3 with a value of 14 ECU. The subject makes no further choices and the 75 seconds runs out. The subjects choices are shown in the following table:

If the time chosen randomly by the system at the end of the Period is anywhere between 2 and 30 seconds, the subject will be paid 10 ECU (for holding Option 1). If it is anywhere between 30 and 40 seconds, they will be paid 12 ECU (for holding Option 2). If it is 40 seconds or higher, they will be paid 14 ECU.

Earnings

You will be paid a guaranteed show-up fee of $7 in addition to your earnings for your decision. The value of the option that is treated as your final choice for a period (i.e. the option held by you at the time chosen by the system) is the sum of the value of its Attributes as given in the table above. Experimental Currency Units (ECUs) will be converted to cash (USD) at a rate of 1 ECU = $1 USD.

Though you will make decisions in each of 31 periods, you will only be paid for 1 of these periods. Which period will be paid will be chosen at random at the end of the experiment, with each period being equally likely to be chosen. Thus, it is in your best interest to behave in each period as if it is the period for which you will be paid.