EXTENSIVE ATTENTION, INTENSIVE ATTENTION AND THE ORIGINS OF RANDOM CHOICE

IRFAN KHAN*, IAN KRAJBICH§, COLLIN RAYMOND†, SAVITAR SUNDARESAN‡, AND ANDRE VEIGA‡

Abstract. Using repeated choices and eye-tracking data across 180 menu instances from 50 subjects we link the randomness of choice to two notions of attention: extensive attention (what options are looked at) and intensive attention (how long options are looked at). We show that models that seek to explain randomness through attention should capture four key facts. First, extensive attention and intensive attention are both related to randomness in choice, although intensive attention, on average, is a better predictor of choice. Second, the mechanism of choice, and the degree of randomness, is very different for large compared to small choice sets. Third, the relative importance of attention in generating randomness in choice is smaller between-person than within-person. Fourth, greater attentional capacity is not associated with reduced randomness at the individual level. Fifth, increased attention is not strongly correlated with better choices, indicating additional attention may be deployed in situations where higher randomness is already likely.

1. Introduction

Heterogeneity in choice is a widely observed fact, whether it be different individuals choosing distinct items from the same choice set, or a single individual choosing differently after repeatedly observing a choice set. In other words, observed choice, either at the individual or aggregate level is typically observed to be random, rather than deterministic. Economists, following the lead of McFadden (1974), have taken seriously the impact of randomness in choice on estimating preferences and conducting welfare analyses (e.g., Bhattacharya (2015)).

* Ohio State University.
§ University of California, Los Angeles.
† Cornell University.
‡ Imperial College London.

Economists have developed a wide variety of approaches to modeling choice randomness, and leveraging these models to estimate preferences. Some approaches, like McFadden (1974), suppose that randomness is due to preference shocks. Other approaches suppose that randomnesss, like Manzini and Mariotti (2014); Brady and Rehbeck (2016); Cattaneo et al. (2020), suppose that it is due to attentional shocks, while still others assume it is due to maximization of a convex utility function Cerreia-Vioglio et al. (2019); Fudenberg et al. (2015). These approaches distinguish between preference based origins of randmoness and attention based origins.[1]

Paralleling recent work on random choice, there has also been a growth in empirical work trying to measure choice process data. Much of this work measures eye movements and relates them to choice Reutskaja et al. (2011); Krajbich et al. (2010); Krajbich and Rangel (2011); Krajbich (2019); Smith and Krajbich (2018); Milosavljevic et al. (2012)). These papers typically relate their findings to drift-diffusion models (Smith (2000)). Drift diffusion models can generate randomness in choice, and have a deep connection to existing economic models of stochastic choice (Webb (2019)).

However, there is little existing work trying to understand how the leverage data on attention in order to improve our understanding of models of random choice. This paper relates stochastic choice data to two key measures of attention — extensive attention (what subjects looked at) and intensive attention (how long subjects looked at an option) — in a lab experiment where subjects chose between risky options. Subjects faced repeated choices of lotteries from choices sets that ranged in size from 2 to 17, allowing us to collect both individual and aggregate measures of random choice. Attention was measured via eye-tracking.

Our data allow us to develop and explore five key facts that any model that seeks to understand choice randomness, and the role of attention in choice, should explain. First, both kinds of attention (extensive and intensive) are related to randomness in choice. However, we find that intensive attention is, on aggregate, a better predictor of choice. Second, choice set size matters: the mechanism of choice, and the degree of randomness, is very different for large compared to small choice sets. Third, both within person and between person choice heterogeneity is important, but they have different characteristics. Extensive attention matters less, but intensive attention more, for

[1]Much of the work on attention parallels related work in deterministic choice trying to understand how attention and product characteristics influence choice in ways not captured by traditional models (e.g., Masatlioglu et al. (2012), Kőszegi and Szeidl (2013),Bordalo et al. (2013), Bushong et al. (2021)).

random choice at the between person level (compared to the within person). Fourth, greater attentional capacity is not associated with lower randomness in choice at the individual level — people who can pay more attention do not choose more consistently. Fifth, increased attention is not strongly correlated with better choices.

Three unique aspects of our data set allow us to conduct analyses that are difficult to do with existing data sets. First, we have data from 50 individuals, each of whom makes choices from many choice sets repeatedly (each choice set is seen either 5 or 10 times by each individual). This means that we can conduct our analyses at either the individual or aggregate level. Second, we have data on menus that range in size from 2 to 17 and which overlap (and nest one another) to varying degrees. This means that we can analyze how attention and choice changes as the menu gets larger or smaller, as well as how choice from a larger menu relates to choice from a given subset. Third, because we have data on both choice and choice process data allows us to examine the extent to which visual attention at both the extensive and intensive margin influence choice.

Section 3 experimental protocol. Our experiment was designed to overcome several key issues. First, we wanted to obtain sufficient data to identify random choice on the individual level, leading to us having individuals make many repeated choices for a given choice set. Second, we want to measure attention in a naturalistic way, with minimal interference. We believe that measuring gaze allows for little intrusion on the deliberation process of subjects, while still allowing us to measure important facets of attention. Moreover, importance of gaze has been emphasized in previous work — e.g., Cattaneo et al. (2020) use eye-tracking data from Reutskaja et al. (2011) as a motivating factor. Third, we wanted to be able to measure multiple forms of attention — not just what people looked at, but how long, and in what order. Fourth, we wanted to have an environment where subjects would not consider all options in all choice sets, thus leading us to choose widely varying choice set sizes. Fifth, we wanted our domain to have a clear relationship to existing work and economic theory — leading us to use risky acts as options.

In order to accomplish these goals, we have subjects choose among (finite) choice sets between lotteries with binary support, where each realization is equally likely. We begin with a state space consisting of two equally likely states. We construct the universe of possible options by taking three budgets sets of lotteries (a la Choi et al.

(2007)) and selecting a number of lotteries (between 7 and 13) from each budget line, giving us a total of 25 options. Each option is thus described as a pair of numbers (e.g., 200,300), denoting the payoff in each of the two states. We then combine these options in different ways to generate 27 possible choice sets, ranging in size from 2 to 17. Each subject (out of 50 total subjects) sees each choice set either 5 (for the choice sets of size 2,3 and 6) or 10 (for the choice sets of size 13, 14 and 17) times — we refer to each time a subject sees a particular choice set as an *instance* of that choice set. This generates 180 observations for each subject. The order of that subjects observe choice sets is randomized. What does an individual see when making their choice from a choice set? The subjects see a screen which has a 5 by 5 grid. Options are randomly assigned to a particular cell in the grid. Cells that contain an option have XXX, XXX written on them. Subjects must direct their gaze at the cell in order to then observe what option is actually in the cell. The cell is only "uncovered" while the subjects' gaze remains fixated on that particular cell. When subjects want to choose the cell they are looking at currently (they can only choose a cell they are currently looking at) they simply press the keyboard. We conduct this procedure with 50 subjects.

We first discuss the related literature and our conceptual appraoch in Section 2.

Section 4 provides a summary of the main variables in our analysis. We initially describe our primary measure of choice randomness — a measure of the probability of a choice reversal between any pair of instances of a choice set. We also describe two other measures of choice randomness — the share of options chosen from a choice set across all instances, and the normalized entropy of the distribution of choices. All three are highly correlated and so we focus on the measure of choice reversals for the rest of the paper.

We then turn to considering our attentional data. We analyze two kinds of data: what is looked at (extensive attention), and for how long items are looked at (intensive attention). We carefully draw a distinction between an option (i.e. a particular lottery) and a fixation (a particular time when a lottery was looked at). An option may have multiple fixations within a single instance of a choice set, as the subject may look at an item, look away, and then re-look at the item. We again describe our primary measure of extensive attention — the share of options looked at in a choice set, averaged across all instances of that choice set. We believe such a notion naturally corresponds to the notion of a consideration set which has been extensively studied in economics and marketing (e.g., Masatlioglu et al. (2012)). However, we also measure the degree to which extensive attention is random, in a way that parallels our measurement of choice randomness — essentially we ask, for a given choice set, how likely is it that across two instances, a given option is either looked at in both, or not looked at in both. In other worse, are they consistent in what they consider across choice sets. For intensive attention, we focus on understanding the average time each item was looked at in a given choice set - what we call dwell time. We also measure the number of fixations per item — the number of times an item was looked at.

In Section 5 documents that attention, both extensive and intensive, is a key predictor of choice randomness. Using our measure of choice reversals, we show that across all our choice sets, extensive attention can rationalize around 30% of the observed choice reversals for an individual. This means that in around 30% of the choice reversals we observe between pairs of instances of a choice set, either the option chosen in the first instance is not looked at in the second, or vice versa. Of the remaining randomness within an individual, intensive choice can rationalize around 80%. If we have a pair of instances of a choice set where the chosen option differs, and both options are looked at in both instances, then 80% of the time the option chosen in a given choice set is looked at for longer *in that choice set* compared to the option that was chosen in the other choice set (and vice versa). Thus, in aggregate, extensive and intensive attention together can rationalize around 85% of choice randomness. Thus, a model of randomn choice that wants to capture its relationship to choice process must allow for the role of both extensive and intensive attention, and the fact that combined they explain the large majority of randomness.

In Section 6 we turn to exploring how randomness in choice and attention vary across choice set size. We show that randomness is much higher for larger choice sets (sizes 13, 14, 17) than smaller choice sets (sizes 2, 3, 6). Moreover, both extensive and intensive attention are much lower for larger choice sets. The relationship between attention and choice also is very different depending on the choice set size. In particular, in small choice sets extensive attention can explain less than 13% of choice randomness, while in large choice sets it is around 43%. Thus, attention is a much larger factor in choice randomness in large choice sets. Interestingly, choice randomness, conditional on attention is very similar in small and large choice sets. A similar pattern is observed with intensive attention, where intensive attention can rationalize a larger fraction of choice randomness in large choice sets, compared to small choice sets (85%compared to 70%). Combining these patterns with additional data on extensive and intensive attention, and the number of fixations, we see that in small choice sets, individuals tend to look at all the options, and look back and forth between options. In contrast, in large choice sets individuals tend to look at any given option once, and stop looking prior to considering all options. Thus, mechanisms that seek to explain random choice should leverage the fact that individuals engage in very different choice processes in small and large sets, and that this ultimately impacts the degree of randomness in choice.

Models of random choice often discuss either randomness across individuals, or randomness within individuals. We leverage our data to discuss the importance of choice heterogeneity both between and within individuals in Section 7. In particular, we can compare what the data looks like when we conduct all analyses on the subject level and then average across subjects, versus pooling all the data as if it came from a representative agent. We find that around 56% of choice randomness at the representative agent level comes from individual heterogeneity, while the rest occurs due to between person differences. Not surprisingly, even conditional on extensive attention, around 50% of choice randomness at the representative agent level comes from between person heterogeneity, indicating that non-extensive attention differences are an important driver of between person differences. This implies that extensive attention explains only about 20% of the choice randomness at the representative agent level (as opposed to around 27% at the individual level). Perhaps surpisingly, we find that intensive attention is a better predictor of choice reversals at the representative agent level, as opposed to the individual level.

Delving deeper, we find that we can characterize individuals by types, both in their choice behavior, as well as patterns of extensive and intensive attention. Subjects' patterns of attention are closely related at the extensive and intensive level — individuals who are likely to look at more things are also more likely to look at items for longer. However, choice randomness and attentional capacity (of either kind) is essentially uncorrelated.

Thus, our data points to the fact that accounting for randomness means accounting for it as both a within and between person phenomena. Moreover, the relative importance of different sources of randomness depend on whether we look at within or between person randomness. However, we want to allow for essentially uncorrelated behavior in terms of attention and randomness.

Last, in Section 8 discuss the relationship between attention and the optimality of choice. Of course, it can be difficult to measure how "good" a choice is, and so we use several approaches. First, we simply consider choices which are first order stochastically dominated by some other available option. We show that violations occur relatively often in our choices — around 16% overall, but are much more prominent in large choice sets relative to small choice sets. However, if we condition on extensive attention, we find that the probability of first order stochastic dominance hardly falls. Thus, it appears that the violations we observe are not driven by a lack of attention, but rather by some other force. One may be concerned that first order stochastic dominance violations in our environment may not be immediately obvious, since it may require comparing payoffs across states. A more obvious form of dominance is statewise dominance. Although individuals engage in far less statewise dominance (indicating that it is easier to observe and avoid) we still find controlling for attention only reduces its occurance marginally. Thus, it appears that attention is not a key driver of "mistakes" in choices. Thus, any account of randomness needs to explain why additional attention seems to not decrease mistakes.

In Section 9 we discuss caveats to our appraoch and provide some takeaways.

2. Conceptual Framework and Related Literature

Our work is inspired by and relates to several strands of existing research in economics and cognitive sciences. We view our contribution as trying to bring together several notions of attention and directly measure and link them to economic models of choice randomness.

Thus, our work draws on the enormous body of work developing the theoretical foundations of random choice, as well as existing empirical tests.

One key explanations for randomness in choice, formalized by Block and Marschak (1959); McFadden (1974); Falmagne (1978), points towards an origin based on utility or preference shocks, i.e., random utility models (RUMs). These papers generally assume that preference heterogeneity occurs because there is a distribution over preferences or utilities from which individuals take a draw from when choosing. Classes of RUMs have been widely used in both theoretical and empirical work (including the multinomial logit model of Luce (1959)). E.g., in industrial organization distribution of population preferences are often estimated using market share data (e.g., Berry et al. (1995)), and so the randomness is typically assume to emerge due to between person preference heterogeneity. Such approaches have also been used to try and test the validity of the RUM approach, using using data from both the field and the lab (e.g., Conte et al. (2011), McCausland et al. (2024)).

There is also a smaller literature focused on understanding random choice (often via the lens of RUMs) using within-person repeated choice. Much, although not all, builds off of Tversky (1969). Examples include Hey (2001), Davis-Stober et al. (2015) (who extend Regenwetter and Davis-Stober (2012)), McCausland et al. (2020), Agranov and Ortoleva (2017), Agranov et al. (2023), Feldman and Rehbeck (2022), and He et al. (2019). These studies tend to focus on choices from binary sets, although some use larger sets (e.g., McCausland et al. (2020) goes up to choice sets of size 5). They also usually to focus on lottery as options (although He et al. (2019) looks at temporal monetary payoff problems). LIke us, these papers typically have subjects make 10 or fewer repetitions (although Davis-Stober et al. (2015) has each individual repeat a choice 24 times). The majority of these papers ask whether individuals' choice distributions can be rationalized by a random utility model.

A distinct preference based motivation for randomness is that individuals deliberately randomize (e.g., Machina (1984); Fudenberg et al. (2015); Cerreia-Vioglio et al. (2019)). Such motivations hae been tested in lab by Agranov and Ortoleva (2017); Agranov et al. (2023); Feldman and Rehbeck (2022).[2] Most of the theoretical approaches assume that violations of first order stochastic dominance should not occur – something we can look for.

That said, our paper cannot precisely identify why people randomize conditional on attention. We can identify with our methods, what kind of residual randomness remains after attention has been accouted for. In almost all existing papers, such residual randomness is attributed to preferences. Although, in line with the literature,

[2]There are a few studies, such as Dwenger et al. (2018) that look at preferences for randomization in a field setting.

also use this heuristic, we want to be clear that this is an assumption, rather than something we can test in our data.

The key difference between our paper and pre-existing work in this literature is our focus on relating random choice to measures of attention. This allows us to identify how much randomness is due to, e.g., limited consideration, and how much is still residual that might be attributed to preference shocks or deliberate randomization. Our design is also substantively different as we have much larger choice sets than existing experiments, and we conduct our our analysis at the withing and between person level.

More recently, an alternative account seeking to explain choice randomness has been developed, which focuses on attention, especially what we call extensive attention, as a key driver of randomness in choice. Papers such as Manzini and Mariotti (2014); Brady and Rehbeck (2016); Cattaneo et al. (2020) posit that individuals have a fixed preference relation, but that their consideration set (i.e. the set of objects they pay attention to) vary across instances of a choice set (and across individuals), leading them to choice heterogeneity. Some work has embedded the randomness in inattention as emerging via a search process (Aguiar et al. (2016); Ishii et al. (2021); Kovach and Ülkü (2020). Much of the recent literature, such as Kashaev and Aguiar (2022); Dardanoni et al. (2020, 2023); Barseghyan et al. (2021a,b); Abaluck and Adams-Prassl (2021) focuses on identifying both what is paid attention to as well as preferences using choice data along. Other papers, often in a field setting have attempted to use some proxy for attention, such as Goeree (2008). Relative to much of this literature we directly measure extensive attention. Our paper also highlights the importance of intensive attention, a factor typically ignored in this literature. Moreover, rather than working solely with aggregate data, we also consider randomness on the individual level.

There are two papers in this literature that are closest to what we do. Aguiar et al. (2023) looks at heterogeneity in choice while manipulating the frame of choice sets (the difficulty in understanding payoffs — building off of ideas in Caplin et al. (2011)). Subjects each made two choices, each from a different choice set, with choice sets varying in size from 2 to 6, with all options being lotteries. They collected data from over 2000 individuals. Thus, although they do not directly observe attention, as the difficulty in understanding outcomes increases, attention should naturally be expected to decline. They document that limited attention plays an important role in choice, but that even conditional on attention heterogeneity in preferences remain. Unlike us, they do not have enough data to conduct an individual level analysis, nor can they directly observe intensive and extensive attention.

Ellis et al. (2024) measure directly both random choice and consideration sets at an individual level. They have a two part experiment. In part 1 subjects face 11 choice sets of sizes 2 to 4, and each set is faced 10 times. Each option is an ambiguous lottery, outcomes are known, but not probabilities. Subjects can choose a subset of the choice set to learn about the probabilities. Within that subset they engage in a a mentally taxing counting task to learn the probabilities. They then have to report the probabilities, and if they report incorrectly, they won't be paid. In the second part they face the same choice sets but where the lotteries are now risky (probabilities are known). They have around 300 subjects.

The way of measuring extensive attention is quite different. In their experiment consideration is a very deliberate choice. In our model, conditional on looking at an additional option, which option is looked at is essentially random, although how long to look at an option is a choice. This lead to very different conclusions — they find most subjects limit their consideration even in doubletons, in constrast to our findings. They can't measure intensive attention. They also find far more deterministic choice as well. About a quarter of the subject are deterministic or nearly so. The choice and attention tasks are de-linked, and so they cannot conduct the same analyses we do.

Outside of economics there is also a growing literature that explicitly models (and measures) intensive attention, primiarily in cognitive science. Often building off of the drift-diffusion model (Ratcliff (1978)) models such as attentional drift diffusion (Krajbich et al. (2010)) and gaze-weighted linear accumulator mode (Thomas et al. (2019)) suppose that dwell times impact information acquisition, which in turn directly impacts choice (see Krajbich (2019) for a review). There is evidence that gaze may be drawn to preferred options (Sepulveda et al. (2020); Westbrook et al. (2020); Gluth et al. (2020), but also that longer dwell times influence preferences directly (Bhatnagar and Orquin (2022)). This body of literature typically combines both theory as well as experimental tests, often leveraging eye-tracking. Distinctly, within economics, there has been recent theoretical work relating time data to notions of random choice (Fudenberg et al. (2018); Webb (2019); Baldassi et al. (2020); Cerreia-Vioglio et al. (2023)). Much of the empirical work in cognitive science focuses on choices from binary menus, but some studies allow for larger choice sets. For example, Reutskaja et al. (2011); Thomas et al. (2019) both consider larger choice sets, ranging in size from 4 to 36. Choice sets were presented either 25 or 50 times respectively. Reutskaja et al. (2011) focuses on matching their data with different notions of search processes. Thomas et al. (2021) find that a gaze-driven evidence accumulation explains their data best. These papers have data quite similar to ours. They different in several ways, both in design and analysis. First, these papers tend to focus on options that are consumption items (e.g., food) rather than lotteries, making it harder to do analysis of things like first order stochastic dominance. Second, they tend not to analyze the properties of random choice and directly relate them to the properties of attention.[3]

3. Experimental Design

3.1. **Options.** We first describe how we constructed the options and choice sets we used in the experiment. The options presented to subjects were binary lotteries (x, y), where x paid off if state 1 was realized and y if state 2 was realized. States 1 and 2 were ex-ante equally likely.

To build the choice sets, we mapped each lottery to the cartesian plane, where each potential option is represented by a point, and each dimension representing the payoff in one of the two states. We constructed three budget lines in this 2-dimensional space, denoted B_1, B_2, B_3, as illustrated in Figure 1. The line B_1 has slope -1, B_2 is shallower (slope > -1) and B_3 is steeper (slope < -1). These budget lines were chosen to intersect at a single point with payoffs (500,500), which can be considered a "focal" or "central" option, behaviourally, while simultaneously acting as a "safe" or "risk-free" option, financially. [4] We decided to choose options along these different budget set in order to ensure that (i) choices were similar to existing experiments inducing agents to choose from menus of risky lotteries, (ii) we have a sufficient number of options to construct multiple large choice sets, (iii) the options were complicated enough (i.e. we have two states) that subjects will eventually tire of looking at options and (iv) the

[3]One recent paper within economics that leverages visual attention carefully is Li and Camerer (2022). However they are focused on understanding the exogenous impact of visual attention on choice, and do not seek to directly relate it to randomness in choice.

[4]This aspect mirrors experiments such as Choi et al. (2007) which induce subjects to choose along a budget set in order to test the revealed preference paradigm.

options are not so complicated that they introduce significant new biases (e.g., we did not want to have non-equal priors over states).

Our aim was to include pairs of options that are ordered by first order stochastic dominance relation between them, to facilitate understanding when dominated choices are made. An option A stochastically dominates another option B if each payoff in B is smaller than at least one payoff in A. First, we chose equally spaced points on B_1: these are points 7-9 and 11-13 in Figure 1. Then, we created points on B_2 (which is shallower than B_1) as follows. For each point in B_1, we created a point in B_2 that is either stochastically dominant (if B_2 is above B_1, in the right portion of the graph) or stochastically dominated (if B_2 is below B_1): these are points 20-25. For instance, option 20 is stochastically dominated by option 7, and 25 stochastically dominates 13.

Then, using a similar procedure, for each point in B_2, we create a new point on B_3: these are points 1-6. Finally, for each point on B_3 we use a similar procedure to build a new set of points on B_1: these are points 14-19. We do this to generate a total of 25 possible options. All 25 points and 3 budget sets are illustrated in Figure 1, which also includes the index for each option. For instance, option 1 corresponds to the payoffs (168,931) while option 24 corresponds to (816,279).

3.2. **Choice Sets.** The aim of our choice set construction was twofold: first to have strict subset and superset relations, as many of the theoretical frameworks we consider have simple and testable implications based on such relations; second, to have a lot of heterogeneity in choice set size, to allow us to understand how choice and attention patterns vary along that dimension.

To that end we constructed twenty-seven choice sets: six of size two, six of size three, six of size six, three of size thirteen, three of size fourteen, and three of size seventeen. The full flow-chart of choice sets and their component options is described in Figure 2. Going forward, we will refer to choice sets 1-18 (those of sizes 2, 3, and 6) as being "small", and choice sets 19-27 (those of sizes 13, 14, and 17) as being "large".

3.3. **Choice Interface.** The subjects were shown 180 slides during the course of the experiment, each of which corresponded to one of the 27 choice sets detailed above. Subjects did not directly observe the options in each choice set. Instead, in each period, subjects saw a 5x5 square grid of blue "cells", with the middle cell missing (so, 24 cells

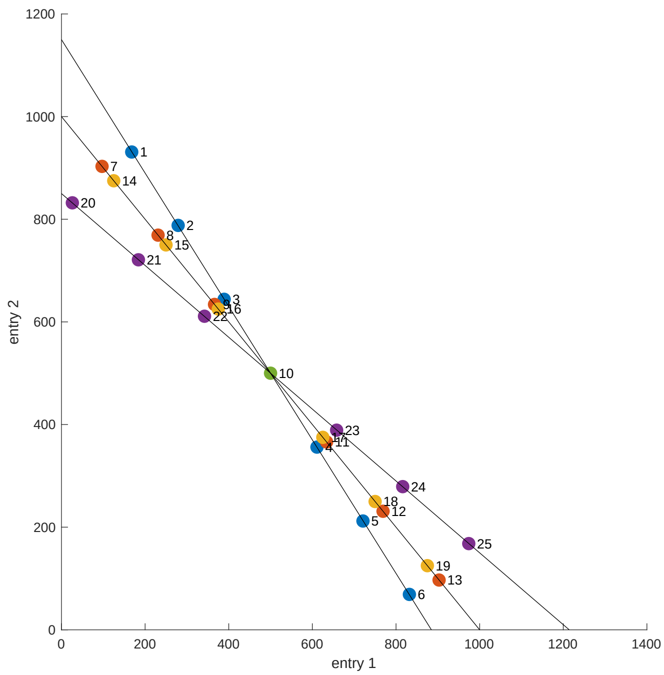


FIGURE 1. This shows the 25 options that were used to create choice sets. Each option is comprised of an (x,y) pair - the x-axis corresponds to the payoff in state one (the first entry), the y-axis corresponds to the payoff in state two (the second entry). Options are referred to by their labels (1-25) from this graph for the rest of the paper. They fall in six groups of four with stochastic dominance relationships (e.g. 24 stochastically dominates 18 and 12, which in turn both stochastically dominate 5).

in total). Each choice set contained a maximum of 17 options. Therefore, the cells that contained lotteries were marked "XXX,XXX". The remaining cells did not contain lotteries and were left blank.

The experiment used a gaze-contingent paradigm. Subjects had to look at (i.e., fixate on) a square for at least 60 milliseconds in order to reveal the lottery payoffs associated with that cell. The lottery's two payoffs remained visible until the subject moved their eyes away, at which point the payoffs reverted to "XXX,XXX". For instance, one pair of payoffs was (342,611) and another was (625, 375). This protocol ensured that subjects could not be aware of lotteries that they didn't overtly attend to. This was done to

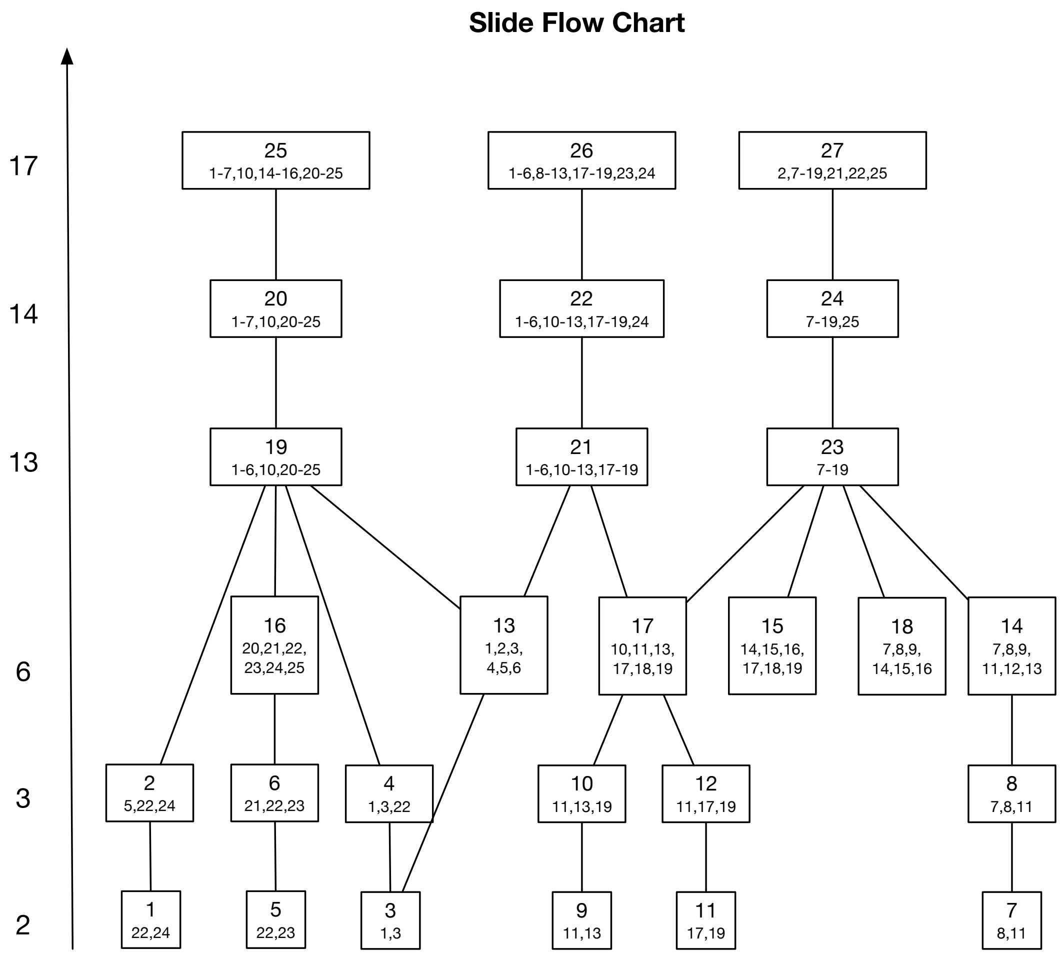


FIGURE 2. The full set of choice sets. The vertical axis denotes the size of the choice set (total number of options). Within each box, the larger number on top is the index for that choice set (i.e., its "name"). The smaller numbers below are the indexes of the options include in this choice set. A line connecting two choice sets indicates a strict subset/-superset relationship.

avoid subjects becoming aware of gambles through their peripheral vision, so that the data is an accurate reflection of which options each subject was aware of.

Subjects had unlimited time to examine the items in each choice set before deciding. Once they were ready to indicate their choice, subjects had to look at their choice on the screen and press the spacebar. Between experimental periods a fixation cross appeared at the center of the screen for 1 second before the next set of lotteries would

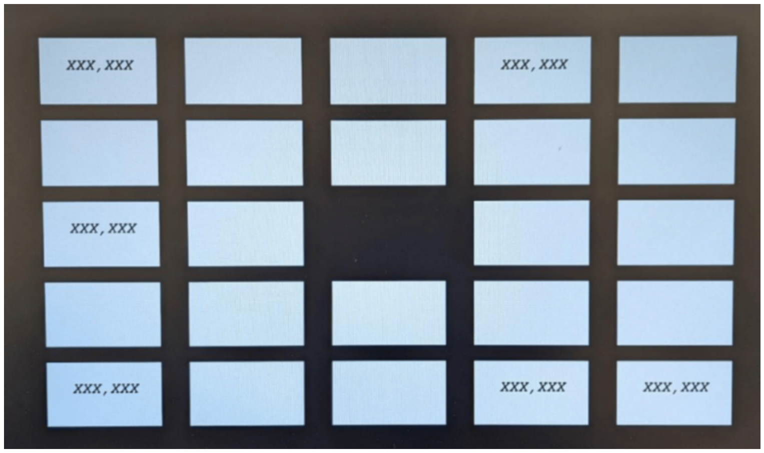


FIGURE 3. Subjects start each trial looking at the center of the screen. Their starting screen will have no gambles revealed

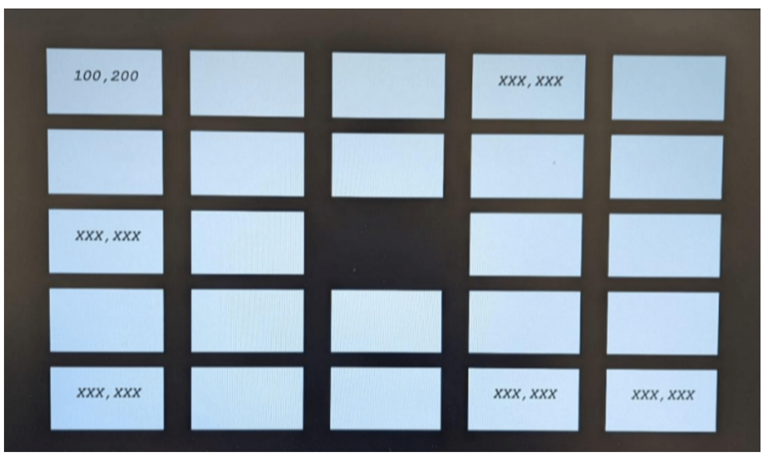


FIGURE 4. Once a subject starts attending to options, they can only look through them one-by-one. The subject in this image is attending to the top left option on the screen which shows the option (100,200).

appear to ensure that the next slide started with the subject staring at the empty spot in the grid.

At the end of the experiment we selected one of the slides at random for each subject (it did not have to be the same slide for each subject). We then took the subject's choice for that randomly selected slide, and flipped a digital coin to determine the subject's payoff (each payoff with a probability of 0.5). Points were converted to dollars at a rate of 50 points per dollar.

Subjects saw 180 choice sets. Choice sets of size 2, 3 and 6 were each seen 5 times (90 total slides). Choice sets of size 13, 14 and 17 were each seen 10 times (90 total slides). Each individual was shown the same choice sets, but in a random order, and with randomized spatial locations of the lotteries (at the individual-trial level).

3.4. **Protocols.** 50 subjects were recruited from The Ohio State University's Experimental Economics Laboratory using ORSEE. Data were collected one subject at a time using an SR Research Eyelink 1000 Plus. Subjects were paid an 5 USD show-up fee plus the payoff from one random period described above. Subjects earned 15.51 USD on average. Subjects ranged in age from 19 to 27 years with an average of 21.56 years. The study was approved by the Ohio State University's institutional review board (Protocol #2013B0583).

Upon arrival at the lab, subjects gave consent to participate. They completed a 9-point calibration and validation procedure with the eye tracker. They read through the instructions on the computer screen, completed a short comprehension quiz, and went through 10 practice trials to become familiar with the interface. After completing the practice trials, subjects faced the choice task, as described previously.

4. Summary Statistics

Here we present summary statistics on choice and some measures of attention to establish the fact patterns that will drive much of the later analysis.

4.1. **Choice Randomness.** We first look at three different measures of choice randomness. The first is *choice reversal.* This measure is constructed at the individual-choice set level. For a given individual and choice set, we calculate the probability that for a randomly selected pair of instances, the two choices were different from one another. Formally, we define for a choice set s and an individual i

$$\text{Choice Reversal}_{is} = \frac{2\sum_j \sum_{-j} \mathbb{I}_{i,s,j,-j}}{n_s(n_s-1)} \tag{1}$$

where n_s is the number of instances that choice set s was shown to subjects, and $\mathbb{I}_{i,s,j,k}$ is an indicator function that takes the value 1 if the choices made by individual s in instances j and k of choice set s are different from one another. It takes the value 0 otherwise. Thus, if an individual has a degenerate distribution of choices, then they exhibit no choice reversals. In contrast, a uniform distribution of choice over a size of set n would imply that if we randomly choose any two instances of a choice set, then if there are enough instances, choice reversals should occur in approximately $\frac{n-1}{n}$ of the pairs.

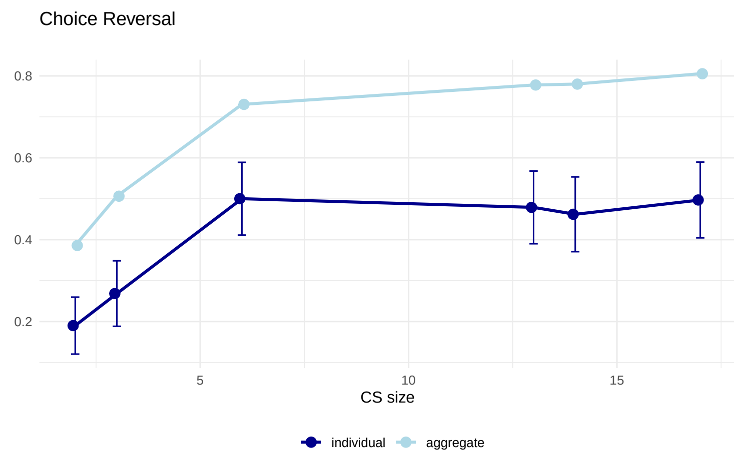


FIGURE 5. Choice Reversals by Choice Set

The cross-sectional mean values for this variable (across subjects) is plotted for each choice set in Figure 5. Choice reversal is a common occurrence across all choice sets,

We can also look at the choice reversal measure across individuals and choice sets. This breakdown is in Figure 6, and we can see that choice reversals are common among most individuals (although there is heterogeneity that will be explored in Section 7 — e.g., there are several subjects who still maintain remarkable consistency across choice sets)

We can compare choice reversal to our two other measures of choice randomness, to ensure that our results are not specifically driven by our choice of measure. The first alternative measure is *share chosen* which is the fraction of options in a particular set that were chosen over all instances of the choice set. Formally, we define for a choice set s and an individual i

$$\text{Share Chosen}_{is} = \frac{\sum_j \mathbb{I}_{i,s,j}}{n_s} \tag{2}$$

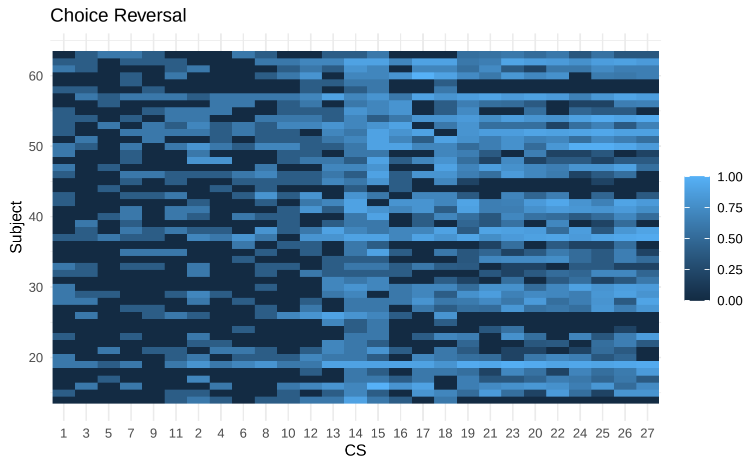


FIGURE 6. Choice Reversals by Choice Set and Individual

where n_s is the number of options in choice set s, and $\mathbb{I}_{i,s,j}$ is an indicator function that takes the value 1 if agent i selected option j in choice set s at least once. It takes the value 0 otherwise.

The second alternative measure is *choice entropy* which is a measure of the dispersion among chosen items in a given choice set. Formally, we define for a choice set s and an individual i

$$\begin{aligned} \text{Choice Entropy}_{is} &= -\sum_j p_{i,j,s} \ln(p_{i,j,s}) \\ &\text{where} \\ p_{i,j,s} &= \frac{\sum_k \mathbb{I}_{i,s,k}}{n_s} \end{aligned}$$

is the probability that agent i selects option j from choice set s across all instances and n_s is the number of options in choice set s.

We plot the correlation of the three measures of randomness in choice at the individual-choice set level in Figure 7. It is clear that all three measures are strongly positively correlated. The details of our alternative measures are in the Appendix 10

4.2. **Attention.** We now consider measures of extensive and intensive measures of attention. Informally, we consider extensive measures of attention to be whether or not an item was looked at, while we consider intensive measures of attention to be related to frequency or duration of how an item was looked at.

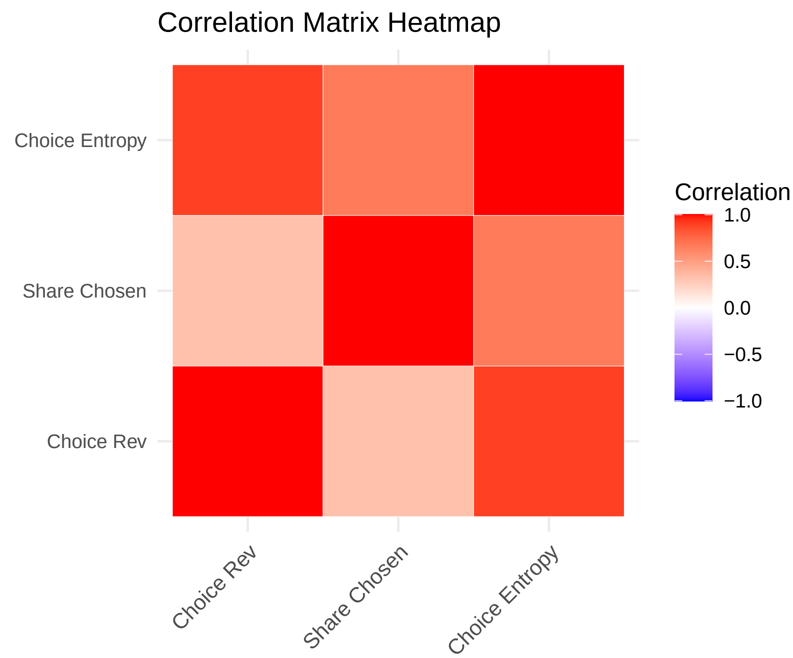


FIGURE 7. Correlation of Choice Measures

4.2.1. *Extensive Attention.* Subjects, on average display strikingly different patterns in extensive measures of attention between small and large choice sets. The first measure we look at is *Attention Share* - what fraction of choices were looked at in a particular instance of a choice set? Formally:

$$\text{Attention Share}_{i,j,s} = \frac{\sum_k \mathbb{I}_{i,j,s,k}}{n_s}$$

For a subject i, and an instance j of a choice set s, $\mathbb{I}$ is an indicator variable that takes a value 1 if item k was looked at in instance j of choice set s, and zero otherwise. n_s is the number of items in choice set s. Overall, subjects tend to look at around 80% of objects in a choice set, although this varies widely by choice set size (which we explore more in Section 6).

As an be seen in Figure 8 individuals typically look at every item in small choice sets, but only about two-thirds of items in large choice sets (aggregating across all individuals, the share looked at is 1 for all sizes, so that line is not included).

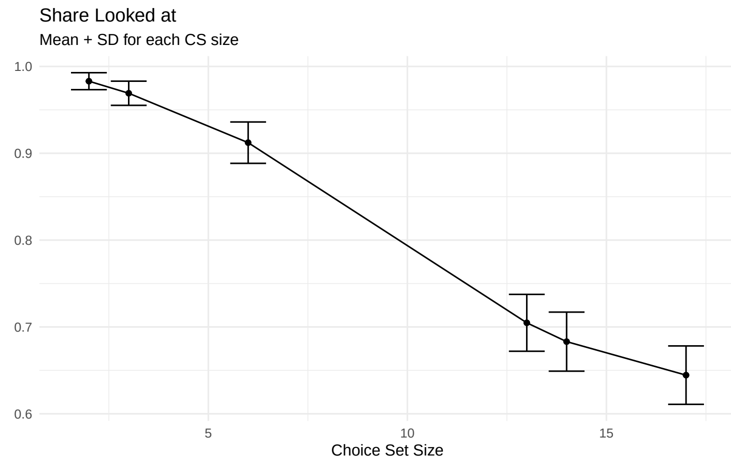


FIGURE 8. Attention Share by Choice Set Size

We also consider a measure of heterogeneity in what is looked at in Appendix 10 - Attention Reversal. Formally, define a variable $A_{i,x,j,s}$ which takes a value of 1 if item x was looked at in instance j of choice set s by agent i and takes a value of 0 otherwise. Then define attention reversal as:

$$\text{Attention Reversal}_{i,s,j,k} = \frac{\sum_x A_{i,x,j,s} \times (1 - A_{i,x,k,s}) + (1 - A_{i,x,j,s}) \times A_{i,x,k,s}}{n_s}$$

It indicates the probability that for a pair of instance j and k of a choice set s, an individual i looked at an item in one instance, but not the other, where n_s is the number of elements in choice set s.

4.2.2. *Intensive Attention.* Our second kind of attention we measure is intensive attention. Here there are two key measurements we look at. The first is dwell time, which is the *amount of time* an agent looks at a particular choice or choice set. The second is fixation, which is the *number of times* an agent looks at option.

For each, we can create two measures of attention randomness: the first is an average per-option level of the measure within a set, while the second is an entropy of the measure across options within a set.

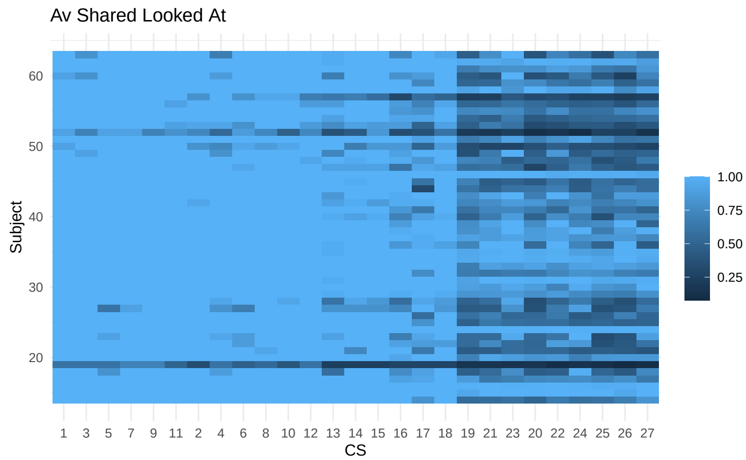


FIGURE 9. Attention Share by Choice Set and Individual

Figures 10 and 11 show the summary statistics for dwell time averages (note that the average dwell time pooling all data looks the same as the average individual here, so the former figure has only a single line). Formally, for a choice set s and an individual i, we define average dwell time as:

$$D_{i,s} = \frac{\sum_k \sum_x d(x|S_{i,k,s})}{n_s}$$

where $d(x|S_{i,k,s})$ is the amount of time spent by subject i on option x in instance k of choice set j and n_s is the number of instances of choice set s. On average, subjects spend more time looking at each item in small choice sets than they do in large one, averaging over 1 second per item in small sets, and under 0.5 seconds per item in large sets. There is a fair amount of heterogeneity in this measure in small sets, but (given the lower mean) quite a bit less in large choice sets.

We have three additional measures of intensive attention, which we define here, but we relegate the discussion of these measures to the appendix, as they are qualitatively similar to the measures already discussed. First, we define Dwell Entropy for an individual i and a choice set s as follows:

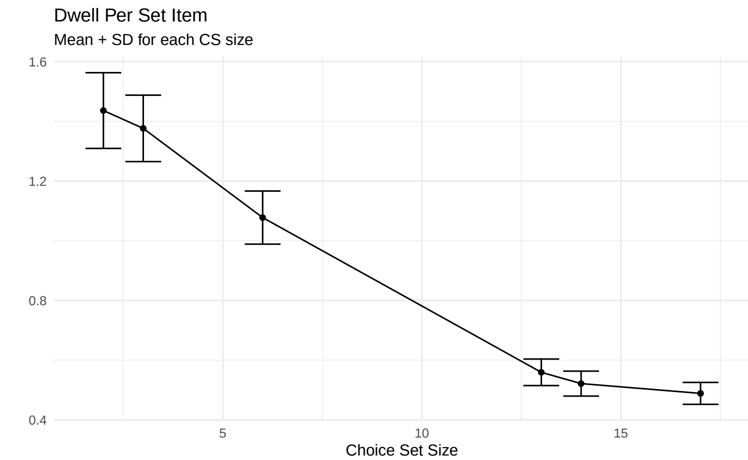


FIGURE 10. Dwell Time by Choice Set Size

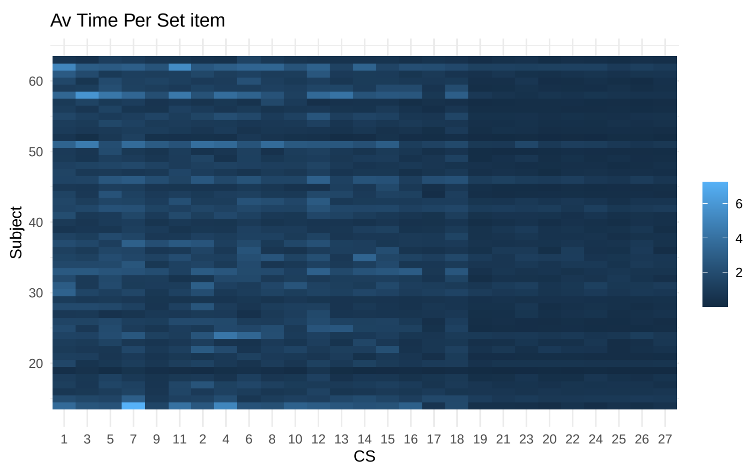


FIGURE 11. Dwell Time by Choice Set and Individual

$$\begin{aligned} \text{Dwell Entropy}_{is} &= \frac{-\sum_j p^d_{x,i,s} \ln(p^d_{x,i,j,s})}{-\frac{1}{n_x}\ln(\frac{1}{n_x})} \\ &\text{where} \\ p^d_{x,i,s} &= \frac{\sum_k d(x|S_{i,k,s})}{\sum_k \sum_x d(x|S_{i,k,s})} \end{aligned}$$

where k indexes the instance of choice set s and x indexes the options in choice set s and n_x is the number of options in choice set s. Observe that to control for natural trends in entropy due to choice set size, we normalize the measure dwell entropy by the maximal possible entropy (a uniform distribution of dwell times).

Second we define Average Fixation as:

$$F_{i,s} = \frac{\sum_k \sum_x f(x|S_{i,k,s})}{n_s}$$

where $f(x|S_{i,k,s})$ is the number of times item x was looked at in instance k of choice set s by individual i.

Finally we define Fixation Entropy as:

$$\text{Fixation Entropy}_{is} = \frac{-\sum_j p^f_{x,i,s} \ln(p^f_{x,i,j,s})}{-\frac{1}{n_x}\ln(\frac{1}{n_x})}$$

where

$$p^f_{x,i,s} = \frac{\sum_k f(x|S_{i,k,s})}{\sum_k \sum_x f(x|S_{i,k,s})}$$

where k indexes the instance of choice set s and x indexes the options in choice set s and n_x is the number of options in choice set s. Observe that to control for natural trends in entropy due to choice set size, we normalize the measure fixation entropy by the maximal possible entropy (a uniform distribution of fixations across items).

As is evident from Figure 12 the various measures of attention described above are all positively correlated, so we will relegate some of the results on attention to the appendix.

One interesting pattern that arises is that subjects tend to look at higher value options for longer. In Figure 13 we plot the total dwell time against the gamble numbers. The average dwell time is in blue (plotted on the left axis), while the expected value is in orange (plotted on the right axis). There is a high correlation between the two. measures, particularly around very high and very low EV options.

5. Can Attention Rationalize Randomness?

We now turn to understanding the relationship between choice randomness and the various notions of attention we measure. In particular, we want to understand

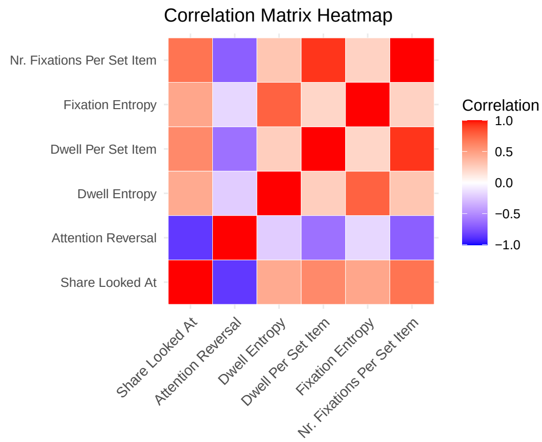


FIGURE 12. Correlation of Attention Measures

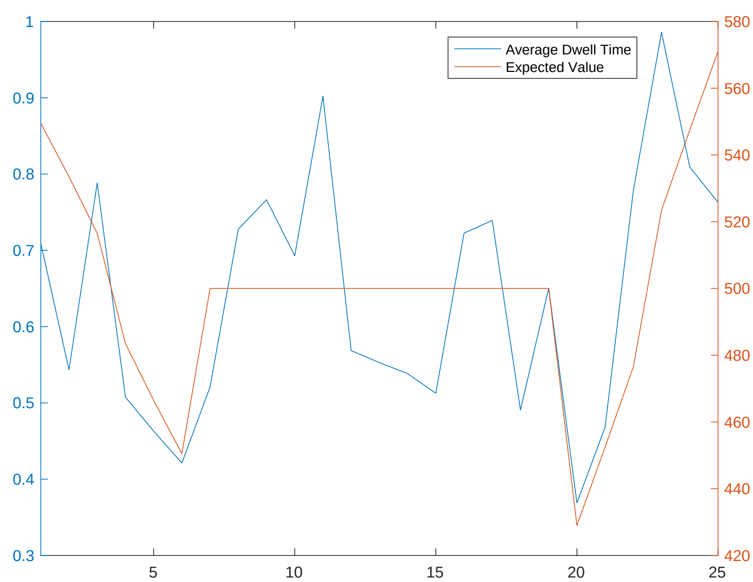


FIGURE 13. Correlation of Attention Measures

to what extent attention can help "rationalize" the randomness we observe in choice. Such an exercise is motivated both by simple intuitions — e.g., individuals cannot choose items they do not look at, and so looking at different options across differences instances of the same choice set can lead to different choice, even in the presence of fixed preference — as well as formal models developed to highlight those intuitions (e.g., Cattaneo et al. (2020)). In doing so, we will focus on one of our three measures of randomness: choice reversals (in the appendices we conduct analogous exercises, with qualitatively similar results, using our other measures of choice randomness).

5.1. **Extensive Attention.** Our first exercise will be to ask what fraction of choice reversals can be rationalized due to the fact that subjects, when choosing differently across different instances of the choice set, may have done so only because the option they choose in Instance 1 was not looked at in Instance 2, or vice versa. If this is true, we can attribute the choice reversal to attention. This assigns the maximal number of choice reversals to attention. It could be the case that x was chosen in Instance 1, and y in Instance 2, and x was not looked at in Instance 2. We would attribute such a reversal to attention. But it could be that even if x had been looked at in Instance 2, it still would have not been chosen. Of course, we cannot identify such counterfactual behavior. Thus, our measure should always be interpreted as attempting to understand the maximal amount of randomness that could be ascribed to extensive attention, given the data.

Recall that we measure choice reversals by considering pairwise combinations of two different instances of a choice set. In order to make our numerical analysis easier, we will reverse code choice reversals. If the choices are the same from the both instances in the pair, then this observation is coded as a 1 (i.e. the choices matched, and there is no choice reversal) otherwise we code the value for this combination as a 0 (indicating a choice reversal). We call a 1 an *unconditional match* (UM) and a 0 an *unconditional reversal* (UR); unconditional here means we do not condition on attention.

We next ask, for a given pairwise combination, whether extensive attention *could* rationalize a reversal. Consider the pairwise combination of two instances of a choice set: $S_{i,j}$ and $S_{i,j'}$ (recall i indexes the choice set, j indexes the instances of that set). Consider $c(S_{i,j})$ and $c(S_{i,j'})$. Denote the set of items that were paid attention to (i.e. looked at) in a given choice set instance as $A(S)$. If $c(S_{i,j}) \in A(S_{i,j'})$ and $c(S_{i,j'}) \in A(S_{i,j})$ then assign a value of 1 to the variable *cross attention* (CA), and otherwise 0. Notice that if we observe cross attention (i.e. the value is 1 for the pair), this means that the choice in instance j was looked at in instance j' and vice versa. Clearly, if the choices in the two instances are the same, CA must take a value of 1. Moreover, if CA is 0, we *must* observe a choice reversal (since agents must have paid attention to an option in order to choose it). If CA is equal to 1 for the pair, yet UR is equal to 0, this means that although the agent paid attention to both choices in both choice sets,

	UM/UR	CM/CR	Fraction Unexplained ($\frac{\text{CR}}{\text{UR}}$)
	Individual Average		
Overall	.6/.4	.71/.29	.725
Small Choice Sets	.68/.32	.72/.28	.875
Large Choice Sets	.48/.52	.7/.3	.577
	Representative Agent		
Overall	.29/.71	.44/.56	.796
Small Choice Sets	.46/.54	.50/.50	.926
Large Choice Sets	.21/.79	39/.61	.787

Table 1. Extensive Attention and Choice Randomness

they still chose differently. In other words, extensive attention cannot rationalize this choice reversal.

We then regress UR on CA (without a constant). The the estimated coefficient has a simple interpretation: it tells us what fraction of the time when both choices were looked at in both choice sets, did we observe the choice matching. If we observe a choice that matches we call this a *conditional match* (CM), while if it doeesn't match we refer to it as a *conditional reversal* (CR).

The following table reports the results for this analysis, where we conduct the analysis for each individual separately, and then average the results across individuals.

Last, we can compute the degree of residual randomness left unexplained by attention. This is simply the ration of conditional reversals to unconditional reversals (i.e. the fraction of reversals remaining after controlling for extensive attention divided by the total fraction of reversals). We refer to this as *fraction unexplained.*

Table 1 reports the results of this analysis. It does so for both all sets, and separately for small and large choice sets. Moreover, it conducts the analysis both at the individual level, and then reports the average across all individuals, as well as at the representative agent level, where we simply pool all choices by all individuals and treat them as coming from a single decision-maker.

First, we focus on the individual analysis. Consistent with the analysis done previously, across all possible choice sets, the average individual is likely to reverse their choice between two instances of a choice set about 40% of the time. Similarly, consistent with our previous analyses, we see far less randomness in small choice sets than in large choice sets. Strikingly though, the rates of conditional matching/conditional reversals are roughly similar. In other words, regardless of the choice set size, if an agent looks at both choices in both choice sets, we observe a reversal at the same rate across choice set size. This is an important fact we will discuss more when we focus on heterogeneity across choice set sizes (Section 6).

Because there are more unconditional reversals in large choice sets this means more reversals in large choice sets are coming from situations where both choices were *not* paid attention to in both choice sets. In other words, attention can rationalize more of the reversals in large choice sets compared to small choice sets. The third column actually provides a numerical value of this — extensive attention can rationalize around 42% of reversals in large choice sets, but only 28% of reversals in small choice sets.

The data indicate that although extensive attention can rationalize a large fraction of choice randomness, especially in large choice sets, the majority of randomness *cannot* be explained by extensive attention.

If we assume the data come from a representative agent, the story is somewhat different. First, we start out with far more randomness — the unconditional reversal rate is .71, 175% of the rate of the average individual — and we see similarly large increases in both small and large choice sets. Moreover, even conditional on attention the proportion of choices which feature reversals is much larger, about 190% larger aggregating across all choice sets. Interestingly enough, although we see more both more unconditional and conditional reversals, the fraction of reversals that cannot be explained by extensive attention for the representative agent is not so different from that of the average individual for small choice sets, but is actually *higher* for large choice sets (and so overall).

5.2. **Intensive Attention.** Of course, even if a subject looks at an option, they may not consider it for very long. As discussed previously, existing evidence indicates that the length of time of that individuals consider an item may also correlate strongly with whether an item is chosen. We will now attempt to understand whether length of time spent considering an item (as measured by dwell time on the item) can help rationalize choice reversals as well, in the same way that measures of extensive attention could.

	$d(c_{i,j'}\vert S_{i,j'}) > d(c_{i,j}\vert S_{i,j'})$	$d(c_{i,j'}\vert S_{i,j'}) < d(c_{i,j}\vert S_{i,j'})$
	Overall ($n = 6897$)	
$d(c_{i,j}\vert S_{i,j}) > d(c_{i,j'}\vert S_{i,j})$	5511=79.9%	646=9.4%
$d(c_{i,j}\vert S_{i,j}) < d(c_{i,j'}\vert S_{i,j})$	670=9.7/%	70=1%
	Small Choice Sets ($n = 2469$)	
$d(c_{i,j}\vert S_{i,j}) > d(c_{i,j'}\vert S_{i,j})$	1746=70.7%	318=12.9%
$d(c_{i,j}\vert S_{i,j}) < d(c_{i,j'}\vert S_{i,j})$	354=14.3%	51=2.1%
	Large Choice Sets ($n = 4428$)	
$d(c_{i,j}\vert S_{i,j}) > d(c_{i,j'}\vert S_{i,j})$	3765=85%	328=7.4%
$d(c_{i,j}\vert S_{i,j}) < d(c_{i,j'}\vert S_{i,j})$	316=7.1/%	19=0.4%

Table 2. Intensive Attention and Choice Randomness: Individual Average

In order to do so, we will now restrict ourselves to considering pairs of choice set instances where i) we observe a choice reversal, and ii) both choices were paid attention to in both choice sets. Denote $d(x|S_{i,j})$ as the dwell time on option x in instance j of choice set S_i. We will say that intensive attention rationalizes the choice reversal if $d(c(S_{i,j})|S_{i,j}) > d(c(S_{i,j'})|S_{i,j})$ and $d(c(S_{i,j'})|S_{i,j'}) > d(c(S_{i,j})|S_{i,j'})$. In other words, when a subject chooses an option in an instance, do they look at it for longer at the instance, compared to alternatives that were chosen in some other instance. If they do, we suppose that intensive attention rationalizes the choice reversal, because individuals choice tracks their dwell time. We do not require that the chosen option in either choice set be the option that was looked at longest.

The following table shows the fraction of times that for a given choice set, for two instances j and j', whether the choice for a given instance was looked at for longer in that instance, than the choice for the other instance. In other words, when a subject chooses an option in an instance, do they look at it for longer at the instance, compared to alternatives that were chosen in some other instance.

Table 2 conducts this exercise separately for each individual, and then averages across individuals. It shows that the vast majority of the time choice reversals mirror dwell time allocations — intensive attention can rationalize around 80% of choice reversals which cannot be rationalized by extensive attention. The patterns across large and small choice sets mimic those found for extensive attention — intensive attention rationalizes fewer choice reversals in small choice sets compared to large choice sets.

	$d(c_{i,j'}\|S_{i,j'}) > d(c_{i,j}\|S_{i,j'})$	$d(c_{i,j'}\|S_{i,j'}) < d(c_{i,j}\|S_{i,j'})$
	Overall ($n = 636752$)	
$d(c_{i,j}\|S_{i,j}) > d(c_{i,j'}\|S_{i,j})$	$543391 = 85.3\%$	$42661 = 6.7\%$
$d(c_{i,j}\|S_{i,j}) < d(c_{i,j'}\|S_{i,j})$	$46277 = 7.3\%$	$4423 = .7\%$
	Small Choice Sets ($n = 258759$)	
$d(c_{i,j}\|S_{i,j}) > d(c_{i,j'}\|S_{i,j})$	$200484 = 77.5\%$	$25086 = 9.7\%$
$d(c_{i,j}\|S_{i,j}) < d(c_{i,j'}\|S_{i,j})$	$29440 = 11.3\%$	$3749 = 1.4\%$
	Large Choice Sets ($n = 377993$)	
$d(c_{i,j}\|S_{i,j}) > d(c_{i,j'}\|S_{i,j})$	$342907 = 91.7\%$	$17575 = 4.6\%$
$d(c_{i,j}\|S_{i,j}) < d(c_{i,j'}\|S_{i,j})$	$16837 = 4.5\%$	$637 = 1.7\%$

Table 3. Intensive Attention and Choice Randomness: Representative Agent

We can conduct the same exercise using the representative agent approach and pooling all the data. Table 3 reports the results. Surprisingly, we find that intensive attention tracks choice reversals even more closely in the aggregate data compared to the individual data.

Thus, once we account for extensive attention, intensive attention can rationalize choice reversals at an extremely high level.

5.3. **Discussion.** The results indicate that that both kinds of attention are important for understanding choice randomness. Moreover, in conjunction, they can rationalize the vast majority of choice reversals.

Result 1. *At the individual level, extensive attention can rationalize around 28% of choice reversals, while intensive attention explains about* $72.5 \times 79.9 = 58$*% of all choice reversals. Combined they can explain around 86%.*

Excluding either form of attention would lead to a significant loss of understanding of the origin of choice reversals. Moreover, between the two forms of attention, approximately 86% of choice reversals can be rationalized. Thus, understanding and properly modeling both forms of attention can help us understand a significant fraction of random choice behavior.

The interpretation of both extensive and intensive attention, and their relationship to choice reversals, will depend on whether we consider them exogenous or endogenous factors. For example, many models of random attention (e.g., Manzini and Mariotti (2014)) assume that attention is allocated randomly across items, and so the formation of the consideration set is exogenous to preferences. In this case, it is natural to ask, once we control for the exogneous allocation of extensive attention, can we explain some of the choice reversals. Similarly, if we think of intensive attention as exogenously assigned, then understanding how they predict randomness is also straightforward.

However, it is likely that the assignment of both extensive and intensive attention is not fully exogenous relative to preferences. This can be most clearly seen from Figure 13, which shows that the expected value and the average dwell time co-vary to a large degree. If we think of preferences as driving dwell time, then our results on intensive attention are hardly surprising at all — they reveal the preference of the individual. This is distinct from dwell time on a choice set revealing the utility differences between items (Alós-Ferrer et al. (2021)).

Similar concerns can be raised for extensive attention as well — if we think of extensive attention as being generated by a search process, encountering high value items early in the search leads to individuals to truncate search early.

6. Heterogeneity Across Choice Sets

We now document the second important pattern that we observe in the data — that the mechanisms, and outcomes, of choice look very different depending on choice set size.

6.1. **Choice Heterogeneity.** Such heterogeneity can clearly be seen in the summary data. For example, one can see in Figure 5 that choices become much more random in larger choice sets (and this is mirrored by our other measures of choice randomness). Recall that the cross-sectional mean values for this variable (across subjects) is plotted for each choice set in Figure 5. Choice reversal is a common occurrence across all choice sets, but is significantly higher in the larger choice sets (size ≥ 6) where choice reversals occur about half the time, as opposed to small choice sets (size ≤ 3) where choice reversals occur about a quarter of the time.

This is also consistent with the numbers reported for UM/CR in Table 1, where we see the probability of an unconditional reversal rise from .32 in small choice sets to

.52 in large choice sets. Interestingly enough, within our data, rather than there being a relatively continuous change in choice randomness with choice set size, we can see that randomness is similar for the smallest choice sets, and then also similar for all the largest choice sets (with the break being between size 3 and size 6). Thus, it seems as if there are two separate choice processes that are occurring.

6.2. **Attentional Heterogeneity.** This is also consistent with attentional data. An increase in choice randomness is mirrored by a decreased ability of the respondents to pay attention to the entire choice set. As is evident from Figure 8 and Figure 9, in choice sets of size 2 and 3, the average attention share is very close to one - every item is typically looked at. In choice sets of size six, the average drops a bit, but is still quite high (well about 0.8). However, for large choice sets the attention share drops precipitously, to around 0.65. Unsurprisingly, with such a high average, there is limited heterogeneity in this measure across individuals for small choice sets. A handful of subjects account for all of the cases where not all items were looked at. In larger choice sets there is substantial heterogeneity across agents, with very few agents still looking at all options. As with the choice data, we see that the share data is quite lumpy, rather than moving continuously with choice set size. It looks as if sets of size 2 and 3 are quite similar, sets of size 6 somewhat lower, and then sets of size 13-17 all similar, in terms of share looked at.

Moreover, the degree of randomness in what is paid attention to in the choice set also grows. As discussed in Appendix 10, we can measure the degree of randomness in attention by looking at attentional reversals — the natural analog of choice reversals. For any pair of instances of a choice set, we can ask what is the probability that an item paid attention to in one is also paid attention to in the other. Figure 20 shows this measurement and demonstrates that the degree of randomness in what is paid attention to also grows with choice set size, and again we see three separate groups — choice sets of size 2-3, those of size 6 and then those of size 13-17, with the data being very similar within each group.

Our measure of intensive attention appears to move a bit more smoothly with choice set size, as seen in Figure 10. at least for the smallest choice sets, where we can see a decrease in dwell time per option as we go from choice sets of size 2 to 3 to 6. However, we observe little change in dwell time per option for the largest choice sets, perhaps indicating that respondents want to devote a minimum amount of time to each option they look at.

As the choice set size grows, attention (both extensive and intensive) explains more and more of the choice randomness. As Table 1 demonstrates, moving from small (size 2-6) to large (size 13-17) choice sets hardly changes the fraction of conditional reversal (which measures, conditional on extensive attention, how random are individuals). However, because individuals are more random in larger choice sets, there is less randomness that cannot be explained by extensive attention (e.g., at the individual level this falls from around 87% in small choice sets to 58% in large choice sets).

In order to unpack these relationships, we will draw on one additional data source - the number of times the respondent looks at a given item. As Figure 22 shows, we also observe a declining number of fixations per option as the choice set gets larger. At small choice sets each item is looked at around twice; while in large choice sets items are looked at only once. Thus, in small choice sets subjects look back and forth between the options, while in larer choice sets subjects will look at an item often only once before moving on.

6.3. **Discussion.** In sum, it seems we can observe two distinct choice processes taking place depending on the choice set size.

Result 2. *Subjects randomness in small choices occurs despite agents paying attention to almost all items in the choice set, and viewing each multiple times. In large choice sets much more of the randomness is due to extensive attention, and subjects fail to consider all options, and the options they do consider they typically only look at once.*

In the two smallest choice set sizes we observe very similar behavior. Subjects typically look at all the options. Their gaze moves back and forth between the options, looking at each option at least twice, and spending between 1.2 and 1.5 seconds on each option. Thus, it appears that subjects observe, understand, and consider each option carefully. Despite this, we still observe quite a bit of randomness — subjects are about 30% likely to change their choice across two instances of a choice set; which is essentially driven by their chance of changing choice conditional on paying attention to both items. Essentially we observe only “residual” randomness. However, this residual randomness is relatively correlated with intensive attention.

In contrast, large choice sets capture a very different mechanic. Here subjects tend to look at only a subset of the options. Moreover of the options they do look at, they look at it at most twice, and typically once, and for a relatively short time (less than half a second). This data is what would emerge from a sequential search situation, where subjects look at each object for a brief time, and only occasionally recalling a past item (often to choose it). Interestingly, in large choice sets, the behavior, conditional on the consideration set, is quite similar to small choice sets. The key difference is that due to the size of the choice set the agent's search process terminates before looking at all options, and so additional randomness occurs due to attention.

Choice sets of size 6 are interesting. In terms of choice randomness they look remarkably similar to the much larger choice sets (of size 13 and up). In contrast, in terms of attentional behavior, they lie closer to choice sets of size 2 and 3.

At the choice set level, it is fairly clear that we observe a negative correlation between randomness and either intensive or extensive attention, and moreover, that across choice sizes we observe a positive correlation between intensive and extensive attention.

7. Heterogeneity Across Individuals

We next turn to documenting the fact that models of random choice should be able to capture both within and across person heterogeneity.

7.1. **Individual Average versus Representative Agent Approaches.** First, the degree of randomness across individuals is almost as large as the degree of randomness within an individual. We can see this in several ways. Looking at figure 5, it appears that in general, the degree of total randomness from the representative agent approach is about twice that of the randomness from the average individual. Table 1 confirms this, with the randomness from the average individual being $\frac{.4}{.71} = 56\%$ of the total choice randomness from the representative agent approach. Thus, between-person randomness form about 45% of the aggregate randomness.

Result 3. *About 56% of aggregate randomness in choice is due to individual randomness, with the rest due to between person randomness*

Looking at Figure 6 provides a more nuanced view of what is happening. In particular, we can observe that in small choice sets, the degree of choice randomness across individuals is fairly similar. However, as the choice sets get larger, we observe far more heterogeneity, with some individuals becoming much more random, while other individuals are still relatively consistent in their choice, even in large choice sets.

However, we actually observe that at the individual average level, attentional randomness takes up a larger fraction of total choice randomness in larger choice sets compared to smaller — 65% versus 59%. The reason is that individual randomness grows when we go from small to large choice sets (at a rate slightly faster than the growth of between-person randomness).[5]

Because we know that attention can rationalize a large fraction of choice randomness, the fact we observe a large amount of choice heterogeneity between people raises the question whether this is matched by attentional heterogeneity. Figure 9 demonstrates that just as with choice randomness, heterogeneity in extensive attention across individuals, as measured by differences in the size of the consideration set, is negligible at small choice sets (because almost all individuals look at all items all the time), but larger in large choice sets; we can observe some individuals still consider almost all items, while others only look at a small fraction. However, Figure 20 shows that the degree of attentional randomness of the representative agent is larger than that of the average individual, but only barely so. Essentially randomness in attention across individuals is fairly close to randomness within a person.

On the other hand, we observe the opposite pattern for intensive attention — heterogeneity falls as the choice sets get larger. Figure 11 provides the intuition for why: in small choice sets some individuals look at each option for a relatively short time, while others look for a relatively long time. In contrast, in large choice sets all individuals look at options for a relatively short amount of time.

The ability of extensive attention to rationalize choices is larger at the average individual level compared to the representative agent approach. Looking at Table 1, across all choice sets, extensive attention explains around 27.5% at the individual average level, while only 21.4% at the representative agent level. This is due in no small

[5]Notice that the representative agent approach doesn't only embed between person randomness. However, because each individual only has 5 or 10 choices, but we have 50 individuals, individual randomness doesn't increase the representative agent approach very much.

part to the fact that there are far more choice reversals, conditional on attention, at the representative agent level compared to the individual average level — we observe almost twice as many reversals conditional on attention at the representative agent level.

The gap between individual average and representative agent approaches is much larger at small choice sets compared to large choice sets. In small choice sets the levels are 12.5% versus 7.4%, while in large choice sets the levels are 43.3% versus 11.3%. Interestingly, just as we don't observe big differences in the conditional choice reversal levels by choice set size, the gap between the individual average and representative agent choice reversals, conditioning on attention, does not change much with choice set size.

In contrast, we actually see that intensive attention can rationalize more choice reversals at the representative agent approach compared to the individual agent level. The gap between the two is relatively constant across choice sets.

7.2. **Individual Types.** Given the fact that across person variation is important in understanding both choice randomness and attention, we now turn to understanding whether there are consistent patterns and types that we can classify individuals into.

We now look at what kind of "types" we observe in choice behavior. We identify five types:

- Individuals who exhibit little randomness in both large and small choice sets (consistent choosers)
- Two types of individuals who exhibit larger amount of randomness in all choice sets (random choosers and extreme random choosers)
- Individuals who exhibit randomness in small choice sets but little randomness in large choice sets (defaulters)
- Individuals who are not random in small sets, but are in large sets (overwhelmed)

These types are easily seen in Figure 14. The horizontal axis classifies individuals by their degree of choice reversals in small choice sets, while the vertical axis does the same for large choice sets. Consistent choosers are those clustered in the lower-left hand area. Defaulters, i.e. those who are random in small choice sets, but in large choice sets default to one option, are those who are close to the horizontal axis but not the vertical axis. Overwhelmed individuals are those who are consistent in small choice sets, but not large choice sets, and so are close to the vertical axis, but not the horizontal. Random choosers are those who are far from either axis, with extreme random choosers being far up and to the left.

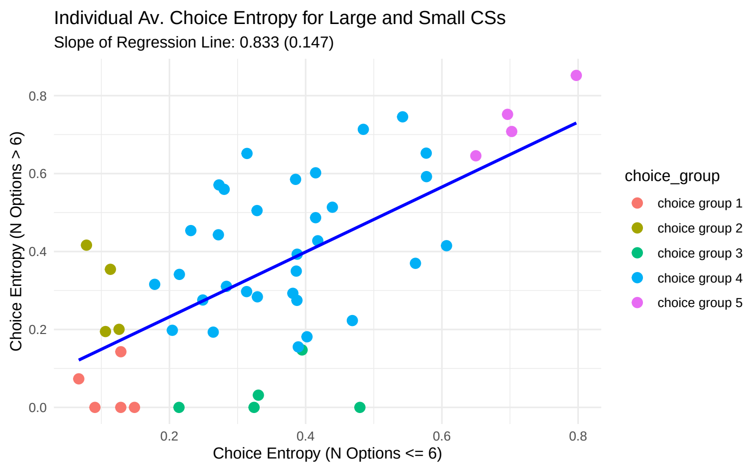


FIGURE 14. “Types” in Choice Reversals

As Figures 25 and 26 show, the choice types are consistent across all types of measurement. In particular, those figures show that using the types, as derived from measures of choice reversals, but then doing the exact same plots using the alternative measures of randomness, leads to almost exactly the same figures.

We can also look at whether there are attentional types. Figure 16 plots the degree of extensive attention (measured as share looked at) at the individual level for small and large choice sets. One can see that behavior in the two kinds of choice sets are highly correlated, but that all points lie below the 45-degree line, indicating that more attention is always paid in smaller choice sets. Three key types emerge. The first are individuals who look at (almost) all the items in small choice sets, and also look at a large fraction of options (greater than about 70%) in large choice sets. The second group are individuals who look at most, but not all options in small choice sets, and look at between roughly 50 and 70% of the options in large choice sets. The last group, which is much smaller, consists of individuals who have much lower attention in both kinds of choice sets.

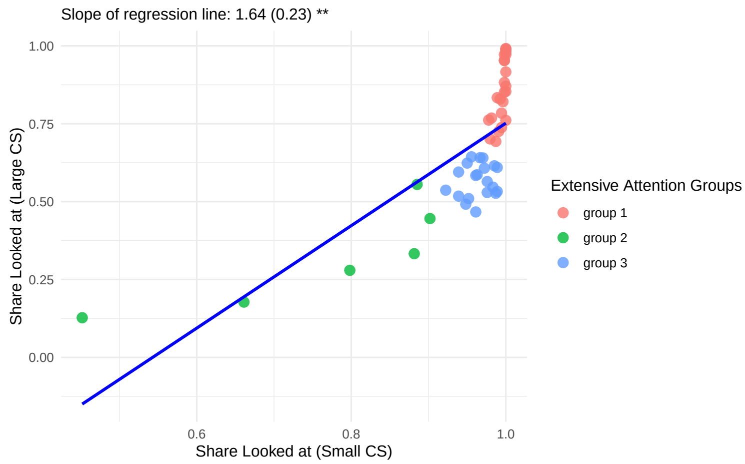


Figure 15. “Types” in Extensive Attention

We can conduct a similar exercise with intensive attention (dwell time per item). First, we wnat to highlight that dwell times in small choice sets is not nearly as concentrated as extensive attention shares. Again, we see a farily strong correlation in behavior across large and small choice sets, but with dwell time in small choice sets always being longer than those in large choice sets. We see roughly four main types. First, there are individuals who tend to not spend much time in either kind of choice for a given item. Second, we have individuals who spend a fairly average amount of time in small choice sets, but are relatively slow in large choice sets. The third group exhibits the reverse behavior, while the fourth group spends lots of time in both kinds of choice sets.

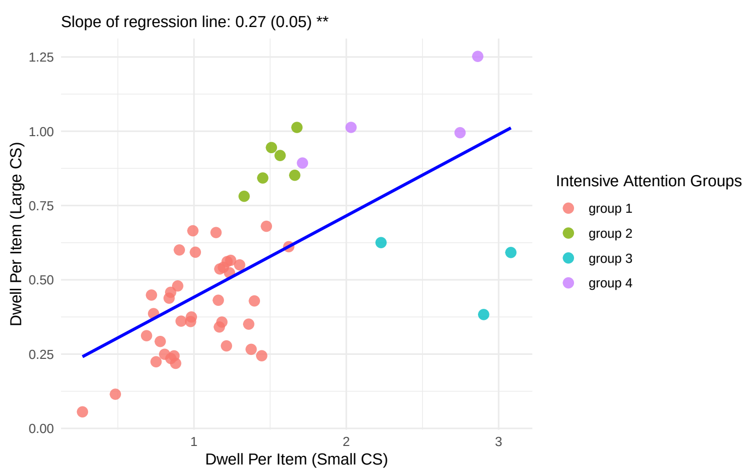


Figure 16. “Types” in Intensive Attention

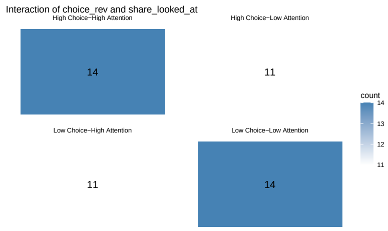
Given that we can classify individuals into both types based both on their choice behavior as well as their attitudes towards attention, We now look at the relationship, at the individual level, between choice behavior and attention. Because we only have 50 subjects, and we will be trying to relate multiple dimensions of behavior to each other, we will use coarser measures for classification than we do for the individual dimensional analysis. In particular, we will classify each individual into high or low randomness, high or low extensive attention, and high or low intensive attention.

We classify individuals as high or low choice randomness by taking their average choice reversal proportion over all choice sets and then doing a median split into those with high choice reversals and low choice choice reversals. For extensive attention, we similarly compute for each individual their average share looked at over all choice sets, and again did a median divide. For intensive attention we conduct the same exercise but now using dwell time per item.

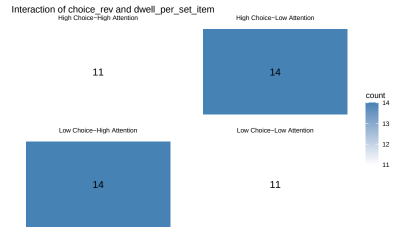
Figure 17 displays the results of such an analysis. Surprisingly, there is very little correlation between choice randomness and either extensive or intensive attention at the individual level. Essentially whether an individual is very random in choice or not is roughly independent of whether they look at many options, or how long they look at options. In other words, just because an individual looks more "rational" in their choice, in that they have less randomness and fewer reversals, does not mean that they satisfy the assumption of considering all options. There are individuals who exhibit little randomness in choice, who nevertheless don't look at all options (they find their favorite and stop consistently). Similarly, there are agents who are quite good at considering all options carefully (both in looking at all of them and looking at them for a longer the average time) who are still quite random in their choice.

In contrast, the final panel of Figure 17 shows that at the individual level intensive and extensive attention are highly correlated. Individuals who tend to look at many item also tend to look at any given item for a long period of time. Thus, attentional capacity of both kinds seems to be an individual trait, but a trait that is unrelated to the degree of randomness that a given individual exhibits.

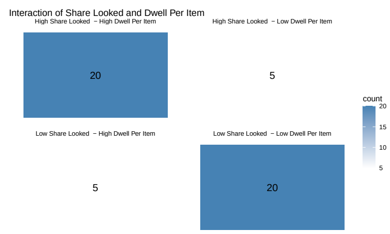
7.3. **Discussion.** Our data indicate that a key feature of any model accounting for random choice much account for the fact that both within and between person heterogeneity are important features to account for. Moreover, such a model needs to



(A) Correlation of Choice and Extensive Attention at Individual Level



(B) Correlation of Choice and Intensive Attention at Individual Level



(C) Correlation of Extensive and Intensive Attention at Individual Level

FIGURE 17. Correlation of Types

plausibly allow for within-person randomness that is not accounted for by extensive attention; but also allow for the fact that non-extensive attention driven randomness is more important at the between-person level. In contrast, intensive attention is more predictive at the representative agent level than the average individual level.

We find that respondents can be classified into types based on choice randomness, extensive attention and intensive attention. Although randomness in large and small choice sets is correlated, we can find specific types of individuals who exhibit relatively more/less randomness in either large or small choice sets. That said, most individuals exhibit choice reversals 20% of the time in both large and small choice sets. Similarly, we can identify different types of individuals using their extensive and intensive attention — although here behavior in small and large choice sets is highly correlated.

However, individuals' choice type is essentially uncorrelated with either of their attentional types. But within attention, extensive and intensive types are correlated, so that individuals who tend to look at many options also tend to look at any given option for a longer time.

8. Value of Attention

We last examine the potential benefits of attention. In other words, is greater attention related to individuals obtaining higher payoffs. To the extent that i)randomness reduces payoffs (because an individual does not always choose the highest utility option) and ii) greater attention reduces randomness (as we have seen); it seems clear that we should expect increased attention to improve payoffs. Of course, this depends on whether, with greater attention, agents more consistently choose the utility maximizing option.

Given that we do not observe the utility function, it is hard to assess whether increased attention increases the chances of choosing the utility maximizing options. Although we cannot necessarily identify the most preferred option, we can conduct several analysis to try and understand these issues.

First, we can identify some options, which, if chosen, cannot maximize any "reasonable" preferences — those preferences which respect monotonicity, i.e., first order stochastic dominance (FOSD). Subjects often make choices from menus which have options that are first order stochastically dominated. We are interested in whether subjects are more likely to make a FOSD violation after we control for attention. We can also look at the variation in this across choice sets of difference sizes — we may expect larger sets to experience more violations because attention is spread more thinly.

For this analysis, we only consider trials in which a FOSD violation can occur. This brings us down from 180 trials to 110 trials for each subject. Across 50 subjects we have 5500 observations in which a FOSD violation can occur. Conditioning on these choices, we report the following table by set size.

Confirming the intuition there is a positive correlation (.911 Pearson's r correlation), between set size and the proportion of FOSD violations. This is suggestive evidence that likelihood of a FOSD violation is related to the size of a menu. However, this analysis leaves out a key covariate: the number of options in which a FOSD violation could be made. If subjects are purely choosing randomly, then a higher proportion of dominated options would lead to a higher violation rate of FOSD. Confirming this notion, the correlation is even stronger between the proportion of options dominated and the rate of FOSD violations (0.962 Pearson's r correlation)

	FOSD Vio.	FOSD Vio. w/ Attention	Ratio	Ratio w/ Attention
Aggregate Choice Sets	16.9%	15.1%	931/5500	710/4701
Small Choice Sets	9%	7.75%	135/1500	108/1394
Large Choice Sets	19.9/%	18.2%	796/4000	602/3307

Table 4. FOSD Violations

	S-Mon. Vio.	S-Mon. Vio. w/ Attention	Ratio	Ratio w/ Attention
Aggregate Choice Sets	6.49%	4.33%	292/4500	154/3551
Small Choice Sets	5.8%	4.59%	29/500	22/469
Large Choice Sets	6.58%	4.28%	263/4000	132/3082

Table 5. Statewise Monotonicity Violations

Next in our analysis, we consider that subjects did not have to pay attention to the entire menu and it is often the case that subjects only paid to a subset of the menu. We can now consider whether a FOSD violation occured, simply because subjects did not explicitly attend to the dominating option. If subjects weren't aware of the dominant option, did they commit a violation of FOSD? Therefore we refine our notion of FOSD to when subjects did pay attention to the dominating option.

When accounting for attention, about one-fourth of all FOSD violations disappear dropping from 931 to 710. However, rate of violation only decreases modestly from 16.9% to 15.1% suggesting that most FOSD violations are not due to unawareness of the dominant option. The strong relation between set size and FOSD remains (.9231 Pearson's r correlation) as well as the strong relation between FOSD violation rates and proportion of dominated options in the attention set (.9043 Pearson's r correlation) indicate that subjects may be engaged in violations of dominance due to pure randomization.

Instead of considering FOSD violations, we can also consider a more transparent form of violation — statewise monotonicity. Not surprisingly, we observe fewer violations.

Although we observe fewer choice mistakes when we control for attention, this is not to say that we observe more attention being correlated with better choices, because the number of potenial violations also goes down.

Result 4. *Controlling for extensive attention only has a negligible impact on the proportion of violations of dominance (either first order or statewise) that we observe.*

This is consistent with models where attention, both extensive and intensive, would be endogenously chosen and so "more difficult" situations would receive more attention, but the additional attention doesn't fully compensate for the added difficulty.

9. Conclusions

9.1. **Generalizing Beyond our Study.** Our study is lab-based with a very specific domain. The benefit of our design is that we can measure multiple dimensions of attention in a clean way using eye-tracking. There are very few environments where we can have this level of detailed measurement. In addition, we believe eye-tracking is a minimally invasive way of measuring attention. Rather than inducing consideration sets through explicit costs or via complexity, in our study we can can measure the natural tendencies of what subjects look at. Because our options are lotteries, the domain of choice corresponds to many other studies. This allows us to more directly compare our findings. Moreover, it allows us to understand mistakes (violations of First Order Stochastic Dominance) in a clean way.

On the other hand, there is always the question of external validity — how might we expect our findings to translate outside of the specifics of our experiment. For example, in the field, options are often not presented so easily as in our study — individuals need to exert much more effort in order to bring items into a consideration set, through search on the internet, walking up and down the aisle at a grocery store, or visiting multiple car dealerships.

We doubt that the exact numerical findings in our study will translate to all other settings. For example, the exact relative importance of extensive versus intensive attention in explaining randomness, or the precise breakdown of within versus between person heterogeneity.

However, we believe key qualitative findings are likely to be true in many other settings:

- Both extensive and intensive attention play important roles in explaining choice randomness
- Randomness in small and large choice sets occur via different processes
- Within and between person randomness are driven to different degrees by attention
- Attentional capacity and choice randomness are not correlated at the individual level
- Attention is not correlated with fewer mistakes

9.2. **Discussion.** Our results reinforce messages from existing empirical work, but also highlight additional areas where additional research needs to be done. Like many existing papers (e.g., Ellis et al. (2024); Barseghyan et al. (2021b); Dardanoni et al. (2020) we document the importance of attention when trying to understand random choice. However, because we jointly study both extensive and intensive attention, our work highlights the fact that joint modelling of both forms of attention is a key area where there is little work. Cerreia-Vioglio et al. (2023) is one of the few papers we know of that tries to jointly accommodate decision time as well as consideration sets.

More generally, although there is a large body of work on extensive attention in economics, much less has been done on intensive attention. In particular the focus has been on total time to decide given a choice set (Fudenberg et al. (2018)) rather than dwell time on an individual item. The tight relationship between dwell time and choice randomness that we document points out the importance of incorporating insights from models that explicitly deal with dwell time at the option level, such as attention drift diffusion models (Krajbich et al. (2010)).

When developing these models, it also appears that we should think about attention as an endogenous outcome, rather than exogenous constraints. We do not observe attention significantly improving choice quality — this could be due to the fact more attention is allocated to more difficult problems.

Similarly, our work highlights the fact that we need to allow for both randomization due to preferences and attentional heterogeneity in order to effectively capture patterns of choice (see Kashaev and Aguiar (2022) for recent theoretical work in this direction). We find that individuals still exhibit a large degree of randomness even after controlling for extensive attention.

We want to highlight that our results do not directly tell us what drives randomness after controlling for extensive attention. The residual randomness could be due to a variety of reasons. We observe violations of stochastic dominance fairly often, which is inconsistent with existing models of deliberate randomization due to preferences (Cerreia-Vioglio et al. (2019)). In this regard we are similar to Agranov et al. (2023), who also find violations of dominance in conjunction with randomization.

The mechanism underlying random choice should be flexible enough to account for the fact we see fairly different patterns of process and choice in large and small choice sets. Smaller choice sets involve all items attracting extensive attention, and the subjects often looking back and forth between them. In contrast, in larger choice sets we observe subjects engaged in a process closer to search with little recall, consistent with Reutskaja et al. (2011).

It appears that extensive and intensive attention are related at the individual level, so that subjects who consider many options also tend to consider options for longer. However, choice randomness seems essentially uncorrelated with how much attention a subject exhibits. Empirical approaches also need to be concsious of the fact that there is additional preference based and attention based randomness across individuals as opposed to within individuals.

References

Abaluck, Jason and Abi Adams-Prassl, "What do consumers consider before they choose? Identification from asymmetric demand responses," *The Quarterly Journal of Economics*, 2021, *136* (3), 1611–1663.

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

______, Paul J Healy, and Kirby Nielsen, "Stable randomisation," *The Economic Journal*, 2023, *133* (655), 2553–2579.

Aguiar, Victor H, Maria Jose Boccardi, and Mark Dean, "Satisficing and stochastic choice," *Journal of Economic Theory*, 2016, *166*, 445–482.

______, ______, Nail Kashaev, and Jeongbin Kim, "Random utility and limited consideration," *Quantitative Economics*, 2023, *14* (1), 71–116.

Alós-Ferrer, Carlos, Ernst Fehr, and Nick Netzer, "Time will tell: Recovering preferences when choices are noisy," *Journal of Political Economy*, 2021, *129* (6), 1828–1877.

Baldassi, Carlo, Simone Cerreia-Vioglio, Fabio Maccheroni, Massimo Marinacci, and Marco Pirazzini, "A behavioral characterization of the drift diffusion model and its multialternative extension for choice under time pressure," *Management Science*, 2020, *66* (11), 5075–5093.

Barseghyan, Levon, Francesca Molinari, and Matthew Thirkettle, "Discrete choice under risk with limited consideration," *American Economic Review*, 2021, *111* (6), 1972–2006.

______, Maura Coughlin, Francesca Molinari, and Joshua C Teitelbaum, "Heterogeneous choice sets and preferences," *Econometrica*, 2021, *89* (5), 2015–2048.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, 1995, *63* (4), 841–890.

Bhatnagar, Roopali and Jacob L Orquin, "A meta-analysis on the effect of visual attention on choice.," *Journal of Experimental Psychology: General*, 2022, *151* (10), 2265.

Bhattacharya, Debopam, "Nonparametric welfare analysis for discrete choice," *Econometrica*, 2015, *83* (2), 617–649.

Block, Henry David and Jacob Marschak, "Random orderings and stochastic theories of response," Technical Report, Cowles Foundation, Yale University 1959.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience and consumer choice," *Journal of Political Economy*, 2013, *121* (5), 803–843.

Brady, Richard L and John Rehbeck, "Menu-dependent stochastic feasibility," *Econometrica*, 2016, *84* (3), 1203–1223.

Bushong, Benjamin, Matthew Rabin, and Joshua Schwartzstein, "A model of relative thinking," *The Review of Economic Studies*, 2021, *88* (1), 162–191.

Caplin, Andrew, Mark Dean, and Daniel Martin, "Search and satisficing," *American Economic Review*, 2011, *101* (7), 2899–2922.

Cattaneo, Matias D, Xinwei Ma, Yusufcan Masatlioglu, and Elchin Suleymanov, "A random attention model," *Journal of Political Economy*, 2020, *128* (7), 2796–2836.

Cerreia-Vioglio, Simone, David Dillenberger, Pietro Ortoleva, and Gil Riella, "Deliberately stochastic," *American Economic Review*, 2019, *109* (7), 2425–2445.

———, Fabio Maccheroni, Massimo Marinacci, and Aldo Rustichini, "Multinomial logit processes and preference discovery: inside and outside the black box," *The Review of Economic Studies*, 2023, *90* (3), 1155–1194.

Choi, Syngjoo, Raymond Fisman, Douglas Gale, and Shachar Kariv, "Consistency and heterogeneity of individual behavior under uncertainty," *American economic review*, 2007, *97* (5), 1921–1938.

Conte, Anna, John D Hey, and Peter G Moffatt, "Mixture models of choice under risk," *Journal of Econometrics*, 2011, *162* (1), 79–88.

Dardanoni, Valentino, Paola Manzini, Marco Mariotti, and Christopher J Tyson, "Inferring cognitive heterogeneity from aggregate choices," *Econometrica*, 2020, *88* (3), 1269–1296.

———, ———, ———, Henrik Petri, and Christopher J Tyson, "Mixture choice data: revealing preferences and cognition," *Journal of Political Economy*, 2023, *131* (3), 687–715.

Davis-Stober, Clintin P, Nicholas Brown, and Daniel R Cavagnaro, "Individual differences in the algebraic structure of preferences," *Journal of Mathematical Psychology*, 2015, *66*, 70–82.

Dwenger, Nadja, Dorothea Kübler, and Georg Weizsäcker, "Flipping a coin: Evidence from university applications," *Journal of Public Economics*, 2018, *167*, 240–250.

Ellis, Keaton, Emel Filiz-Ozbay, and Erkut Ozbay, "CONSISTENCY AND HETEROGENEITY IN CONSIDERATION AND CHOICE," 2024.

Falmagne, Jean-Claude, "A representation theorem for finite random scale systems," *Journal of Mathematical Psychology*, 1978, *18* (1), 52–72.

Feldman, Paul and John Rehbeck, "Revealing a preference for mixtures: An experimental study of risk," *Quantitative Economics*, 2022, *13* (2), 761–786.

Fudenberg, Drew, Philipp Strack, and Tomasz Strzalecki, "Speed, accuracy, and the optimal timing of choices," *American Economic Review*, 2018, *108* (12), 3651–3684.

______ **, Ryota Iijima, and Tomasz Strzalecki**, "Stochastic choice and revealed perturbed utility," *Econometrica*, 2015, *83* (6), 2371–2409.

Gluth, Sebastian, Nadja Kern, Maria Kortmann, and Cécile L Vitali, "Value-based attention but not divisive normalization influences decisions with multiple alternatives," *Nature human behaviour*, 2020, *4* (6), 634–645.

Goeree, Michelle Sovinsky, "Limited information and advertising in the US personal computer industry," *Econometrica*, 2008, *76* (5), 1017–1074.

He, Lisheng, Russell Golman, and Sudeep Bhatia, "Variable time preference," *Cognitive psychology*, 2019, *111*, 53–79.

Hey, John D, "Does repetition improve consistency?," *Experimental economics*, 2001, *4*, 5–54.

Ishii, Yuhta, Matthew Kovach, and Levent Ülkü, "A model of stochastic choice from lists," *Journal of Mathematical Economics*, 2021, *96*, 102509.

Kashaev, Nail and Victor H Aguiar, "A random attention and utility model," *Journal of Economic Theory*, 2022, *204*, 105487.

Kőszegi, Botond and Adam Szeidl, "A model of focusing in economic choice," *The Quarterly journal of economics*, 2013, *128* (1), 53–104.

Kovach, Matthew and Levent Ülkü, "Satisficing with a variable threshold," *Journal of Mathematical Economics*, 2020, *87*, 67–76.

Krajbich, Ian, "Accounting for attention in sequential sampling models of decision making," *Current opinion in psychology*, 2019, *29*, 6–11.

______ **and Antonio Rangel**, "Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions," *Proceedings of the National Academy of Sciences*, 2011, *108* (33), 13852–13857.

______ **, Carrie Armel, and Antonio Rangel**, "Visual fixations and the computation and comparison of value in simple choice," *Nature neuroscience*, 2010, *13* (10), 1292–1298.

Li, Xiaomin and Colin F Camerer, "Predictable effects of visual salience in experimental decisions and games," *The Quarterly Journal of Economics*, 2022, *137* (3), 1849–1900.

Luce, R Duncan, "Individual choice behavior.," 1959.

Machina, Mark J, "Temporal risk and the nature of induced preferences," *Journal of economic theory*, 1984, *33* (2), 199–231.

Manzini, Paola and Marco Mariotti, "Stochastic choice and consideration sets," *Econometrica*, 2014, *82* (3), 1153–1176.

Masatlioglu, Yusufcan, Daisuke Nakajima, and Erkut Y Ozbay, "Revealed attention," *American Economic Review*, 2012, *102* (5), 2183–2205.

McCausland, William J, AAJ Marley, and Clintin P Davis-Stober, "A population choice experiment for testing axioms of stochastic discrete choice," 2024.

———, Clintin Davis-Stober, Anthony AJ Marley, Sanghyuk Park, and Nicholas Brown, "Testing the random utility hypothesis directly," *The Economic Journal*, 2020, *130* (625), 183–207.

McFadden, Daniel, "The measurement of urban travel demand," *Journal of public economics*, 1974, *3* (4), 303–328.

Milosavljevic, Milica, Vidhya Navalpakkam, Christof Koch, and Antonio Rangel, "Relative visual saliency differences induce sizable bias in consumer choice," *Journal of consumer psychology*, 2012, *22* (1), 67–74.

Ratcliff, Roger, "A theory of memory retrieval.," *Psychological review*, 1978, *85* (2), 59.

Regenwetter, Michel and Clintin P Davis-Stober, "Behavioral variability of choices versus structural inconsistency of preferences.," *Psychological review*, 2012, *119* (2), 408.

Reutskaja, Elena, Rosemarie Nagel, Colin F Camerer, and Antonio Rangel, "Search dynamics in consumer choice under time pressure: An eye-tracking study," *American Economic Review*, 2011, *101* (2), 900–926.

Sepulveda, Pradyumna, Marius Usher, Ned Davies, Amy A Benson, Pietro Ortoleva, and Benedetto De Martino, "Visual attention modulates the integration of goal-relevant evidence and not value," *Elife*, 2020, *9*, e60705.

Smith, Philip L, "Stochastic dynamic models of response time and accuracy: A foundational primer," *Journal of mathematical psychology*, 2000, *44* (3), 408–463.

Smith, Stephanie M and Ian Krajbich, "Attention and choice across domains.," *Journal of Experimental Psychology: General*, 2018, *147* (12), 1810.

Thomas, Armin W, Felix Molter, and Ian Krajbich, "Uncovering the computational mechanisms underlying many-alternative choice," *Elife*, 2021, *10*, e57012.

_______, _______, _______, Hauke R Heekeren, and Peter NC Mohr, "Gaze bias differences capture individual choice behaviour," *Nature Human Behaviour*, 2019, *3* (6), 625–635.

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

Webb, Ryan, "The (neural) dynamics of stochastic choice," *Management Science*, 2019, *65* (1), 230–255.

Westbrook, Andrew, Ruben Van Den Bosch, Jessica I Määttä, Lieke Hofmans, Danae Papadopetraki, Roshan Cools, and Michael J Frank, "Dopamine promotes cognitive effort by biasing the benefits versus costs of cognitive work," *Science*, 2020, *367* (6484), 1362–1366.

10. Appendix: Summary Statistics

10.1. **Choice Randomness.** Here we provide further details of the two alternative measures we use for choice randomness. Recall that these are share of options chosen in a choice set, and the normalized (by the maximum possible) entropy of the choice set distribution. Figure 18 shows the share of options chosen by choice set size, and Figure 19 does the same for entropy.

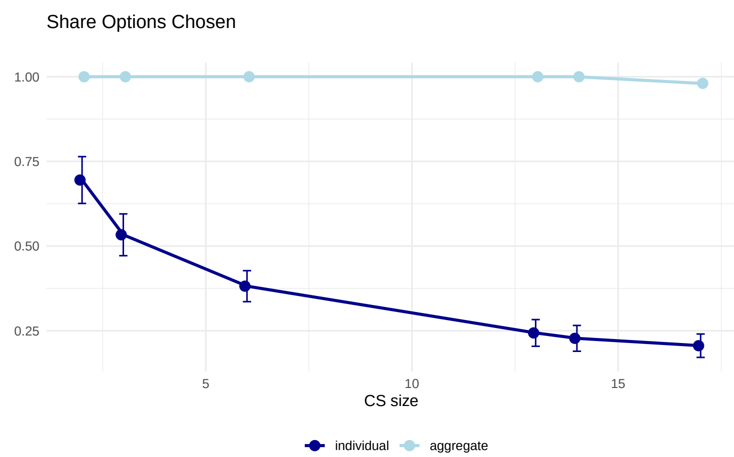


Figure 18. Choice Shares by Choice Set

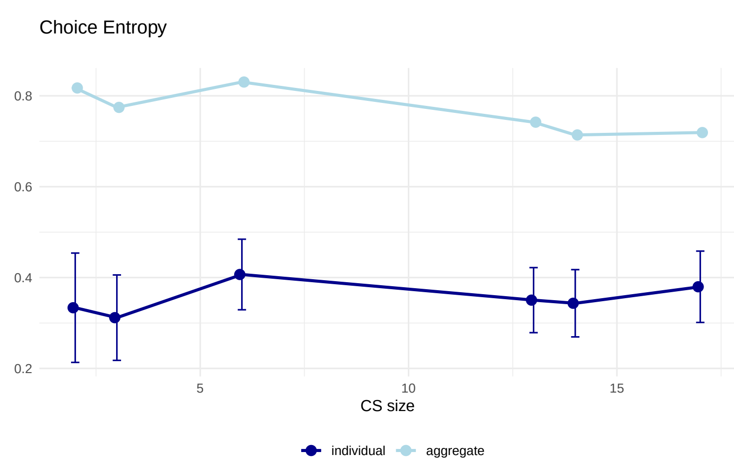


Figure 19. Choice Entropy by Choice Set

10.2. **Attention.** In this subsection we present some summary statistics for the alternative definitions of attention. In terms of extensive measures, attention reversal is presented in Figure 20. Attention reversal is increasing across choice set size, both at the aggregate and the individual level (with standard error bands). The change is statistically significant between small and large choice sets, with small sets averageting below 0.1, while large choice sets average above 0.3.

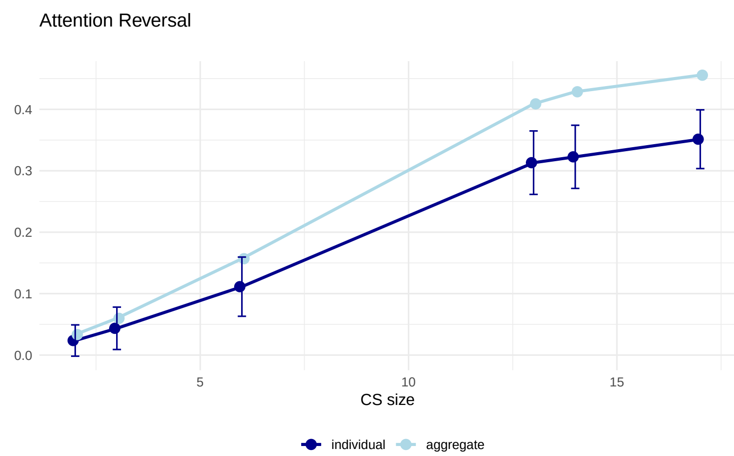


FIGURE 20. Attention Reversal by Choice Set

The remaining attentional measures are intensive. In figure 21 we plot dwell entropy both at the aggregate level, and at the individual level (along with standard error bands). While the two measures are positively correlated with one another, there appears to be a slight positive trend in the aggregate with choice set size, but a slight negative trend in the individual. The differences between small and large choice sets are not stasticically significant.

In figure 22 we plot the average number of fixations per choice set item, by choice set size across individuals (with standard errors). The number of fixations per set item is sharply decreasing from over 2 at small choice sets to only slightly above 1 at larger choice sets.

Finally in figure 23 we plot fixation entropy at the aggregate level, and at the individaul level (with standard errors). The patterns here are not dissimilar to the dwell entropy figure. There are not statistically significant differences in entropy across choice set size, and while the aggregate and individual measures are positively correlated, they again have slight trends away from one another.

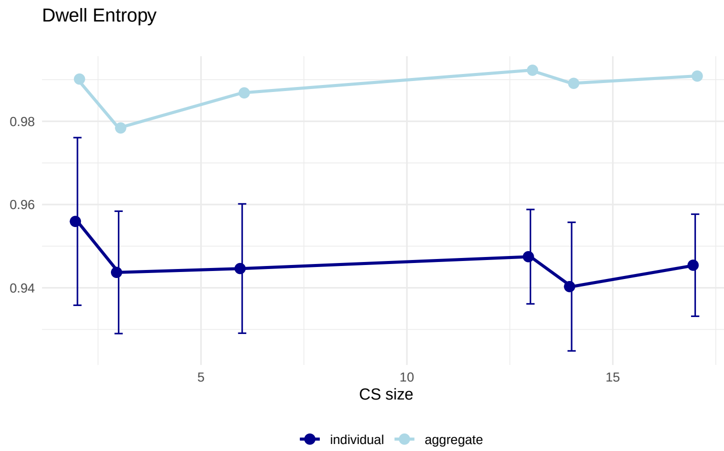


FIGURE 21. Dwell Entropy by Choice Set

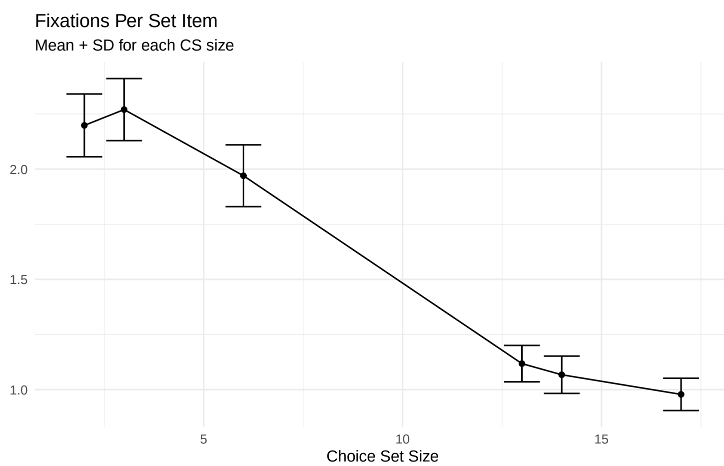


FIGURE 22. Fixations by Choice Set Size

11. Appendix: Can Attention Explain Randomness

We essentially replicate our analysis of trying to understand to what extent attention can explain randomness, but now using a distinct measure of choice randomness — normalized entropy of the choice distribution. Recall we quantify the randomness in choice for a given choice set as as the observed entropy in the distribution of choices from that choice set compared to the maximal possible entropy of the set. We then take the average of this fraction across choice sets, weighted by how often the choice set is seen by subjects. We can do this both at the aggregate level (such that for a given

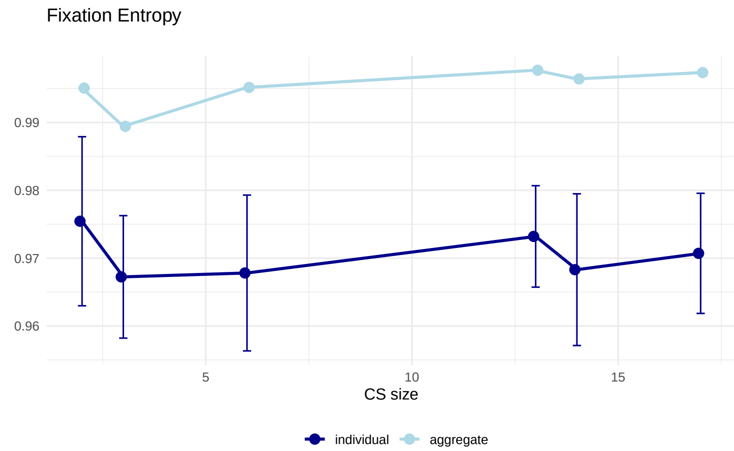


FIGURE 23. Fixation Entropy by Choice Set Size

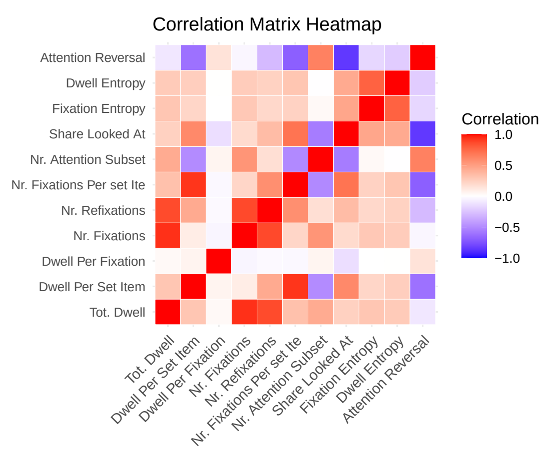


FIGURE 24. Correlation of Attention Measures: Extended

choice set we look at the distribution of choices made by everyone) or at the individual level (where we look only at the choices made by a given individual). To obtain a single statistic in the individual case, we then average this number across individuals.

Maximal entropy is occasionally constrained by the number of observations for consideration sets. We correct for this by defining the full randomness benchmark as $\min\{\log \frac{1}{n}, \log \frac{1}{x}\}$ where n is the number of choices in the choice set and x is the number of times the choice/attention set was seen.

We then do the same thing, but rather than looking at the choice set, we instead consider the set of items that was paid attention to (the consideration set). Thus, the maximum amount of entropy will typically change, since the set that is paid attention to is smaller than the choice set. By comparing the degree of randomness in consideration sets to the degree of randomness in choice sets, we can discover, if we control for attention, how much residual randomness remains. If the only reason individuals change choices is because they consider different items, then our measure of randomness should be 0% for the consideration set.

	Aggregate	Individual
Choice Set	76%	32%
Consideration Set	55%	20%

TABLE 6. Degree of Randomness in Choice from Choice Sets and Consideration Sets

Table 6 shows the results of the analysis. Considering the randomness of choice, we see that choice at the individual level is significantly less noisy than that at the aggregate level — about half as much. It is also true that choice conditional on the consideration set is also about twice as random more random at the aggregate level as at the individual level. This implies that we observe far more heterogeneity in choice at the aggregate level, suggesting an important role for heterogeneity across individuals, although this doesn't pin down whether it is attentional or preference heterogeneity.

Moreover, at the aggregate level, attention explains about $1 - \frac{55}{72} = 23.6\%$ of the randomness in observed choice. At the individual level, randomness explains $1 - \frac{20}{32} = 37.5\%$ of randomness in observed choice. Thus, in both cases, attention explains the minority of observed randomness, although it is higher at the individual level. This is what we would expect if we thinking randomness at the aggregate level is more driven by utility shocks compared to individual choice. Thus, it is likely that most of the additional randomness at the aggregate is driven by preference heterogeneity.

12. APPENDIX: HETEROGENEITY BY INDIVIDUAL

Here we show some patterns in individual heterogeneity. In Figure 25 we compare the share of options chosen in small and large choice sets. Each dot corresponds to an individual, and we add a blue trendline. The same individuals are also shown in Figure 26 which plots choice reversals in small and large choice sets.

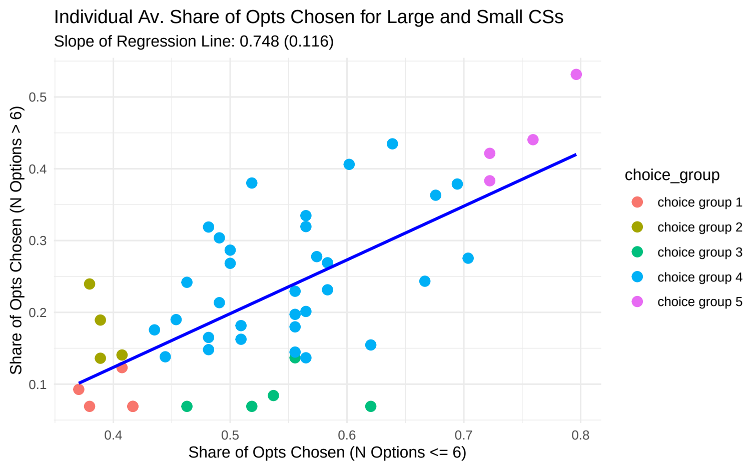


FIGURE 25. "Types" in Choice Reversals Share of Options Chosen

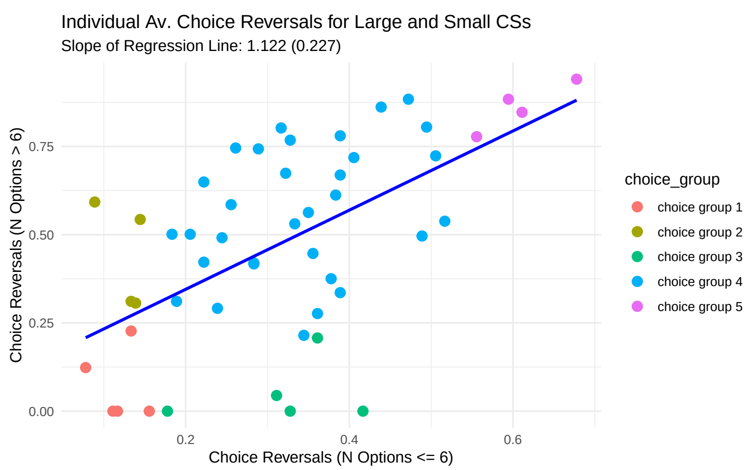


FIGURE 26. "Types" in Choice Reversals Normalized Entropy

We divide the individuals into five groups. The first (or red group) show low levels of randomness at small and large choice sets. They make consistent choices across instances of the same choice set. The second (or yellow group) make consistent choices in small choice sets, but exhibit randomness at larger choice sets. One possibility is that these subjects get overwhelmed by larger choice set sizes. The third (or green group) show high (almost total) consistency in large choice sets, but a fair amount of randomness in small choice sets. These subjects exhibit satisficing behavior - always selecting the safe option in large choice sets. The fourth (or blue group) show middling levels of consistency across all choice set sizes. The last (or purple group) show high levels of randomness across all choice sets. This last group tends to spend very little time per slide.