STOCHASTIC CHOICE WITH LIMITED MEMORY

ECE YEGANE

ABSTRACT. We model a decision maker who observes available alternatives according to a list and stochastically forgets some alternatives. Each time the decision maker observes an item in the list, she recalls previous alternatives with some probability, conditional on those alternatives being recalled until this point. The decision maker maximizes a preference relation over the set of alternatives she can recall. We show that if every available alternative is chosen with strictly positive probability, the preference order and the list order must coincide in any limited memory representation. Under the full support assumption, the preference ordering, the list ordering and the memory parameters are uniquely identified up to the ranking of the two least preferred alternatives. We provide conditions on observable choice probabilities that characterize the model under the full support assumption. We then apply our model to study the pricing problem of a monopolist who faces consumers with limited memory. We show that when the probability of forgetting is high, the monopolist is better off charging a lower price than the optimal price in the perfect memory case.

JEL classification: D11; D91
Keywords: Stochastic choice; memory; list

DEPARTMENT OF ECONOMICS, UNIVERSITY OF MARYLAND, TYDINGS HALL 3114, 7343 PREINKERT DRIVE, COLLEGE PARK, MD 20742
E-mail address: eyegane@umd.edu.
Date: September 1, 2022.
I am indebted to my advisor Yusufcan Masatlioglu for his support and guidance throughout this project. I thank Lawrence Ausubel, Christopher Chambers, Paul Cheung, Shanglyu Deng, Keaton Ellis, Agata Farina, Emel Filiz-Ozbay, Semih Kara, Tarik Kara, Sueyoul Kim, Mario Leccese, Zhenxun Liu, Rachel Nesbit, Elif Osun, Erkut Ozbay, John Shea, Daniel Vincent, Kemal Yildiz as well as conference participants at BRIC 2022 for their helpful comments and suggestions.

1. Introduction

Decision makers often rely on information that is not externally present at the time of decision but rather is recalled from memory. For example, when deciding where to have lunch, a decision maker may recall nearby restaurants from memory and choose from the restaurants she recalled. In this choice problem, the decision maker is effectively constructing a consideration set from the alternatives that were successfully retrieved from memory and then making a choice from this set. Even if the available alternatives are externally present during the time of choice, consumers may not search externally present products intensively but instead rely on memory when making their choice. For example, Dickson and Sawyer [1990] find that grocery shoppers buying food items such as coffee and cereal spend on average 12 seconds between arriving and departing at a product display. Hoyer [1984] finds that 72 percent of consumers look at only one item when deciding. In fact, only 11 percent looked at more than two different products. The minimal time spent considering visible options suggests that even when decisions are made in the store, memory factors might be at play in product selection (Lynch et al. [1991]). Therefore, whether an alternative is successfully retrieved from memory can have important consequences on the decision maker's choices.

Memory is incredibly complex and it can affect choices in various ways. In this paper, we focus on how forgetting available alternatives can shape the consideration set of a decision maker and affect her choices through the effect on the consideration set. We model a rational decision maker with memory limitations. We assume the decision maker uses a deterministic preference ordering over available alternatives, but she may not recall some of those alternatives due to random memory errors. Hence, the observed behavior of the decision maker is stochastic due to random memory errors. To model forgetting, we focus on a well-studied mechanism behind forgetting that is retroactive interference (first documented by Müller and Pilzecker [1900], see Dewar et al. [2007] for a review of interference effects). The key observation is that when additional materials intervene between the presentation of an item and the recall of an item, the additional materials may disrupt the retrieval process of the to-be-remembered item. Retroactive interference occurs when recent memories cause the forgetting of older memories. Thus, as the amount of intervening materials increase, recalling a to-be-remembered item becomes more difficult.

A decision maker behaving according to our model uses the following procedure when making a decision. First, she observes available alternatives one by one, which is captured by a *list*. When the list is exhausted, she chooses the most preferred alternative among the ones she remembers. Memory errors occur in the model because later alternatives in the list, which intervene between the presentation and the recall of previously observed alternatives, disrupt the retrieval process of previously observed alternatives. In particular, each time the decision maker observes an item in the list, she recalls previous alternatives with some probability, conditional on those alternatives being recalled until this point. We refer to this unobservable, alternative-specific probability as the *period recall probability*. To illustrate, suppose that in the list associated with some choice problem, alternative x is listed just before alternative y. When the decision maker observes alternative y, it interferes with the recall of x and may cause x to be forgotten. Thus, after observing y, the decision maker recalls x with probability $q(x)$ and forgets x with probability $1-q(x)$. Since some alternatives can be more memorable than the others, we allow the period recall probabilities to vary across alternatives. Now suppose that there is a third alternative z and the list ordering is such that x is listed first, y is listed second, and z is listed last. Since the decision maker chooses after observing z, there are two alternatives in between observing x and the time of choice. Thus, two alternatives interfere with the recall of x, and x is recalled at the time of choice with probability $q(x)^2$. We refer to the probability of recalling x at the time of choice as the *final recall probability of* x. Since there is only one alternative that follows y, the final recall probability of y is $q(y)$. As z is the last alternative in the list, so there are no alternatives that interfere with recalling z, its final recall probability is 1.

Our main goal is to understand how the limited memory model relates to observable choices. We now describe what we assume to be observable in this paper. We consider a decision maker who chooses from the same set of available alternatives repeatedly, but she does not always choose the same alternative.[1] We assume to observe the limiting distribution of the decision maker's repeated choices from each set of available alternatives.[2] We assume that the underlying list is unobservable and fixed, and the analyst only observes how the distribution of choices changes with the variation in the availability of alternatives. This is the standard type of choice data that is assumed to be observable in the random choice literature (e.g. Luce [1959], Block and Marschak [1960], Gul et al. [2014], Manzini and Mariotti [2014], Cattaneo et al. [2020]. We first show that, when every available alternative is chosen with strictly positive probability, in any SCLM representation of the choice data, the preference ordering and the list induce the same ranking of alternatives. Under the full support assumption, we show that the revealed preference/list and the period recall probabilities of the model are uniquely identified up to the ranking of the decision maker's two least preferred alternatives from the choice data. Our identification results disentangle the effects of preferences and memory on choices, which allow us to distinguish between an alternative that is not chosen due to a memory error and an alternative that is not chosen due to the presence of a better alternative. However, our identification results are applicable if the random choice data is generated from the limited memory model. We also provide a characterization of the model under the full support assumption, thereby providing testable conditions for the model.

We apply our representation to the pricing problem of a monopolist who faces consumers with limited memory. As a benchmark case, we consider the situation in which consumers have perfect memory, so they always recall every good with probability one. Our main result is that when consumers have limited memory and the probability of forgetting is sufficiently low, the monopolist is better off charging a lower price than the perfect memory case. This

[1]Choosing different alternatives in repeated choice settings is a common finding in experimental settings. Tversky [1969] and more recently Agranov and Ortoleva [2017]) document this type of behavior. Hey [2001] finds that the variability of responses of a large proportion of the subjects does not decay with experience and repetition.

[2]Random choice can also be interpreted as coming from a population of individuals choosing from each choice problem once. In our setting, this corresponds to a population of individuals who choose once among vertically differentiated products, but each individual may recall different set of alternatives at the time of choice.

is because forgetting causes the consumers to choose the outside option even when there is a good supplied by the monopolist that they would prefer over the outside option. Thus, in contrast with the perfect memory benchmark, the monopolist has an incentive to decrease the price, in order to capture the consumers choosing the outside option. Moreover, due to decreased prices, consumer surplus is strictly higher in the limited memory case compared to the benchmark case. Hence, when the monopolist responds optimally to the cognitive constraints consumers face, this makes consumers better off.

Models of a decision maker with limited attention have gained a lot of interest in economics (e.g. Sims [2003], Masatlioglu et al. [2012]). In the stochastic choice literature, randomness of attention has been proposed as a channel that causes the decision maker's choices to appear stochastic (e.g. Manzini and Mariotti [2014], Cattaneo et al. [2020]). Our paper is similar in this regard to the random attention literature: It is the random forgetting that causes the decision maker's choices to seem stochastic. Considering this relation to the random attention literature, and the fact that memory and attention form the basis for cognition, we now discuss how they relate to each other. Memory and attention are distinct but interdependent cognitive processes. Memory is the faculty of retaining and recalling previous experiences, and it has limited capacity (Cowan [1998]). On the other hand, attention is the concentration of consciousness, and it determines what information is selected for memory encoding (Cowan [1998]). The consensus in psychology and neuroscience is that increasing attention improves memory encoding (Chun and Turk-Browne [2007]), but forgetting can still occur even when a substantial amount of selective attention is devoted to a particular alternative. A surprising relationship between memory and attention is the potential for what is retrieved from memory to affect what attracts attention (Hannula [2018]). For example, Fan and Turk-Browne [2016] show that information that was encoded in long-term memory influences the subsequent visual search, and guides attention according to the information that was encoded in memory. Models with consideration sets in economics literature usually take why a decision maker pays attention to a particular subset of alternatives as given. Therefore, our model can also be interpreted as providing a foundation to the problem of what affects consideration set formation. That is, encoded memories that are affected by the order in which alternatives were observed affect which alternatives are considered at time of choice.

In our model, as the number of alternatives that follow an alternative x increases, the final recall probability of x decreases. Thus, our model generates the recency effect, [3] in the sense that an alternative is recalled with a higher probability if its location in the list is modified so that it appears later. However, alternatives that appear later in the list do not necessarily have a higher final recall probability than the alternatives that appear earlier in the model. For example, consider a three alternative list in which the period recall probability of the first alternative is 0.75, the second is 0.5 and the last is 0.5. Then the final recall probability of the first alternative is equal to 0.5625 which is greater than the final recall probability of the second alternative, which is equal to 0.5. Because the period recall probabilities can vary across the alternatives, if the period recall probabilities of earlier alternatives in the list are higher than the later alternatives in the list, the final recall probabilities of earlier alternatives may be higher. While such recall probabilities can be accommodated, this is not a general feature of the model, as it only occurs in some lists. Note that the last alternative in the list, is always recalled with probability 1. We relax this assumption in Appendix B where we introduce a default/no-purchase option to our model and consequently allow for the possibility of the decision maker forgetting every available alternative in her choice set, including the last alternative she observed.

In Section 4, we compare the choice data generated from our model with related models of stochastic choice (Block and Marschak [1960], Manzini and Mariotti [2014], Aguiar et al. [2016] Brady and Rehbeck [2016], Aguiar [2017], Echenique et al. [2018], Cattaneo et al. [2020]) and related models of choice from lists (Rubinstein and Salant [2006], Yildiz [2016],

[3]The recency effect has been documented in many situations such as performance appraisals (Steiner and Rain [1989]), auditor evaluations (Ashton and Ashton [1988]), interviewer impressions (Farr [1973]; Farr and York [1975]), judgments of innocence and guilt in trials (Furnham [1986]), and food consumption (Garbinsky et al. [2014]).

Kovach and Ülkü [2020]). We show that our model is independent from any of the aforementioned models that make a prediction about stochastic choice functions. Therefore, our model is descriptively and observationally distinct from these models.

The rest of the paper is organized as follows. Section 2 introduces the notation, the model, and provides identification and characterization results. In Section 3, we apply our representation to the pricing problem of a monopolist who faces consumers with limited memory. Section 4 compares our model with related work from the literature. Section 5 concludes.

2. Stochastic Choice with Limited Memory

Let X be a finite set of alternatives that may be available to the decision maker to choose from. Let $\mathcal{X}$ denote the set of all nonempty subsets of X. We refer to each element of $\mathcal{X}$ as a *choice set*. We consider a decision maker who chooses from the same choice set repeatedly. We assume to observe the limiting distribution of the decision maker's repeated choices from each choice set, which we refer to as a *random choice rule*. Formally, a random choice rule ρ maps each choice set $S \in \mathcal{X}$ to a probability distribution over its elements. We denote the probability of choosing alternative x from choice set S with $\rho(x, S)$.

In our model, the decision maker observes the alternatives in her choice set one by one before she chooses an alternative. The order in which the decision maker observes the alternatives in the choice set is captured by a *list*. Formally, a list $\mathcal{L}$ is a finite sequence of alternatives of X such that every alternative in X appears exactly once in $\mathcal{L}$. Let $\mathbb{L}$ denote the set of all possible lists. Let $\mathcal{L}(S)$ denote the subsequence of a list $\mathcal{L}$ that is obtained from $\mathcal{L}$ by deleting the alternatives in $X \setminus S$. For any $x \in S$, $\mathcal{L}(x, S)$ denotes the set of alternatives that follow x in $\mathcal{L}(S)$. When the cardinality of the choice set is equal to three, we use the notation $[x, y, z]$ to denote a list that is ordered such that x is listed first, y is listed second and z is listed last. Throughout the paper, whenever we refer to an alternative as first or last, we are referring to that alternative's position in the associated list. When the list is exhausted, the decision maker chooses her most preferred alternative among the ones she remembers. The decision maker's preferences are represented by a linear order[4] on X, denoted by $\succ$.

A decision maker behaving according to our model forgets an alternative x because the subsequently observed alternatives in the list interfere with the recall of x. To illustrate, suppose that x is listed just before y. When the decision maker observes alternative y, she recalls x with some probability $q(x)$ and she forgets x with probability $1 - q(x)$. We refer to this probability of recalling an alternative when the decision maker observes the next alternative in the list as the *period recall probability*. Since some alternatives can be more memorable than others, we allow period recall probabilities to be alternative-specific. We also would like to rule out the situation in which some of the alternatives are always forgotten or always remembered. Therefore, we do not allow period recall probabilities to be 0 or 1. Formally, period recall probabilities are captured by the *period recall probability function* $q : X \to (0, 1)$.

Interference is captured through the number of alternatives that follow x in the list. For example, in list $[x, y, z]$, there are two alternatives that follow x, alternatives y and z. Since the decision maker chooses after observing the last element in the list, both y and z interfere with the recall with x. Thus, x is recalled at the time of choice with probability $q(x)^2$. There is only one alternative that follows y in the list, so y is recalled with probability $q(y)$ at the time of choice. Since z is observed last, none of the alternatives interfere with the recall of z, so it is recalled with probability 1. We refer to these probabilities as the *final recall probability* of $x/y/z$ in list $[x, y, z]$. Thus, the final recall probability refers to the probability of remembering an alternative at the time of choice, whereas the period recall probability captures the probability of recalling an alternative when the next alternative in the list is observed [5]. We now formally define our model.

[4]A linear order is a complete, transitive, asymmetric binary relation.

[5]We would like to point out an alternative interpretation of our model. Throughout the paper, we focus on the interpretation that the recall probability function reflects the retrieval of alternatives from memory. However, it can also capture unobservable availability variations in a physical sense, which may be of interest when modeling consumer substitution behavior (e.g. Blanchet et al. [2016]). Consider the following: Every day a new product arrives in a store. The arrival of each product over time is captured by a list. After a product arrives in a store, in each of the following days it can become out-of-stock with some probability. Then the

Definition 1. A random choice rule ρ has a *stochastic choice with limited memory (SCLM)* representation if there exists a preference ordering $\succ$ on X, a list $\mathcal{L} \in \mathbb{L}$, and a period recall probability function $q : X \to (0,1)$, such that for any $x \in S$ and $S \in \mathcal{X}$,

$$\rho(x,S) = q(x)^{|\mathcal{L}(x,S)|} \prod_{\{y \in S \ | \ y \succ x\}} (1 - q(y)^{|\mathcal{L}(y,S)|})$$

An assumption of our model is that the last alternative in any list is recalled with probability 1. Thus, the decision maker never ends up in a situation in which she does not remember any of the available alternatives. We assume that there is a fixed list that generates the random choice rule, and the analyst does not observe any information about the list. We also assume the decision maker observes every alternative in her choice set prior to choosing. We make this assumption because in our model forgetting an alternative is the only reason for the exclusion of an available alternative in the consideration set. Forgetting an alternative means that alternative was accessible to the decision maker at a previous point but it is not accessible at the time of choice. If the decision maker was never aware of an alternative, for example, because she stopped in the middle of the list and therefore never observed some of the alternatives, then there is another factor different than memory that affects the consideration set formation process. We now provide a simple example to illustrate a random choice rule generated by the model.

Example 1. Suppose that the decision maker's choices are generated by the SCLM model with preferences $x \succ y \succ z$, list $[y,x,z]$ and the period recall probability function $q : X \to (0,1)$. The following table demonstrates the resulting random choice rule in the associated choice problems.

In choice set $\{x,y,z\}$ under list $[y,x,z]$, there are two alternatives that follow y, so y is recalled at the time of choice with probability $q(y)^2$. There is one alternative listed after x, so x is recalled at the time of choice with probability $q(x)$. The alternative z is the last one

random choice rule generated from the model captures the choices of a population of consumers arriving on a single day.

S	$\rho(x,S)$	$\rho(y,S)$	$\rho(z,S)$
$\{x,y,z\}$	$q(x)$	$q(y)^2(1-q(x))$	$(1-q(y)^2)(1-q(x))$
$\{x,y\}$	1	0	-
$\{x,z\}$	$q(x)$	-	$1-q(x)$
$\{y,z\}$	-	$q(y)$	$1-q(y)$

in the list, so z is recalled with probability 1. Since x is the most preferred alternative, it is chosen whenever it is recalled, which happens with probability $q(x)$. For y to be chosen, y must be recalled at the time of choice, which happens with probability $q(y)^2$, and x must be forgotten at the time of choice, which happens with probability $1-q(x)$. Therefore, y is chosen with probability $q(y)^2(1-q(x))$. The least preferred alternative z is recalled with probability 1. So z is chosen whenever x and y are forgotten, which happens with probability $(1-q(x))(1-q(y)^2)$. Now consider the choice set $\{x,y\}$, x is the last alternative in the list $[y,x]$ so it is recalled with probability 1, and it is preferred to y. Hence, x is chosen with probability 1 and y is chosen with probability 0. In choice set $\{x,z\}$, z is the last alternative in the list $[x,z]$, and x is preferred to z. Thus, z is chosen when x is forgotten, so with probability $1-q(x)$, and x is chosen with probability $q(x)$.

2.1. **Identification.** Suppose a random choice rule has a stochastic choice with limited memory representation. In this section, we inquire whether it is possible to identify the preference relation, the list and the period recall probability function of a decision maker who behaves according to the SCLM model from the observed random choice rule. We show that under the assumption that each available alternative is chosen with strictly positive probability, we can uniquely identify the preference ordering, the list and the period recall probability function up to the ranking of the two least preferred alternatives. We start with the following definition.

Definition 2. A random choice rule has *full support* if for all $S \in \mathcal{X}$, and all $x \in S$, $\rho(x,S) > 0$.

We now show that the full support assumption has strong behavioral implications for the SCLM model. The following lemma shows that, if a random choice rule has full support, in any rationalization of the random choice rule with the SCLM model, the preference ordering and the list ordering must induce the same ranking.

Lemma 1. *Let ρ be generated by $(\succ, \mathcal{L}, q)$. If ρ has full support, then $y \in \mathcal{L}(x, X)$ if and only if $x \succ y$.*

Proof. For a contradiction, suppose that there exists alternatives $x, y \in X$ such that $x \succ y$ and $y \notin \mathcal{L}(x, X)$. Since ρ is generated by $(\succ, \mathcal{L}, q)$ and $x \succ y$, $\rho(y, \{x, y\}) = q(y)(1 - q(x)^{|\mathcal{L}(y,\{x,y\})|}) = q(y)(1 - q(x)^0) = 0$, which contradicts that ρ has full support. $\square$

Note that Lemma 1 is an implication of the full support assumption, as Example 1 demonstrates that if ρ does not have full support, two alternatives can be ranked differently in the preference order and in the list order as part of a SCLM representation. We assume that the observed random choice rule has full support in the rest of Section 2. Whenever the preference ordering and the list coincide, we drop the list notation and denote a SCLM model with $(\succ, q)$.

We now describe our identification strategy. Consider a full support random choice rule generated by the SCLM model, and consider a set S with at least two alternatives, and denote the most preferred alternative in S with x. How does adding an alternative y, which is inferior to x, to set S affect the choice probability of x? By Lemma 1, it must be that y follows x in the list. Thus, adding y to set S causes an additional alternative to interfere with the recall of x, which decreases the choice probability of x by $q(x)$. Note that this decrease is exactly equal to the choice probability of x in the binary set $\{x, y\}$. Adding y to S can also affect the final recall probabilities of the alternatives in S other than x. However, as x is the most preferred alternative in S, its choice probability in S is independent from the final recall probabilities of other alternatives in S. Therefore, we can deconstruct the choice probability of the most preferred alternative x in $S \cup \{y\}$ to its choice probability in S multiplied by it's choice probability in $\{x, y\}$. We now provide a definition that formalizes this idea.

Definition 3 (Stochastic Expansion)**.** For any $S \in \mathcal{X}, |S| \geq 3$, an alternative x satisfies *stochastic expansion in* S if for all $T \subseteq S$ with $x \in T$,

$$\rho(x, T) = \prod_{y \in T} \rho(x, \{x, y\})$$

Following the intuition that is described in the paragraph above, the most preferred alternative in set S must satisfy stochastic expansion in S. We define the revealed preference relation of the SCLM model as follows.

$x \succ_\rho y$ for all $y \in S \setminus \{x\}$ if there exists $S \in \mathcal{X}$ in which x satisfies stochastic expansion in S.

To obtain the revealed preference ordering, we first identify the alternative that satisfies stochastic expansion in the grand set X, say alternative x_1. Next, we identify the alternative x_2 that satisfies stochastic expansion in $X \setminus \{x_1\}$. We continue this process, until there are two alternatives left.

If x is revealed to be preferred to y, then x precedes y in the underlying list. Therefore, the choice probability of x in the set $\{x, y\}$ must equal to the period recall probability of x. Thus, we can obtain the period recall probabilities in the SCLM model as,

$$q(x) = \rho(x, \{x, y\}) \text{ if } x \succ_\rho y.$$

2.2. **Uniqueness.** We now inquire whether multiple representations are possible in the SCLM model when we apply this identification strategy. First, the decision maker's least preferred alternative in any set S is only chosen if it is the last alternative in the list *i.e.* when it is recalled with probability 1. Therefore, the period recall probability of the least preferred alternative in X never appears in any of the choice probabilities. Hence, the SCLM model can be described by the period recall probability of the most preferred $|X| - 1$ alternatives. Second, since stochastic expansion requires a set of at least three alternatives, the revealed preference relation in the SCLM model is silent about the ranking of the decision maker's two least preferred alternatives in X. The following example demonstrates that two representations differing in the ranking of the two least preferred alternatives are possible.

Example 2.

S	$\rho(x,S)$	$\rho(y,S)$	$\rho(z,S)$
$\{x,y,z\}$	0.64	0.144	0.216
$\{x,y\}$	0.8	0.2	-
$\{x,z\}$	0.8	-	0.2
$\{y,z\}$	-	0.4	0.6

Note that ρ has full support so in each SCLM representation, the preference order and the list order has to coincide. Alternative x is the only alternative that satisfies stochastic expansion in S, so in every representation, it must be that $x \succ_\rho y$, $x \succ_\rho z$ and $q(x) = 0.8$. Consider the following two representations: (1) the preferences and the list are given by $x \succ_1 y \succ_1 z$, the period recall probability function is $q_1(x) = 0.8$, $q_1(y) = 0.4$, $q_1(z) = 0.5$; (2) the preferences and the list are given by $x \succ_2 z \succ_2 y$, the period recall probability function is $q_2(x) = 0.8$, $q_2(y) = 0.5$, $q_2(z) = 0.6$. Both $(\succ_1, q_1)$ and $(\succ_2, q_2)$ represent ρ.

Two representations are possible because stochastic expansion is applicable to sets of at least three alternatives. Therefore, it does not provide any information about the ranking of alternatives in the set $\{y,z\}$. Since the period recall probability of the least preferred alternative is a free variable of the model, we can either set $q(y) = \rho(y, \{y,z\})$ and $y \succ_\rho z$ or $1 - q(z) = \rho(y, \{y,z\})$ and $z \succ_\rho y$. Hence, we end up with two representations that are the same except the preference ordering, list ordering and period recall probabilities of the two least preferred alternatives.

Next, we show that only the most preferred alternative in S satisfies stochastic expansion in S. Thus, the preference ordering, the list and the period recall probability function is uniquely identified up to the ranking of the two least preferred alternatives. To see why alternatives other than the most preferred one in S must violate stochastic expansion in S, suppose that the decision maker's choices are generated by the SCLM model with preferences and list $x \succ y \succ z$, the period recall probability function $q : X \to (0,1)$. The following table demonstrates the resulting random choice rule in the associated choice sets.

We compare the choice probability of y in $\{x,y,z\}$ with its choice probability in $\{x,y\}$ multiplied by the choice probability of y in $\{y,z\}$. Because x is preferred to y, the final

S	$\rho(x,S)$	$\rho(y,S)$	$\rho(z,S)$
$\{x,y,z\}$	$q(x)^2$	$q(y)(1-q(x)^2)$	$(1-q(y))(1-q(x)^2)$
$\{x,y\}$	$q(x)$	$1-q(x)$	-
$\{x,z\}$	$q(x)$	-	$1-q(x)$
$\{y,z\}$	-	$q(y)$	$1-q(y)$

recall probability of x affects the choice probability of y whenever x is available in addition to y. Two alternatives follow x in list $[x,y,z]$, so the final recall probability of x in $\{x,y,z\}$ is $q(x)^2$, whereas the final recall probability of x in $\{x,y\}$ is $q(x)$. The change in the final recall probability of x when we add z to the set $\{x,y\}$ results in an increase in the choice probability of y with magnitude $\frac{1-q(x)^2}{1-q(x)}$. Due to this indirect effect on the choice probability of y, $\rho(y,\{x,y,z\}) \neq \rho(y,\{x,y\}) \cdot \rho(y,\{y,z\})$. Note that the most preferred alternative in any set is unaffected by such auxiliary final recall probability effects, so that the choice from smaller sets is in line with the choice from the large set.

Lemma 2. *Let ρ be a random choice rule with full support that is generated by a SCLM model $(\succ, q)$. If there exists $x \in S$ such that $x \succ y$, then y violates stochastic expansion in S.*

Proof. See the appendix.

2.3. **Characterization.** In the identification section, we showed that we can infer the decision maker's preferences, list ordering and period recall probability function from the observed choice data. However, the identification results are only applicable if the decision maker's choice behavior is generated by the SCLM model. In this section, we provide axioms which ensure that a full support random choice rule has a SCLM representation.

In the identification section, we showed that if a random choice rule has an SCLM representation, then an alternative is the $\succ_\rho$-maximal alternative in some set S in the associated representation if and only if it satisfies stochastic expansion in S. The first axiom guarantees that there exists a $\succ_\rho$-maximal alternative in every set by utilizing the stochastic expansion definition.

Axiom 1. For any $S \in \mathcal{X}, |S| \geq 3$, there exists an alternative $s \in S$ such that s satisfies stochastic expansion in S.

Axiom 1 is related to the stochastic path independence condition in Yildiz [2016]. In his model, the stochastic path independence condition is satisfied by alternative x in $S \cup \{y\}$ if y follows x in the list. We discuss the relation between Yildiz [2016] and the SCLM model in detail in the literature review section.

Stochastic expansion states how the choice probability of the decision maker's most preferred alternative x in some set S relates to the choice probability of x in subsets of S. However, stochastic expansion does not restrict how x's binary choice probabilities with two different alternatives from S must be related to each other. The following example demonstrates the role of binary choices in the SCLM model.

Example 3.

S	$\rho(x, S)$	$\rho(y, S)$	$\rho(z, S)$
$\{x, y, z\}$	0.72	0.112	0.168
$\{x, y\}$	0.8	0.2	-
$\{x, z\}$	0.9	-	0.1
$\{y, z\}$	-	0.4	0.6

The random choice rule in Example 3 does not have a SCLM representation. To see why, notice that x is the only alternative that satisfies stochastic expansion in $\{x, y, z\}$, so x must be the decision maker's most preferred alternative in $\{x, y, z\}$ and the first alternative in the list. This implies that $\rho(x, \{x, y\}) = \rho(x, \{x, z\}) = q(x)$ which is violated in the example. Therefore, we need an additional axiom ensuring the existence of a period recall probability function.

Axiom 2. If s satisfies stochastic expansion in S, then for any $x, y \in S \setminus \{s\}, \rho(s, \{s, x\}) = \rho(s, \{s, y\})$.

Within the context of the SCLM model, this axiom says that diversion from the to-be-remembered item s is the only cause of forgetting, and two different alternatives that follow s in the list divert from s in the same proportion. Therefore, Axiom 2 rules out factors contributing to forgetting based on the similarity between the to-be-remembered alternative and the interfering alternatives. One such example is inability to recall due to confusion caused by similar items. When alternatives that are similar to the to-be-remembered alternative are introduced closer to the retrieval stage, this may cause confusion in retrieval and decrease the probability of recall (Dewar et al. [2007]). Hence, if alternative x is similar to item s but y is not, then we may expect the final recall probability of s to be lower in the presence of x than the presence of y. Axiom 2 rules out such situations.

The next axiom tells us when Luce's [1959] Independence of Irrelevant Alternatives (IIA) axiom is satisfied in the model. First, recall Luce's IIA,

Luce's IIA. For any subset $\{x, y\} \subseteq S, T \in \mathcal{X}$,

$$\frac{\rho(x, S)}{\rho(y, S)} = \frac{\rho(x, T)}{\rho(y, T)}$$

Luce's IIA states that the ratio of choice probabilities of any two alternatives is independent from the other alternatives in the choice set. In the SCLM model, the alternative that satisfies stochastic expansion in S, say s, is revealed to be the most preferred and first alternative in S. Thus, if we remove s from set S, then the choice probability of all the remaining alternatives in $S \setminus \{x\}$ is boosted by the same amount, that is the probability of forgetting s in S. Moreover, because s is the first alternative in list, removing it does not change the final recall probability of any other alternative. Thus, the Luce ratios remain the same. Our next axiom formalizes this observation.

Axiom 3. If s satisfies stochastic expansion in S, then for any for any $x, y \in S \setminus \{s\}$,

$$\frac{\rho(x, S)}{\rho(y, S)} = \frac{\rho(x, S \setminus \{s\})}{\rho(y, S \setminus \{s\})}$$

We are now ready to provide the characterization theorem.

Theorem 1 (Characterization). *A random choice rule ρ with full support has a stochastic choice with limited memory representation if and only if ρ satisfies Axioms 1-3.*

Proof. See the appendix.

3. Application: Monopoly Pricing

In this section, we apply our representation to the pricing problem of a profit-maximizing monopolist who faces consumers with limited memory. We focus on a simple setting with two goods and two types of consumers such that the first type (type 1) prefers good 1 and the second type (type 2) prefers good 2. The list ordering can be in either order with equal probability. We assume that the monopolist does not observe the list ordering, but it knows the true distribution of the list ordering followed by the consumers. There is an outside option which corresponds to not purchasing either of the goods the monopolist supplies. We normalize the cost of producing each good to zero.

We start with summarizing our results. As a benchmark case, we consider the situation in which consumers have perfect memory, so they always recall both goods with probability one. In this case, the monopolist can extract the highest possible amount from each type of consumer by setting the price of each good equal to the willingness to pay of the consumer type who prefers that good, so the consumer surplus is zero. Our main result is that when consumers have memory limitations, so they do not always recall all the available goods, and the probability of forgetting is high, the monopolist is better off charging a lower price than the perfect memory case. Therefore, consumer surplus becomes strictly positive when the probability of forgetting is sufficiently high. Hence, memory limitations according to our model do not provide an opportunity through which the monopolist can exploit the consumers. In fact, when the monopolist responds optimally to the cognitive constraints consumers face, this makes consumers better off.

The monopolist is better off charging a lower price when the probability of forgetting is high because when consumers are forgetful, they make memory errors and opt for the outside option with some probability even when there is a good supplied by the monopolist that they prefer. The monopolist can capture the forgetful consumers by decreasing the price of the good that the consumer always remembers, which is the last good in the list. This causes the consumers to choose the always-recalled good instead of the outside option.

The monopolist can increase the volume of consumers, at the expense of decreased prices. Thus, the monopolist faces a trade-off between increasing the profit margin and increasing the volume. When the recall probabilities are low, the portion of consumers opting for the outside option is large. Hence the volume channel dominates, and the monopolist is better off charging a lower price than the perfect memory case. If the recall probabilities are high, then the portion of consumers choosing the outside option becomes negligible, and the profit margin channel dominates the volume channel. Thus, the monopolist maximizes profits by charging the same prices in the perfect memory case.

When consumers have perfect memory, the monopolist knows that both goods are recalled by the consumers, so the monopolist knows what is in the consumers' consideration sets. When consumers have limited memory, the consideration sets become stochastic and the monopolist faces uncertainty about the composition of the consideration sets. As a result of this uncertainty, the monopolist cannot perfectly discriminate between different consumer types. Thus, our results are in line with the classical results of Maskin and Riley [1984].

We would like to focus on the margin-volume trade-off that is driving our results, so we first start with a simplified situation in which the share of type 1 and type 2 consumers are equal, and each type of consumer uses each list ordering with probability 0.5, and the preferences of the two types of consumers are symmetric, so the utility indices of the two types of consumers are given by, $u_1(1) = u_2(2) = 1$ and $u_1(2) = u_2(1) = \alpha$, $0 < \alpha < 1$ where $u_k(t)$ denotes the utility index of consumer k of good t for $k, t = 1, 2$. If consumer k purchases good t from the monopolist at price p_t, she gets utility $U_k(t, p_t) = u_k(t) - p_t$. We normalize the utility of the outside option to 0 for both types of consumers. We assume that both types of consumers have the same period recall probability function $q(1), q(2) \in (0, 1)$. In Appendix A.3, we introduce a parameter for the consumer type distribution, a parameter for the list ordering distribution and a parameter that allows the utility indices for the less-preferred good of the two consumers to be different. Our qualitative results from the simplified case still holds.

In the SCLM model, we consider a decision maker who can rank the available alternatives according to a linear order. We make the following assumptions about how the consumer chooses when she gets the same utility from two (or three) goods. First, if a consumer is indifferent between good t and the outside option, conditional on recalling t, the consumer purchases t. Second, if a consumer is indifferent between good 1 and good 2, conditional on recalling both goods, she breaks the indifference in favor of her favorite good according to the above utility indices. We write $\rho_k(t,p)$ to denote consumer k's expected demand for good t when the monopolist charges p_1 for the first good and p_2 for the second good $p=(p_1,p_2)$, and consumer k uses each list with probability 0.5.

Optimal Pricing with Perfect Memory: As a benchmark case, we assume that both types of consumers recall each good with probability 1 (*i.e.* $q(1)=q(2)=1$). In the benchmark case, consumers always recall both goods, so the monopolist maximizes profits by setting $p_1=p_2=1$. Consumer 1 purchases only good 1 and consumer 2 purchases only good 2. The consumer surplus equals zero.

Optimal Pricing with Limited Memory: We now assume that $q(1),q(2)\in(0,1)$. Thus, consumers make memory errors with some probability, which may cause them to opt for a less-preferred option. The goal of our analysis is to understand how the monopolist can use pricing as a tool to respond to memory errors. In what follows, we calculate the profit-maximizing prices for different values of $q(1),q(2),\alpha$. We start with calculating the probability of consumers recalling good 1 and/or good 2 at the time of choice. Consumers use the list in which good 1 is listed first and good 2 is listed last with probability 0.5. Conditional on the consumers using this list, good 1 is recalled at the time of choice with probability $q(1)$ and good 2 is recalled with probability 1. Similarly, the list in which good 2 is listed first and good 1 is listed last is used with probability 0.5, so conditional on using this list, good 1 is recalled with probability 1 and good 2 is recalled with probability $q(2)$. Therefore, the probability of only recalling good 1 is $\frac{1-q(2)}{2}$, the probability of only recalling good 2 is $\frac{1-q(1)}{2}$ and the probability of recalling both good 1 and good 2 is $\frac{q(1)+q(2)}{2}$.

We continue with the following observation: Given the utility indices of the consumers, it is sub-optimal for the monopolist to choose p such that $\alpha < p_1 < 1$ and/or $\alpha < p_2 < 1$. To see why, suppose that p is optimal and $\alpha < p_1 < 1$. Under p, consumer 1's demand is $q(1)$ if good 2 is listed last and 1 if good 1 is listed last. Consumer 2's demand for good 1 is 0 as $u_j(1) - p_1 = \alpha - p_1 < 0$. If the monopolist increases price of good 1 to 1, consumer 1's demand and consumer 2's demand remain the same, but the profit margin of good 1 increases. By the same argument, $p_1 < \alpha$ and/or $p_2 < \alpha$ also cannot be optimal. Therefore, to solve the monopolist's optimal pricing problem, it is sufficient restrict our attention to prices with $p_1, p_2 \in \{\alpha, 1\}$. Table 1 shows the expected demand of consumer 1 and 2 for good 1 and good 2 as well as the proportion of consumers who opt to the outside option.

p	$\rho_1(1,p)$	$\rho_2(1,p)$	$\rho_1(2,p)$	$\rho_2(2,p)$	Outside Option
$(1,1)$	$\frac{1+q(1)}{2}$	0	0	$\frac{1+q(2)}{2}$	$\frac{2-q(1)-q(2)}{4}$
$(1,\alpha)$	$\frac{1+q(1)}{2}$	0	$\frac{1-q(1)}{2}$	$\frac{1+q(2)}{2}$	$\frac{1-q(2)}{4}$
$(\alpha,1)$	$\frac{1+q(1)}{2}$	$\frac{1-q(2)}{2}$	0	$\frac{1+q(2)}{2}$	$\frac{1-q(1)}{4}$
(α,α)	$\frac{1+q(1)}{2}$	$\frac{1-q(2)}{2}$	$\frac{1-q(1)}{2}$	$\frac{1+q(2)}{2}$	0

TABLE 1. Expected consumer demand.

Suppose that the monopolist sets $p_1 = p_2 = 1$, then consumer 1 is indifferent between good 1 and the outside option, and good 2 yields strictly less utility. Therefore, consumer 1 chooses the outside option when she only recalls good 2, which happens with probability $\frac{1-q(1)}{2}$. Consumer 1 chooses good 1 whenever she recalls it, which happens with probability $\frac{1+q(1)}{2}$. Similarly, consumer 2 chooses the outside option if she only recalls good 1, which happens with probability $\frac{1-q(2)}{2}$, and chooses good 2 with probability $\frac{1+q(2)}{2}$. Now suppose that the monopolist sets $p_1 = 1$, $p_2 = \alpha$. The price decrease in good 2 does not affect consumer 2's demand but consumer 1 is now indifferent between good 1, good 2 and the outside option. Thus, consumer 1 chooses good 1 whenever it is recalled, and she chooses good 2 when only good 2 is recalled. Therefore, by decreasing the price of good 1, the monopolist can capture the portion of consumer 1 who is opting for the outside option when

$p_1 = p_2 = 1$, which is given by $\frac{1-q(1)}{4}$. If α is sufficiently high, then the monopolist can be better off charging α for good 2.

The monopolist's incentives for increasing the volume of good 2 strengthen as the probability of recalling good 1 (*i.e.* $q(1)$) decreases. When $q(1)$ is low, the monopolist does not benefit from decreasing the price of good 1. This is because consumers who are opting for the outside option are doing so because they are unaware of this product. Decreasing the price of good 1 does not affect the recall probabilities, so it does not help the monopolist. Instead, the monopolist compensates for the loss due to low $q(1)$ by decreasing the price of good 2, as good 2 is recalled by the consumers. By the same reasoning, the monopolist's incentives for decreasing the price of good 1 strengthen as $q(2)$ decreases.

When the monopolist sets the price of a good to α, the consumer type who prefers that good gets a surplus of $1 - \alpha$. Therefore, that consumer type is better off than the perfect memory case. To demonstrate, suppose that $q(1)$ is sufficiently low and $q(2)$ is sufficiently high so that the monopolist optimally sets $p_1 = 1$ and $p_2 = \alpha$. In this case, consumer 2 purchases good 2 whenever she recalls it and gets a utility of $1 - \alpha$. Whenever she forgets good 2, she opts for the outside option. Consumer 1 always gets a utility of 0. Thus, the consumer surplus under $p_1 = 1$ and $p_2 = \alpha$ is $\frac{(1-\alpha)(1+q(2))}{4}$.

In Figure 1, we fix the α parameter at 0.75 and graphically show which pricing is optimal for each $q(1)$ and $q(2)$ combination. The graph captures the intuition described above. When $q(1)$ and $q(2)$ are both sufficiently high, then $p_1 = p_2 = 1$ is optimal. If $q(2)$ is high but $q(1)$ is low, then the monopolist compensates for the loss in good 1 by setting $p_1 = 1$ and $p_2 = \alpha$. Similarly, when $q(1)$ is high and $q(2)$ is low, the monopolist sets $p_1 = \alpha$ and $p_2 = 1$. When both $q(1)$ and $q(2)$ are sufficiently low, the monopolist sets $p_1 = p_2 = \alpha$.

Finally, in Proposition 1 we characterize the optimal pricing choice of the monopolist for any given $\alpha, q(1), q(2) \in (0, 1)$.

Proposition 1. If $\frac{\alpha}{1-\alpha} \leq \frac{1+q(i)}{1-q(j)}$, then the price of good i is 1, otherwise it is α.

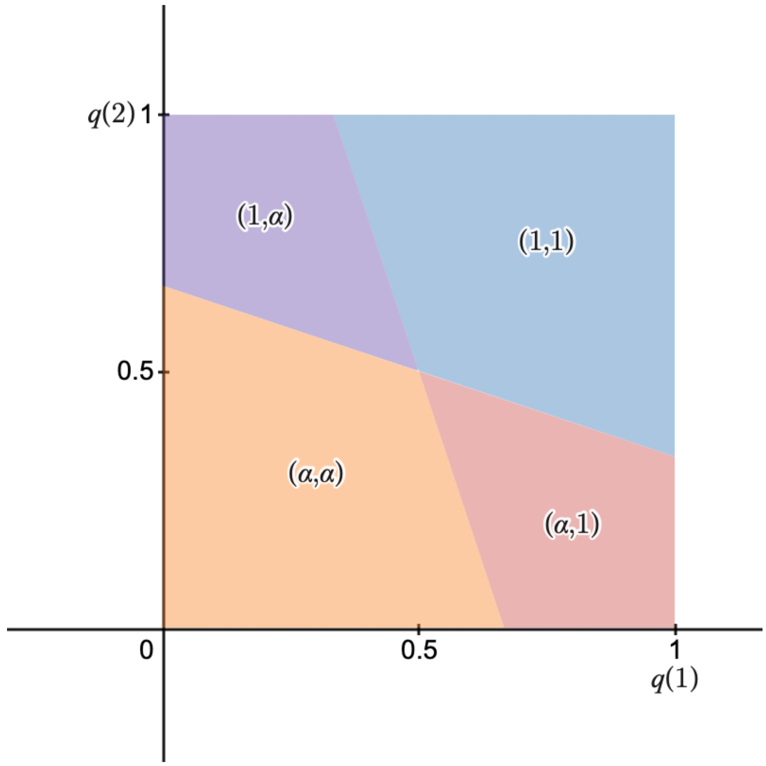

Figure 1. Optimal pricing when $\alpha = 0.75$.

4. Literature Review

We now discuss how our paper is related to other works in the literature. In particular, we compare the set of random choice rules generated by the SCLM model with the set of random choice rules generated by related models of stochastic choice and choice from lists. We show that in terms of the observed choice probabilities, the SCLM model does not nest any model in this section, and it is not nested by any model we discuss below as long as the compared model imposes some restriction on random choice rules. We start our discussion by considering models of stochastic choice in which the domain of the random choice rule is $X \times \mathcal{X}$, the same domain with the SCLM model.

Cattaneo et al. [2020] characterize a class of stochastic choice rules referred to as the random attention model (RAM). RAM features a decision maker with stochastic attention, which means that in a given choice set $S \in \mathcal{X}$, the decision maker may pay attention to different subsets of S with some probability. Conditional on considering the subset $T \subseteq S$, the decision maker chooses her most preferred alternative in T. In RAM, the revealed preference relation is defined as follows: If $\rho(a, S) > \rho(a, S \setminus \{b\})$, then a is revealed to be preferred to b. In the random choice rule generated by the SCLM model with preferences $x \succ a \succ b \succ y$, list $[x, b, a, y]$ and period recall probability function $q(a) = q(y) = 0.2$, $q(b) = q(x) = 0.9$, we have $\rho(a, \{x, y, a, b\}) > \rho(a, \{x, y, a\})$ and $\rho(b, \{x, y, a, b\}) > \rho(b, \{x, y, b\})$. Since the revealed preference relation has a cycle, ρ does not have a RAM representation and the SCLM model is not a subset of the RAM.

Two papers that are closely related to our model are Manzini and Mariotti [2014] and Brady and Rehbeck [2016]. The SCLM model is independent from both of these models. Manzini and Mariotti [2014] characterize random consideration set rules (RCSR) in which each available alternative x is considered by the decision maker with some probability $\gamma(x)$. The decision maker chooses her most preferred alternative among the ones she considers. Brady and Rehbeck [2016] characterize random conditional choice set rules (RCCSR) with unobservable feasibility variation. In their model, in any choice problem $S \in \mathcal{X}$, the decision maker faces each subset T of S with some positive probability $\pi(T)$. This is interpreted as availability variations that are not observable to the analyst. The probability of choosing $x \in S$ is given by the sum of feasibility probabilities of sets in which x is the most preferred alternative. Brady and Rehbeck [2016] show that their model generalizes RCSR. Moreover, the stochastic consideration set rule in RCCSR satisfies monotonic attention (see Supplemental Appendix to Cattaneo et al. [2020]) and so it is nested in RAM. Therefore, the SCLM model is not a subset of RCSR or RCCSR.

Next, we show that RCSR is not a subset of SCLM. Because RCSR is nested in RCCSR, which is nested in RAM, we conclude that the SCLM model is independent from all three models. Firstly, RCSR and RCCSR assume there is a default option, denoted by x^*, is available as part of any choice set. In other words, when the decision maker faces set S, she is effectively choosing from the set $S^* = S \cup \{x^*\}$. If the decision maker's consideration set is empty, which happens with strictly positive probability, then the decision maker opts for the default option. Therefore, $\sum_{x\in S}\rho(x,S)<1$ and the choice probability of the default option is given by $\rho(x^*,S^*)=1-\sum_{x\in S}\rho(x,S)$. Clearly, we can always create a trivial example showing RCSR is not an SCLM model by utilizing the fact that $\sum_{x\in S}\rho(x,S)<1$ in RCSR. However, the choice behavior implied by the SCLM and RCSR are distinct beyond the fact that $\rho(.,S)$ is not a probability function. To provide a more accurate comparison of these models, we follow the standardization process in Horan [2019] to remove the default option in RCSR, and compare RCSR without the default option with the SCLM model. Consider the random choice rule generated by the RCSR model with $x\succ y\succ z$ and $\gamma(x)=\gamma(z)=0.5$, $\gamma(y)=0.2$. Applying the standardization process to this choice rule yields a random choice rule in which none of the alternatives satisfy stochastic expansion in $\{x,y,z\}$. Thus, the SCLM model does not nest RCSR, RCCSR or RAM. In Appendix B, we add a default option to the SCLM model, and compare the SCLM model with default option with RCSR and RCCSR. Our independence result still holds.

The canonical model of stochastic choice behavior in economics is the random utility model (RUM) (Block and Marschak [1960]). RAM nests RUM (see Cattaneo et al. [2020]), therefore the SCLM model is not nested in RUM. We next show that the RUM is not a subset of the SCLM model. Since the Luce model is a RUM, we first show that the Luce model is not a subset of the SCLM model and conclude that the RUM is not a subset of the SCLM model. This is due to the fact that in the SCLM model with full support random choice rules, Luce ratios between two sets that involve the alternative that satisfies stochastic expansion (in this case, y) necessarily violate Luce's IIA condition. Hence, the Luce model is not a subset of the SCLM model. Therefore, the RUM is not a subset of the SCLM model.

Aguiar et al. [2016] characterize a satisficing choice procedure in which the decision maker has deterministic preferences but the search order is random, causing the choices to appear stochastic. They consider two restrictions on the search order distributions. First, they assume that any alternative will be searched with strictly positive probability and characterize the full support satisficing model (FSSM). Second, in addition to the full support assumption, they assume that the search order distribution is the same for any choice set and characterize the fixed distribution satisficing model (FDSM). They show that the FDSM is the exact intersection of the FSSM and the random utility model. Their characterization of the FSSM shows that FSSM does not impose any restrictions on full support random choice rules. In other words, any random choice rule with full support has a FSSM representation. Therefore, any full support random choice rule that violates Axioms 1-3 has a FSSM representation but it does not have an SCLM representation. When the random choice rule has full support, FDSM is equivalent to the random utility model. Therefore, FSSM and FDSM are not nested in the SCLM model.

However, when we consider random choice rules that FSSM makes a prediction about, *i.e.* random choice rules in which some available alternatives are chosen with 0 probability, FSSM does not nest the SCLM model. Any FSSM or FDSM satisfy the following axiom: If there exists $S \in \mathcal{X}$ such that $x \in S$ and $\rho(x, S) = 0$, then for all $T \in \mathcal{X}$, either $\rho(x, T) = 0$ or $\rho(x, T) = 1$. A SCLM rule with preferences $x \succ y \succ z$, list $[y, x, z]$ and period recall probability function $q(x) = q(y) = q(z) = 0.5$ violates this condition and therefore does not have a FSSM or FDSM representation.

Aguiar [2017] characterizes random choice rules in which alternatives are exogenously bundled into categories, and the decision maker considers each of these categories with some probability (RCG). The decision maker chooses her most preferred alternative that is available and belongs to the chosen category. An alternative a is revealed to be preferred to alternative b in the RCG if for some $S \subseteq \mathcal{X}$ with $b \in S$, $\rho(b, S \cup \{a\}) \neq \rho(b, S)$. The RAM example shows that it is possible to create a cycle in the revealed preference of RCG in a SCLM rule, therefore RCG is not nested in SCLM. Moreover, RCG is a subset of RUM so, RCG does not nest the SCLM model.

Echenique et al. [2018] consider the perception-adjusted Luce model (PALM) that builds on Luce [1959]. In their model, the decision maker perceives alternatives according to a perception ordering. An alternative's choice probability depends on its Luce ratio and the

Luce ratios of previously perceived alternatives. PALM also features a default option. If we consider random choice rules in which the default option is never chosen, PALM reduces to the Luce model. We argued above that the SCLM and Luce models are independent. Hence, the PALM and the SCLM models are independent. In Appendix B, we compare PALM in which the default option is chosen with strictly positive probability with the SCLM model with a default option. The independence result still holds.

We now discuss how the SCLM model is related to the literature on choice from lists. To provide an accurate comparison with the SCLM model, we only consider models in which the domain of the random choice rules is $X \times \mathcal{X}$. Rubinstein and Salant [2006] (RS) first consider a deterministic model of choosing from lists (choice function from list). Then, they consider random choice rules defined over choice sets in which the decision maker deterministically chooses from a list, as outlined in the deterministic model, but the list is stochastic. They characterize all choice functions from lists that result in random choice rules that satisfy regularity and preserves inequalities. A random choice rule ρ preserves inequalities if for any set $S \subseteq X$ and for every $x, y \in S$ either (i) $\rho(x, S) = \rho(y, S) = 0$ or (ii) $\rho(x, S) \geq \rho(y, S)$ if and only if $\rho(x, \{x, y\}) \geq \rho(y, \{x, y\})$. We already showed that the SCLM model violates regularity in the RAM example. Moreover, the SCLM rule generated by $x \succ y \succ z$, $[x, y, z]$, $q(x) = 0.6$, $q(y) = 0.8$ and $q(z) = 0.5$ does not preserve inequalities, as $\rho(x, \{x, y, z\}) < \rho(y, \{x, y, z\})$ while $\rho(x, \{x, y\}) > \rho(y, \{x, y\})$.

Yildiz [2016] considers a decision maker who evaluates a choice set by recursively comparing two alternatives until the choice set is exhausted. The decision maker only makes binary comparisons and these comparisons are random. The order of comparison is captured by a list. Yildiz [2016] characterizes random choice rules that are rationalizable by this process, referred to as list rational (LR) random choice rules. He defines the following stochastic path independence condition on random choice rules:

Definition 4 (Stochastic Path Independence, Yildiz [2016]). The random choice rule ρ satisfies stochastic path independence if for each $T \in \mathcal{X}$ with $x \in T$ and $y \notin T$,

$$\rho(x, T \cup \{y\}) = \rho(x, T)\rho(x, \{x, y\})$$

Stochastic path independence is closely related to our stochastic expansion definition. In LR, stochastic path independence captures the decision maker's recursive comparisons: If an alternative $x \in S$ precedes another alternative y, then adding y to the choice set S decreases the choice probability of x by the amount equal to how often y beats x in binary comparisons. Therefore, the choice probability of x in $S \cup \{y\}$ can be decomposed into the choice probability of x in S and in $\{x, y\}$. In the SCLM model, only the decision maker's most preferred alternative in some set S, say x, satisfies stochastic path independence, because the choice probability of x is equal to its final recall probability. Thus, the choice probability of x in $S \cup \{y\}$ can be decomposed into the choice probability of x in S and in $\{x, y\}$.

In LR, x is revealed to follow y in the list (xfy) if there exists some set $S \in \mathcal{X}$ with $x \in S$, $y \notin S$, in which stochastic path independence is violated. In the SCLM model, it is possible to create cycles in the revealed to follow relation. LR model is not nested in the SCLM model, because stochastic path independence is necessarily violated by alternatives that are not the most preferred alternative in any set S (see Lemma 2). Therefore, the f relation obtained from the random choice rules generated by the SCLM model must be cyclical. Moreover, acyclicity of f characterizes LR, so it is easy to see that LR is not a subset of SCLM.

Kovach and Ülkü [2020] characterize a satisficing model with random thresholds (RSR). In their model, for each choice problem, the DM draws a random threshold alternative x and then constructs a stochastic consideration set by considering only the set of available alternatives that are at least as good as x. If the consideration set is nonempty, the DM compares the alternatives in her consideration set by following a list and choosing the first alternative in the list that is in her consideration set. If the consideration set is empty, or if the drawn threshold alternative is the default option, the DM chooses the default option.

RSR is a random utility model, so it does not contain SCLM. SCLM does not contain RSR, as the following RSR does not have a SCLM representation: consider a RSR with preferences $x \succ y \succ z$, list $[z, y, x]$ and random threshold function $\pi(y) = 0.4$, $\pi(z) = 0.6$. Note that in this random choice rule, the default option is never chosen in binary and ternary sets. In Appendix B, we compare RSR in which the default option is chosen with strictly positive probability with the SCLM model with a default option. The independence result still holds.

5. Conclusion

In this paper, we investigated the choices of a decision maker who suffers from limited memory. The decision maker has a deterministic preference relation over available alternatives and she observes the available alternatives according to a list. As the decision maker goes through the alternatives in the list, she may forget some of the alternatives that were observed. The probability of forgetting decreases if an alternative's position is altered so that it appears later in the list. We showed that we can identify the underlying preference relation, the memory parameters and the list from the decision maker's choice probabilities. We showed that the identification is unique up to the ranking of the two least preferred alternatives. We provided conditions on observable choice probabilities that fully characterize the model. We then applied our representation to study the pricing problem of a monopolist who faces consumers with limited memory. We show that if forgetting is severe, the monopolist is better off charging a lower price than the optimal price in the perfect memory case. Thus, our model can generate new predictions in applications.

We underlined forgetting as a phenomenon that can affect consideration set formation. The parametric structure we imposed on the model allows for precise estimates of memory parameters and allows us to disentangle the effects of memory and preferences on choice probabilities. Considering the interaction between memory and attention, this paper is complementary to previous work on limited attention. Future work may focus on understanding this interaction in economically relevant environments.

Appendix A. Proofs

A.1. **Proof of Lemma 2.** Consider $T \subseteq S$ with $T = \{x, y, z\}$. By Lemma 1, the preference and the list induce the same ranking so there are three cases to consider with respect possible preference rankings and list orderings over $\{x, y, z\}$.

(1) $x \succ y \succ z$: Then $\rho(y, \{x, y, z\}) = (1 - q(x)^2) \cdot q(y) \neq (1 - q(x)) \cdot q(y) = \rho(y, \{x, y\}) \cdot \rho(y, \{y, z\})$.
(2) $x \succ z \succ y$: Then $\rho(y, \{x, y, z\}) = (1 - q(x)^2) \cdot (1 - q(z)) \neq (1 - q(x)) \cdot (1 - q(z)) = \rho(y, \{x, y\}) \cdot \rho(y, \{y, z\})$.
(3) $z \succ x \succ y$: Then $\rho(y, \{x, y, z\}) = (1 - q(z)^2) \cdot (1 - q(x)) \neq (1 - q(z)) \cdot (1 - q(x)) = \rho(y, \{x, y\}) \cdot \rho(y, \{y, z\})$.

Hence, y does not satisfy stochastic expansion in S.

A.2. **Proof of Theorem 1. (Necessity)** The necessity of all axioms is already outlined in the main text.

(Sufficiency) Suppose that ρ is a full support random choice rule satisfying Axioms 1 - 3. Note that by Lemma 1, if a representation exists, the preference order and the list in this representation must be the same ranking. To simplify notation, we denote a SCLM model in which the preference order and the list coincide and is equal to $\succ$ with $(\succ, q)$. By Axiom 1, in any set $S \in \mathcal{X}$, there exists an alternative $x \in S$ such that x satisfies stochastic expansion in S. Define a binary relation $\succ_\rho$ on X as follows: If x satisfies stochastic expansion in S, then $x \succ_\rho y$ for all $y \in S \setminus \{x\}$. If x satisfies stochastic expansion in S, define $q : X \to (0, 1)$ as $q(x) = \rho(x, \{x, y\})$ for $y \in S \setminus \{x\}$. Axiom 2 ensures that this probability is well-defined.

We construct the revealed preference/list $\succ_\rho$ as follows: First, identify the alternative that satisfies stochastic expansion in X, call it x_1. Then, identify the alternative that satisfies stochastic expansion in $X \setminus \{x_1\}$, call it x_2. If x satisfies stochastic expansion in S, then x satisfies stochastic expansion in any set $T \subseteq S$ with $x \in T$ and $|T| \geq 3$. Therefore, the revealed preference relation does not have any cycles. Since stochastic expansion in set S requires S to have at least three alternatives, we can repeat this process until there are two alternatives left. Letting $|X| = n$, we refer to the last two alternatives as x_{n-1} and x_n. Define $x_{n-1} \succ_\rho x_n$ and set $q(n - 1) = \rho(x_{n-1}, \{x_{n-1}, x_n\})$. The period recall probability of the least preferred alternative never appears in any of the choice probabilities. Thus, setting $q(n)$ equal to any number in $(0, 1)$ is consistent with the model.

Consider any $S \in \mathcal{X}$ with $|S| \geq 3$, and denote the alternative that satisfies stochastic expansion in S with x. For any $T \subseteq S$ with $|T| \geq 3$,

$$\begin{aligned}\rho(x, T) &= \prod_{y \in T \setminus \{x\}} \rho(x, \{x, y\}) \\ &= \rho(x, \{x, y\})^{|T|-1}\end{aligned}$$

By Axiom 2 $\rho(x, \{x, y\}) = q(x)$ for all $y \in T \setminus \{x\}$.

$$\rho(x, T) = q(x)^{|T|-1}$$

Therefore, $(\succ_\rho, q(.))$ represents the choice probability of each alternative in each of the sets in which they satisfy stochastic expansion. For x_1, this corresponds to all subsets of X, for x_2 it is all subsets of $X \setminus \{x_1\}$ and so on. What remains to show is that the choice probabilities of alternatives are consistent with the representation in sets in which those alternatives don't satisfy stochastic expansion.

First consider binary sets: $\{x_i, x_j\}$. If $x_i = x_{n-1}$ and $x_j = x_n$ then $\rho(x_i, \{x_i, x_j\}) = q(n-1)$, $x_i \succ_\rho x_j$ by construction. If $x_i \notin \{x_{n-1}, x_n\}$, then there exists $S \in \mathcal{X}$ with $x_j \in S$ such that x_i satisfies stochastic expansion in S. Thus, $x_i \succ_\rho x_j$ and $\rho(x_i, \{x_i, x_j\}) = q(i)$.

For any set $S \in \mathcal{X}$ with $|S| \geq 3$, we provide a proof by induction on the cardinality of the set S. Suppose that $|S| = 3$, $x_i, x_j \in S$ and x_i, x_j do not satisfy stochastic expansion in S. By Axiom 1, there exists $x_k \in S$ such that x_k satisfies stochastic expansion in S, so $\rho(x_k, \{x_k, x_i, x_j\}) = q(k)^2$. Without loss of generality, let $i < j$. By Axiom 3,

$$\frac{\rho(x_i, \{x_k, x_i, x_j\})}{\rho(x_j, \{x_k, x_i, x_j\})} = \frac{\rho(x_i, \{x_i, x_j\})}{\rho(x_j, \{x_i, x_j\})} = \frac{q(i)}{1 - q(i)}$$

Combining this with $\rho(x_i, \{x_k, x_i, x_j\}) + \rho(x_j, \{x_k, x_i, x_j\}) = 1 - q(k)^2$ yields,

$$\rho(x_i, \{x_k, x_i, x_j\}) = (1 - q(k)^2)q(i)$$
$$\rho(x_j, \{x_k, x_i, x_j\}) = (1 - q(k)^2)(1 - q(i))$$

We now extend the same idea to sets with $|S| > 3$. The induction hypothesis is that the representation holds for $x_i, x_j \in T$ and x_i, x_j, do not satisfy stochastic expansion in T and $|T| = n$. We show that the representation holds for $x_i, x_j \in S$, where x_i, x_j do not satisfy stochastic expansion in S and $|S| = n + 1$.

By Axiom 1, there exist an alternative x_k that satisfies stochastic expansion in S. Therefore,

$$\rho(x_k, S) = q(k)^{|S|-1}$$

By Axiom 3, for any $x_i, x_j \in S \setminus \{x_k\}$,

$$\frac{\rho(x_i, S)}{\rho(x_j, S)} = \frac{\rho(x_i, S \setminus \{x_k\})}{\rho(x_j, S \setminus \{x_k\})}$$

Since $S \setminus \{x_k\}$ is a set of size n

$$\frac{\rho(x_i, S)}{\rho(x_j, S)} = \frac{\rho(x_i, S \setminus \{x_k\})}{\rho(x_j, S \setminus \{x_k\})} = \frac{q(i)^{|\mathcal{L}(x_i, S\setminus\{x_k\}|)} \prod_{y \in U_\succ(x_i, S\setminus\{x_k\})}(1 - q(y)^{|\mathcal{L}(y, S\setminus\{x_k\})|})}{q(j)^{|\mathcal{L}(x_j, S\setminus\{x_k\}|)} \prod_{y \in U_\succ(x_j, S\setminus\{x_k\})}(1 - q(y)^{|\mathcal{L}(y, S\setminus\{x_k\})|})}$$

Since $\sum_{x\in S\setminus\{x_k\}} \rho(x,S) = 1 - q(k)^{|S|-1}$

$$\rho(x_i,S) = (1-q(k)^{|S|-1})q(i)^{|\mathcal{L}(x_i,S\setminus\{x_k\})|}\prod_{y\in U_\succ(x_i,S\setminus\{x_k\})}(1-q(y)^{|\mathcal{L}(y,S\setminus\{x_k\})|})$$

By Lemma 1, $\mathcal{L}(x_k,S) = S\setminus\{x_k\}$ so $\mathcal{L}(x,S\setminus\{x_k\}) = \mathcal{L}(x,S)$ for all $x\in S\setminus\{x_k\}$.

$$\rho(x_i,S) = (1-q(k)^{|S|-1})q(i)^{|\mathcal{L}(x_i,S)|}\prod_{y\in U_\succ(x_i,S\setminus\{x_k\})}(1-q(y)^{|\mathcal{L}(y,S)|})$$

$$\rho(x_i,S) = q(i)^{|\mathcal{L}(x_i,S)|}\prod_{y\in U_\succ(x_i,S)}(1-q(y)^{|\mathcal{L}(y,S)|})$$

Repeating the same argument for x_j,

$$\rho(x_j,S) = q(j)^{|\mathcal{L}(x_j,S)|}\prod_{y\in U_\succ(x_j,S)}(1-q(y)^{|\mathcal{L}(y,S)|})$$

A.3. **Application Extension.** We extend the monopoly pricing application in Section 3 by adding three more parameters: one parameter to distinguish the preference intensity of two types of consumers (β), one parameter for the consumer type distribution (θ) and one parameter for the list distribution (λ). The consumer tastes are given by $u_1(1) = u_2(2) = 1$ and $u_1(2) = \beta$ $u_2(1) = \alpha$, where $\alpha,\beta\in(0,1)$. The monopolist faces consumer 1 with probability θ, and consumer 2 with probability $1-\theta$. Consumers use the list $[x_1,x_2]$ with probability λ and the list $[x_2,x_1]$ with probability $1-\lambda$. Table 2 shows the expected demand of consumer 1 and j for good 1 and good 2. Table 3 shows the expected share of consumers who opt for the outside option.

p	$\rho_i(1,p)$	$\rho_j(1,p)$	$\rho_i(2,p)$	$\rho_j(2,p)$
$(1,1)$	$\lambda q(1)+1-\lambda$	0	0	$\lambda+(1-\lambda)q(2)$
$(1,\beta)$	$\lambda q(1)+1-\lambda$	0	$\lambda(1-q(1))$	$\lambda+(1-\lambda)q(2)$
$(\alpha,1)$	$\lambda q(1)+1-\lambda$	$(1-\lambda)(1-q(2))$	0	$\lambda+(1-\lambda)q(2)$
(α,β)	$\lambda q(1)+1-\lambda$	$(1-\lambda)(1-q(2))$	$\lambda(1-q(1))$	$\lambda+(1-\lambda)q(2)$

Table 2. Expected consumer demand.

p	Outside Option
$(1,1)$	$\theta\lambda(1-q(1))+(1-\theta)(1-\lambda)(1-q(2))$
$(1,\beta)$	$(1-\theta)(1-\lambda)(1-q(2))$
$(\alpha,1)$	$\theta\lambda(1-q(1))$
(α,β)	0

Table 3. Expected share of consumers who opt for the outside option.

The following proposition characterizes the optimal pricing choice of the monopolist for any given $\theta,\lambda,\alpha,\beta,q(1),q(2)\in(0,1)$. Let $K_1 = \lambda q(1)+1-\lambda$ and $K_2 = \lambda+(1-\lambda)q(2)$.

Proposition 2. (1) If $q(2) \geq 1 - \frac{\theta(1-\alpha)K_1}{\alpha(1-\theta)(1-\lambda)}$, then $p_1 = 1$; otherwise, $p_1 = \alpha$.
(2) If $q(1) \geq 1 - \frac{(1-\theta)(1-\beta)K_2}{\beta\theta\lambda}$, then $p_2 = 1$; otherwise, $p_2 = \beta$.

Appendix B. Stochastic Choice with Limited Memory with Default Option

In this section, we extend the SCLM model to accommodate a default option. We then compare the choice behavior implied by our model to other models that feature a default option, namely the stochastic choice models of Manzini and Mariotti [2014] (RCSR), Brady and Rehbeck [2016] (RCCSR), Echenique et al. [2018] (PALM) and Kovach and Ülkü [2020] (RSR). Suppose that the decision maker does not always recall the last alternative in the list. We now assume that after observing the last alternative x in the list, she can forget x at the time of choice with probability $1 - q(x)$. Whenever the decision maker forgets every available alternative, she chooses the default option x^*, which is interpreted as choosing nothing. The default option is always available as part of every choice set. To indicate the choice sets augmented with the default option, we use the notation, $X^* = X \cup \{x^*\}$, $S^* = S \cup \{x^*\}$. Formally,

Definition 5. A *random choice rule with default option* is a mapping $\pi : X^* \times \mathcal{X} \to (0,1)$ such that for all $S \in \mathcal{X}$, $\sum_{x \in S^*} \pi(x, S) = 1$.

We formally define a SCLM model with a default option as follows,

Definition 6. A random choice rule with default option π has a *stochastic choice with limited memory with default option (SCLM*)* representation if there exists a preference ordering $\succ$ on X, a list $\mathcal{L} \in \mathbb{L}$, and a period recall probability function $q : X \to (0,1)$, such that for any $S \in \mathcal{X}$ and $x \in S$,

$$\pi(x, S) = q(x)^{|\mathcal{L}(x,S)|+1} \prod_{\{y \in S \ | \ y \succ x\}} (1 - q(y)^{|\mathcal{L}(y,S)|+1})$$

The random choice rule with default option generated by the RCSR model with $x \succ y \succ z$ and $\gamma(x) = \gamma(z) = 0.5$, $\gamma(y) = 0.2$, does not have an SCLM* representation. Therefore, RCSR is not a a subset of our model. As RCCSR generalizes RCSR, it is also not a subset of the SCLM* model.

We now show that the SCLM* model is not a subset of RCCSR. The first axiom in the characterization of RCCSR is as follows,

Axiom (Asymmetric Sequential Independence)**.** For all distinct $x, y \in X$, exactly one of the following holds,

$$\pi(x, \{x,y\}) = \pi(x, \{x\})(1 - \pi(y, \{x,y\})) \text{ or } \pi(y, \{x,y\}) = \pi(y, \{y\})(1 - \pi(x, \{x,y\}))$$

Consider the random choice rule with default option, generated by the SCLM* with $x \succ y$, $\mathcal{L} = [y, x]$, $q(x) = q(y) = 0.5$. Neither $\pi(x, \{x,y\}) = \pi(x, \{x\})(1 - \pi(y, \{x,y\}))$ nor $\pi(y, \{x,y\}) = \pi(y, \{y\})(1 - \pi(x, \{x,y\}))$ is satisfied. Therefore, π does not have a RCCSR representation. Hence, we conclude that the SCLM* model is not nested in RCCSR, and therefore, it is not nested in RCSR.

We now show that SCLM* and PALM are independent from each other. Consider the example generated by PALM with $u(x) = 0.75$, $u(y) = 0.5$, $u(x^*) = 0.25$ and perception order $x \sim y$: $\pi(x,\{x\}) = 3/4$ $\pi(y,\{y\}) = 2/3$, $\pi(x,\{x,y\}) = 1/2$ $\pi(y,\{x,y\}) = 1/3$. π does not have an SCLM* representation, so PALM is not a subset of SCLM*. Next, we show that SCLM* model is not a subset of PALM. Consider the example, generated by the SCLM* model with $x \succ y \succ z$, $\mathcal{L} = [y,x,z]$, $q(x) = q(z) = 0.5$, $q(y) = 0.8$. In PALM, x is revealed to be perceived at the same time as y or $x \sim^0 y$ if

$$\frac{\pi(x,\{x,y\})}{\pi(y,\{x,y\})} = \frac{\pi(x,\{x,y,z\})}{\pi(y,\{x,y,z\})}$$

for all $z \in X$. Alternative x is revealed to be perceived before y or $x \succ^0 y$ if

$$\frac{\pi(x,\{x,y\})}{\pi(y,\{x,y\})} > \frac{\pi(x,\{x,y,z\})}{\pi(y,\{x,y,z\})}$$

for all $z \in X$ such that $z \nsim^0 x$ and $z \nsim^0 y$, and if there exists at least one such z.

It can be shown that π has the following cycle: $x \succ^0 y \succ^0 z \succ^0 x$. Therefore, we conclude that π does not have an SCLM* representation and PALM does not nest SCLM*.

We now compare RSR and SCLM*. RSR is a random utility model, so SCLM* is not a subset of RSR. Consider the RSR with preferences $x \succ y \succ z$, list $[z,y,x]$ and random threshold function $\pi(y) = 0.4$, $\pi(z) = 0.5$, $\pi(x^*) = 0.1$. This random choice rule with default option does not have SCLM* representation. Therefore, RSR and SCLM* are independent.

References

M. Agranov and P. Ortoleva. Stochastic choice and preferences for randomization. *Journal of Political Economy*, 125(1):40–68, 2017.

V. H. Aguiar. Random categorization and bounded rationality. *Economics Letters*, 159: 46–52, 2017.

V. H. Aguiar, M. J. Boccardi, and M. Dean. Satisficing and stochastic choice. *Journal of Economic Theory*, 166:445–482, 2016.

A. H. Ashton and R. H. Ashton. Sequential belief revision in auditing. *Accounting Review*, pages 623–641, 1988.

J. Blanchet, G. Gallego, and V. Goyal. A markov chain approximation to choice modeling. *Operations Research*, 64(4):886–905, 2016.

H. Block and J. Marschak. Random orderings and stochastic theories of responses. In *Economic Information, Decision, and Prediction*, pages 172–217. Springer, 1960.

R. L. Brady and J. Rehbeck. Menu-dependent stochastic feasibility. *Econometrica*, 84(3): 1203–1223, 2016.

M. D. Cattaneo, X. Ma, Y. Masatlioglu, and E. Suleymanov. A random attention model. *Journal of Political Economy*, 128(7):2796–2836, 2020.

M. M. Chun and N. B. Turk-Browne. Interactions between attention and memory. *Current Opinion in Neurobiology*, 17(2):177–184, 2007.

N. Cowan. *Attention and memory: An integrated framework.* Oxford University Press, 1998.

M. T. Dewar, N. Cowan, and S. Della Sala. Forgetting due to retroactive interference: A fusion of müller and pilzecker's (1900) early insights into everyday forgetting and recent research on anterograde amnesia. *Cortex*, 43(5):616–634, 2007.

P. R. Dickson and A. G. Sawyer. The price knowledge and search of supermarket shoppers. *Journal of Marketing*, 54(3):42–53, 1990.

F. Echenique, K. Saito, and G. Tserenjigmid. The perception-adjusted luce model. *Mathematical Social Sciences*, 93:67–76, 2018.

J. E. Fan and N. B. Turk-Browne. Incidental biasing of attention from visual long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(6): 970, 2016.

J. L. Farr. Response requirements and primacy-recency effects in a simulated selection interview. *Journal of Applied Psychology*, 57(3):228, 1973.

J. L. Farr and C. M. York. Amount of information and primacy-recency effects in recruitment decisions. *Personnel Psychology*, 1975.

A. Furnham. The robustness of the recency effect: Studies using legal evidence. *The Journal of General Psychology*, 113(4):351–357, 1986.

E. N. Garbinsky, C. K. Morewedge, and B. Shiv. Interference of the end: why recency bias in memory determines when a food is consumed again. *Psychological Science*, 25(7): 1466–1474, 2014.

F. Gul, P. Natenzon, and W. Pesendorfer. Random choice as behavioral optimization. *Econometrica*, 82(5):1873–1912, 2014.

D. E. Hannula. Attention and long-term memory: Bidirectional interactions and their effects on behavior. In *Psychology of Learning and Motivation*, volume 69, pages 285–323. Elsevier, 2018.

J. D. Hey. Does repetition improve consistency? *Experimental economics*, 4(1):5–54, 2001.

S. Horan. Random consideration and choice: A case study of "default" options. *Mathematical Social Sciences*, 102:73–84, 2019.

W. D. Hoyer. An examination of consumer decision making for a common repeat purchase product. *Journal of Consumer Research*, 11(3):822–829, 1984.

M. Kovach and L. Ülkü. Satisficing with a variable threshold. *Journal of Mathematical Economics*, 87:67–76, 2020.

R. D. Luce. *Individual choice behavior: A theoretical analysis.* New York, Wiley, 1959.

J. G. Lynch, J. Alba, and J. W. Hutchinson. Memory and decision making. *Handbook of Consumer Behavior*, pages 1–9, 1991.

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

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

E. Maskin and J. Riley. Monopoly with incomplete information. *The RAND Journal of Economics*, 15(2):171–196, 1984.

G. E. Müller and A. Pilzecker. *Experimentelle beiträge zur lehre vom gedächtniss*, volume 1. JA Barth, 1900.

A. Rubinstein and Y. Salant. A model of choice from lists. *Theoretical Economics*, 1(1):3–17, 2006.

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

D. D. Steiner and J. S. Rain. Immediate and delayed primacy and recency effects in performance evaluation. *Journal of Applied Psychology*, 74(1):136, 1989.

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

K. Yildiz. List-rationalizable choice. *Theoretical Economics*, 11(2):587–599, 2016.