Consumer Memory and Competitive Interference: The Case of Auto Insurance Advertising *

Navdeep S. Sahni† Yifan Yang‡

November 16, 2024

Abstract

Insights from psychology suggest competitive interference: when a firm advertises, consumers are less likely to think of its competitor, benefiting the advertiser by keeping it top of mind and suppressing competitor recall. We empirically investigate this phenomenon in the context of auto insurance, a sector featuring some of the most advertised brands in the US. Using a large-scale randomized field experiment, we examine the dynamic effect of advertising on an individual's consideration of the advertised insurer. Despite our advertiser being well-known and one of the largest spenders on advertising, the magnitude of our observed effect surpasses those documented in prior literature. The immediate and carry-over effects of advertising depend heavily on competitive actions, consistent with psychological predictions; the advertising impact increases when competitors advertise. Together with other findings, our data support interference as a driver of the marginal effects of ads. We build and estimate a cognitive psychology-based consumer decision-making model, demonstrating that it predicts both in- and out-of-sample data better than the flexible standard advertising stock model.

Keywords: advertising, memory, cognitive psychology, field experiment, structural model, causal inference, information retention

*We thank Matt Gentzkow, Sam Goldberg, Yewon Kim, Sanjog Misra, K. Sudhir, Anthony Wagner; seminar participants at Chicago Booth, Stanford Memory Lab at the Psychology Dept, Amazon Ad science, University of Miami, University of Rochester, Marketing Science Conference, University of Toronto, AI and ML conference at Temple University, Yale conference on Customer Insights, Northwestern University, Columbia University for thoughtful comments and suggestions.

We thank our data sponsors, Stanford GSB for sponsoring access to the comScore data, the Kilts Center at The University of Chicago for sharing the Nielsen Ad Intel data.

Disclaimer: Researcher(s)' own analyses calculated (or derived) based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

The conclusions drawn from the Nielsen data are those of the researcher(s) and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

†Associate Professor of Marketing, Stanford University Graduate School of Business. navdeep.sahni@stanford.edu

‡Doctoral Candidate, Stanford University Graduate School of Business. yifany@stanford.edu

1 Introduction

Consumers' decisions crucially depend on what they know. Hence, researchers have extensively studied how exposure to information affects consumer choices (e.g., Stigler, 1961). However, the impact of information hinges on its recall during the decision-making process, which is not guaranteed. The prevalent view in psychology suggests information recall is subject to interference, which implies that exposure to information can hinder the recall of previously known information. Overlooking this can lead to an incomplete understanding of consumer decision-making. This paper examines this understudied phenomenon by analyzing the impact of exposure to advertising on consumer online shopping decisions using a unique randomized field experiment.

While advertising's effects have been extensively studied, traditional theories fail to fully explain how advertising influences consumer decisions. A primary view is that ads disseminate information about the advertised product (Nelson, 1974). However, this perspective does not address how long the information is retained, making it insufficient to explain the widespread practice of repetitive advertising by well-known brands. An alternative perspective suggests that advertising enhances the perceived utility of the advertised product (Bagwell, 2007). However, this view is overly broad and does not specify the conditions under which advertising will impact consumer behavior or the duration of this impact. To address these limitations and comprehensively analyze the impact of advertising, we focus on its influence on the information recall process.

In this paper, we first build a consumer decision-making model that incorporates memory interference and is inspired by the computational cognitive psychology literature (Anderson, 1996; Kahana, 2012). We draw testable implications of the model. We then gather individual-level data from a field experiment randomizing auto insurance advertising exposures and consumer behavior to test our model's implications. Overall, we find significant interference, which is important to explain the prevalent advertising: one firm's advertising diminishes the recall of competitors' offerings during consumer decision-making, thus motivating competitive advertising.

The consumer in our model stores product options in a memory network in which the product options are linked to the product category. When the consumer considers a purchase in the category, she draws the associated product options from memory, and different options compete to be drawn. This conceptualization of memories being organized in a network and competing for salience is well-established in the literature on memory models (Kahana, 2012). Recalling a product strengthens its association with the category thereby increasing the likelihood of its future recall and decreasing the likelihood of a competitor becoming salient. Over time, the strength of the association decays unless another recall occurs. Advertising exposure strengthens the association of the advertiser and the category, decreasing the likelihood that the consumer will consider a competitor. Hence, advertising by well-known firms may make competing brands less likely to come to mind – effectively causing consumers to "forget" about them – thereby incentivizing competitive advertising.

To empirically investigate this effect, we focus on auto insurance advertising. This is an appealing

context for our study because it is characterized by well-known and recognized brands that have existed for generations but are also among the most heavily advertised. Annually, the most prominent brands spend over a billion dollars each on advertising, equivalent to over a thousand impressions of premium digital ads per adult in the US, on average. Notably, their advertising campaigns are characterized by repetition of the same message, often calling for individuals to "get a quote" and feature mascots that have achieved near-ubiquitous recognition (Advertising Age, 2022).

To empirically examine our model's implications we need to estimate the effects of advertising which is known to be challenging because, in practice, ads are highly targeted based on many factors that remain unobserved to researchers. Hence, researchers have to overcome a significant selection problem. Moreover, our objective of studying memory interference requires us to go beyond typical ad effects measurement and estimate the effects of ads across competitors, which necessitates linking a brand's ad exposure at the individual level with behavior concerning competing brands. This is hard to do, given the data typically available to advertisers. The need to estimate time-varying effects adds to the challenge because it requires stable long-term individual identifiers that are challenging to obtain.

We overcome these challenges by partnering with a large media platform. We analyze a field experiment conducted on the platform in which exposure to ad banners of one of the largest U.S. auto insurance companies was randomized across individuals, enabling us to overcome the potential selection problems caused by ad targeting. We focus on three hundred thousand individuals who also had installed a browser plugin owned by the same platform, which enables us to directly link randomized experimental assignment with online consumer behavior over time, both in relation to the advertiser and its competitors. This setup also enables us to identify the individual's geographic location, which we use to isolate additional exogenous variation in competitive advertising.

As the main outcome, we focus on whether consumers visit a firm's website which signifies them recalling *and* wanting to consider the firm. In our data we do not observe consumer recalls about a product option, which would be closer to our model's predictions but is generally hard to measure in a natural setting at scale. In addition to the focal outcome, we consider customers switching their insurance providers to assess the economic significance of the focal ad campaign.

We find that despite being a marginal campaign featuring our focal advertiser's standard ads, it significantly impacts consumer behavior. The likelihood of individuals visiting the advertiser's website increases from 0.2% to 0.8% on the day of experimental ad exposure, representing a substantial 300% relative increase. Moreover, we detect a significant effect persisting for up to two months after the campaign's conclusion. Although smaller in magnitude per day compared to the effect during the campaign, this long-lasting impact accumulates over time, resulting in a total effect comparable to that observed during the campaign period.

The observed effect suggests that, although the advertiser had been extensively advertised previously, the experimental ad that contained little new information, was necessary for a significant portion of the interested population to consider the advertiser. The ad effects we find are remarkably large;

our effects consistently rank high, at least in the top quartile, compared to similar effects analyzed in recent meta-analyses (Johnson et al., 2017; Gordon et al., 2023).

Our model predicts that experimental ad exposure makes it less likely that consumers recall competitors, and this negative effect on competitors may increase in magnitude after the experiment is over: this is because ad exposure highlights the category which may benefit the competitor, but this benefit vanishes when consumers recall options without the help of ads. Consistent with this prediction, we find that the visits to competitors decreases more on the day after the experimental ad exposure, when consumers are more likely to think of the auto insurance category on their own. This moment directly contradicts the standard advertising paradigm that models ads as increasing preference for the advertised product, which predicts the highest negative effects on competitors on the day of exposure.

We next examine how competitive actions change the experimental treatment effect, and find that the treatment effect increases at times when competitors advertise in the future. We use the fact that the Major League Baseball (MLB) season started three weeks after our experimental campaign and our focal advertiser's competitors prominently advertised in the MLB games.[1] We find that the effect of our experimental treatment increased in the week after MLB season started relative to the week before, and this effect is greater for individuals in DMAs where MLB is more popular. This finding is at odds with the standard advertising models that assume advertising effect monotonically decreases over time. However, this finding is aligned with our model: the treatment group individuals more strongly associate our focal advertiser with auto insurance in their memory, relative to control, so they are more likely to consider our focal company after competitive advertising occurs.

To examine this notion in a different cut of the data, we analyze how our experimental treatment effect changes with our focal advertiser's and its competitors' pre-experimental TV advertising. Based on the observed location of the individuals in the experiment we merge TV advertising conducted by our focal advertiser and its competitors and study the interaction between past TV advertising and the effect of experimental ads. We find that the effect of experimental advertising increases with competitor TV advertising but decreases with our focal advertiser's past TV advertising. A simple extrapolation implies that if all competitor TV advertising was turned off, the experimental banners would have no effect on the day of the experiment. Overall, these and other heterogeneous treatment effects we report below support our memory based model.

We then specify an empirical version of a memory based model of consumer website visits. Estimating this model we are able to predict several in- and out-of-sample moments in the data better than the standard flexible ad stock model commonly used in the advertising literature. While the standard model does not predict the interaction between own and competitive ads, ours does.

On the whole, we find evidence supporting a mechanism that suggests advertising interferes with consumers' information recall process which affects the options they consider while making decisions. Advertising doesn't just convey information and make consumer aware of the advertised message.

[1] In our sampling 16 of the 17 opening day games featured prominently visible auto insurance banners from our advertiser's competitors; our focal advertiser appeared once.

Instead, it deters competitors by making it less likely that they become salient in the consumer's mind. A competitor's advertising can incentivize a firm to advertise repetitively. In our setting, competitive advertising makes it less likely that the consumers will recall and consider our focal firm without advertising, thereby increasing the effect of advertising.

Our work draws heavily from cognitive psychology where the *consensus* view is that forgetting is driven not just by the passage of time, but rather by interference. This is a textbook view, with experimental studies dating back to Müller and Pilzecker (1900) (reviewed in Lechner et al., 1999). On this, (Kahana, 2012, p. 9) notes "rather than reflecting the loss of information (forgetting) represents the interference of later learning with earlier learning".

Our model of recall of previously learned information is inspired by the literature on models of human memory that operationalize recall as a process of making draws from memory in which individual memories compete for salience. This is a feature of the standard models (e.g., ACT model Anderson, 1993; SAM model Raaijmakers and Shiffrin, 1980) that have been tested to predict recall in lab-based recall tasks. We extend these findings to model actual behavior and quantify their implications in the real world.

Our work is related to the literature that has systematically analyzed the role of limited consumer memory in economic outcomes (Mullainathan, 2002). Relatively recent research incorporates memory interference in the formation of consumer beliefs and explicitly models it via memory sampling, as we do in our model (Bordalo et al., 2020; Sial et al., 2023). This literature explains consumer stated intentions and beliefs in several domains including financial decision making (Bordalo et al., 2023). We contribute to it by demonstrating the real-world behavioral significance of the interference phenomenon in the context of advertising. Additionally, memory associations in our model are strengthened with rehearsals, which is essential for accounting for repetitive advertising.

Research in marketing views advertising effects as being driven by memory mechanisms (Keller, 1987). Several studies have examined interference in the lab. Alba and Chattopadhyay (1986) show that making one brand salient decreases the likelihood of consumers thinking about competing brands. Consistent with the notion that cues interfere in the recall process, Burke and Srull (1988) show that advertising of a brand interferes with the other brand's advertising and changes its effect. Danaher et al. (2008) show evidence of positive competitive advertising interactions in a grocery setting, consistent with our findings. Our work extends this literature by showing the significance of interference phenomenon using a field experiment, and provides a model to assess its impact on dynamic consumer behavior. The mechanism we focus on complements other work that has built psychology-based models of advertising (Sahni, 2015; Naik et al., 1998) but has not incorporated interference.

Our research sheds light on when advertising can be combative (Bagwell, 2007). Researchers analyzing the consequences of advertising on various market outcomes have assumed various consumer decision making models. The most common approach is based on Nerlove and Arrow (1962) conceptualizes advertising exposure builds an advertising "stock", sometimes referred to as goodwill stock. An increase in this advertising stock increases the expected utility the consumer gets from the advertised

option, which then implies an increase in the consumer choosing the advertised option (e.g., Dubé et al., 2005; Doraszelski and Markovich, 2007). Although variations of this model can be flexible, they do not fully comply with our data. For example, these models predict that a firm's advertising makes the consumers substitute away from the competitors and, hence, competitors suffer the most when the advertiser's gains are the highest. This prediction does not hold in our data, which shows patterns consistent with our memory-based model. These utility increasing models also do not say much about the role of competition in the decay of the advertising stock. In section 8 we estimate these models for detailed comparison.

Another stream of research theoretically studies the effects of advertising using models that incorporate features of interference (e.g., Haan and Moraga-González, 2011; Anderson and De Palma, 2013; Chen et al., 2009). These models allow competitors to displace each other in the consumer's consideration but preclude the potential for positive memory spillovers, effects that are both observed in our findings and noted in other studies on advertising (Sahni, 2016; Shapiro, 2018). Our model provides a psychology based approach to extend these models to allow for such spillovers.

Finally, our paper is related to the empirical work in area of advertising that has documented its effects and highlighted the difficulty in obtaining credible effects due to selection on unobservable characteristics (Gordon et al., 2023) and low statistical power (Lewis and Rao, 2015). Consistent with the recommendations of this literature, we use randomization to estimate the effect of advertising and extend it by examining long-term effects, which is rare, and showing evidence for a mechanism driving ad effects.[2]

The rest of the paper is organized as follows. Section 2 presents our model in a simple two product setting and derives empirical predictions to guide the empirical analysis. Section 3 provides background on the setting of auto insurance advertising and 4 describes our experiment design. We describe our data in Section 5 and discuss our estimated advertising effects in 6. Section 7 shows competitive interactions and 8 estimates the our memory based models to compare its predictions with those of a standard advertising stock model. Finally, Section 9 concludes the paper.

2 Model

We present a model to precisely describe interference and its effects, and use it to guide our empirical analysis in the following sections. Our model is inspired by the literature in cognitive psychology that has developed models of information retrieval from human memory. We incorporate the basic ideas that have made their way into textbook versions of the standard models, for example, Anderson (2000) and Kahana (2012).[3]

[2]Focusing on auto insurance, Tsai and Honka (2021) use survey data from JD Power to show that people recall 30% of the auto insurance brands they are aware of, on average, and TV advertising increases advertisers' recall. Their estimate of TV advertising elasticity implies a smaller effect than ours. This could be due to differences in research designs – we observe individual-level ad exposure while they use aggregate metrics – or the fact that, unlike TV advertising, consumers are able to click and act at the time of exposure to online ads.

[3]These models are built to explain patterns of associations formed in the lab. In a typical experiment demonstrating interference, participants are asked to memorize a list of words, then given a second list of words, and finally asked to

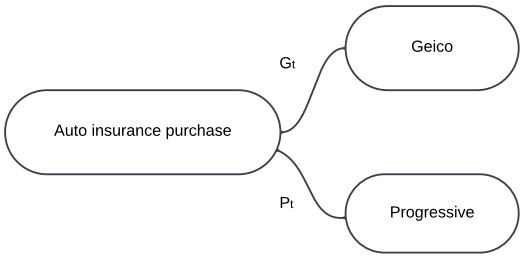


Figure 1: Memory network

Our model is close in its form to the Search of Associative Memory (SAM) model in Raaijmakers and Shiffrin (1981). Essentially, the consumer in our model stores information in his memory, which is an associative network of memory images. During recall, he draws a subset of this information from memory, and different pieces of information compete to be drawn. The recalled information influences his decision-making. Below, we specify our model for the context of auto insurance shopping. We first describe the memory structure, then the process that drives salience and the role of advertising, followed by implications for consumer decision-making. Appendix A discusses the parallels with the SAM model in more detail.

2.1 Information stored in an associative network of images

Consider an individual browsing the internet. He stores memories related to auto insurance in the form of a linked network of images, which may not be salient at any given time. For our context, one memory image is the "Auto Insurance Purchase Context," and others linked to it are about auto insurance options. For simplicity of exposition in this section, we assume just two auto insurance options, Geico and Progressive, shown in Figure 1. There could be other associations and images, for example, related to past experiences which we assume away for simplicity. This star-like structure of the network is similar to how consumer psychologists have conceived the organization of brand memories (e.g., Nedungadi, 1990).

2.2 Strength of association

The existence of a link between memory images implies their association in the consumer's memory. The links are characterized by a strength that signifies the degree of association. The association strengths are time-specific. We denote the strengths of association at the beginning of day t between "Auto Insurance Purchase Context" and "Progressive" and "Geico" by P_t and G_t, respectively.

The association strength increases when the images are recalled together. So, if "Auto Insurance Purchase Context" and "Progressive" are drawn together their association becomes stronger. Oth-

recall the first list. The second list of words can be either related or unrelated to the first. Participants' accuracy in recalling the first list decreases when the second list is related to it compared to when it is unrelated, suggesting the presence of interference.

erwise, it decays over time with a survival rate of δ per day. Hence, $P_t = \delta\,(rp_{t-1} + P_{t-1})$ where $0 < \delta < 1$, r is the increase in memory strength due to a recall, and p_{t-1} is the number of times Progressive was recalled with the purchase context on day $t-1$. Similarly, $G_t = \delta\,(rg_{t-1} + G_{t-1})$.

2.3 Drawing information in the absence of advertising

The consumer may not recall the auto insurance purchase context and forgets it with probability ϕ on any given day. With probability $1-\phi$, the purchase context recall gets triggered, possibly due to various events such as a conversation about auto insurance.

On recalling the context, the consumer sequentially draws options from memory with replacement. The likelihood of drawing Progressive when the consumer first queries his memory on day t is given by

$$h(P_t, G_t) = \frac{P_t}{P_t + G_t}.$$

This functional form aligns with textbook models of recall from memory (c.f., Kahana, 2012, chapter 7). Since $h(P_t, G_t)$ decreases with G_t, Progressive is less likely to be drawn if Geico is more strongly associated with the purchase context, which happens if it has been recalled more often previously.

Once an option is drawn, its strength of association increases by r which makes it more likely to be drawn in the next attempt.

The consumer continues drawing options from memory and stops when he discovers no new information from the last draw. This aspect of our model, along with the fact that the consumer draws with replacement, aligns with foundational principles in psychological models (e.g., Rundus, 1973).[4] In our analysis below, we assume that the individual ceases to draw when a retrieval yields no new information. Employing an alternative stopping criterion, such as a fixed number of draws, does not alter the fundamental implications of the interference phenomenon we study.

2.4 From Recall to Behavior

Before making a purchase, the consumer may consider a recalled option by visiting its website and evaluating the relevant information. The decision to visit a website results from a tradeoff between the cost of processing the information on the website and the benefit of potentially buying better auto insurance. The purchase itself is an outcome of a comparison of the considered options.

We focus on website visits as the primary outcome of interest because it is the closest observable behavior to the recall process, which is our focus. However, besides recall, a website visit indicates a preference above a certain threshold. Additionally, recalls may not lead to website visits if they occur on inconvenient occasions, such as when the consumer is preoccupied.

[4]Describing this model, Alba and Chattopadhyay (1985) state "The cessation rule states that attempts to recall new items may end once a criterion number of consecutive samples is made from which no new items are obtained."

Overall, the consumer visits Progressive's website with probability

$$\begin{aligned}\text{Visit}_{P,t}(\text{No Ad}_t) &= \Pr(\text{Recalling context}) \times \Pr(\text{Drawing } P \mid \text{Recalling context}) \times \Pr(\text{Visit } P|\text{Drawing } P) \\ &= \underbrace{(1-\phi)\left(1-h(G_t,P_t)h(G_t+r,P_t)\right)}_{\Pr(\text{Drawing } P \text{ at } t|\text{No Ad}_t)} \times \Pr(\text{Visit } P|\text{Drawing } P).\end{aligned}$$

2.5 Effect of advertising

Advertising causes the consumer to simultaneously think of the purchase context and the advertised brand. Since the auto insurance purchase context is salient in the consumer's mind, he considers shopping for it. At this time, the advertiser's salience makes it a primary option for consideration. The simultaneous recall of both the advertiser and the purchase context reinforces their association within the consumer's memory, leading to a stronger future likelihood of the consumer recalling the advertiser and overshadowing other options.

2.5.1 Contemporaneous effect

Since Progressive's advertising at time t makes it salient in the consumer's mind, the consumer visits its website with probability $\text{Visit}_{P,t}(\text{Ad}_t) = \Pr(\text{Visit } P|\text{Draw } P)$. Therefore, the effect of advertising is

$$\text{Visit}_{P,t}(\text{Ad}_t) - \text{Visit}_{P,t}(\text{No Ad}_t) = \Pr(\text{Visit } P \mid \text{Draw } P) \times \left[1 - \Pr(\text{Draw } P \text{ at } t \mid \text{No Ad}_t)\right], \quad (1)$$

which says that the impact of advertising depends on the likelihood of the consumer thinking of Progressive in the absence of the ad. In a world where the consumers always recall the options they would consider and only fail to recall options they would not consider, advertising would not affect visits because $\Pr(\text{Draw } P \text{ at } t \mid \text{No Ad}_t) = 1$ when $\Pr(\text{Visit } P \mid \text{Draw } P) \neq 0$.

If conditional on recalling the purchase context, consumers certainly recall options they would consider, advertising of any brand will impact P similarly. The advertiser's identity would not matter because any advertiser makes the purchase context salient.

However, advertising will distinctly affect the advertiser's visits if the consumer does not recall the advertiser even though he would consider it if he did.[5]

Role of past advertising Since $\frac{\partial \Pr(\text{Draw P at } t|\text{No Ad}_t)}{\partial G_t} < 0$, our model and equation (1) predict that Progressive's advertising will have an unambiguously larger contemporaneous effect on its visits if G_t is higher, which could be a consequence of past exposure to Geico ads that increases G_t. Similarly, since $\frac{\partial \Pr(\text{Draw P at } t|\text{No Ad}_t)}{\partial P_t} > 0$, Progressive's advertising will have a smaller impact on its visits at t if the consumer has seen more Progressive ads previously.

[5] Additionally, an ad may have a larger effect on visits if it is displayed when information processing is more convenient, such as when a consumer is browsing the internet with time on hand compared to when he is busy watching a TV show.

2.5.2 Dynamic effects over time

The ad's effect persists over time through its impact on the memory association strengths that persist. At $t+1$, the consumer may recall the purchase context due to exogenous factors. Since Progressive's advertising at t increases P_{t+1}, it increases Progressive's likelihood of being drawn and considered for purchase at $t+1$, holding other factors constant. However, this carryover effect also depends on G_{t+1}, which may get updated due to factors such as competitive advertising. Further, this carryover effect of advertising will depend on the salience of the auto insurance purchase context, which may vary due to unrelated reasons. Finally, since the advertiser's association with the purchase context decays over time, the effect of advertising will decrease over time, holding everything else the same. Appendix E elaborates on the pattern of dynamic advertising effects over time.

2.5.3 Effects on competitors

At time t, the effect on competitors occurs through two channels. Firstly, the ad makes the purchase context salient, which could benefit the competitor because it might get recalled via its association with the purchase context. Secondly, since advertising strengthens the advertiser's association with the purchase context, it reduces the likelihood of the competitor being drawn, conditional on the purchase context being salient. The overall effect of advertising on competitors at time t is the net effect through both these channels. We show in Appendix E.3 that Geico may experience a positive impact of Progressive's advertising, especially when the likelihood of the consumer shopping for auto insurance in the absence of ads, ϕ, is low.

While the contemporaneous effect on competitors can be positive, the future effect is not. After t, the consumer exogenously recalls the purchase context, and the competitor must compete for salience with an opponent whose association with the purchase context is relatively strengthened due to ad exposure at t. Thus, relative to contemporaneous effect at t, a negative effect of advertising on competitors is more likely at $t+1$. Appendix E.2 illustrates in detail the future effect of advertising on competitors.

2.6 Discussion

Overall, we show the implications of two well-researched features of human memory: first, memory comprises associations that weaken over time and get stronger as they are rehearsed; second, memories compete with each other for salience. Together, these features predict a "fixation" – by recalling an option, the consumer gets drawn into it and becomes less likely to think of other options.

The above description shows the key role of competitive actions. Even if a consumer is provided no new information, by repeating the same ad, a firm lowers the likelihood of the consumer thinking of the competitors, which may incentivize them to advertise. One can draw additional implications, such as the effects of multiple firms advertising simultaneously, an area we do not explore here as our focus is on explaining the specific setting below.

2.6.1 Comparison with the Ad stock based model

A common way to incorporate the effect of advertising into consumer choice is to assume that the consumer obtains utility from a stock of "goodwill" towards the advertiser that accumulates with advertising exposure. This increase in utility due to the ad stock makes it more likely that the consumer will consider and purchase the advertised option. The stock decays over time, often assumed as an exponential decay similar to a capital stock (Nerlove and Arrow, 1962). This workhorse model has also been used in non-advertising information retention contexts (e.g., Bronnenberg et al., 2012).

This model predicts a monotonic decay of the advertising stock, which leads to a decaying effect over time if the product utility remains the same. Contrary to the predictions of the memory-based model, the standard ad-stock model predicts that competitive advertising decreases advertising effects, as detailed in Appendix C. Further, since the advertiser's gain comes at the competitor's expense, this model also predicts the competitor's loss is the highest at the time when ad exposure occurs when the advertiser's gains are the highest. These predictions contradict ours.

Empirical applications of the ad-stock model have estimated flexible extensions of it, allowing consumer behavior to depend on the stocks of both the advertiser and its competitors without imposing restrictions on the sign of the effects (e.g., Slade, 1995). In section 8 we estimate these versions of the model and show that they still get the predictions off, even when they have more free parameters than our model.

3 Auto insurance advertising

In the US, auto insurance is mandatory for every driver to buy. Many consumers shop for insurance each year, and when they do, most consumers search for information online (comScore, 2015; Insurance Research Council, 2015). Our analysis of browsing data from comScore, detailed in Appendix B, shows that many buyers navigate to websites by directly typing the company's name into their browser. This behavior indicates some buyers remember the company from memory, which is also supported by publicly available aggregated search data from Google.

There are several auto insurance sellers in the market, with the top four sellers – StateFarm, Geico, Progressive and Allstate – accounting for more than 50% of the share of premiums (MarketWatch, 2024). These companies, having been established in the first half of the twentieth century, demonstrate a long-standing presence. Progressive, as the newest among the top four, has been in operation for over eighty years, indicating that even the most recent of these industry leaders has been around for generations.

At the same time, these companies are some of the highest spenders on advertising. According to Advertising Age, Geico spent $1.5 billion on advertising in 2022 and has consistently been one of the most advertised brands in the US since 2009 (Advertising Age, 2020, 2023). Progressive, which is the second most advertised auto insurance brand spent $1.4 billion, and the top ten auto insurance companies together spent $6.5 billion on advertising in 2022. Other sources report even higher spending

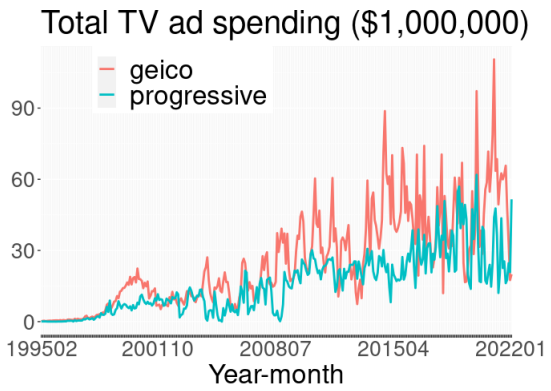


Figure 2: Ad spending for Geico and Progressive

Notes: This figure presents the total monthly TV advertising spending for Geico and Progressive over time, measured in millions of dollars. The data spans from February 1995 to January 2022 and highlights an increasing trend in advertising spending. The dataset was obtained from Kantar Media.

(e.g., S&P Global Market Intelligence, 2020).

This spending implies substantial consumer exposure to auto-insurance ads. To put these numbers into perspective in terms of the intensity of ad impressions, Geico's spending of $1.5 billion in a year is equivalent to more than a thousand impressions of social media banner ads to every adult in the US, assuming an average cost of half a cent per impression which is the average cost according to Statista (2024).

To analyze the evolution of ad spending in this industry, we obtained monthly TV ad spending data for Geico and Progressive from Kantar Media.[6] Figure 2 plots this data and shows that while TV advertising spend for both Geico and Progressive was low in the 1990s, both these companies increased their spending several fold to a level where, in some months we observe them spending close to $100 million in 2022.

While auto insurance is a complex product and the intensity of its advertising is high, it is widely noted that auto insurance ads tend to be repetitive and superficial mostly feature a funny character (e.g., NPR, 2022; Policygenius, 2024). At the time of our study Geico's ads for years had been featuring a gecko urging customers to shop at Geico which could lower their price. Progressive did the same through a character named Flo. These and other mascots in auto insurance ads have near *universal* recognition in the US (Advertising Age, 2022).

4 Setting and Experiment Design

We study the impact of advertising by a major auto insurance company on consumer online search behavior for auto insurance. Our analysis centers on a campaign that was run on the homepage of one of the top five most visited websites worldwide, which publishes general content across multiple topics

[6]This dataset is provided by Vivvix, formerly AdSpender (`https://www.vivvix.com/mediaradar-acquires-vivvix`).

such as finance, entertainment, and sports. Our partner platform manages this site and operates a widely adopted internet browser plugin extension designed to enhance user experience by facilitating easier access to platform services, including website navigation, bookmark management, blocking pop-ups, etc. Our investigation focuses on individuals who have installed this extension and who visited the platform's website on the day of our focal advertising campaign was launched, with the campaign itself being active for a single day in March 2015.

Our focal ad campaign displays the advertiser's standard, long-standing character along with the brand name. The message conveyed to the viewers that they could save money and invited them to "get a quote". Clicking on this banner would take the individual to the advertiser's website where she could get a quote after filling in a form.

Among the users who visited the website on the day of the campaign, 75% were assigned to a "treatment" group, and were displayed our focal advertiser's auto-insurance banner ad each time they visited the website on that day. The remaining 25% were assigned to a control group and shown a Public Service Announcement unrelated to auto-insurance.

5 Data

We analyze data on 326,952 individuals who have the browser extension installed, and visited the platform's website on the day of the experimental campaign. We compile several pieces of information in our dataset.

Internet browsing data Each individual is identified by a unique identifier which is specific to the browser. This identifier is likely to be more persistent than browser cookies which are frequently deleted. We observe the click-stream data of individuals from February 2015 to June 2015, capturing both the websites they visit and the times when these visits occurred. We do not observe the full web page addresses; instead, most web visits in our data are identified with the sub-domain level. For example, blog.example.com where blog is the sub-domain.

Other individual characteristics data Additionally, we observe the individual's location at the Designated Market Area (DMA) level. This allows us to merge census-based average characteristics of the individual's location. For a subset of 230,709 individuals we also observe their self-reported age which we use for analyzing heterogeneity in effects below. Self-reported age is likely to be noisy but Weber and Castillo (2010) provide evidence of statistical validity of such data. Supporting this notion, we find a significant correlation between the average reported age across DMAs and average DMA ages in the Nielsen data ($F(1, 207)$=68.63; p-value<.001).

Advertising exposure data We track whether and when each individual was exposed to the banner ads of the experimental campaign. Additionally, by using each individual's DMA location, we integrate our individual-level data with market-level TV advertising information from Nielsen AdIntel, accessed

	(1) Avg. in Control	(2) Avg. in Treatment	(3) p-value testing (1) and (2) are equal
Focal advertiser visits	0.4274	0.4184	0.6245
Competitor visits†	2.7020	2.6544	0.8766
Any Auto-insurance Brands Visited	0.1611	0.1627	0.3198
Number of Auto-insurance Brands Visited	0.1974	0.1981	0.7303
Focal advertiser's identified customer††	0.0178	0.0182	0.4270
Competitor's identified customer††	0.0547	0.0549	0.8440
Total TV ad GRPs pre-experiment Month	125.9588	125.9406	0.4730
Observations	81,881	245,071	

Table 1: Individual behavior and exposure to TV ads for auto insurance during pre-experiment time (Feb 1st 2015 - Mar 10th 2015)

Notes: The table reports individual behavior and exposure to TV ads for auto insurance during the pre-experiment period (Feb 1st 2015 - Mar 10th 2015). It includes averages in columns (1) and (2), and p-value from a t-test with H_0: averages in columns (1) and (2) are equal.
† We pulled data for ten major auto insurance companies, including our focal advertiser. Here, competitor means all competitor companies combined. For consistency with the analysis below, we exclude USAA, which does not serve the general public.
†† We are able to identify a subset of customers using the browsing data. See Appendix D for more details.

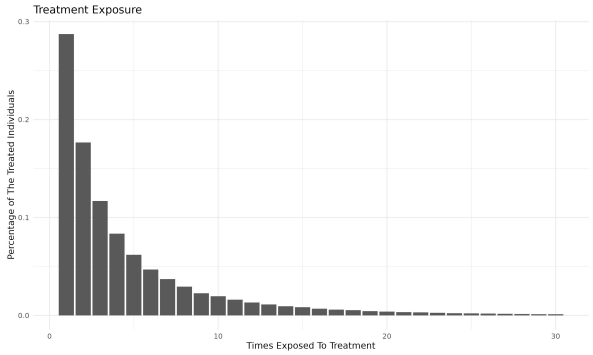


Figure 3: Number of impressions to the experimental banner

Notes: This figure shows the distribution of the number of impressions to the experimental banner for the treatment group. The median individual in the treatment group was exposed to the banner three times.

through the Kilts Center at the University of Chicago. This integration enables us to link each individual to the TV advertising impressions – both in volume and timing – delivered by auto insurance companies in their specific location.

Table 1 shows average statistics related to our sample's browsing behavior and TV ad exposure prior to the experiment. It also shows that these behaviors are similar across the experimental groups, which is expected due to randomization.

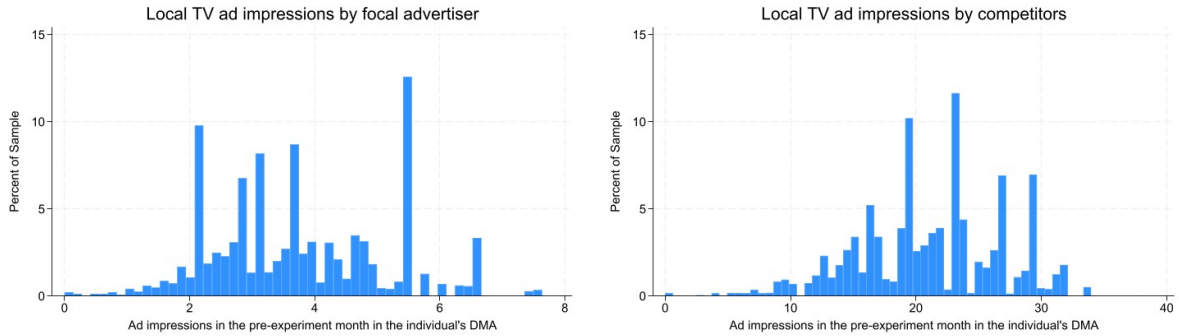
5.1 Ad impressions

The median treatment group individual got exposed to our experimental banner three times. However, there is variation across individuals as can be seen in Figure 3 which shows the distribution.

Turning to TV advertising, we pulled data for ten auto insurance companies including our focal

advertiser. Our data shows that, in the month prior to our experiment, these companies together delivered 125 GRPs, which is the average number of TV ad impressions per household. Our experimental advertiser delivered 14 GRPs, while the competitors together delivered 101 GRPs. Overall, 20% of impressions are delivered through local advertising, which varies across DMAs while the rest is through national advertising.

To examine the variation in the intensity of local advertising in our sample, we plot a histogram of local advertising in each individual's DMA in Figure 4. We observe significant variation; individuals range from being in areas with almost none, to significant advertising from both our focal advertiser and its competitors. TV ad impressions in a given DMA are highly correlated over weeks, making it hard to assess their impact using temporal variation. For example, within a DMA, the coefficient of correlation between TV GRPs in two successive weeks is 0.48.



(a) Focal advertiser's local TV ad impressions (b) Competitors' local TV ad impressions

Figure 4: Distributions of local TV ad impressions in the month before the experiment

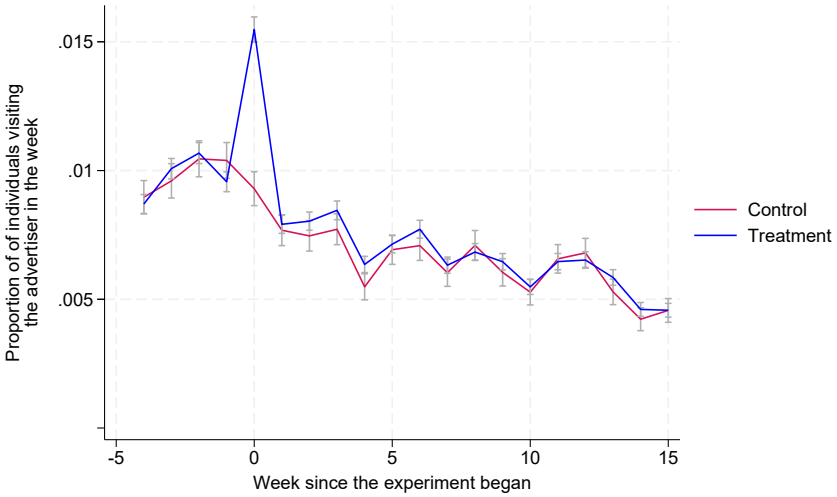


Figure 5: Visits to the advertiser's website.

Notes: Weekly visits to the advertiser's website from February to June 2015; for each week we plot the proportion of individuals who visit the advertiser in that week. Week 0 starts on the experimental day. The shaded areas represent 95% confidence intervals around the estimated number of visits.

6 Effect of the experimental campaign

Figure 5 plots the proportion of individuals in our sample visiting the advertiser's website at least once in different weekly time intervals, split by the experimental group assignment. It shows that individuals in the treatment group visit the advertiser at a higher rate in the week of the experimental campaign, and in the subsequent two months. Figure 6a does the same at a daily level focusing on the two months post experiment, and Figure 6b shows the corresponding differences along with their 95% confidence intervals.

These figures show that, while the fraction of individuals visiting the advertiser's website and the trend in visits are similar across the two experimental groups up to the day of the experiment, there is a 300% increase in the treatment group on the day of the experiment.[7] There is a smaller visible effect on the day after, and the effect is visibly smaller after that. At the same time, we see significant differences appearing in later days. For example, on March 27^{th} we observe an effect which is smaller relative to the first day, but statistically significant even after Bonferroni adjustment for multiple hypotheses testing.

To estimate the overall effect, Table 2 regresses the number of days individuals visited the advertiser's website at least once during various time intervals on the experimental treatment assignment. The analysis shows that the treatment increases website visits up to two months after the campaign ends. Beyond that, the treatment effect is undetectable. Comparing the size of the coefficients across the columns, we note that the incremental visits on the experiment day are almost equal to the incremental visits made in the next two months. This shows that almost 50% of the ad's effect occurs on the day when it is shown.[8]

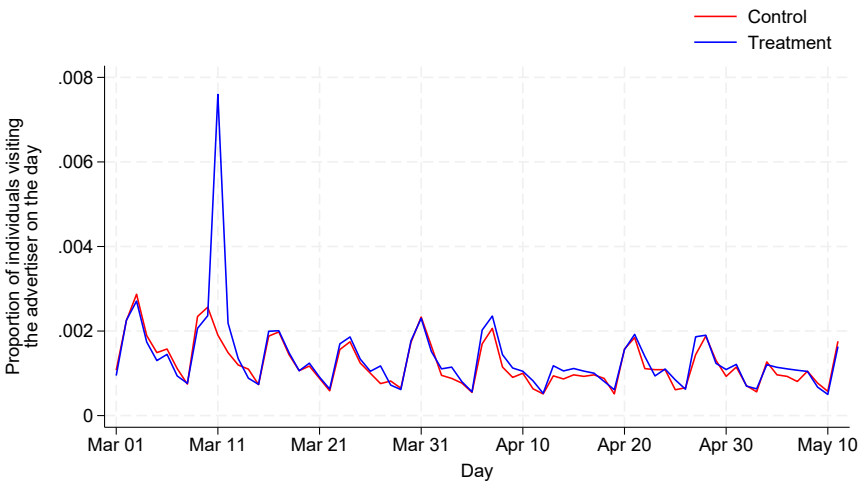
While our primary behavior of interest is website visits, which represent consideration, we also examine the significance of our experimental campaign in terms of whether it causes individuals to switch to the advertiser. We identify whether individuals in our sample are the advertiser's customers based on their browsing urls as detailed in Appendix D. Table 3 shows evidence that is consistent with individuals switching to the advertiser from its competitors.

6.1 Discussion

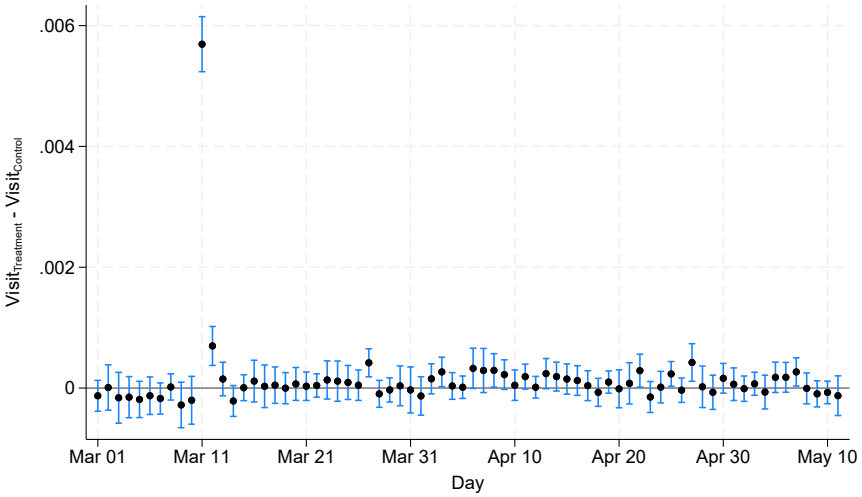
The effect of advertising we show is high compared to similar effects documented in the literature. We compare our numbers with others documented in the online advertising literature. Johnson et al. (2017) analyze the distribution of effects of 432 advertising campaigns on Google Display Network in terms of site-visit "lift" which is the treatment effect on advertiser website visits divided by the average visits in the control group and report a median lift of 0.17. Gordon et al. (2023) present a similar distribution for 601 upper funnel advertising campaigns on Facebook with median lift of 0.29

[7] All visits to the website within the same day are consolidated and treated as a single visit throughout the analysis.

[8] An analogous Table 16 in the appendix shows that visits to any insurance website significantly increase only on the day of the experiment, which is also consistent with our model.



(a) Daily visits to the advertiser's website



(b) Treatment Effect on daily Visits

Figure 6: Trend in visits and treatment effects

Notes: In 6 (a), we plot the daily proportion of individuals visiting the advertiser's website for each experimental group starting from March 1, 2015, to two months post-experiment, which is May 11, 2015. In 6 (b), we plot the difference in visit proportions between the treatment and control groups, along with the 95% confidence interval, assuming each day is independent.

Table 2: Effect of treatment on visits to advertiser's website at different times.

	(1) Before	(2) Day of Experiment	(3) Month after	(4) Month 2 after	(5) Month 3 after	(6) Month 4 after†
Treatment	-0.0006	**0.0057**	**0.0034**	**0.0021**	-0.0001	0.0010
	(0.0014)	**(0.0002)**	**(0.0011)**	**(0.0010)**	(0.0010)	(0.0006)
Intercept	**0.0596**	**0.0019**	**0.0376**	**0.0306**	**0.0315**	**0.0139**
	(0.0012)	**(0.0002)**	**(0.0009)**	**(0.0008)**	**(0.0009)**	**(0.0005)**
Observations	326,952	326,952	326,952	326,952	326,952	326,952

Notes: Each column represents a regression of the dependent variable – the count of unique days an individual visited the advertiser's website during the specified time interval – on an indicator variable for assignment to the treatment group. Multiple visits by an individual on the same day are counted as a single visit. Robust standard errors are shown in parentheses. Coefficients are bolded to indicate statistical significance at the 5% level.
† Month 4 after the experiment is partially observed in the data window.

Table 3: Effect of treatment on individual being a customer.

	(1) Advertiser's customer	(2) Competitor's customer
Treatment	**0.0017 (0.0006)**	-0.0019 (0.0012)
Intercept	**0.0260 (0.0006)**	**0.0899 (0.0010)**
Observations	326,952	326,952

Notes: Each column presents results from a regression on an indicator of assignment to the treatment group. In column (1), the dependent variable is a binary indicator for whether the individual is identified as the focal advertiser's customer during the post-experiment period. In column (2), the dependent variable indicates whether the individual is identified as a competitor's customer during the same period. Robust standard errors are shown in parentheses. Coefficients are bolded to indicate statistical significance at the 5% level.

(and 75^{th} percentile of 0.9). Lewis and Nguyen (2015) report similar effect sizes in advertising format similar to ours. All these numbers are smaller than our estimated lift during the campaign.[9]

Overall, we conclude that our estimated effects are on the higher side of the distribution, even though our focal advertiser is one of the highest spenders on advertising in the US.

Our Interpretation We view the significant effect of experimental advertising as a consequence of many consumers not recalling the focal firm without seeing its ad, despite wanting to consider it.

6.2 Direct effect of ads on future visits

Our model suggests that advertising exposure shifts consumer associations between brands and the auto-insurance purchase context, which favors the advertiser and makes it more likely that consumers will think of and visit the advertiser in the future. This is a *direct* effect on future visits, as opposed to indirect effects that could occur through short-term visits; for example, exposure to ads makes people visit the advertiser's website, where they might get exposed to new information, which causes them to come back in the future.

[9]Note is that Gordon et al. (2023)'s lift estimates vary across campaigns of different durations and lift may include days after the end of the campaign. If we estimate our lift including 14 days after the campaign, our lift is 58%, which is also higher than their median.

To establish the memory-based mechanism, we investigate the data to check for evidence of this direct effect of exposure to the experimental campaign on post-experimental visits. In general, direct effects are hard to isolate because typical mediation analysis requires conditioning on endogenous intermediate outcomes. However, we can test the existence of direct effects under the (monotonicity) assumption that exposure to advertising does not deter any individual from visiting the advertiser's website on the experiment day.

The idea is to focus on individuals who did not visit the advertiser during the experiment, and contrast the future visits among the treatment and control group individuals. Conditioning on no visit during the experiment creates the possibility that the two comparison groups inherently differ in their future visit propensities. However, we show below that, relative to the control group, the treatment group individuals who did not visit during the experiment have more future visits than possible without direct effects.

To see this formally, let the set of treatment and control group individuals that did *not* visit the advertiser's website on the experiment day be denoted by T_0 and C_0, respectively. Assuming advertising did not deter anyone from visiting on the experiment day, we can split C_0 into two subsets $\bar{C}_0$ and $\tilde{T}_0$, where $\tilde{T}_0$ represents individuals who would not have visited the advertiser on the experiment day if they were in the treatment group; $\bar{C}_0 = C_0 \setminus \tilde{T}_0$. Let Y_i^{Post} denote the number of visits by i in the two months post experiment day. Then

$$\begin{aligned} E(Y^{Post}|C_0) &= \frac{|\tilde{T}_0|}{|C_0|}E(Y^{Post}|\tilde{T}_0) + \frac{|\bar{C}_0|}{|C_0|}E(Y^{Post}|\bar{C}_0) \\ &\geq \frac{|\tilde{T}_0|}{|C_0|}E(Y^{Post}|\tilde{T}_0) \\ \implies \frac{|C_0|}{|\tilde{T}_0|}E(Y^{Post}|C_0) &\geq E(Y^{Post}|\tilde{T}_0). \end{aligned}$$

Therefore, the direct effect of advertising on post-experimental visits for treatment group individuals who did not visit the advertiser on the experimental day is

$$E(Y^{Post}|T_0) - E(Y^{Post}|\tilde{T}_0) \geq E(Y^{Post}|T_0) - \frac{|C_0|}{|\tilde{T}_0|}E(Y^{Post}|C_0). \tag{2}$$

The RHS of (2) is a lower bound on the direct effect, and can be estimated from the data. $E(Y^{Post}|C_0)$ and $E(Y^{Post}|T_0)$ can be estimated as sample averages in C_0 and T_0. Given the experimental randomization, $\frac{|C_0|}{|\tilde{T}_0|}$ can be estimated by $\frac{|C_0|}{|T_0|}$.

Table 4 shows that among the individuals who did not visit the advertiser on the experiment day, the treatment group individuals have 7.4% more visits in the two months post-experiment, relative to control, which is higher than what is possible in the absence of direct effects. The lower bound in (2) is positive and comparable to the total post-experiment two-month impact in Table 2, indicating significant direct effects of advertising.

Table 4: Examining the direct effect of treatment on post experiment behavior.

	Estimate	Standard error
Avg. post-experiment visits from those who did *not* visit on the experiment day		
(A) Control group: $E(Y^{Post}\|C_0)$	.0661	.0013
(B) Treatment group: $E(Y^{Post}\|T_0)$	.0710	.0008
Difference (B) - (A)	.0049	.0016
$E(Y^{Post}\|T_0) - \frac{\|C_0\|}{\|T_0\|}E(Y^{Post}\|C_0)$	.0043	.0017†

Notes: The table estimates the quantities required to assess the direct effect using the bound derived in inequality (2). It shows that the direct effect of the experimental advertising is significantly positive for the individuals T_0: the treatment group individuals who did not visit the advertiser's website on the experiment day.
† Bootstrapped standard errors.

6.3 Dynamic effect on competitors

As discussed in section 2.5.3, a unique prediction of our model is that competitors experience a larger negative effect of the ad, not at the time of exposure, but later when consumers think of the category on their own unaided by the ad.

To empirically examine this implication, we consider the treatment effect of the experiment on advertiser and competitor visits on the day and the day after the experiment.[10] We chose these two days for this comparison because these are the two days when the experimental treatment effect is the highest, and the day after the experiment gives little time for the competitors to overcome the effects of our ad through their own advertising.

Columns (1) and (2) in Table 5 show that going from the experiment day to the next, the experimental ad's effect on competitors increases in magnitude. On the day of the experiment, competitors do not experience any significant effects. This suggests that the advertiser's gain in consumer attention was not at the competitor's expense on the experimental day. However, competitors experience a 10% decrease in visits the day after the experiment, which is similar in magnitude to the positive effect experienced by the advertiser on that day (from Figure 6b). Column (3) shows that the estimated next-day effect on competitors is not sensitive to adding experiment-day outcomes.[11] These findings are consistent with our model but contradict models predicting that advertising solely transfers consumers from competitors, as the experiment day effect in Column (1) is much smaller than the advertiser effects. They also contradict models predicting no competitor impact, given the next-day effects on competitors are significant and similar to the advertiser's effect in magnitude.

[10] We exclude USAA from the analysis because it is unavailable to the general public.
[11] For example, if the negative next-day competitor effects are due to consumers clicking on the advertiser and getting enrolled in a different ad campaign being run by the advertiser, we expect to weaken the estimated negative effect by including experiment day outcomes in the regression, which does not happen in Table 5.

Table 5: Examining the dynamic effect on competitors.

	(1) Experiment day Num. competitor visits	(2) Next day Num. competitor visits	(3) Next day Num. competitor visits
Treatment	-0.0001	**-0.0009**	**-0.0009**
	(0.0004)	**(0.0004)**	**(0.0004)**
Experiment day: Num. competitor visits			**0.1057**
			(0.0081)
Experiment day: Advertiser visit			0.0045
			(0.0031)
Intercept	**0.0078**	**0.0085**	**0.0076**
	(0.0003)	**(0.0003)**	**(0.0003)**
Observations	326,952	326,952	326,952

Notes: Table presents regressions: (1) the number of competitors visited on the experiment day, regressed on assignment to the Treatment group; (2) the dependent variable changed to the number of competitors visited on the day *after* the experiment campaign; (3) a specification that includes controls for experiment day behavior for robustness check. Robust standard errors are shown in parentheses. Coefficients are bolded to indicate statistical significance at the 5% level.

7 Interaction with competitive actions

We examine how changes in competitive advertising impacts our treatment effect. Since our experiment varies impressions to one advertiser, we look for sources of natural variation in competitive advertising to examine its interaction with experimental treatment effects. To delve into this interaction we present below analysis of two sources of competitive advertising variation.

7.1 Interaction with pre-experimental competitive TV advertising

We analyze how the experimental treatment effect varies with the degree of TV advertising that took place in the month before the experiment. In this analysis, we aim to test the implications drawn in section 2.5.1, which is about the impact of past advertising on the contemporaneous effect of experimental advertising. Hence, we focus on experiment day visits as the outcome of interest.

As noted in section 5, we do not observe exposure to TV advertising at the individual level, which is typically unobserved to researchers. Instead, we use variation in brand-DMA level local TV advertising GRPs as a proxy for variation of intensity of exposure to pre-experimental local advertising on top of national advertising.

Table 6 shows estimates from the following regression:

$$\begin{aligned} y_i =& \beta_1 T_i + \beta_2 T_i \times \text{pre-experiment TV Ads}_m + \beta_3 T_i \times \text{competitors' pre-experiment TV Ads}_m \\ &+ \gamma_m + \epsilon_i, \end{aligned} \tag{3}$$

where T_i is and indicator of whether the individual i is in the treatment group; pre-experiment TV

Ads_m is standardized local TV ad GRPs delivered in the month before the experiment in i's DMA m; competitors' pre-experiment TV Ads_m is the same for the focal advertiser's competitors; γ_m are DMA specific intercepts; y_i is an indicator of whether the individual visited the advertiser's website on the experiment day.

Estimates in column (1) of Table 6 show that while the effect of our experimental campaign ads on visits to the advertiser is lower in regions where our focal advertiser recently advertised, it increases with competitors' advertising.

Column (2) adds controls for DMA size and DMA-specific spending by the advertiser and competitors in the previous year, to ensure the coefficients β_2 and β_3 are not capturing across-DMA heterogeneity in advertising effects which the advertiser might use to target TV ads. It also controls for individual-level pre-experimental behavior regarding auto insurance.

Importantly, adding these controls does not change the estimates for the coefficients β_2 and β_3. Note that the treatment effect on individuals who had been to the advertiser's website previously is not significantly lower than average. The effect is larger for individuals who had previously searched for competitors.

7.1.1 Discussion

These findings are consistent with our model's predictions. Increase in competitor advertising makes less likely that the advertiser is recalled without ads, which increases the effect of the experimental advertising. This effect is quantitatively significant. The variables in the regression in Table 6 are standardized. The standard deviation of competitor's pre-month local GRP is 1.4. Extrapolating based on this model, removing all competitor advertising including national ads which amounts to 122 GRPs, would negate the experimental treatment effect.

If individuals have seen the focal advertiser's TV ad in the previous month, the incremental effect of the experimental campaign is lower because the recent replenishment of the advertiser's association with the auto insurance category makes the control group's recall more likely.

The interaction effects with pre-experimental visits are also consistent with the notion that an individual who had previously been to the competitor did not recall our focal advertiser and *needed* a reminder in the form of an ad. For this reason, the pre-experiment competitor visitors respond more to the experimental ad.

7.2 Post experiment competitive activity

To understand how competitor activity post experiment impacted the longer-term effect of the experimental ad, we focus on a specific shock in exposure to auto insurance that occurred with the start of the Major League Baseball (MLB) season in the US, which began three weeks after our experimental campaign.

MLB is the second most watched sports league in the US, and is one of the most viewed local

Table 6: Examining the effect of pre-experiment competitive advertising.

	(1) Experiment day visit	(2) Experiment day visit
Treatment	**0.0057**	**0.0055**
	(0.0002)	**(0.0002)**
Treatment × Advertiser's last months local TV Ads*	-0.0004	**-0.0004**
	(0.0002)	**(0.0002)**
Treatment × Competitors' last months local TV Ads*	**0.0004**	**0.0006**
	(0.0002)	**(0.0002)**
Treatment × Visited advertiser previously		-0.0018
		(0.0067)
Treatment × Visited competitor previously		**0.0037**
		(0.0014)
Treatment × previous customer		0.0059
		(0.0057)
Treatment × previous competitors' customer		-0.0008
		(0.0013)
Visited advertiser previously		**0.0258**
		(0.0058)
Visited competitor previously		0.0006
		(0.0010)
Previous customer		**0.0138**
		(0.0046)
Previous competitors' customer		-0.0003
		(0.0007)
Intercept	**0.0019**	**0.0012**
	(0.0002)	**(0.0002)**
DMA fixed effects included?	Yes	Yes
Interaction of Treatment and DMA characteristics included?		Yes
Observations	326,952	326,952

* TV advertising is included as standardized GRPs delivered in the individual's DMA in the month prior to the experiment. A GRP is the total number of ad impressions per household.

Notes: Table presents regression results with the outcome variable being whether individuals visited the advertiser's website on the day of the experiment. Column (1) regresses the outcome variable on the treatment assignment and its interaction with local TV advertising by the focal advertiser and its competitors. Column (2) includes other interactions to control for other systematic heterogeneity in effects. DMA characteristics included are its size in terms of the number of households; spending in the previous year by the focal advertiser and its competitors. Robust standard errors are shown in parentheses. Coefficients are bolded to indicate statistical significance at the 5% level.

TV program (Nielsen, 2023; MLB.com, 2024). Its viewership is particularly high during the season's opening week, which is supported by a spike in Google trends for MLB. Auto insurance companies are known to be major MLB advertisers (Sports Business Journal, 2021, 2022) and did so in 2015 when Esurance and Hartford were among the official sponsors. Auto insurance companies are commonly featured, either in the form of banners on the field, behind the batter, or in the form of visible permanent fixtures such as dugouts / pavilions covered with auto insurance brand names.

Figure 7 shows example screenshots from the 2015 opening week, that were dominated by advertising from our focal company's competitors.[12]

[12]In our quick viewing of the 17 games that occurred during the first two days of the season 16 of the 17 games

Figure 8a shows the average visits to our focal advertiser's website by treatment and control group individuals in our sample in the week before, versus the week after the start of the MLB season, which happened on April 5th 2015. From figure we note that while the treatment effect is small and insignificant in the week before the season starts, it is significant in the week after. The difference in the treatment effects is significant at the 10% significance level.



(a) Atlanta Braves vs Miami Marlins April 6, 2015



(b) Chicago White Sox vs KC Royals April 6, 2015



(c) Wash Nationals vs NY Mets April 6, 2015



(d) San Diego Padres vs LA Dodgers April 6, 2015

Figure 7: Screenshots from MLB season opening day games

Note: Figure shows examples of auto insurance companies advertising during four opening weekend games of the 2015 MLB season. Out of the 17 opening games we found our focal advertiser's competitor auto insurance companies advertising in 16.

We next use the fact that there is significant variation in MLB's popularity across the US. Some cities (e.g., San Antonio, Austin, TX) do not have local MLB teams, while other similar-sized cities (e.g., Milwaukee, WI) do, and the MLB's popularity varies accordingly. We expect the MLB game viewership to be correlated with this. Hence, we obtain data from Google Trends on the popularity of the search term MLB across the 210 US DMAs and correlate the change in treatment effects over time with MLB popularity.

The bottom two rows in Figure 8 split the sample based on whether, or not the individual is in a DMA with above median popularity of MLB. They show that in areas where MLB is popular the control group visits to our focal advertiser decreased over time, while the treatment group held on to its level. Overall, the treatment effect increases in the week after MLB starts, and this effect is larger and more significant in DMAs with higher MLB popularity. These findings are consistent with our model's prediction that the effect of advertising may increase over time due to competitor's advertising.

prominently displayed auto-insurance brands that were our focal advertiser's competitors.



(a) Across the full sample



(b) Above median MLB markets



(c) Below median MLB markets

Figure 8: Visits to the advertiser before (left) and after (right) MLB season began.

Note: The above figures show the proportion of individuals who visit the focal advertiser's website in the week before (left column) and after (right column) the beginning of the 2015 MLB season, separately for the treatment and control groups. Figure (a) plots the numbers for the overall sample, and figures (b) and (c) show the same for markets above and below the median in MLB popularity, respectively. The plots show that while the treatment effect is insignificant the week before the MLB start, it becomes significant in the week after. The difference is mainly occurring in the markets where MLB is more popular.

7.3 Interaction with age

Memory researchers have long studied the variation in the functioning of memory with age (Light, 1991). One relevant and widely recognized implication occurring within the age range of our sample is that an individual's ability to form associations between bits of information declines with age, which leads to a decline in episodic memory (e.g., Old and Naveh-Benjamin, 2008).[13] In our context, age might affect the impact of advertising in strengthening the association between the advertiser and the auto insurance category. If this association is less likely to be remembered, the long term effects of advertising could be lower for older individuals. We examine this implication in our data.

We have age information for 232,750 individuals, and we split our sample at the median.[14]

Table 7 shows the immediate and longer term effects of our experimental advertising across younger and older individuals. We first note that in the control group, both younger and older individuals visit the website with similar likelihoods on the experiment day. However, the pattern of advertising effects is strikingly different. On the day of the experiment, the treatment effect is significantly higher for the older sample, while the longer term effect is higher for the younger sample. This variation in the

[13] On this point, Old and Naveh-Benjamin (2008) note that "a major contributor to older adults' deficiencies in episodic memory is their relative inability to form and retrieve links among single bits of information."

[14] The average ages of sub-samples below and above the median are 43 and 67 respectively. The psychology literature on aging effects, for example, Old and Naveh-Benjamin (2008), consider a sample of younger and older individuals to have average ages of 35 and 60, respectively. Hence, the variation in age in our sample is comparable to theirs.

treatment effect over time is consistent with the notion that association between the advertiser and the auto insurance category is stronger for the younger sample, which leads to higher longer term effects.

Overall, our data patterns are consistent with the predictions of a memory-based mechanism.

Table 7: Examining the effect of pre-experiment competitive advertising.

	(1)	(2)	(3)	(4)
	Visit on experiment day		Visits in two months after the experiment	
	Age below median	Age above median	Age below median	Age above median
Treatment	**0.0044**	**0.0081**	**0.0088**	0.0030
	(0.0004)	**(0.0004)**	**(0.0031)**	(0.0028)
Constant	**0.0024**	**0.0020**	**0.0810**	**0.0703**
	(0.0003)	**(0.0003)**	**(0.0027)**	**(0.0024)**
Observations	118,932	113,818	118,932	113,818

Notes: We split the sample into individuals who are above or below the median age. The table presents regressions of visits to the focal advertiser on the treatment assignment: columns (1) and (2) use as dependent variable visits on the experiment day while columns (3) and (4) do so for the two months post experiment day. The table shows that while younger and older cohorts have similar visits in the control group, their visits differ in the treatment group. The older group has a larger effect on the day of the experiment but their longer-term effect is undetectable. Robust standard errors are shown in parentheses. Coefficients are bolded to indicate statistical significance at the 5% level.

8 Model estimation and its implications

Researchers often analyze the impact of information exposure through the lens of a consumer demand model. In this section, we do so with our model. We specify an empirical version of the memory-based model presented in section 2 and calibrate it using our experiment dataset. We then compare its in- and out-of-sample predictions with our empirical findings, as well as the predictions of flexible specifications of the standard ad stock model to assess the predictive gains from incorporating the psychology theory of memory. Overall, this comparison below shows that the memory-based model uniquely predicts the directions of important effects and fits the data better despite being relatively parsimonious.

8.1 Model Setup

We include data from the day of the experiment and the day after for the purpose of this calibration. Also, we include two firms, our experimental advertiser, denoted as P, and a similar-sized competitor, denoted as G. Our goal is to predict the probability of each consumer in our sample visiting P's and G's websites on the experiment day ($t = 0$) and on the next day ($t = 1$). Each individual i has a treatment status D_i, which is 1 if i is assigned to the experiment's treatment group, and 0 otherwise. Following the notation from section 2, P_{it} and G_{it} represent the strengths of association between the auto insurance purchase context and P and G, respectively, in i's memory at the start of day t. For simplicity, we assume $P_{i0} = G_{i0} = 1$ for all i, which is consistent with P and G being similar in their

traffic in the pre-experimental data. The $\Pr(\text{Visit } J \mid \text{Drawing } J) = \alpha$ for $J \in \{P, G\}$.

We assume our experimental banner is the only ad available in these two periods. Every treatment group individual is exposed to the banner, and none of the control group individuals are.

8.1.1 Likelihood of the observed data

Given our model's memory draw process, memory recalls are correlated across brands, so we consider their joint probability. The detailed working of the model estimation is presented in Appendix F. We present a summary here.

The distribution of recalls and website visits for P and G on day 0 depend on the treatment assignment, and the parameters represented by $\boldsymbol{\theta}$, which includes the following: the increase in memory strength from each draw (r), the survival rate of memory strength (δ, with $1-\delta$ as the daily depreciation rate), the probability of forgetting the auto insurance category when a consumer did not see any ads (ϕ), the probability of visiting a brand after recalling it is assumed to be a constant, denoted by α.

Recalls on day 1 are a function of the parameters, along with the outcomes of day 0. The choices of website visits are, in turn, a function of recalls. Hence, the likelihood of observing visits to websites is given by:

$$\begin{aligned}
\boldsymbol{L}(\boldsymbol{\theta}) &= \prod_{i=1}^{N} \Pr\left(\text{Predicted visits} = \text{Observed visits} \mid D_i, P_{i0}, G_{i0}; \boldsymbol{\theta}\right) \\
&= \prod_{i=1}^{N} \left[\Pr\left(\text{Predicted visits}_0 = \text{Observed visits}_0 \mid D_i, P_{i0}, G_{i0}; \boldsymbol{\theta}\right)\right. \\
&\quad \left. \times \Pr\left(\text{Predicted visits}_1 = \text{Observed visits}_1 \mid D_i, P_{i0}, G_{i0}, \text{Day 0 visits}; \boldsymbol{\theta}\right)\right]
\end{aligned}$$

We estimate the parameter values, $\hat{\boldsymbol{\theta}}_{ML}$, that maximize this likelihood. The precision of our estimates is given by:

$$\text{Var}(\boldsymbol{\theta}) = [-H(\hat{\boldsymbol{\theta}}_{ML})]^{-1}$$

where $H(\hat{\theta}_{ML})$ is the Hessian of the likelihood evaluated at $\hat{\theta}_{ML}$.

8.2 Estimation Results and Implications

Table 8 shows our estimates and confidence intervals. As expected, the model estimates suggest that recalling a brand significantly increases its association with the auto insurance purchase context, relative to the baseline strength of one. About 45% of the strength of association survives to the next day. The estimates also indicate that fewer than 1% of the individuals who recall the brand actually visit its website, which is consistent with the data presented earlier.

Based on these estimates, we predict six important moments to assess the alignment of our model predictions with the empirical findings. This includes four in-sample moments: τ_1 to τ_4, representing the experimental assignment's treatment effects on advertiser and competitor visits on the experimental

Table 8: Estimation Result of Memory Interference model, Advertiser and A Top Competitor

Parameter	Description	Estimate [95% Confidence Interval]
r	The increase in memory strength from each draw	1.9113 [1.5627, 2.3375]
δ	The survival rate of memory strength	0.4580 [0.2500, 0.6817]
ϕ	The probability of forgetting the auto insurance category	0.6444 [0.6206, 0.6675]
α	The probability of visiting a brand conditional on recalling it	0.0076 [0.0073, 0.0080]

Note: This table presents the point estimates and confidence intervals for the parameters of the memory interference model. Our estimation used relevant data of two firms: the advertiser in our experiment and a top competitor. To ensure that the parameters satisfy the boundary conditions ($r > 0$, $\delta \in [0,1]$, $\phi \in [0,1]$, $\alpha \in [0,1]$), we estimate a transformed version of these parameters (e.g. we estimate the free parameter $\tilde{r}$ where $r \equiv \exp(\tilde{r})$). We transformed the estimated $\tilde{r}$ and its confidence interval back to the original parameters.

and the next day. They also include two out-of-sample moments that capture the interaction between the treatment effect and past advertising: τ_5 represents the impact of an increase in past exposure to the focal advertiser's ads on the treatment effect observed on the same day; τ_6 measures the impact of past exposure to competitor's ads on the treatment effect for the advertiser on the experiment day.

We denote model M's prediction of the effect of the experimental assignment on visits to j's website on day t as:

$$\Delta_{jt}^{M}(P,G) = \overline{\hat{\text{Visit}}_{jt}^{M}}(D=1,P,G) - \overline{\hat{\text{Visit}}_{jt}^{M}}(D=0,P,G), \tag{4}$$

where $\overline{\hat{\text{Visit}}_{jt}^{M}}$ is the average of the predicted likelihood of visit to j at time t across our sample, given the arguments. Based on this, the four in-sample moments mentioned above can be calculated as follows:

$$\begin{aligned}
\hat{\tau}_1^{\text{M}} &= \Delta_{P,0}^{\text{M}}(1,1) \\
\hat{\tau}_2^{\text{M}} &= \Delta_{G,0}^{\text{M}}(1,1) \\
\hat{\tau}_3^{\text{M}} &= \Delta_{P,1}^{\text{M}}(1,1) \\
\hat{\tau}_4^{\text{M}} &= \Delta_{G,1}^{\text{M}}(1,1)
\end{aligned} \tag{5}$$

To predict the interaction with past advertising, τ_5, we assume that the consumer viewed a focal advertiser's ad on a randomly chosen day in the week prior to the experiment.[15] We predict τ_6 in a similar manner, except that it involves a competitor's ad instead of the focal advertiser's.

The results of the predictions are shown in column (2) in Table 9 and compared with the empirical findings in Section 6. For all statistically significant moments, the memory interference model successfully predicts the signs, and the magnitudes of the effects are similar.

[15] If the pre-experiment ad exposure occurred at day s, we assume that the memory strengths P and G equal to 1 prior to s. The memory strengths update on s as per the model parameters and carry over to the experiment day. Hence, $\hat{\tau}_5^{M} = \frac{\sum_s \Delta_{P,0}^{\text{M}}(P_{\text{Ad on } s}, G_{\text{Ad on } s}) - \Delta_{P,0}^{\text{M}}(1,1)}{7}$, where $P_{\text{Ad on } s}, G_{\text{Ad on } s}$ represent the memory strengths on day 0 given an ad impression at s.

Table 9: Comparing Empirical Findings and Predictions of The Memory Interference Model

Moment	Description	(1) Empirical Finding	(2) Prediction of Memory Interference Model
In-sample Moments			
τ_1	day 0 treatment effect on advertiser visits	$\mathbf{0.0057}$	0.0059
τ_2	day 0 treatment effect on competitor visits	-1.9×10^{-4}	2.46×10^{-4}
τ_3	day 1 treatment effect on advertiser visits	$\mathbf{6.97 \times 10^{-4}}$	6.95×10^{-4}
τ_4	day 1 treatment effect on competitor visits	$\mathbf{-8.62 \times 10^{-4}}$	-6.81×10^{-4}
Out-of-sample Moments			
τ_5	Past advertiser ads effect on τ_1	$\mathbf{-4.4 \times 10^{-4}}$	-1.45×10^{-4}
τ_6	Past competitor ads effect on τ_1	$\mathbf{3.8 \times 10^{-4}}$	1.38×10^{-4}

Note: This table compares the empirical findings and our model's predictions of six important moments: τ_1 through τ_6. The empirical estimates for τ_1 through τ_4 are obtained by regressing whether an individual visited the firm j or not on day t on the treatment status, controlling for DMA fixed effects.
The empirical estimates for τ_5 and τ_6 are obtained by regressing whether an individual visited the advertiser P or not on the day t on the treatment status, the treatment status interacted with P's TV advertising in the month before the experiment, and the treatment status interacted with G's TV advertising in the month before the experiment. The coefficient on the first interaction gives the empirical estimate of τ_5, and the coefficient on the second interaction gives the empirical estimate of τ_6. We show the estimates from the interaction of treatment with TV ads in the past month to match the previously shown estimates; our analysis shows that the effect is mainly driven by TV ads received in the previous week.
This table shows that the memory interference model correctly predicts the signs and magnitudes of all five statistically significant moments in our data. Coefficients are bolded to indicate statistical significance at the 5% level.

8.3 Predictive Gains Relative to Alternative Models

We consider alternative models to assess the predictive gains from using our model. As a benchmark, we focus on flexible specifications of the standard ad stock model. We estimate each of these models using the same data, and assess their fit in terms of log-likelihood and their predictions for the six moments mentioned above.

8.3.1 Ad stock models

The main idea here is that advertising exposure builds an advertising "stock" for each product option. Individual i's utility from visiting j's ($j \in \{P, G\}$) website is given by:

$$u_{ijt} = a + b_1 S_{ijt} + b_2 S_{i(-j)t} + \epsilon_{ijt} \tag{6}$$

where a is the baseline utility, b_1 is the marginal contribution of a firm's ad stock on consumer utility, b_2 is the marginal contribution of competitor's ad stock on consumer utility, ϵ_{ijt} is individual-brand-day utility shock which is i.i.d., following the Extreme Value Type-I distribution. Individual i visits j's website at t if and only if $u_{ijt} \geq 0$. This allows for the possibility of a consumer visiting multiple brands' websites.

Stock S_{ijt} is updated as follows:

$$S_{ij(t+1)} = d \cdot S_{ijt} + D_i \cdot 1\{t = 0, j = P\}, \tag{7}$$

meaning that the ad stock is built by the experimental ad exposure captured by D_i. The initial memory strengths are set to $S_{ij0} = S_0$.

Using the same data as above in section 8.2, we apply a maximum likelihood method detailed in Appendix F to estimate the five parameters: the baseline utility a, the marginal utilities from ad stock b_1, b_2, the ad stock survival rate d, and initial stock S_0.

We also consider a constrained version of this model by restricting $S_0 = 1$ to keep the number of model parameters the same as the memory-based model.

8.3.2 Results of Model Estimates

Table 10: A Comparison with Predictions of Alternative Models

	(1) Empirical Finding	(2) Ad Stock Model (4-parameters)	(3) Ad Stock Model (5-parameters)	(4) Memory Model (4-parameters)
Log likelihood		-24837	-24837	-24811
In-sample Moments				
τ_1	$\mathbf{0.0057}$	0.0056	0.0057	0.0059
τ_2	-1.9×10^{-4}	-7.89×10^{-5}	4.70×10^{-6}	2.46×10^{-4}
τ_3	$\mathbf{6.97 \times 10^{-4}}$	7.05×10^{-4}	7.23×10^{-4}	6.95×10^{-4}
τ_4	$\mathbf{-8.62 \times 10^{-4}}$	-1.59×10^{-5}	9.60×10^{-7}	-6.81×10^{-4}
Out-of-sample Moments				
τ_5	$\mathbf{-4.4 \times 10^{-4}}$	8.52×10^{-4}	7.59×10^{-4}	-1.45×10^{-4}
τ_6	$\mathbf{3.8 \times 10^{-4}}$	-2.09×10^{-5}	1.10×10^{-6}	1.38×10^{-4}

Note: This table compares different models' predictions of six important moments: τ_1 through τ_6 and examines how well they match the empirical findings. We compare our memory model with 4 parameters in column (4) with the Ad Stock model with 4 parameters in column (2) and the Ad Stock model with 5 parameters in column (3). The results show that among these alternative models, our model is the only one that correctly predicts the signs of the five statistically significant in- and out-of-sample moments in our data. Coefficients are bolded to indicate statistical significance at the 5% level.

The results of the predictions from the considered models are shown in Table 10. In terms of the log-likelihood, the advertising stock models are worse at fitting our data relative to the memory model, even when the memory model has fewer parameters. This results from the inability of the ad stock model to justify the ad's "next-day effect" on competitors being lower than the "same-day effect", which is the moment τ_2 in the table. This is consistent with our discussion in section 6.3.

As for the out-of-sample moments, none of the alternative models can predict their signs correctly. Unlike the memory based model, they predict that the past advertising by the advertiser would increase the impact of the experimental campaign, which contradicts our empirical findings.

Overall, this analysis shows the predictive gains from using the memory-based model relative to the standard estimation frameworks in empirical research.

9 Conclusion

A consumer's decisions depend on the information recall during decision making. This aspect is non-trivial but often overlooked. This paper suggests a method to incorporate this factor into frameworks

for assessing consumer decisions. We build a model based on the fundamental principles of cognitive psychology to explain how advertising influences consumer behavior. We evaluate this model's implications in the context of auto insurance, an industry with some of the highest advertising spending in the US, despite being regulated and comprising old, well-known brands. Using unique data from multiple sources merged with a randomized field experiment we provide evidence for competitive interference and other unique implications of our model. Overall, we demonstrate significant predictive gains from adopting the psychology-based model. Our model offers a framework for understanding the impact of repeated exposure to the same information on consumer decision-making. Future research can use this framework to understand prevalent advertising patterns and evaluate their welfare effects.

References

Advertising Age (2020). The 10 most-advertised brands today vs. a decade ago. https://adage.com/article/year-end-lists-2019/10-most-advertised-brands-today-vs-decade-ago/2221686. Accessed April 3, 2024.

Advertising Age (2022). Most popular insurance ad characters, ranked. https://adage.com/article/marketing-news-strategy/most-popular-insurance-ad-characters-ranked/2392531. Accessed April 3, 2024.

Advertising Age (2023, June). The 15 most-advertised brands in the us, ranked—leading national advertisers 2023. https://adage.com/article/datacenter/15-most-advertised-brands-us-include-amazon-walmart-and-geico/2497391. Accessed April 3, 2024.

Alba, J. W. and A. Chattopadhyay (1985). Effects of context and part-category cues on recall of competing brands. *Journal of Marketing Research 22*(3), 340–349.

Alba, J. W. and A. Chattopadhyay (1986). Salience effects in brand recall. *Journal of Marketing Research 23*(4), 363–369.

Anderson, J. R. (1993). Problem solving and learning. *American psychologist 48*(1), 35.

Anderson, J. R. (1996). Act: A simple theory of complex cognition. *American psychologist 51*(4), 355.

Anderson, J. R. (2000). *Learning and memory: An integrated approach*. John Wiley & Sons Inc.

Anderson, S. P. and A. De Palma (2013). Shouting to be heard in advertising. *Management Science 59*(7), 1545–1556.

Bagwell, K. (2007). The economic analysis of advertising. *Handbook of industrial organization 3*, 1701–1844.

Bordalo, P., J. J. Conlon, N. Gennaioli, S. Y. Kwon, and A. Shleifer (2023). Memory and probability. *The Quarterly Journal of Economics 138*(1), 265–311.

Bordalo, P., N. Gennaioli, and A. Shleifer (2020). Memory, attention, and choice. *The Quarterly journal of economics 135*(3), 1399–1442.

Bronnenberg, B. J., J.-P. H. Dubé, and M. Gentzkow (2012). The evolution of brand preferences: Evidence from consumer migration. *American Economic Review 102*(6), 2472–2508.

Burke, R. R. and T. K. Srull (1988). Competitive interference and consumer memory for advertising. *Journal of consumer research 15*(1), 55–68.

Chen, Y., Y. V. Joshi, J. S. Raju, and Z. J. Zhang (2009). A theory of combative advertising. *Marketing Science 28*(1), 1–19.

comScore (2015). Online auto insurance shopping report. https://www.comscore.com/Insights/Presentations-and-Whitepapers/2015/2015-Online-Auto-Insurance-Shopping-Report. Last retrieved April 3, 2024.

Danaher, P. J., A. Bonfrer, and S. Dhar (2008). The effect of competitive advertising interference on sales for packaged goods. *Journal of Marketing Research 45*(2), 211–225.

Doraszelski, U. and S. Markovich (2007). Advertising dynamics and competitive advantage. *The RAND Journal of Economics 38*(3), 557–592.

Dubé, J.-P., G. J. Hitsch, and P. Manchanda (2005). An empirical model of advertising dynamics. *Quantitative marketing and economics 3*, 107–144.

Gordon, B. R., R. Moakler, and F. Zettelmeyer (2023). Close enough? a large-scale exploration of non-experimental approaches to advertising measurement. *Marketing Science 42*(4), 768–793.

Haan, M. A. and J. L. Moraga-González (2011). Advertising for attention in a consumer search model. *The Economic Journal 121*(552), 552–579.

Insurance Research Council (2015, June). Aided by internet, consumers shop for auto insurance to get better price. https://insurance-research.org/sites/default/files/downloads/Shopping%20For%20Insurance%20NR%202015.pdf. Last retrieved April 3, 2024.

Johnson, G., R. A. Lewis, and E. Nubbemeyer (2017). The online display ad effectiveness funnel & carryover: Lessons from 432 field experiments. *Available at SSRN 2701578*.

Kahana, M. J. (2012). *Foundations of human memory*. OUP USA.

Keller, K. L. (1987). Memory factors in advertising: The effect of advertising retrieval cues on brand evaluations. *Journal of Consumer Research 14*(3), 316–333.

Lechner, H. A., L. R. Squire, and J. H. Byrne (1999). 100 years of consolidation—remembering müller and pilzecker. *Learning & memory 6*(2), 77–87.

Lewis, R. and D. Nguyen (2015). Display advertising's competitive spillovers to consumer search. *Quantitative Marketing and Economics 13*, 93–115.

Lewis, R. A. and J. M. Rao (2015). The unfavorable economics of measuring the returns to advertising. *The Quarterly Journal of Economics 130*(4), 1941–1973.

Light, L. L. (1991). Memory and aging: Four hypotheses in search of data. *Annual review of psychology 42*(1), 333–376.

MarketWatch (2024). 10 largest car insurance companies in the u.s. https://www.marketwatch.com/guides/insurance-services/largest-car-insurance-companies/. Accessed March 20, 2024.

MLB.com (2024). Baseball continues to be most-viewed local tv programming. Accessed April 12, 2024.

Mullainathan, S. (2002). A memory-based model of bounded rationality. *The Quarterly Journal of Economics 117*(3), 735–774.

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

Naik, P. A., M. K. Mantrala, and A. G. Sawyer (1998). Planning media schedules in the presence of dynamic advertising quality. *Marketing science 17*(3), 214–235.

Nedungadi, P. (1990). Recall and consumer consideration sets: Influencing choice without altering brand evaluations. *Journal of consumer research 17*(3), 263–276.

Nelson, P. (1974). Advertising as information. *Journal of political economy 82*(4), 729–754.

Nerlove, M. and K. J. Arrow (1962). Optimal advertising policy under dynamic conditions. *Economica*, 129–142.

Nielsen (2023, December). Most watched sports leagues in the united states in 2023, by minutes watched (in billions). Chart. Accessed April 12, 2024.

NPR (2022). The gecko effect - why are insurance ads everywhere? `https://www.npr.org/2022/06/01/1102496094/the-gecko-effect`. Accessed April 3, 2024.

Old, S. R. and M. Naveh-Benjamin (2008). Differential effects of age on item and associative measures of memory: a meta-analysis. *Psychology and aging 23*(1), 104.

Policygenius (2024). Best auto insurance commercials – policygenius. `https://www.policygenius.com/auto-insurance/auto-insurance-commercials-by-company/`. Accessed March 22, 2024.

Raaijmakers, J. G. and R. M. Shiffrin (1980). Sam: A theory of probabilistic search of associative memory. In *Psychology of learning and motivation*, Volume 14, pp. 207–262. Elsevier.

Raaijmakers, J. G. and R. M. Shiffrin (1981). Search of associative memory. *Psychological review 88*(2), 93.

Rundus, D. (1973). Negative effects of using list items as recall cues. *Journal of Verbal Learning and Verbal Behavior 12*(1), 43–50.

Sahni, N. S. (2015). Effect of temporal spacing between advertising exposures: Evidence from online field experiments. *Quantitative Marketing and Economics 13*, 203–247.

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

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

Sial, A. Y., J. R. Sydnor, and D. Taubinsky (2023). Biased memory and perceptions of self-control. Technical report, National Bureau of Economic Research.

Slade, M. E. (1995). Product rivalry with multiple strategic weapons: An analysis of price and advertising competition. *Journal of Economics & Management Strategy 4*(3), 445–476.

S&P Global Market Intelligence (2020). Insurers that spent the most on advertising in 2020. `https://www.spglobal.com/marketintelligence/en/news-insights/latest-news-headlines/insurers-that-spent-the-most-on-advertising-in-2020-64864585`. Accessed March 22, 2024.

Sports Business Journal (2021, October). Sbj atlas: State farm the most active brand during mlb season.

Sports Business Journal (2022, January). Insurance, telecom companies set pace across 2021 ad spend.

Statista (2024). Online advertising metrics - statistics & facts. `https://www.statista.com/topics/5063/online-advertising-metrics/#topicOverview`. Last accessed March 20, 2024.

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy 69*(3), 213–225.

Tsai, Y.-L. and E. Honka (2021). Informational and noninformational advertising content. *Marketing Science 40*(6), 1030–1058.

Weber, I. and C. Castillo (2010). The demographics of web search. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, pp. 523–530.

A Relation with SAM Model

We discuss the similarities between our model and the Search of Associated Memory model (Raaijmakers and Shiffrin, 1980), from which it is inspired. The SAM model has been used to explain memory recall in controlled lab environments, we extend it to consumer decision-making in the field. In brief, the SAM model conceptualizes the memory as a system comprising a short- and long-term storage, analogous to the architecture of computer memory. The salient memory images are in the short term storage (STS) whereas the long-term storage (LTS) contains all the known memory images in the form of a linked associative network.

Memories may get drawn into the STS through external contextual events, for example, someone talking about auto insurance. If the external environment is the same, the likelihood of drawing and, in our context, a person thinking about auto insurance remains the same. Our model reflects this aspect through a parameter ϕ that reflects the individual's likelihood of drawing auto insurance related memories through external contextual changes. Recall to STS in SAM can also occur through contagious recall based on what is already present in the STS: salience of the purchase occasion in the STS may lead to draw of associated product options. This aspect is consistent with our memory drawing process. The strengths of association between memories that are recalled to the STS at the same time gets stronger, which is a manifestation of the classical rehearsal effect in psychology. Our model incorporates this through an increase in association strength between a brand and the purchase context when the brand gets recalled.

B Path to the auto insurance websites

To assess the importance of brand recall in the context of auto insurance purchase, we conducted multiple analyses.

Google Trends We use publicly available data on Google Trends (`trends.google.com`) to examine the number of brand searches on Google, in which the user types in the name of the company relative to other search keywords, which may signify consumers selecting from a list of options. Figure 9 shows that more people searched for "geico" relative to generic keywords such as "auto insurance" and "car insurance", supporting the notion that consumers do use their own recall for navigating to auto insurance company websites.

comScore data analysis of buyers We also analyze the percentage of buyers who visit auto-insurance websites by recalling the brand. Specifically, we analyze consumer shopping sessions - defined as the days when consumers made a purchase from our focal advertiser - and explore the role that brand recall plays on their arrival to the website. We distinguish consumers by whether they arrive at the website directly, searching for the particular brand on Google, or generally searching for auto-insurance and clicking through a result. The results are shown in Table 11.

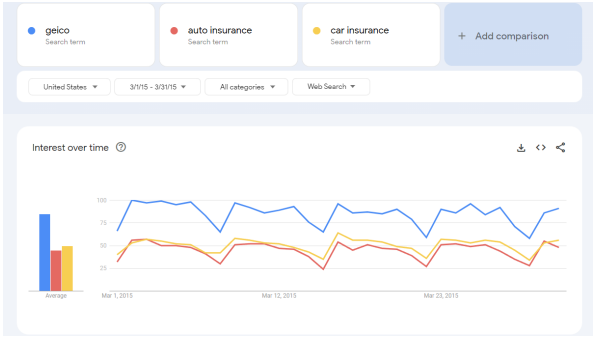


Figure 9: Google trends data: comparisons of searches for the term "Geico" vs. "auto insurance".

For this purpose we use comScore data that reports all url visits and searches conducted by a panel of millions of users. Based on the click-stream data, we obtain the days that consumers made a purchase. We then determine the first visit the buyer made to the website, and see if this visit is initiated by a search; and if so, we document the associated search keywords.

The tricky part is that even if consumers didn't arrive through searches, they could be arriving from other websites (e.g., aggregators). We assume consumers to be typing the brand's website directly if the first page is the website's home page (e.g., www.geico.com).

Based on this exercise, around 29% of consumers who purchased an auto-insurance product from our focal advertiser arrived either directly or by searching for the advertiser in a major search engine. This indicates that brand recall plays a significant role in attracting auto-insurance customers.

Table 11: 2015 Traffic Sources for Our Focal Auto-Insurance Advertiser

	# of observations	% of observations
Focal Advertiser's buyers	1,119	
Arrived directly	221	19.75%
Arrived from searching the advertiser's brand name	103	9.20%
Arrived from other search keywords	38	3.40%
Arrived from other sources	757	67.65%

One concern with the validity of these estimates is the potential inaccuracy in the number of consumers who arrived directly. To validate our procedure, we compared the estimated results with the actual numbers using the 2019 comScore data, which is the first year when information from referring webpages was recorded by comScore. Therefore, for 2019 data we know the actual number of individuals who arrived directly to the auto-insurance website.

Based on our comparison using 2019 data, we conclude that our estimate of the percentage of direct arrivals is conservative. While our method estimates 10.98% of 2019 buyers being direct arrivals, the actual number is 49.52%. This indicates that our estimate above is likely to be a lower bound on the importance of brand recall in driving the traffic of online customers.

C Predictions of the Ad-stock model

In most applications of the ad-stock model advertising is added to the utility $u_{Pt} = f(S_{Pt})$, where S_{Jt} is firm J's ad stock at time t. The stock builds as $S_{Jt} = j_t + dS_{Jt-1}$ where j_t is the number of ad impressions at time t and $0 < d < 1$ is the rate at which the stock decays. In the standard implementation f takes a logit form and the likelihood of the consumer choosing P is

$$\begin{aligned}\Pr(P) &= f(p_t + \delta S_{Pt-1}, g_t + \delta S_{Gt-1}) \\ &= \frac{\exp(\alpha_1 + \beta_1 S_{Pt})}{1 + \exp(\alpha_1 + \beta_1 S_{Pt}) + \exp(\alpha_2 + \beta_2 S_{Gt})},\end{aligned}$$

where $\beta_1, \beta_2 > 0$.

Effect of P's advertising at time t is

$$\begin{aligned}\frac{\partial f}{\partial p_t} &= \frac{\beta_1 \exp(\alpha_1 + \beta_1 S_{Pt})}{1 + \exp(\alpha_1 + \beta_1 S_{Pt}) + exp(\alpha_2 + \beta_2 S_{Gt})} - \frac{\beta_1 \left(\exp(\alpha_1 + \beta_1 S_{Pt})\right)^2}{(1 + \exp(\alpha_1 + \beta_1 S_{Pt}) + \exp(\alpha_2 + \beta_2 S_{Gt}))^2} \\ &= \beta_1 \Pr(P)(1 - \Pr(P))\end{aligned}$$

How does this effect of p_t change with G's past advertising?

$$\begin{aligned}\frac{\partial^2 f}{\partial p_t \partial S_{Gt-1}} &= \frac{\partial^2 f}{\partial p_t \partial S_{Gt}} \delta \\ &= \delta\beta_1 \left((1 - \Pr(P)) - \Pr(P)\right) \frac{\partial f}{\partial S_{Gt}} \\ &= \delta\beta_1 \left(1 - 2\Pr(P)\right) \left(\frac{-\exp(\alpha_1 + \beta_1 S_{Pt})\beta_2 \exp(\alpha_2 + \beta_2 S_{Gt})}{(1 + \exp(\alpha_1 + \beta_1 S_{Pt}) + \exp(\alpha_2 + \beta_2 S_{Gt}))^2} \right) \\ &= -\delta\beta_1\beta_2 \left(1 - 2\Pr(P)\right) \Pr(P) \Pr(G),\end{aligned}$$

which is negative unless $\Pr(P) > 1/2$. This is unlikely in a setting like auto insurance in which no firm dominates, hence past competitor advertising would be predicted to decrease a firm's advertising effect.

D Identifying customers

In this section, we explain how we identify whether an individual in our sample is a customer of our focal advertiser or one of its competitors. Since we only observe browsing URLs in the experimental dataset, we infer customer status by examining keywords within the URLs of the web pages they visit. For example, keywords such as "service", "login", or "my account", suggest that the individual is interacting with a customer portal rather than browsing as a non-customer. We identified these keywords based on their frequent use in customer portals across multiple insurance providers, indicating areas where users typically log in to access services.

To verify our inference, we manually visit the most visited URLs to determine if they redirect to a login page. This indicates that login credentials are required, supporting the notion that the individual is an existing customer. For example, the URL "service.geico.com/insite/iEntry" is frequently visited by participants for the brand Geico. Visiting this URL manually redirects us to a login page, suggesting that it is a customer interaction portal for Geico users. When a URL does not direct us to a login page, we do not consider its visitors as customers.

In addition to identifying existing customers, we classify visitors as potential new customers when they visit pages associated with obtaining quotes. We consider a page to be associated with quotes if the URL contains keywords such as "quote" or "sales", which suggest interest in purchasing a new policy rather than servicing an existing one. We manually visited these pages to check if they redirected to obtaining quotes. If they did, we used the associated keywords for the brand to identify potential customers.

In Table 12, we summarize the keywords that help identify individuals as existing or potential customers based on their browsing URLs. Using this keyword mapping, we classify individuals in our sample as existing or potential customers of top auto insurers. Our approach is deliberately conservative to ensure higher precision in classification: we limit our analysis to keywords found within the top ten most frequently visited URLs.

Table 12: Keywords and Webpages That Identify Existing and Potential Consumers

Brand	Visited by the brand's customer	Visited by the brand's potential customer
Progressive	servicing, service, login, myaccess, claim	progressivedirect.com
Geico	service, login, policy	quote, sales
Allstate	policy, myaccount, myportal	quote
Statefarm	login, service, claim, login, renewal	quote
Nationwide	myaccount, login	quote
Liberty Mutual	service, login, home	quote
Farmers	service, login	quote
Esurance	policy, login, customer, payment	quote
AAA	home, login, service, payment	quote

Notes: We use comScore data to examine the urls visited by customers of top auto insurance companies to find words in the urls that identify customers using the procedure explained in Appendix section D.

E Advertising Carryover Effects

E.1 Advertiser's Carryover

What is the effect of advertising on the following day visits for the advertiser?

On the next day, call it day 1, the consumer needs to invoke the auto insurance, which occurs due to exogenous events, regardless of any day 0 advertising. To visit the advertiser, the consumer also needs to draw it from memory.

The likelihood of visiting P's website on the next day is

$$\begin{aligned}\text{Visit}_P^1 &= \Pr(\text{Recalling the context}) \times \Pr(\text{Drawing P}|\text{Context being salient}) \times \Pr(\text{Visiting P}|\text{Drawing P}) \\ &= \alpha \times (1-\phi) \times \Pr(\text{Drawing P at } 1|P_1, G_1),\end{aligned}$$

assuming the notation from section 2; the likelihood of visiting P given a draw is α; P_1, G_1 are the strength of memory associations at the beginning of day 1. Assuming no advertising on day 1, P_1 and G_1 depend on the draws (p_0, g_0) made on day 0. On day 0, the consumer may not think of auto insurance, and if he does, he stops at getting two draws of a brand. So, the combination (p_0, g_0) can take five different values: $(0,0); (0,2); (1,2); (2,0); (2,1)$. Hence, we can rewrite

$$\text{Visit}_P^1 = \alpha \times (1-\phi) \times \left[\sum_{\{p_0, g_0\}} \underbrace{\Pr(p_0, g_0)}_{\text{Weight}} \times \underbrace{\Pr(\text{Drawing P at } 1|p_0, g_0)}_{\text{Dependence of draw at } t \text{ on draws at } t\text{-}1} \right].$$

So the effect of advertising at time 0 on visits at time 1 is

$$\Delta\text{Visit}_P^1 = \alpha(1-\phi) \left[\sum_{\{p_0, g_0\}} \underbrace{[\Pr(p_0, g_0|\text{P Ad at } 0) - \Pr(p_0, g_0|\text{No ad at } 0)]}_{\text{Change in weight}} \Pr(\text{Drawing P at } 1|p_0, g_0) \right].$$

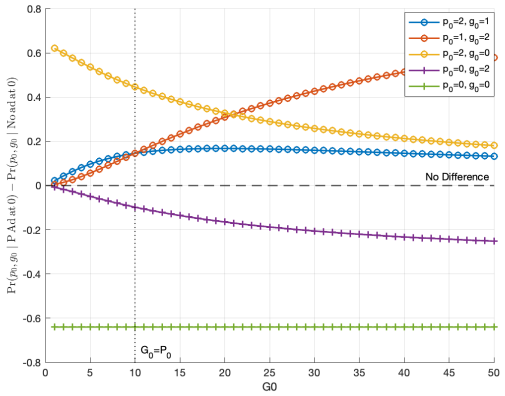
In other words, advertising changes the weights taken by the five states of (p_0, g_0) at time 1.

Since advertising at time 0 causes the consumer to draw P, it forces $p_0 \neq 0$. Hence, it decreases the weights on the states $(0,0)$ and $(0,2)$, where P was not drawn, and shifts it to other states $(2,0)$, $(2,1)$ and $(1,2)$ where there is at least one draw of P. Note that because $\Pr(\text{Drawing P at } 1|p_0, g_0)$ is lowest in $(0,2)$ among all five states, a shift away from $(0,2)$ always benefits the advertiser at time 1. However, a shift away from $(0,0)$ does not necessarily do so (e.g., if P's day 0 ad makes the consumer draw G twice). A shift away from $(0,0)$ is more harmful to the advertiser carryover effect as G_0 decreases. However, as G_0 decreases, the weights on states $(0,2), (1,2), (2,1)$ do not change much due to day 0 advertising which mitigates this effect. Overall, the advertiser benefits at time 1 from time 0 advertising.

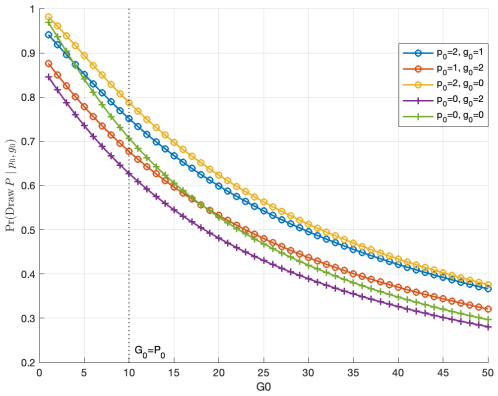
In Figure 10, we illustrate this with specific parameter values from the structural estimation[16]. For $P_0 = 10$ and G_0 from 1 to 50, we examine how advertising changes the weights of five states in (a), and the probability of drawing P given each of these five states in (b). (a) shows that P's advertising at time 0 increases the probability of consumers recalling P at time 0, and decreases the probability of consumers not recalling P: the weights on the states $(0,0)$ and $(0,2)$ decrease while the weights on other states $(2,0)$, $(2,1)$ and $(1,2)$ increase. (b) shows that $(0,2)$ always leads to the smallest probability of drawing P compared to other states, and $(2,0)$ always leads to the highest probability of drawing P. $(0,0)$ can lead to a higher probability of drawing P compared to $(1,2)$ and $(2,1)$ as G_0 decreases, but this is mitigated by the weight changes: we can see in (a) that as G_0 decreases, the

[16] We use $\alpha = 0.0076$, $\phi = 0.64$, $\delta = 0.46$, $r = 1.91$.

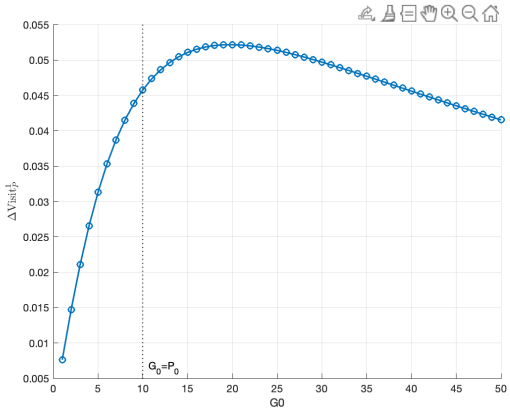
weight of $(0,0)$ shifts more to $(2,0)$ instead of $(1,2)$ or $(2,1)$.



(a) $\Pr(p_0, g_0|\text{P Ad at 0}) - \Pr(p_0, g_0|\text{No ad at 0})$



(b) $\Pr(\text{Drawing P}|p_0, g_0)$



(c) Advertiser carryover effect

Figure 10: Decomposing the carryover effect on advertiser visits

To comprehensively examine the positivity of the carryover effects, we numerically check for the lowest value taken by the carryover effects within a wide range of parameters. Since $\alpha > 0$ is a common multiplier, we set $\alpha = 1$ without loss of generality. We restrict ϕ and δ to be between 0 and 1, and we restrict P_0, G_0, and r to be greater than 0.

To ensure a global minimum is reached, we randomly sample 100 starting points within these ranges and find the minimum carryover effect for each using the quasi-newton method. This approach reduces the risk of convergence to local minima. The carryover effect is globally minimized at $-5 * 10^{-17}$, which is effectively 0 within the machine precision limits.[17]

This analysis supports that the advertiser's carryover from day 0 advertising is positive.

E.2 Dynamic effect on the competitor

What is the effect of advertising on the day 1 visits for the competitor?

[17]The parameter values that minimize the carryover effect are: $\phi = 0.4849$, $\delta = 0.6889$, $r = 0.0199$, $P_0 = 1.3122 * 10^4$, $G_0 = 0.0272$.

Following the same line of thinking as above, on day 1, the consumer needs to invoke the auto insurance category, draw competitor in their memory, and visit the competitor conditional on drawing.

The likelihood of visiting G's website on the next day can be written as

$$\text{Visit}_G^1 = \alpha \times (1-\phi) \times \sum_{\{p_0,g_0\}} \Pr(p_0,g_0) \times \Pr(\text{Draw G at } 1 | p_0,g_0).$$

So the effect of advertising at time 0 on competitor visit at time 1 is

$$\Delta\text{Visit}_G^1 = \alpha \times (1-\phi) \times \left[\sum_{\{p_0,g_0\}} \left[\Pr(p_0,g_0|\text{P Ad at } 0) - \Pr(p_0,g_0|\text{No ad at } 0) \right] \Pr(\text{Draw G at } 1 | p_0,g_0) \right].$$

Advertising changes the weights taken by the five states of (p_0, g_0). It decreases the weights on the states $(0,0)$ and $(0,2)$, where P was not drawn, and shifts them to other states $(2,0)$, $(2,1)$ and $(1,2)$. Note that the state $(2,0)$ is the worst for the competitor, and the weight shift to this state is zero when G_0 is low, so the competitor faces a smaller negative effect when G_0 is low. The negative impact gets more intense as G_0 increases from a low value. However, as G_0 gets large, the difference between the competitive outcomes at various states is small, so the negative carryover is small. Still, the competitor loses on day 1 from day 0 advertising even though this loss varies depending on initial memory strengths.

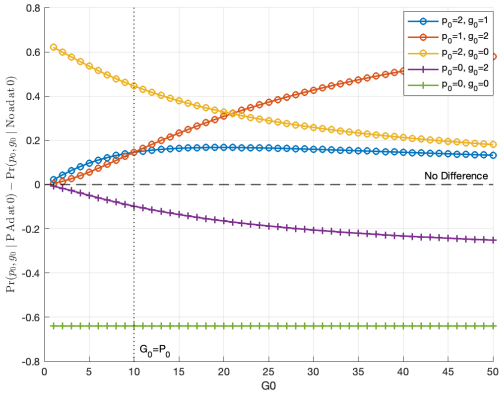
In Figure 11, we illustrate this with specific parameter values from the structural estimation[18]. For $P_0 = 10$ and G_0 from 1 to 50, we examine how advertising changes the weights of five states in (a), and the probability of drawing G given each of these five states in (b). (a) is the same as Figure 11 (a): treatment decreases the weights on the states $(0,0)$ and $(0,2)$ and increases the weights on states $(2,0)$, $(2,1)$ and $(1,2)$. (b) shows that as G_0 decreases, a shift from $(0,0)$ to $(2,0)$ is always harmful for the carryover effect, and the size of this is maximized at intermediate level of G_0. As G_0 decreases, a shift from $(0,0)$ to $(2,1)$ or $(1,2)$ becomes more harmful. This is mitigated by the decreasing weight-shifting from $(0,0)$ to $(2,1)$ or $(1,2)$ as G_0 decreases.

To comprehensively examine the negativity of the carryover effects, we numerically check for the highest value taken by the carryover effects within a wide range of parameters. Since $\alpha > 0$ is a common multiplier, we set $\alpha = 1$ without loss of generality. We restrict ϕ and δ to be between 0 and 1, and we restrict P_0, G_0, and r to be greater than 0. To ensure a global maximum is reached, we randomly sample 100 starting points within these ranges and find the minimum carryover effect for each. The carryover effect is globally maximized at $6 * 10^{-18}$, which is effectively 0 within the machine's precision limits. [19]

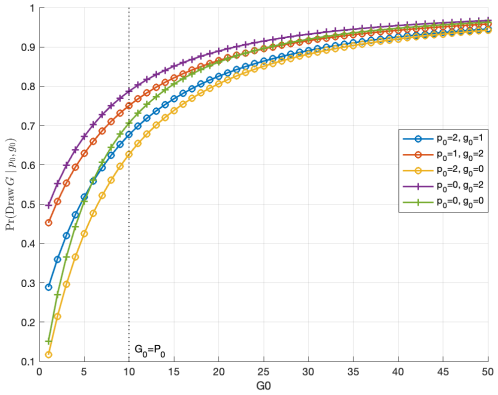
This analysis supports the notion that the carryover on the competitor is negative.

[18] We use $\alpha = 0.0076$, $\phi = 0.64$, $\delta = 0.46$, $r = 1.91$.

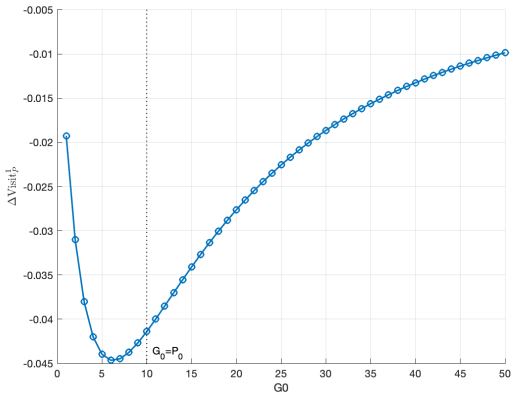
[19] The parameter values that minimize the carryover effect are: $\phi = 0.9970$, $\delta = 0.7830$, $r = 0.0000$, $P_0 = 0.0562$, $G_0 = 1.6047 * 10^6$.



(a) $\Pr(p_0, g_0|\text{P Ad at 0}) - \Pr(p_0, g_0|\text{No ad at 0})$



(b) $\Pr(\text{Drawing G}|p_0, g_0)$



(c) Competitor carryover effect

Figure 11: Decomposing the carryover effect on competitor visits

E.3 Competitor's contemporaneous effect

Note that even though the competitor carryover is negative, the competitor effect on day 0 can be positive. The likelihood of visiting G's website on day 0 is

$$\text{Visit}_G^0 = \Pr(\text{Recalling context}) \times \Pr(\text{Drawing } G \mid \text{Recalling context}) \times \Pr(\text{Visit } G|\text{Drawing } G).$$

So, the effect of advertising at time 0 on competitor visits at time 0 is

$$\begin{aligned}\Delta\text{Visit}_G^0 &= [\Pr(\text{Draw } G \text{ at } 0 \mid \text{P Ad}_0) - \Pr(\text{Draw } G \text{ at } 0 \mid \text{No P Ad}_0)] \times \Pr(\text{Visit } G \mid \text{Draw } G) \\ &= \alpha \times [h(G_0, P_0 + r) - (1-\phi)\left[h(G_0, P_0) + h(P_0, G_0)h(G_0, P_0 + r)\right]].\end{aligned}$$

In a scenario where the consumer always forgets the category ($\phi = 1$) in the absence of ads, he will not visit any auto-insurance brand without advertising. In this scenario, when the advertiser advertises, the consumer will visit the competitor with a positive probability, resulting in a positive day 0 effect on competitor visits. In contrast, in a scenario where the category is always salient to the consumer ($\phi = 0$), advertising does not benefit the competitor and instead reduces the likelihood of recalling it, leading to a negative day 0 effect on competitor visits. When $\phi = 0$, $\Delta\text{Visit}_G^0 = -h(G_0, P_0)h(P_0 + r, G_0)$, which is always negative and decreases in magnitude as G_0 increases. This implies that the negative effect is small for a large competitor.

F Details on Model Working and Estimation

This section provides details to the model estimation presented in section 8.

F.1 Memory Model Setup

F.1.1 Drawing Brands From Memory

Here we discuss how consumers draw brands from their memory in our model depending on ad exposure. Based on this analysis, we obtain the probability distribution of consumer memory strengths given their ad exposure.

Each day t, consumer i updates his memory strengths based on the brands he draws. When exposed to advertising, he recalls the auto-insurance purchase context; when not exposed to any ads, he forgets about the purchase context with probability ϕ. Once the consumer recalls the purchase context, he draws brands from memory. Following the assumption in section 2, we assume consumers stop drawing brands from their memory as soon as they retrieve a brand that has been drawn before.

If a consumer recalls one brand, the memory strength of that brand increases by r, which reduces the probability that the consumer recalls others. Therefore, the probability of recalling one brand and

the probability of recalling others are correlated, so we need to consider the joint probability.

Given our stopping rule, there are 5 possible combinations of recalls: we use (R_{iPt}, R_{iGt}) to denote the combinations, where each element takes the value of 0, 1, or 2. One and only one out of these values gets to take the value of 2, since the consumer stops retrieving more brands once he recalls the same brand twice.

The matrix in Table 13 shows the probability distribution of (R_{iPt}, R_{iGt}) given each possible ad exposure combination (A_{iPt}, A_{iGt}). Each row of the matrix sums up to 1. In the context of our experiment, each consumer's ad exposure can be expressed by $(A_{iPt}, A_{iGt}) = (D_i \cdot 1\{t = 0\}, 0)$, where $1\{t = 0\}$ indicates the day of the experiment.

Table 13: The Probability Distribution of Recalling P and G Conditional on Ad Exposure

A_{iPt}, A_{iGt}	$\{R_{Pt}, R_{Gt}\}$				
	$\{2,1\}$	$\{1,2\}$	$\{2,0\}$	$\{0,2\}$	$\{0,0\}$
$\{1,1\}$	$h(P_t + r, G_t + r)$	$h(G_t + r, P_t + r)$	0	0	0
$\{1,0\}$	$h(G_t, P_t + r)h(P_t + r, G_t + r)$	$h(G_t, P_t + r)h(G_t + r, P_t + r)$	$h(P_t + r, G_t)$	0	0
$\{0,1\}$	$h(P_t, G_t + r)h(P_t + r, G_t + r)$	$h(P_t, G_t + r)h(G_t + r, P_t + r)$	0	$h(G_t + r, P_t)$	0
$\{0,0\}$	$(1-\phi)h(P_t + r, G_t + r)\cdot$ $(h(P_t, G_t)h(G_t, P_t + r)+$ $h(G_t, P_t)h(P_t, G_t + r))$	$(1-\phi)h(G_t + r, P_t + r)\cdot$ $(h(G_t, P_t)h(P_t, G_t + r)+$ $(h(P_t, G_t)h(G_t, P_t + r))$	$(1-\phi)\cdot$ $h(P_t, G_t)h(P_t + r, G_t)$	$(1-\phi)\cdot$ $h(G_t, P_t)h(G_t + r, P_t)$	ϕ

Notes: For each possible ad exposure combination (A_{iPt}, A_{iGt}), the table shows the probability distribution of (R_{iPt}, R_{iGt}).

Take $\Pr(2,1 \mid 1,0)$ for an example: since $(R_{iPt}, R_{iGt}) = (2,1)$, $(A_{iPt}, A_{iGt}) = (1,0)$, the second row, first column shows the probability of consumer i recalling P twice and recalling G once conditional on consumer i being exposed to P's advertising and not exposed to G's advertising. As i is exposed to ads from P but not G, he recalls P with certainty. After recalling P, his association between purchase context and P increases by r, leading to memory strengths $(G_t, P_t + r)$. At this point, the stopping rule is unsatisfied, so he draws a brand from memory. With probability $h(P_t + r, G_t)$, he draws brand P; with probability $h(G_t, P_t + r)$, he draws brand G. In the former case, he draws P for the second time, which satisfies the stopping rule, so he stops drawing. In the latter case, he continues to draw another brand from memory. His memory strength between the purchase context and G increases by r, leading to memory strength $(G_t + r, P_t + r)$. In the next draw, the probability that he draws G is $h(G_t + r, P_t + r)$, while the probability that he draws P is $h(P_t + r, G_t + r)$. In either of these cases, he is drawing a brand that he has drawn before, so he stops drawing. Based on this, we can calculate the probability of drawing P twice and drawing G once conditional on seeing P's ad but not G's ad:

$$\Pr\Big((R_{iPt}, R_{iGt}) = (2,1) \mid (A_{iPt}, A_{iGt}) = (1,0) \Big) = h(G_t, P_t + r) \cdot h(P_t + r, G_t + r) \tag{8}$$

The path of recalls given $(A_{iPt}, A_{iGt}) = (1,0)$ is shown in Figure 12. Based on the analysis above, we obtain the conditional probability distribution of consumer recalls given their ad exposure.

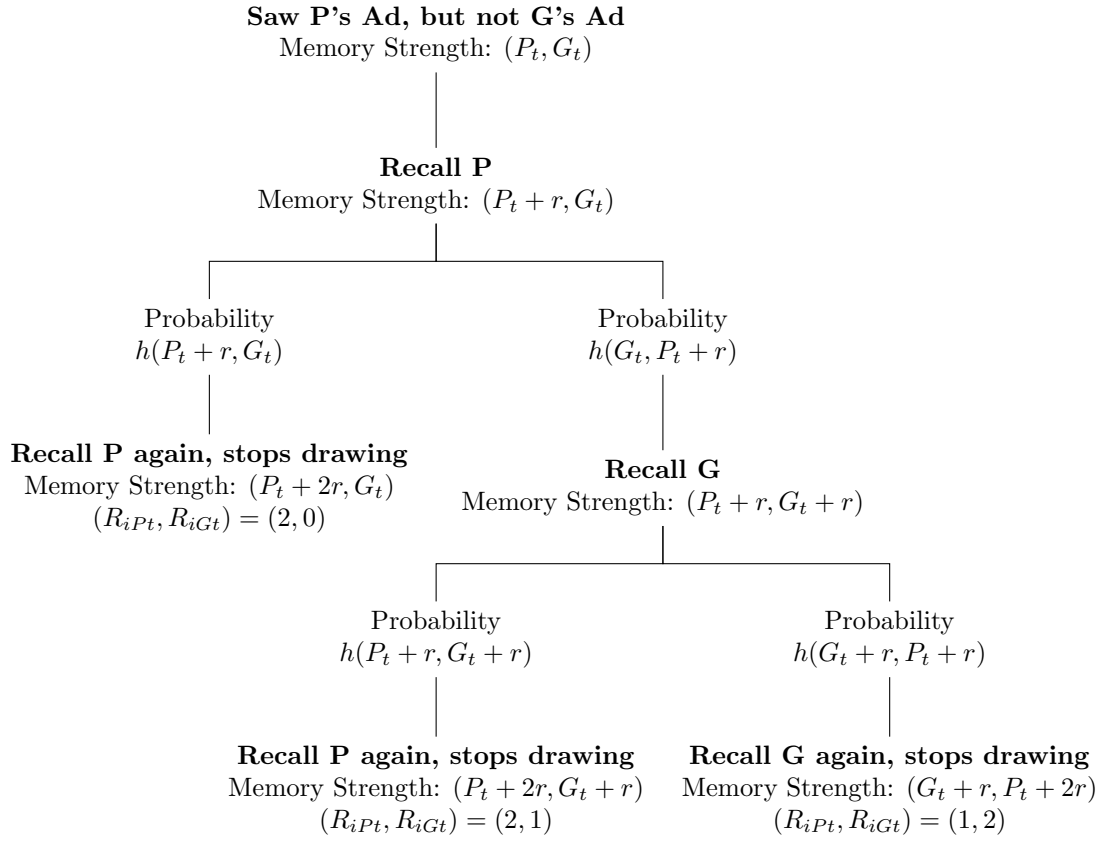


Figure 12: The Path of Recalls for $(A_{iPt}, A_{iGt}) = (1, 0)$

Note: This figure shows the paths of consumer recalls when consumer is exposed to P's advertising but not G's advertising on day t. Specifically, we illustrate consumer's sequence of recalls, the probability associated with each path of recall, and the resulting memory strengths and times of recalls.

F.1.2 Visiting Brands Based on Recalls

After the consumer stops recalling brands from memory, he decides whether to visit them or not. He can only visit the brands that he successfully recalls.

For simplicity, we assume 1) consumers visit a brand that they recall with a fixed probability α, which is an outcome of the trade-off discussed in section 2; 2) the conditional probability of visiting one brand given recalls is independent of the conditional probability of visiting another brand given recalls. Specifically, using $\hat{c}_{ijt}$ to represent the model's prediction on whether i visits brand j at day t, we have:

$$\Pr(\hat{c}_{ijt}, \hat{c}_{i(-j)t}|R_{ijt}, R_{i(-j)t}) = \Pr(\hat{c}_{ijt}|R_{ijt}, R_{i(-j)t}) \cdot \Pr(\hat{c}_{i(-j)t}|R_{ijt}, R_{i(-j)t}) \tag{9}$$

for $j \in P, G$.

The probability that consumer i visits brand j or not at day t only depends on whether he recalls j. For $j \in P, G$, the probability is given by:

$$\Pr(\hat{c}_{ijt} = x|R_{ijt} = y, R_{i(-j)t}) = \begin{cases} \alpha & \text{if } x = 1, y \geq 1 \\ 1-\alpha & \text{if } x = 0, y \geq 1 \\ 0 & \text{if } x = 1, y = 0 \\ 1 & \text{if } x = 0, y = 0 \end{cases} \tag{10}$$

F.1.3 Updating Memory Strength

At the end of each period, each consumer's memory strengths between the purchase category and the brands change.

To ensure that the $h(\cdot, \cdot)$ function is well-behaved (the sum of the two entries can't be 0), we assume that the memory strengths for each brand is at least 1 in the estimation version of the model. A consumer's memory strengths at the start of time t depend on the number of recalls of this brand in the past, how each retrieval increases the memory strengths (r) and how the effect of past retrievals decay over time (with a survival rate of δ):

$$\begin{aligned} P_{it} &= \begin{cases} 1 & \text{if } t = 0 \\ 1 + \sum_{z=0}^{t-1} (R_{iPz} r \delta^{t-z}) & \text{if } t \geq 1 \end{cases} \\ G_{it} &= \begin{cases} 1 & \text{if } t = 0 \\ 1 + \sum_{z=0}^{t-1} (R_{iGz} r \delta^{t-z}) & \text{if } t \geq 1 \end{cases} \end{aligned} \tag{11}$$

The following equation gives how the next period's memory strength is updated depends on the current period's memory strengths:

$$(P_{i(t+1)}, G_{i(t+1)}) = (\delta(P_{it} - 1) + r\delta \cdot R_{iPt} + 1, \delta(G_{it} - 1) + r\delta \cdot R_{iGt} + 1) \tag{12}$$

F.2 Likelihood-based Estimation of the Memory Model

Based on the model above, we calculate the joint probability of consumers visiting P and G. We then maximize the likelihood of our observed consumer visits based on the set of parameters $\boldsymbol{\theta}$, including the following: the increase in memory strength from each draw (r), the survival rate of memory strength (δ, with $1 - \delta$ as the depreciation rate), the probability of forgetting the auto insurance category when a consumer did not see any ads (ϕ), the probability of visiting a brand after recalling it (α). We use $\{c_{ijt}\}_{t=\{0,1\},j\in\{P,G\}}$ to represent actual consumer visits we observe in the data, in contrast to $\{\hat{c}_{ijt}\}_{t=\{0,1\},j\in\{P,G\}}$ predicted by our model. The likelihood function is given by:

$$\boldsymbol{L}(\boldsymbol{\theta}) = \prod_{i=1}^{N} \Pr(\{\hat{c}_{ijt}\}_{t\in\{0,1\},j\in\{P,G\}} = \{c_{ijt}\}_{t\in\{0,1\},j\in\{P,G\}} \mid D_i, \{A_{ijt}\}_{t=\{0,1\},j=\{P,G\}}, P_{i0}, G_{i0}; \boldsymbol{\theta})$$

F.2.1 Decomposing The Likelihood Function

The likelihood function can be decomposed into the likelihood of experiment day visits and the conditional likelihood of the next-day visits:

$$\begin{aligned}\boldsymbol{L}(\boldsymbol{\theta}) =& \prod_{i=1}^{N} \Pr(\{\hat{c}_{ij0}\}_{j\in\{P,G\}} = \{c_{ij0}\}_{j\in\{P,G\}} \mid D_i, \{A_{ij0}\}_{j=\{P,G\}}, P_{i0}, G_{i0}; \boldsymbol{\theta}) \times \\ & \prod_{i=1}^{N} \Pr(\{\hat{c}_{ij1}\}_{j\in\{P,G\}} = \{c_{ij1}\}_{j\in\{P,G\}} \mid \{c_{ij0}\}_{j\in\{P,G\}}, D_i, \{A_{ijt}\}_{t=\{0,1\},j=\{P,G\}}, P_{i0}, G_{i0}; \boldsymbol{\theta})\end{aligned}$$

F.2.2 Conditional Likelihood of Day 1 Visits

Calculating the likelihood of day 0 visits is straightforward. Calculating the likelihood of day 1 visits is trickier since the likelihood is conditional on day 0 visits. To obtain day 1 likelihood, we need consumers' day 1 memory strengths, which depend on consumers' day 0 recalls of brands. Although we do not observe consumer recalls, we observe day 0 visits, which allows us to update the probability distribution of day 0 recalls.

We can derive the conditional probability of day 0 recalls using the Bayes' Rule.

$$\Pr(\{R_{ij0}\}_{j\in\{P,G\}} \mid \{c_{ij0}\}_{j\in P,G}) = \frac{\Pr(\{R_{ij0}\}_{j\in\{P,G\}}, \{c_{ij0}\}_{j\in P,G})}{\sum_{\{\{R_{ij0}\}_{j\in\{P,G\}}\}} \Pr(\{\{R_{ij0}\}_{j\in\{P,G\}}, \{c_{ij0}\}_{j\in P,G})}$$

where $\{\{R_{ij0}\}_{j\in\{P,G\}}\}$ is the set of all possible combinations of R given in Table 13.

We now have all the elements we need to calculate the likelihood function.

F.3 Estimates of the Alternative Models

We present here the estimates of the parameters for the alternative models discussed in section 8. Tables 14 and 15 show the parameter estimates of the ad stock model specifications with four and five parameters, respectively.

Table 14: Estimation Result of 4-Parameter Ad Stock Model, Advertiser and A Top Competitor

Parameter	Description	Estimate [95% Confidence Interval]
a	The Baseline Utility	-1.8917 $[-1.9119, -1.8714]$
b_1	The Marginal Utility of Own Brand's Ad Stock	0.2458 $[0.2344, 0.2573]$
b_2	The Marginal Utility of Competitor Brand's Ad Stock	-0.0066 $[-0.0216, 0.0084]$
d	The Surviving Rate of Ad Stock Per Period	0.2534 $[0.1839, 0.3382]$

Note: This table presents the point estimates and confidence intervals for the parameters of the 4-parameter ad stock model. Our estimation used relevant data of two firms: the advertiser in our experiment and a top competitor (same as the data used for estimation in Table 8). To ensure that the parameters satisfy the boundary condition $d \in [0, 1]$, we estimate a transformed version of parameter d. We estimate parameter $\tilde{d}$ such that $d = \frac{\exp(\tilde{d})}{\exp(\tilde{d})+1}$. We transformed the estimates $\tilde{d}$ and its confidence interval back to the original parameter d. The results show that advertising increases consumer utility of the advertised brand. While there is no detectable effect of advertising on consumer utility of competitors, the sign is overall negative. At each period, around 25% of the ad stock in the previous period is retained.

Table 15: Estimation Result of 5-Parameter Ad Stock Model, Advertiser and A Top Competitor

Parameter	Description	Estimate [95% Confidence Interval]
a	The Baseline Utility	-1.8886 $[-1.9092, -1.8681]$
b_1	The Marginal Utility of Own Brand's Ad Stock	0.2526 $[0.2327, 0.2725]$
b_2	The Marginal Utility of Competitor Brand's Ad Stock	0.0004 $[-0.0221, 0.0228]$
d	The Surviving Rate of Ad Stock Per Period	0.2511 $[0.1839, 0.3327]$
S_0	The Initial Ad Stock	0.8015 $[0.4546, 1.4128]$

Note: This table presents the point estimates and confidence intervals for the parameters of the 5-parameter ad stock model. Our estimation used relevant data of two firms: the advertiser in our experiment and a top competitor (same as the data used for estimation in Table 8). To ensure that the parameters satisfy the boundary condition ($d \in [0, 1]$, $S_0 \geq 0$), we estimate a transformed version of these parameters. We then transformed the estimates and their confidence interval back to the original parameters. The results show that advertising increases consumer utility of the advertised brand, and there is no effect of advertising on consumer utility of competitors. At each period, around 25% of the ad stock in the previous period is retained.

Table 16: Effect of Treatment on visits to any auto insurance company's website at different times.

	(1) Before	(2) Day of Experiment	(3) Month after	(4) Month 2 after	(5) Month 3 after	(6) Month 4 after†
Treatment	-0.0004598	**0.006006**	0.003627	-0.003538	-0.002499	0.0001245
	(0.0041)	**(0.000467)**	(0.003328)	(0.00301)	(0.002944)	(0.001737)
Intercept	**0.2977**	**0.01025**	**0.2166**	**0.1765**	**0.1689**	**0.08133**
	(0.003536)	**(0.0003797)**	**(0.002867)**	**(0.002622)**	**0.1689 (0.002562)**	**(0.001502)**
N	326,952	326,952	326,952	326,952	326,952	326,952

Notes: Each column represents a regression of the dependent variable – the count of unique days an individual visited any auto insurance company's website during the specified time interval – on an indicator variable for assignment to the treatment group. Multiple visits by an individual to a website on the same day are counted as a single visit. Robust standard errors are shown in parentheses. Coefficients are bolded to indicate statistical significance at the 5% level.

† Month 4 after the experiment is partially observed in the data window.