Ads as Cues

Pedro Bordalo, Giovanni Burro, Nicola Gennaioli, Gad Nacamulli and Andrei Shleifer[1]

October 13, 2025

Abstract

Why do we see both advertising and powerful consumer habits for well-known, intrinsically similar brands? We start from the idea that a consumer is more likely to demand a good if she recalls pleasurable experiences with it. In turn, recall is stronger for frequently consumed goods and those similar to cues. Our model yields context-dependent habits where ads work as memory cues. It predicts ads are more effective for habitual consumers, and exhibit spillovers that are stronger for goods with similar ads. NielsenIQ consumer and advertising data supports these predictions in 20 undifferentiated, highly advertised product categories.

[1] The authors are from Oxford Said Business School, Binghamton University, Università Bocconi, London School of Economics and Harvard University, respectively. We are grateful to Bart Bronnenberg, Jean-Pierre Dubé, Matthew Gentzkow, Thomas Graeber, Alex Imas, Jesse Shapiro, Josh Schwartzstein and Tosten Persson for very helpful comments. Researchers' own analyses are calculated based in part on (i) retail measurement/consumer data from Nielsen Consumer LLC ("NielsenIQ"); (ii) media data from The Nielsen Company (US), LLC ("Nielsen"); and (iii) marketing databases provided through the respective NielsenIQ and 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 NielsenIQ and 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. Gennaioli thanks the European Research Commission for an Advanced Grant (GA 101097578). Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

1. Introduction

Consumers exhibit strong habits in their demand for brands such as Coke and Pepsi even though these drinks are very similar and their relative prices vary significantly over time. These brands are also highly advertised, despite being well known and heavily purchased. Neoclassical theories view habits as "brand capital" built through consumption (Becker and Murphy 1988, henceforth BM) and uninformative ads as a utility-complement of consumption (Becker and Murphy 1993). Such utility functions account for some puzzles but also face problems. For instance, they predict that consuming a good should increase the utility of seeing its ad, which is unintuitive. More important, they do not make predictions on how habits and ads interact. Is a habitual Coke drinker more or less influenced by a Coke ad than an infrequent one? Would a Coke ad trigger a habitual Pepsi drinker to demand more or less Pepsi? These questions are key to understanding how advertising works but are not addressed by utility models.

We develop a theory of consumer demand based on well-established regularities of human memory. The theory unifies habits, ads, and their interaction. Its intuition is captured by a story recounted to one of us by George Stigler almost 40 years ago. Stigler supposedly asked Mr. Wrigley, then running his family's chewing gum company in Chicago, why he kept spending on advertising when everyone already knew his brand. "See that electric train passing by", replied Wrigley. "What would happen if they turned off electricity?" Like trains, people often follow habitual consumption paths. They would however stop without ads. The energy in ads, we contend, is memory: it renders specific consumer wants salient, cueing their satisfaction.

In our model, as in Bordalo et al. (2020), retrieved experiences affect predicted utility: a consumer recalling the pleasure from consuming a brand is more likely to buy it today, especially if her past pleasure was high. As in Bordalo et al. (2020, 2023) and standard psychology (Kahana

2012), recall favors brands that are frequently consumed and similar to cues, subject to interference from other stimuli. Memory then yields habits: a more frequent Coke drinker is more likely to recall its pleasure and buy it again. It also yields a role for ads: a Coke ad cues pleasurable past experiences with Coke, boosting demand. Ads are uninformative yet powerful cues.

Memory also disciplines the *interaction* of habits with ads, yielding several new predictions. The marginal effect of ads for a consumer is increasing in past purchase frequency, but at a slower rate for very frequent buyers. A Coke ad is a stronger cue for frequent Coke drinkers, for whom many Coke sensations come to mind compared to sporadic drinkers. These memories may also be more pleasurable due to a higher intrinsic taste for Coke. Eventually, though, very habitual drinkers need few cues, so sensitivity to ads declines.

Memory yields new testable predictions based on the similarity between ads and the interaction of similarity with habits. First, an ad for a brand should create positive spillovers on brands with similar ads, especially for habitual buyers of those brands. A Pepsi ad cues frequent Pepsi buyers to recall and buy Pepsi, but surprisingly it also cues frequent Coke buyers to recall and buy Coke. Spillovers also work across product categories: ads for "colas" should cue frequent consumers of similar soft drinks such as Sprite to recall and buy those. Unlike in BM (1993), an ad does not increase the marginal utility of the advertised brand, because consumers with different habits represent the same ad differently: a Pepsi ad reminds a Pepsi drinker of Pepsi, but it reminds a Coke drinker of Coke. Similarity differs sharply from economic substitutability: it reinforces prior habits and reduces consumer mobility across goods, and more so precisely when goods are more similar (either in their ads or intrinsically) and thus more economically substitutable.

The second similarity prediction is that an ad for a product should create *negative* spillovers for products with dissimilar ads, due to memory interference. An ad of frozen pizza may cue a

snack like nachos, but cause consumer forgetting of dissimilar fruit or yogurt. Unlike in BM (1993), where ads are optimally demanded by the consumer, in our model they cue specific wants and interfere with other wants involuntarily, bottom up, possibly reducing consumer welfare.

We test the model's predictions by jointly using, for the first time, proxies for habits, cues, similarity, and choices. From the NielsenIQ Homescan panel and the Nielsen Ad Intel databases, we measure a household's i) past purchases, our proxy for habits, and ii) its weekly exposure to TV ads in each Designated Market Area (DMA), our proxy for cues. We also observe iii) short descriptions of ad content, from which we construct proxies of similarity across brands. We study household demand for 40 products, identified as the top two brands in 20 ad intensive categories, ranging from colas to diapers (Bronnenberg et al 2012). To capture standard drivers of demand, we control for household characteristics (income, size, etc.), goods' prices, and discounts. We also control for unobserved cross sectional and time series demand shifters such as seasonality using a rich set of dummies, and implement the Shapiro et al. (2021) strategy to isolate quasi random variation in ads. In addition, we control for a product's aggregate weekly demand in the DMA itself and run a large battery of other tests.

Our results are stable across these tests and highlight memory's explanatory power. First, there is a strong and concave habit-ads interaction. An extra 2.75 minutes of Coke ads lead an *average* Coke drinker to buy an extra can of Coke in a week, but a one standard deviation *more frequent* Coke drinker needs only 1 extra minute of Coke ads to buy an extra can. We do not separate whether this effect arises because the regular Coke drinker retrieves many Coke experiences or has a higher taste for Coke, because both effects are predicted by memory. On the other hand, we find that the marginal effect of ads is smaller for very frequent drinkers, an effect due in our model to their spontaneous retrieval of many Coke experiences, not to tastes.

Second, the data support our new predictions regarding similarity-based spillovers. The same one standard deviation more frequent Coke drinker consumes an extra Coke can if exposed to an extra 80 seconds of Pepsi ads. We also detect negative spillovers. Exposure to ads for colas or beer boosts demand for other drinks and reduces demand for non-drink (and therefore dissimilar) products in our sample. Measured similarity helps explain the surprising effect of an ad on the demand for non-advertised goods.

Regularities in selective recall offer new explanations for long-standing issues in consumer demand. Consumers tend to buy the same brand repeatedly (Frank 1962, Bronnenberg and Dubé 2017). Heckman (1981) argues that this may reflect stable tastes, while Dubé et al. (2010) show that state-dependent "brand loyalty" is more consistent with the data. Bronnenberg et al. (2012) link brand loyalty to having in the past lived in a city where that brand was predominant, while Atkin (2013), Bronnenberg, Dubé, and Syverson (2022) tie it directly to past consumption. Our memory model yields habits which, consistent with cognitive science (see Webb et al. 2025 for a review), depend on irrelevant context. Compared to Webb et al. (2022), who model habits using reinforcement learning, we trace the role of context to memory cues and similarity, which yields a new theory of advertising. It also yields new predictions that allow us to test state dependence using changes in context, without relying solely on positive autocorrelation in demand.

Marketing scholars have long recognized the role of memory. Laboratory experiments show that ads exhibit "reminder" effects (e.g. Keller 1987, 1993), which are also documented for state lotteries (He and Klein 2023). Ehrenberg (1974), Tellis (1988), and Keller (1993) show for some products that ads can be more effective with loyal customers. Sahni (2016) shows positive spillovers for restaurants, Shapiro (2018) for antidepressants. We offer systematic evidence across 20 major categories and show how these effects follow a similarity structure. Shapiro (2006) offers

an insightful memory model in which ads remind consumers of positive vs negative experiences with a product or convert negative experiences into positive ones. We build on broader recall patterns whereby ads make specific wants salient and produce similarity-based spillovers.

As we argue in Section 6, memory has key implications compared to utility or information-based models of ads (Stigler 1968, Nelson 1974). It helps explain why firms spend a lot on uninformative advertising despite a small short run effect (Lodish 1995, Shapiro et al. 2021). It allows a systematic study – through similarity – of whether and how advertising content matters. It offers a nuanced welfare perspective on how ads or other cues such as social media can benefit consumers by offering useful reminders (even if uninformative) but can also cause costly neglect of non-cued wants and promote market power. Ads cue specific wants "bottom up", and interfere with others. They thus act like an externality and are unlikely to be socially efficient.

The paper proceeds as follows. Section 2 presents our model and derives its predictions. Section 3 presents the data. Section 4 tests the habits-ads interaction within product categories. In Section 5 we measure similarity based on ad content and test its role, also across product categories. Section 6 discusses implications for consumer choice and welfare. Section 7 concludes.

2. A Memory Model of Habits and Advertising

There are K product categories (e.g. cola, beer, or nappies) indexed by $k = 1, \dots, K$. In each category there is a leader and a follower brand (e.g. Coke and Pepsi), indexed by $b = 1$ and $b = 2$, respectively. A good $g \in G \equiv \{kb\}_{k=1,\dots,K;\, b=1,2}$ is a product in category k and brand b. It yields deterministic consumption utility $q_g > 0$ and its price is p_g. Tastes q_g can vary across

consumers, but may entail indifference within a category $q_{k1} = q_{k2}$, e.g. some goods cannot be recognized in blind tasting. We collect a consumer's tastes in vector $\mathbf{q} = (q_{11}, \dots q_{K2})$.

Total utility is separable across goods. Compared to a standard setting, our consumer may forget some brands and recall others. Recall of g retrieves its pleasure q_g, which boosts the consumer's marginal utility from g compared to not recalling its pleasure. Bushong, King, Camerer, and Rangel (2010) show that the physical presence of a good increases the willingness to pay for it, consistent with the idea that we better imagine the pleasure from consuming g if the good is top of mind.[2] Ads affect demand by acting as a cue for q_g. The source of our new predictions lies in formalizing recall based on memory regularities.

2.1 Cues, Similarity and Recall

Following memory research (Kahana 2012, Bordalo et al. 2023, 2024), cued recall brings experiences to mind based on their past frequency, their similarity to the cue, and interference from other experiences. The memory database is a vector $\mathbf{f} = (f_{11}, \dots, f_{K2}, f_n)$ listing the relative frequency f_g of the consumer's experiences with each good $g = kb \in G$, and the relative frequency f_n with which she experienced a category of "neutral" events n, so $\sum_x f_x = 1$. Neutral events do not involve brands (talking to friends, taking a nap, etc.), so we should expect n to be much higher than specific goods $f_n \gg f_g$. A role of ads is to cue a good against the "fog" of non-branded activities. The categories of experiences are $X = G \cup \{n\}$.

A cue c is a signal pointing to a category in X, for simplicity $c \in X$, affecting retrieval. A physiological signal, e.g. a craving for Coke, is a cue. Ads are also cues. An ad of g is a cue $c =$

[2] Memory can also shape predicted utility by promoting the simulation of future experiences from past experiences (Bordalo, Burro et al 2025). In this way, memory can also be relevant for ads of new goods, see Section 6.

g. A non-branded activity, e.g. hiking, is a neutral cue $c = n$. A symmetric function $S: X \times X \to [\underline{S}, \overline{S}]$ measures similarity between a cue and a category of experiences. $S(c, x)$ is positive, increasing in the feature overlap between c and x, and satisfies:

Assumption 1. *For each good g, similarity $S(g, x)$ with $x \in X$ is maximal with itself, $S(g, g) = \overline{S}$, and minimal with neutral experiences, $S(g, n) = \underline{S}$.*

The ad for a brand is most similar to consuming the same brand because it shares many features with it: the name of the good, its color, shape, etc. A brand is however also similar to other brands, especially in the same product category, e.g. Pepsi shares the features "drink", "sparkling", "sweet", and so on, with Coke. Similarity decays as we move to goods with fewer shared features. Ads can also include context features that affect similarity because the brand is consumed in similar situations, e.g. a party with friends. In general, features associated with more pleasurable experiences are more similar to consumption pleasure, promoting retrieval of q_g. We summarize shared features with a similarity function $S(c, x)$ between cues and brands, which we later measure. This strategy allows us to capture our effect of interest: ads describe a context that cues a brand's pleasure, with potential spillovers to other brands.

Neutral activities $x = n$ have the lowest similarity to experiencing brands because they do not directly concern consumption and occur in a wide variety of contexts that are in general different from consumption or ads.[3] The probability $r_x(c, \mathbf{f})$ of recalling one instance of $x \in X$ based on cue c is proportional to $S(c, x)$. As a result:

[3] A neutral activity may be similar to a brand, e.g. jogging and Nike, but the average similarity of neutral activities with a brand is low. We allow $S(n, n) \in [\underline{S}, \overline{S}]$ but arguably $S(n, n) < \overline{S}$ because n includes very different experiences.

$$r_x(c, \mathbf{f}) = \frac{S(c, x) \cdot f_x}{\sum_{x' \in X} S(c, x') \cdot f_{x'}}. \quad (1)$$

The numerator captures associativity and frequency in recall. Recalling $x = Coke$ is more likely after seeing a Coke ad, $c = Coke$, than a Pepsi or Sprite ad due to higher similarity of $Coke$ memories with a *Coke* ad, $S(Coke, Coke) > S(Coke, Pepsi) > S(Coke, Sprite)$. The same cue $c = Coke$, however, acts differentially across people: retrieving Coke experiences is *ceteris paribus* more likely for a habitual Coke drinker, for whom f_{Coke} is higher.

The denominator in (1) captures interference. Experiences compete in recall. Thus, if $x' \neq x$ is very similar to the cue or frequent in the database, high $S(c, x')$ or $f_{x'}$ respectively, it inhibits the retrieval of x. A consumer may fail to retrieve Coke, even when faced with a Coke ad, because the similarity of other soft drinks $S(Coke, Sprite)$ is high. This is especially likely to happen if the consumer is a frequent Sprite drinker, f_{Sprite} is high. Neutral events $x = n$ may be retrieved even if dissimilar, given their high frequency f_n. Interference captures distraction.

One key assumption in our model is that even if two brands are intrinsically similar (in the limit their utilities can be identical), their experiences are stored separately in memory due to irrelevant features, such as "branding", which are salient during consumption. Cues then prompt selective recall of one brand or the other, not of a homogenous product. (Goods that are very different are of course encoded separately, but our focus is on branded undifferentiated products, such as Coke vs Pepsi, for which loyalty and uninformative advertising are most puzzling.)

Exposure to different cues in a time unit (a week) is a vector $\mathbf{e} = (e_{11}, \ldots, e_{K2}, e_n)$, where e_c is the absolute frequency of exposure to $c \in X$ and $e = \sum_x e_x$ is the time budget in that time unit. Ads do not carry utility. Crucially, exposure to ads is exogenous, unlike in BM (1993) where

it is optimally chosen. Surely, consumers try to avoid unpleasant ads and watch others (e.g., Superbowl ads), but much exposure is due to "bottom up" attention. We spontaneously look at a billboard on the street or at a brand's name on an athlete's shirt.

The recall probability of g in the time unit is then the average across cues, so that:

$$r_g(\mathbf{e}, \mathbf{f}) = \sum\nolimits_{c \in X} r_g(c, \mathbf{f}) \cdot \eta_c, \quad (2)$$

where $\eta_c \equiv e_c/e$ is relative exposure to c. Equation (2) predicts that the recall of goods, which leads to purchase decisions, is based on: i) differential exposure to ads/cues **e**, ii) past experiences **f**, and iii) similarity between ads/cues and brands.

3.2 Cues and Consumer Demand

Given quasilinear and separable utility, the consumer trades off her predicted utility from g with its numeraire cost. Critically, memory affects predicted utility. A consumer who recalls g simulates its pleasure q_g and buys g if and only if $q_g + \varepsilon_g - \theta \cdot p_g > 0$, where θ transforms dollars into utils. Absent recall, predicted marginal utility is lower: the consumer buys g if and only if $\varepsilon_g - \theta \cdot p_g > 0$. The shock ε_g is i.i.d, mean zero, and uniform.[4] It captures a transient taste shock orthogonal to q_g (e.g. feeling thirsty) or a "choice utility" shock (e.g., Coke is in a convenient supermarket location).

With exposure **e** and database **f**, the consumer recalls consumption of g with the probability given in (2) and chooses whether to buy taking the ε_g shock into account. Over the time unit, total demand for g is an affine, increasing function of:

[4] The assumption of uniformity is made for simplicity, but would be straightforward to generalize.

$$d_g(\mathbf{e},\mathbf{f}) = q_g \cdot r_g(\mathbf{e},\mathbf{f}) - \theta \cdot p_g. \quad (3)$$

In the rational model, recall is full, $r_g(\mathbf{e},\mathbf{f}) = 1$, so demand is marginal utility minus the opportunity cost of money.[5] In this neoclassical benchmark the consumer has a stable demand for each good g that increases in her taste q_g for it and decreases in its price p_g. Consumers have habits, some buy more Coke than others, but these are stable because they are due to stable tastes.

Memory affects demand through recall $r_g(\mathbf{e},\mathbf{f})$ in two ways. First, it creates habits due to mere past consumption: higher past frequency f_g causes good g to be more top of mind, boosting its demand, keeping tastes fixed. Our main tests do not distinguish this source of habits from the taste-based one. More importantly, habits are context dependent: seeing more Coke ads increases the consumer's retrieval of past pleasure from Coke, boosting her current demand, and the more so for people with stronger habits.[6] Variation in uninformative context $\mathbf{e}$ can help understand whether, unlike in neoclassical models, habits are context dependent due to memory.

Proposition 1. *(Habits versus Instability) There is a threshold share of neutral experiences* $\overline{f_n} \in (0,1)$ *such that, for* $f_n > \overline{f_n}$ *the consumer's demand for good* g*:*

a) increases in its past consumption frequency f_g *or taste* q_g*,* $\partial d_g(\mathbf{e},\mathbf{f})/\partial f_g, \partial d_g(\mathbf{e},\mathbf{f})/\partial q_g > 0$.

b) increases with exposure e_g *to its ads,* $\partial d_g(\mathbf{e},\mathbf{f})/\partial e_g > 0$,

[5] Equation (3) can be derived under separable quadratic utility across goods, so that y units of g give utility $u_g(y) = \theta^{-1} \cdot \left[\left(q_g \cdot \mathbb{I}_g + \varepsilon_g\right)y - y^2/2\right]$ where $\mathbb{I}_g$ is a dummy equal to one if and only if past consumption is recalled. In this case, we recover (3) assuming $\varepsilon_g > \theta \cdot p_g$, so consumption is positive. We could alternatively assume that a person only considers g if she recalls it, as in consideration set models (Masatlioglu et al. 2012, Manzini and Mariotti 2014). In this case, she chooses by comparing q_g (and possibly a utility shock) with the price p_g. These formulations give similar results. They differ only in the interaction between ads and price elasticity, from which we abstract here.

[6] Recalling a brand upon seeing an ad on TV promotes consumption in two ways. It renders this brand more available when the consumer next visits the supermarket, consistent with the slow decay of cued thoughts and with the rehearsal effects documented in memory research (Kahana 2012). Alternatively, upon seeing a Coke ad, the consumer feels an urge to drink a can from the fridge, which is later replaced.

c) increases more with exposure e_g for more frequent or high taste consumers, but the effect of frequency is concave, $\partial^2 d_g(\mathbf{e},\mathbf{f})/\partial e_g \partial f_g, \partial^2 d_g(\mathbf{e},\mathbf{f})/\partial e_g \partial q_g > 0$ and $\partial^3 d_g(\mathbf{e},\mathbf{f})/\partial e_g \partial f_g^2 < 0$.

Memory is a force toward habits (point *a*). Frequent past consumers of g recall the pleasure of g more easily today, and those with higher taste q_g recall more pleasurable experiences. Both effects increase demand of g for any cue. Through the frequency channel, mnemonic habits arise also due to idiosyncratic events such as being born in a city where a brand is prevalent. Dynamically, the habit fades when moving to a new city, consistent with Bronnenberg et al. (2012). We micro-found such state-dependence with memory.

Memory also yields effects of uninformative ads. In *b)*, increasing a consumer's exposure to an ad for g cues recall of its pleasure, promoting demand for g. Conversely, low exposure entails forgetting, depressing the demand for g, as in Stigler's story. Ads increase the marginal utility of buying the good, but unlike in BM (1993) consuming a good does not encourage consumption of its ads (which yield no utility).[7]

Critically, memory implies a habit-ads interaction: the role of habits is particularly strong in contexts that cue the good. The effect of ads should thus be stronger for more habitual past consumers, either because they recall higher pleasure (higher q_g) or because they have more of the pleasurable experiences to recall (higher f_g).[8] Instability based on irrelevant advertising context is at odds with models of habits such as BM (1988). Note also that in our model the marginal effect

[7] Laibson (2001) develops a model of state-dependent preferences where a cue can increase the marginal utility of a good, through learned associations between the two. This model does not generate habits, nor an interaction between habits and cues driven by recall and similarity.

[8] If $f_n > \overline{f_n}$ is violated, a cue $c \neq g$ may foster retrieval of g itself compared to $c' = g$ by reducing interference from a category $y \neq g$ for which $S(g,y) \ll S(c,y)$. By Assumption 1, this cannot happen for $y = n$ because $c = g$ maximizes similarity of g and dissimilarity from n. A large f_n then implies that each good g "mostly" competes with neutral activities n, so that recall of any g increases in similarity with the cue c, and thus is maximized at $c = g$.

of the ad increases with frequency f_g in a concave way, and it can eventually diminish for the most frequent past buyers, who have the good already "top of mind" (so cues are unnecessary for them). Concavity does not hold with respect to tastes: higher recalled consumption pleasure q_g uniformly boosts predicted future pleasure and hence consumption.

A third key set of new predictions arises from similarity-based spillovers. The first kind of spillovers operates within a product k. Here the two competing brands b (e.g. Pepsi and Coke) are similar to each other compared to other products (e.g. beers or detergents), leading to *positive* spillovers.

Proposition 2. *(Spillovers within k). Suppose that similarity within a product is higher than across products,* $S(k1, k2) > S(kb, k'b')$, $k' \notin k$. *If* $f_n > \overline{f_n}$, *and* $e_n > e^*$, *higher exposure to* $g' = kb'$:

d) increases demand for the competing good $g = kb$, $b \neq b'$, $\partial d_g(\mathbf{e}, \mathbf{f})/\partial e_{g'} > 0$,

e) the effect increases in its high past frequency f_g *or taste* q_g, *but the effect of higher frequency is concave,* $\partial^2 d_g(\mathbf{e}, \mathbf{f})/\partial e_{g'}\partial f_g$, $\partial^2 d_g(\mathbf{e}, \mathbf{f})/\partial e_{g'}\partial q_g > 0$ *and* $\partial^3 d_g(\mathbf{e}, \mathbf{f})/\partial e_{g'}\partial^2 f_g < 0$,

f) increases demand for product k as a whole, $\partial\left[d_g(\mathbf{e}, \mathbf{f}) + d_{g'}(\mathbf{e}, \mathbf{f})\right]/\partial e_{g'} > 0$, *more so for more frequent buyers of good g or g′,* $\partial^2\left[d_g(\mathbf{e}, \mathbf{f}) + d_{g'}(\mathbf{e}, \mathbf{f})\right]/\partial e_{g'}\partial f_x > 0$, $x = g, g'$.

Associative memory creates a positive spillover within a product k from ads of brand b' to the demand for its competing brand b (point *d*). This spillover comes from similarity: seeing an ad for Pepsi cues experiences of consuming the similar Coke, boosting demand for Coke. Shapiro (2018) documents positive ad spillovers between antidepressants and argues that it may be due to a client-doctor interaction (a prescription is needed). Associative recall creates these effects more generally, consistent with Sahni (2016) who finds positive spillovers for restaurants.

Critically, spillovers should be stronger for more habitual consumers of g, see point e). These consumers either associate the product category more strongly with g, due to high frequency f_g, or they retrieve higher pleasure due to higher taste q_g. As a result, cuing the competing brand boosts their predicted utility by more. A Pepsi ad is a stronger cue of Coke for a regular Coke buyer than for a sporadic one. Again, this happens when g suffer from significant interference from neutral experiences, and the effect of higher past frequency diminishes and eventually reverses. As we argue in Section 6, spillovers have important implications: they reduce consumer mobility between brands, affecting competition.

Point *f*) shows another key implication of positive spillovers. While the ad of brand g' does not necessarily shift the consumer toward g' itself, it does increase overall consumer demand for product k by making the pleasure from this product more salient in consumers' mind. This effect is stronger for more habitual buyers of the product. Rather than "stealing business", ads for a brand make the product category more popular. As we argue in Section 6, this effect can explain why negative advertising is uncommon for brands competing in the same product category.

Habit-dependent spillovers arise because memory causes different consumers to represent the same ad differently: a Coke ad cues Coke for habitual Coke drinkers, but it cues Pepsi for habitual Pepsi drinkers. Such spillovers also work across product categories: seeing ads for beer may cue habitual cola drinkers to recall and demand Coke or Pepsi.

Spillovers can occur also across product categories but, crucially, in this case spillovers can be *negative*, depending on similarity. Let $S(k,-k)$ be the average similarity between the brands in product k and the brands in another product $-k$, where the average is taken using the database **f**.

Proposition 3. *(Similarity versus Interference across Products). For a product k there is a threshold S_k such that higher exposure e_{-k} to ads of goods in another product $-k \neq k$:*

g) boosts demand for k if similarity is high enough and dampens it otherwise, $\partial[d_{k1}(\mathbf{e},\mathbf{f}) + d_{k2}(\mathbf{e},\mathbf{f})]/\partial e_{-k} > 0$ if and only if $S(k,-k) > S_k$. The positive effect is stronger for more similar products $\partial^2[d_{k1}(\mathbf{e},\mathbf{f}) + d_{k2}(\mathbf{e},\mathbf{f})]/\partial e_{-k}\partial S(k,-k) > 0$.

h) the effect is stronger for more frequent k consumers, $\partial^2|d_{k1}(\mathbf{e},\mathbf{f}) + d_{k2}(\mathbf{e},\mathbf{f})|/\partial e_{-k}\partial f_k > 0$.

Spillovers from the ads for product $-k$ on the demand for product k reflect a tradeoff between similarity and interference. On the one hand, $-k$ cues k if the two are similar, boosting the demand for k. On the other hand, $-k$ interferes with recall of k if the two are dissimilar, reducing the demand for k. When two products are very different, thinking about one of them distracts the consumer from thinking about the other. Forgetting here generates a "business stealing" effect of advertising. As we show in Section 6, this sheds light on when goods are perceived as substitutes or as complements, with implications for the kinds of products for which habits and advertising effects are expected to be strong.

Our model shows that spillovers between categories of products are more "diluted" than the within-category spillovers of Proposition 2. The two brands in a category are typically highly similar, so cuing one also cues the other. Cross-product spillovers spread over a range of categories (e.g. colas may cue lemon colas and beers and interfere with dissimilar fruit, bread, detergents, etc.). This principle guides our empirical analysis: detecting cross-category spillovers is easier when product categories are aggregated. We perform this test in Section 5, using our quantitative proxy for similarity.

Similarity based spillovers and their interaction with habits offer key tests of our memory mechanism. They also shape welfare effect of ads and their economic implications.

3. Data

We use the NielsenIQ Homescan household panel to measure quantities bought and prices paid by more than 60,000 U.S. households.[9] We measure advertising and viewership using Nielsen Ad Intel. To study brand loyalty and advertising in markets with low intrinsic differentiation, we focus on purchases of the top two brands, in terms of market share, in 20 product categories which Bronnenberg et al. (2012) identify as having low differentiation and high advertising intensity.[10]

Table 1
Product categories, brands' purchase share and ad exposure

Product	Market leader	Mean ad exposure	Market follower	Mean ad exposure	Aggregate market share
Beer	Bud	25.11	Miller	12.53	0.64
Candy chocolate	M&M	12.79	Reese	29.30	0.52
Cereal	Cheerios	29.62	Post	8.82	0.54
Coffee	Folgers	4.72	Maxwell	2.81	0.5
Colas	Coke	32.88	Pepsi	22.56	0.52
Cheese Crackers	Pepperidge	5.67	Sunshine	6.82	0.65
Detergent heavy duty	Tide	17.19	Purex	1.78	0.57
Diapers	Pampers	32.40	Huggies	15.44	0.53
Dog food dry	Purina	26.74	Iams	3.32	0.52
Facial tissue	Kleenex	4.21	Puffs	5.01	0.63
Frozen Pizza	DiGiorno	9.13	Red Baron	3.85	0.51
Lemoncolas	Sprite	4.54	Sevenup	3.06	0.66
Oral rinse / antiseptic	Listerine	17.95	Crest	5.83	0.74
Other carbonated soft drink	Mountain Dew	13.82	Dr Pepper	11.24	0.53

[9]Bronnenberg et al. (2012) and Einav, Leibtag, Nevo (2010) offer detailed analyses of NielsenIQ Homescan data.
[10] A brand includes all possible variants sold by the firm in the same product and all possible packages in which it is sold. For example, "Bud" includes both regular and low calories beer and purchases ranging from a single can to the 24-bottles case. We exclude products for which one of the two brands' average weekly exposure is below 1 second.

Pain remedy headhache	Aleve	33.02	Tylenol	7.30	0.52
Shampoo	Suave	5.06	Pantene	7.50	0.53
Soap specialty	Suave	5.06	Dove	12.11	0.52
Soup canned	Campbell's	11.82	Progresso	6.83	0.8
Toilet tissue	Charmin	22.84	Angel Soft	8.18	0.54
Yogurt	Dannon	15.18	Yoplait	11.41	0.62

Notes: the table reports the 20 product markets in our sample, together with the two leading brands in terms of market share, the mean expected advertising exposure to each brand in seconds across the whole population recorded in the Nielsen dataset, and the aggregate market share of the two brands. See the text for the definition of ad exposure. Bronnenberg et al. (2012) label a product category as 'ad intense' if it lies in the top 25% of modules by advertising expenditure based on 2008 data. Among these products, we keep only those for which both of the top two brands have average weekly expected exposure above 1 second.

In Equation (3), consumer demand depends on exposure to ads **e**, frequency of past purchases **f**, similarity $S(c, x)$ and tastes **q**. We now construct proxies for these drivers.

Exposure. The Nielsen Ad Intel data measures advertising in designated market areas (DMAs). There are 210 DMAs covering the U.S., mostly capturing metropolitan areas. Advertising 'occurrences' report: duration in seconds (ranging from 10 to 120), advertised brand, channel of broadcast, DMA, date and time.[11] Advertising "Impressions" estimate the number of people from a given age or gender group watching each TV station in each DMA, within each month, day of the week and time, at 5 minutes intervals. From occurrences and impressions, we compute the number of viewers from different age or gender groups of an ad and their exposure to the ad in seconds. We estimate a household's weekly exposure to a brand's ads in four steps. First, for each household member we compute the probability of being exposed to an ad in a given day and DMA by dividing the number of viewers of the ad with the same gender and age[12] by the total number of people with those characteristics in the DMA. We then multiply this probability by the seconds of the ad in the day, giving expected daily exposure. Third, we compute daily exposure of

[11] Four different TV media types are covered in the data: Cable, Network, Syndicated, and Spot. We extract Cable TV occurrences from the national data, and both Spot TV occurrences and Network Clearance Spot TV occurrences from the local data. Network Clearance Spot TV refer to the local broadcast of national Network TVs. These local broadcasts may override nationally purchased ads. Spot TV channels instead refer to local channels.

the household by averaging the expected exposure of its members. Fourth, we aggregate daily household exposure for the seven days of the week. This *Exposure* index is our main proxy for **e**.

Exposure may in principle capture information about promotions. Our data on ad content shows that this is nearly never the case: the word "price" is never included in Nielsen ad descriptions and "sale" is present in only 0.05% of the ads. To be conservative, we nevertheless exclude giveaways from our dataset and perform robustness tests by controlling for promotions. Another concern is that the supply of ads, and hence *Exposure*, may be endogenous to demand. We tackle this issue by extensively controlling for predictable demand drivers such as seasonality and, among other things, by using in Section 4.3 the Shapiro et al. (2021) method for isolating quasi random variation in the supply of ads based on the institutional setting of advertising markets.

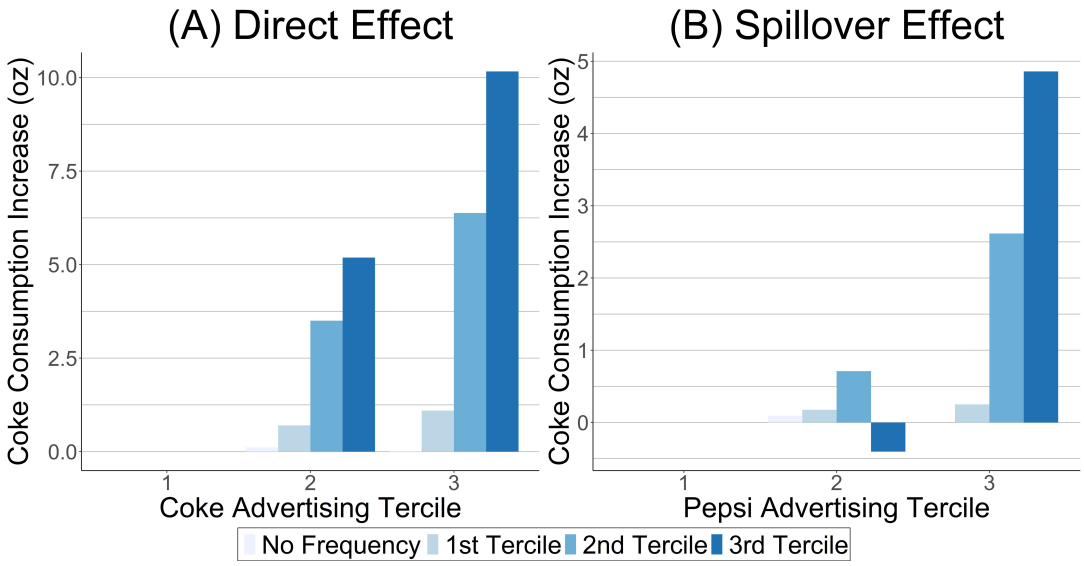
Frequency. We proxy experiences with a brand by the number of shopping trips in the prior calendar year in which the brand was purchased by the household. Thus, when assessing purchases in 2016, *Frequency* is the number of purchases of a brand in 2015. Higher *Frequency* of g captures both more "repetititon" f_g in the database and higher stable tastes q_g that led to more purchases. Both forces increase the effect of ads in our model but only f_g produces concavity. In our main analysis we interpret *Frequency* as a proxy for **f** and **q**. In Section 4.3 we try to disentangle these channels using our panel dataset. We prefer *Frequency* to past consumption as a measure of experiences because the latter may occur at special events such as parties that may not be retrieved in a standard purchase context, but our key results are robust to using past consumption. We focus on purchases in the previous year because, being more recent, they are easier to recall.

Similarity. We measure similarity using the descriptions of the ads of goods g and g' in our dataset, as detailed in Section 5. Measuring similarity from ad content follows our model of

ads as memory cues, and the results align with intuition. For instance, it supports the assumption of Proposition 2 that the similarity between competitor brands in the same product category (e.g. Coke and Pepsi) is higher than similarity across products, such as colas and beers.

For intuition, we use our *Frequency* and *Exposure* proxies to illustrate the ads-habits interaction for Coke and Pepsi. These brands are well known and famously undifferentiated (many consumers fail to distinguish them in blind taste tests, Mcclure et. al 2004). How do their ads work? Figure 1 reports on the y-axis the raw change in a household's weekly Coke consumption when its *Exposure* to Coke or Pepsi ads is in the second and third tercile compared to the first, split across different terciles of the household's purchase *Frequency* of Coke. The patterns are consistent with our model. In Panel A, higher exposure to Coke ads boosts consumption, and more so for more frequent Coke buyers. In Panel B, Pepsi ads spill over to Coke consumption, and more so for more frequent Coke buyers. We next systematically test our theory for all sample products.

Figure 1: The role of ads for sporadic and habitual Coke consumers.



Notes: The bars display the mean weekly change in consumption of Coke in different terciles of *Exposure* to Coke (left panel) or Pepsi (right panel ads), compared to in the first *Exposure* tercile, grouped by different levels of Coke *Frequency*. For each group, the average consumption of Coke at the lowest tercile is normalized to 0. The Frequency

groups include the observations for households in the three terciles of Coke, among those that bought Coke at least once in the previous year. We include also a column for those who did not purchase Coke in the previous year.

4. Habits, advertising, and spillovers within a product category

Section 4.1 presents tests of Propositions 1 and 2 on the effects of habits, ads, their interaction, as well as tests of similarity-based spillovers within a product category. Section 4.2 shows the results of our tests across our 20 products. Section 4.3 presents robustness tests. Our key tests using measured similarity are in Section 5.

4.1 The Regression Specification

To assess Proposition 1 and within-product spillovers in Proposition 2, the dependent variable is a household j's demand for a target good $\tau = kb$ in week t, $d_{j,\tau,t}$. We measure demand from January 4th 2016, when data on advertising becomes sizable, to December 22nd 2019, before the Covid pandemic. Weekly windows are a plausible timescale for ads to affect demand, but we perform extensive robustness.[13] We approximate Equation (3) by specifying $d_{j,\tau,t}$ as a linear function of its *own* past purchase *Frequency* of the same brand $f_{j,\tau,t}$, its *Exposure* to the ad of the same brand in the week, $e_{j,\tau,t}$, the interaction between the two. We also include the purchase frequency $f_{j,-\tau,t}$ of the competing brand $-\tau = kb'$, exposure $e_{j,-\tau,t}$ to its ads, and their interaction.

$$d_{j,\tau,t} = \sum_{s\in\{\tau,-\tau\}} \alpha_s \cdot f_{j,s,t} + \beta_s e_{j,s,t} + \sum_{s,s'\in\{\tau,-\tau\}} \gamma_{s,s'} \cdot e_{j,s,t} \cdot f_{j,s',t} + \delta_{s,s'} \cdot e_{j,s,t} \cdot f^2_{j,s',t} + X'_{j,\tau,-\tau,t} \cdot \xi + \varepsilon_{j,\tau,t}, \quad (4)$$

[13] Among others, we consider monthly rather than weekly purchases (Table D3), which on average triples the share of observations with positive demand, and additionally confirm that our results hold at the intensive margin of demand, namely whether the consumer purchases good g in a given week (Table E.2).

where $X'_{j,\tau,-\tau,t}$ is a vector of household and good-specific controls capturing standard drivers of demand such as: the average prices of own and other brands in (4) paid in the DMA in the week, the (log of) income of household j, the number of household members, whether the household has children, and whether the head of household is unemployed, has college education or not, is male or female, its marital status, race, age. To compute the average price, we sum the total dollar expenditure of all consumers in a given week in a given DMA for a given brand, and divide it by the total quantity (defined as number of units or in ounces as appropriate) purchased.[14]

The linear specification in Equation (4) fails to capture weeks in which a household does not purchase good τ. To mitigate this concern, we exclude households who dropped out of the sample or bought the products only in some years.[15] We rule out that our results are an artifact of linearity or outliers by looking at the data non-parametrically in Section 4.3.

Coefficients α_τ and $\alpha_{-\tau}$ capture the impact of past frequencies of buying the target and its competing brand, respectively, on future consumption of the target. By Prediction 1.a), past purchase frequency of τ should increase current demand for that brand, so $\alpha_\tau > 0$. Empirically, this coefficient could capture habits based on memory f_τ or higher tastes q_τ.

Coefficients β_τ and $\beta_{-\tau}$ capture the effect of higher weekly exposure to ads of the target and the competing brand, respectively, on weekly target consumption. Critically, ads interact with habits as captured by coefficients $\gamma_{s,s\prime}$ and $\delta_{s,s\prime}$. Stronger exposure to own ads $e_{j,\tau,t}$ should promote demand for the product, but especially so for more frequent buyers, $\gamma_{\tau,\tau} > 0$, and in a concave way, $\delta_{\tau,\tau} < 0$ (Predictions 1.b and 1.c). The habits-ads interaction could arise from higher

[14] Section 3 and Appendix D provide a more detailed analysis on the role of prices and discounts.
[15] As we show in the Appendix, our results go through if we keep any consumer who has purchased one of the leading brands at least once over the previous year, or once in the full database.

Frequency capturing either more experiences, higher f_τ, or more pleasurable ones, higher q_τ. In both cases, the ad memory cue is stronger. The full impact of higher exposure to ads on demand thus combines the baseline effect β_τ with the experience-dependent terms $\gamma_{\tau,\tau}$ and $\delta_{\tau,\tau}$.

Consider spillovers next. Because similarity between competing brands is high, Prediction 2 implies positive spillovers on brand τ from greater exposure to $-\tau$ that are especially strong for regular buyers of g, namely $\gamma_{-\tau,\tau} > 0$, but in a concave way, $\delta_{-\tau,\tau} < 0$. The stronger similarity-based spillovers for more habitual buyers are a key new prediction of memory. We also test the related prediction 2.f) on total demand for category k by regressing the household's weekly total purchased quantity of a product's brands, $\left(d_{j,\tau,t} + d_{j,-\tau,t}\right)$, on its past purchase frequency of the two $f_{j,\tau,t}$ and $f_{j,-\tau,t}$, on its weekly exposure to their ads $e_{j,\tau,t}$ and $e_{j,-\tau,t}$, and on their interactions. Memory then implies that category consumption should increase in past purchase frequency and ad exposure, especially so for frequent past buyers, again in a concave way.

The supply of ads may respond to the demand for brands, e.g. due to seasonality or geography (soft drinks may face higher demand, but also more ads, in the summer or in the South etc.). To capture these factors, our baseline specification controls for month of the year, year, and state dummies, and their interaction with product dummies (and for product dummies on their own). In Section 4.3 we perform extensive robustness checks using more granular dummies. We start by replicating Shapiro et al.'s (2021) strategy to isolate quasi random variation in ad exposure, but then expand their analysis by controlling for even more granular variation. We also address the related concern that ads target consumers who are more prone to purchasing, even controlling for geography and seasonality. Our results are robust to these and other concerns suggesting that endogeneity is unlikely be the source of our results.

The variables used for Equation (4), namely demand, frequency of past purchase, and ad exposure, are summarized in Table A.1. Tables A.2, A.3 and A.4 report descriptive statistics and correlations between variables. Across product markets, average exposure to ads ranges from 2.6 to 35.2 seconds per week, with an average of 16.1 seconds for the market leader. The (unconditional) average frequency of purchases in a year ranges from 0.82 to 8.08 across products, with an average of 3.98 for the market leader. Weekly variation in demand is large: the average coefficient of variation (standard deviation of weekly demand divided by its mean) is 6.

4.2 Ads, Experiences, and Spillovers in the 20 Product Markets

We estimate Equation (4) by pooling all brands. To do so, we standardize Total Purchases, *Frequency* and *Exposure* at the product level, and cluster standard errors at the product-week-year-DMA level, which captures potential sources of correlation. Table 2 reports the coefficients of theoretical interest (omitted coefficients on prices or household income have qualitative and quantitatively reasonable values). Column (1) tests for Predictions 1 and 2. Column (2) tests for the effect of experiences and ads on *total* category consumption (in column 2 subscript 1 refers to the market leader, 2 to the follower).

Table 2: Predictions 1 and 2, pooled product categories

	(1)		(2)
VARIABLES	d_τ	VARIABLES	$d_1 + d_2$
f_τ	0.223***	f_1	0.153***
	(0.00175)		(0.00131)
$f_{-\tau}$	-0.0174***	f_2	0.126***
	(0.000302)		(0.00148)
e_τ	0.00757***	e_1	0.00309***
	(0.000586)		(0.000644)
$e_{-\tau}$	-0.00159**	e_2	0.00417***
	(0.000640)		(0.000693)
$e_\tau \cdot f_\tau$	0.0113***	$e_1 \cdot f_1$	0.00618***

	(0.00129)		(0.00152)
$e_{-\tau} \cdot f_{\tau}$	0.00999***	$e_2 \cdot f_1$	0.00548***
	(0.00133)		(0.00148)
$e_{\tau} \cdot f_{-\tau}$	-0.000598*	$e_1 \cdot f_2$	0.00299**
	(0.000309)		(0.00141)
$e_{-\tau} \cdot f_{-\tau}$	-0.00115***	$e_2 \cdot f_2$	0.00667***
	(0.000323)		(0.00142)
$e_{\tau} \cdot f_{\tau}^2$	-0.000567***	$e_1 \cdot f_1^2$	-0.000379**
	(0.000140)		(0.000158)
$e_{-\tau} \cdot f_{\tau}^2$	-0.000933***	$e_2 \cdot f_1^2$	-0.000741***
	(0.000139)		(0.000137)
$e_{\tau} \cdot f_{-\tau}^2$	-4.41e-05	$e_1 \cdot f_2^2$	0.000236**
	(4.07e-05)		(0.000120)
$e_{-\tau} \cdot f_{-\tau}^2$	0.000282***	$e_2 \cdot f_2^2$	9.63e-05
	(4.72e-05)		(0.000126)
Observations	72,116,888	72,116,888	36,058,444
R-squared	0.054	0.054	0.044

Notes: the sample for Table 2 is a vertical merge of the 40 samples for each brand within our analysis. For each brand consumption observation, "τ" variables contain the information for the brand associated with the same observation, "$-\tau$" variables refer to the competing brand in the product, for the same observation. Total Purchases, *Frequency* and *Exposure* are standardized at the product level. Total consumption in Column (2) is computed after standardization. In Column (2), for each product, the variables associated with brand 1 (e.g. e_1 and f_1) refer to the market leader. Regressions include the usual controls, including month of the year, state, year and product dummies, and interaction of product dummies with the previous dummies. Column (1) estimates Equation (4), Column (2) estimates Equation (5) for product demand. Superscripts: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the product-week-year-DMA level.

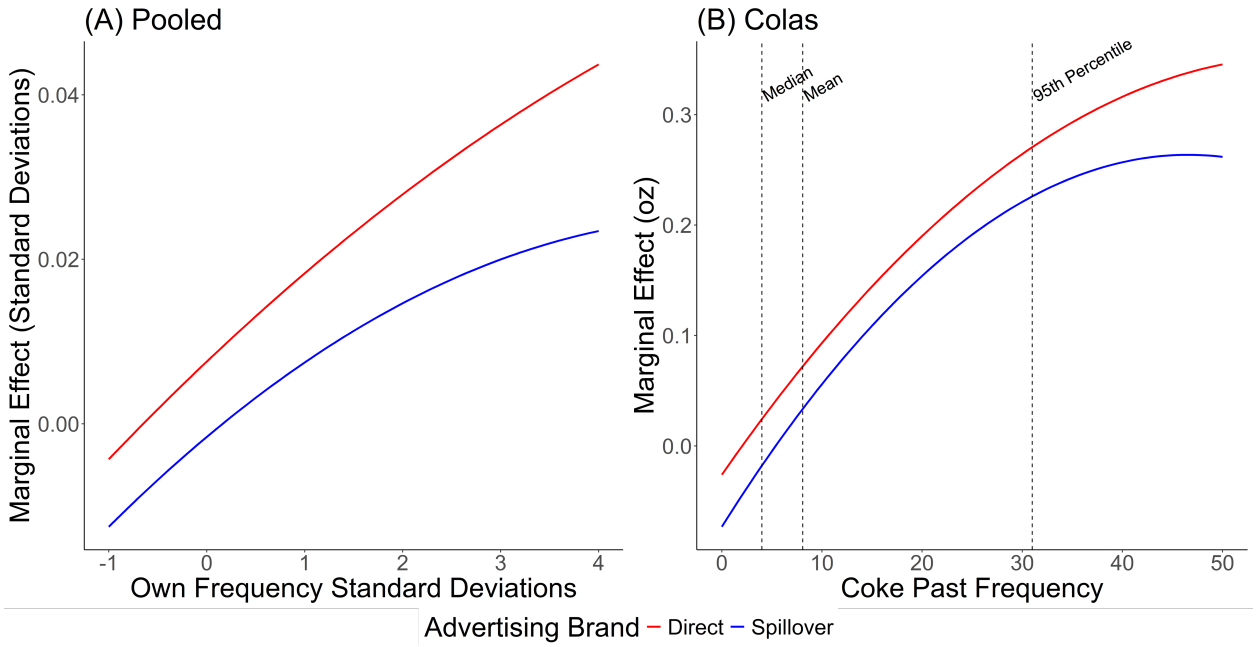
Across the 20 product categories, Predictions 1 and 2 are robustly confirmed. The large positive coefficient on f_{τ} shows that habitual past consumers exhibit sharply higher consumption today. The positive coefficient on e_{τ} confirms that above average exposure to the ads of a brand raises its consumption, but the effect is an order of magnitude smaller than for habits. Critically, the positive coefficient on $e_{\tau} \cdot f_{\tau}$ confirms that, consistent with memory, the effect of ads is stronger for more frequent buyers, but in a concave way (negative coefficient on $e_{\tau} \cdot f_{\tau}^2$).

The estimates also confirm similarity-based spillovers, which are strong and work through habits. The positive coefficient on $e_{-\tau}$ confirms that exposure to the ads of the *other* brand boosts consumption of the target brand, and the positive coefficient on $e_{-\tau} \cdot f_{\tau}$ confirms that spillovers

are especially strong for frequent past buyers of the target brand. Also consistent with Prediction 2, in Column (2) ads boost total product demand.[16]

Figure 2 illustrates the estimated direct and spillover effect of ads, for all products in Panel A, and for Colas in Panel B (the latter is computed using the estimates of Equation (4) for colas only, see Table B.5 in the Appendix).

Figure 2: Full marginal effect of advertising on pooled sample and colas.



Notes: Panel A shows the marginal effect (in standard deviations) of *Exposure* to τ and $-\tau$ for different levels of τ *Frequency*. We estimate the effect of ads by adding the baseline standardized *Exposure* coefficient with its linear and second order interaction with standardized τ *Frequency*, keeping standardized *Frequency* of $-\tau$ at 0. For the τ advertising line (in red) in panel A, the plotted equation is: $me_\tau = \hat{\beta}_1 + \hat{\gamma}_1 * f_\tau + \hat{\delta}_1 * f_\tau^2$ and similarly for the $-\tau$ advertising line (in blue). The estimates are taken from Column (1) of Table 2. Panel B displays the marginal effect (measured in ounces) of *Exposure* to Coke and Pepsi for different levels of Coke *Frequency*. The marginal effect of ads sums the baseline advertising exposure coefficient with its linear and second order interaction with Coke frequency, keeping Pepsi frequency at 0. For the Coke advertising line (in red) in panel A, the plotted equation is $me_{Coke} = \hat{\beta}_1 + \hat{\gamma}_1 * f_{coke} + \hat{\delta}_1 * f_{coke}^2$, and similarly for the Pepsi advertising line (in blue). The estimates are taken from column (2) of Panel A of Table B.5

[16] We also estimate Equation (4) in the 20 separate product markets. The role of experiences, advertising cues, and their interaction including positive and frequency-based spillovers arise in most of these markets (see Table C.3 in Appendix C). For 80% of brands ads are stronger for more frequent past buyers, for 63% of brands the positive spillovers are stronger for more frequent buyers. Most of these coefficients are statistically significant at conventional levels. Variation in the strength of these effects is partly due to similarity, as we show in Section 5.

In Panel A, a one standard deviation higher exposure to the ads for a brand is associated with 0.008 standard deviations increase in the demand for the same brand for the average household, and the effect increases to 0.018 standard deviations when moving to a one standard deviation higher frequency f_τ household. For the spillovers, a one standard deviation greater exposure to the $-\tau$ brand decreases demand of τ brand by 0.002 standard deviations. This effect turns to a positive spillover of 0.007 standard deviation for a one standard deviation higher frequency f_τ household. Thus, spillovers are sizable but direct advertising effects are larger, which preserves a brand's positive incentive to advertise: the difference between the direct and spillover effect at the average frequency is of 0.009 standard deviations, and it increases to 0.011 at the one standard deviation higher frequency f_τ household.

Panel B shows the quantitative effects for Colas. At the average Coke frequency ($\approx$ 8.08 yearly trips), each extra second of exposure to Coke ads increases weekly demand for Coke by roughly 0.07 ounces. The average Coke consumer needs roughly 166 extra seconds of ads in a week to increase her demand for Coke by an extra can (12 oz). On the other hand, a 95th Coke frequency percentile household, who purchases Coke 31 times in a year, needs only 44 seconds of Coke ads in a week to buy an extra can.

Spillovers in the Colas market are also large. It takes 362 seconds of Pepsi ads to induce a 12oz shift toward Coke for the average Coke consumer, but only 53 seconds for the 95th percentile Coke consumer. Looking at the difference between direct and spillover effects, the average Coke consumer needs 190 more seconds of Pepsi ads vs Coke ads exposure to consume one extra can of Coke, while at the 95th percentile of frequency this difference goes down to 8.75 seconds. The effects of ads for Pepsi consumption are comparable and again confirm that spillovers are sizable but direct effects are stronger.

4.3 Robustness Tests

We discuss robustness for the linear functional form assumption on demand, Equation (4), for the possibility that ads may co-occur with price reductions or promotions, and for the endogeneity of the supply of ads, i.e. ad targeting. We also propose a test to separate the role of higher frequency f_g from higher tastes q_g in the habits-ad interaction. Appendices C and D present several other robustness tests.[17]

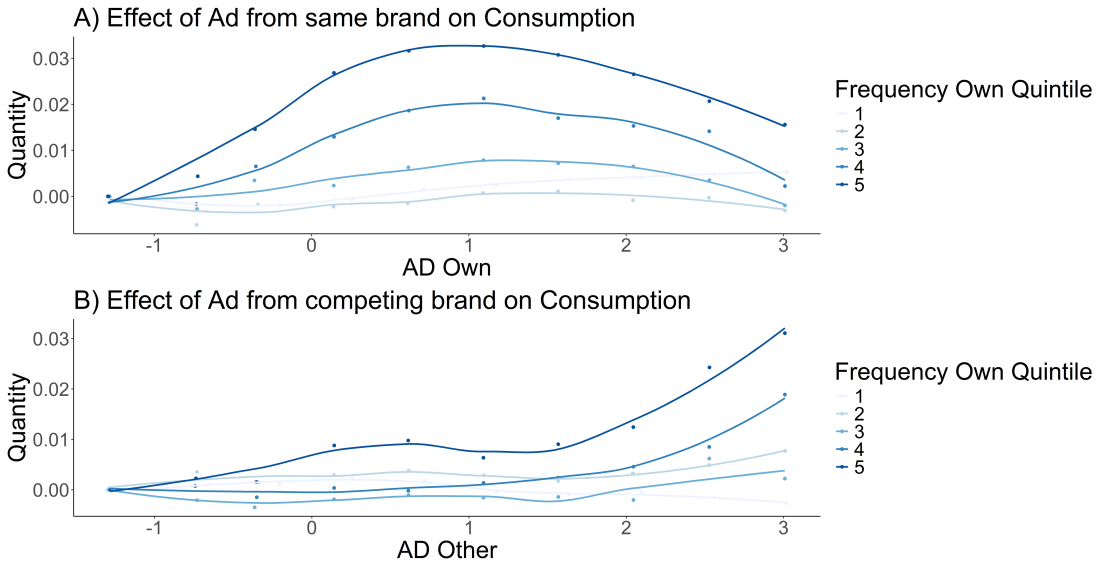
Functional Form Assumptions. Our linear regressions do not capture households who consume zero units of a good (e.g. they may dislike or be allergic to it), which could mechanically affect the estimated habits-ads interaction. Linear coefficients may also be sensitive to outliers. Figure 1 assuages these concerns for Coke: the habit-ads interaction is not created by particular frequency bins (the same is true for Pepsi, see Figure B.2). To address these issues for our 20 products, we use a binned-scatterplot following the algorithm in Cattaneo et al. (2024). We first residualize the quantity purchased by a household on the regressors in Table 2 except for purchase frequencies, ad exposures, and their interaction. We then split the data into quintiles based on frequency of purchase in the previous year. For each frequency quintile, we further split the data into bins for exposure to advertising.

Figure 3 reports a household's exposure to the ad of a generic brand on the x-axis and the mean of the residualized quantity on the y-axis, for each bin of exposure and each frequency quintile (of the same brand in Panel A, of the competing brand in Panel B). Panel A thus captures

[17] Appendix C considers a different measure of ad exposure (the total weekly count of ad occurrences for each brand in each DMA) and a different sample of households (those who purchase a product at least once in the previous year). Appendix D considers whether ads are correlated with, and possibly highlight, price discounts and whether ads prompt consumers to upfront their purchase but have no effect on overall consumption. Our results are robust in these tests.

the direct effect of an ad, and Panel B its spillover. To ease comparability across advertising changes, we normalize demand in each frequency bucket by the demand in the same bucket when advertising is in the lowest bin. The results confirm the habit-ads interaction: the quantity purchased increases with ads, directly and through spillovers, but more so for more habitual consumers, as shown by the increasing curves at higher frequency buckets.

Figure 3: Residualized Consumption as a function of Exposure and Memory



Notes: The figure reports the value of residualized quantity purchased of a generic brand as a function of past frequency consumption of the same brand and as a function of exposure to advertising of the same brand (Panel A) or the competing brand (Panel B). For each frequency quintile, we subtract the value of consumption corresponding to the lowest bin of exposure. We adopt the method in Cattaneo et al. (2024) to optimally define advertising bins and to residualize quantity with respect to regressors in Table 2, excepts for frequencies, and ad which are being plotted, and their interaction.

This and other tests confirm that Predictions 1 and 2 hold robustly are not an artifact of functional form assumptions or outliers.[18]

[18] Concavity is not evident in Figure 4, but it is captured by an alternative non-parametric exercise in which we regress quantity on ads exposure for different frequency percentiles. To further assess the robustness check of the role of "zeroes", we separately analyze the extensive and intensive margins of buying (i.e. whether the brand is purchased, and quantity conditional on a purchase, Tables E.1 and E.2). We perform tests using different samples (i.e. including consumers who purchased a product in the previous year, or at any point in the past, Tables C3 and C4).

Price Reductions/Discounts. Ads may boost demand because they occur with, and possibly highlight, price discounts. Appendix D.1 shows that ads for our brands do not, as a rule, mention prices or discounts, making this unlikely. However, to explicitly control for this possibility, we interact prices with frequency of purchase of each product in our baseline regression. If purchases from frequent consumers were moved mostly by price cuts, adding these controls would cause the coefficients on exposure-frequency interactions to fall. Table D.2 shows that this is not the case: our estimates of interest barely change in this specification. Interestingly, the effect of ads is comparable to price reductions. A one standard deviation reduction in own price leads to a 0.013 standard deviations increase in own consumption for the average consumer. This compares to a 0.005 increase from a one standard deviation increase in own ads. At a 1 standard deviation higher frequency of consumption, these effects move to 0.024 for price and 0.015 for ads.

Finally, we control for the interactions between frequency and discounts, defined as the share of products of a brand whose price in a week has dropped by more than 5% relative to the long-term average. Table D.3 shows that discounts boost demand for the brand after controlling for price, consistent with nonlinearities, but again our coefficients of interest are stable.[19]

Ad targeting. The supply of ads is endogenous, as firms may want to increase advertising at times when consumer demand is high (e.g. advertising ice cream more during the summer), or to consumers with higher propensity to purchase. Firms could plan advertising expenditures ex-ante, to target specific times or households, but they may also try to target ex post, upon unpredictable shocks to demand. It is however difficult to systematically engage in fine ex-post targeting, due to the institutional setting of advertising markets, a feature exploited by Shapiro et al (2021) to isolate

[19] Nonlinearity is consistent with discounts drawing attention to the discounted product or the category as a whole (an effect we revisit in Section 6). In our model this could also lead to spillovers. Consistent with this, Figure D.2 shows that a discount for brand $-\tau$ exerts a positive spillover on demand for brand τ for frequent consumers of the latter.

quasi-random variation in ad exposure. Based on the institutional structure of advertising markets, they control for the main predictors of ex-ante ad targeting: week of the year dummies (which track highly granular seasonality), month-year dummies (which control for granular changes in demand over time), and state dummies. Here we follow a similar strategy but we sequentially add finer controls to assess the stability of the estimated interaction between habits and ads.

We next show the robustness of our results to controls that capture consumer targeting at different levels of granularity. If our estimates remain stable as we add these controls, we can be confident that they are not distorted by the ad-targeting these controls capture. It also suggests that our results are robust to finer ex-post or household specific targeting, which is harder to implement and likely entails smaller variations in tastes than what we control for.

Table 3 reports the estimated context dependence of habits, as captured by the coefficients on ad *Exposure* and its interaction with *Frequency* from Equation (4), for both direct and spillovers effects (see Table C.9 in the Appendix for full results). Column (1) reports our baseline estimates, which control for month of the year, state, year and product dummies, and interaction of product dummies with the previous dummies. Column (2) implements the main identification strategy in Shapiro et al. (2021), by including both week of the year dummies (which track highly granular seasonality), month-year dummies (which control for trends), and replacing state with DMA dummies, which controls for finer seasonal and geographical variation. We include product dummies and their interaction with all of the above. Column (3) further saturates the specification by allowing for DMA-specific seasonality: we include the interaction between week of year and DMA dummies, and their triple interaction with the product dummies. Column (4) reports again the baseline specification by switching product dummies with brand dummies. In the remaining two columns we go beyond Shapiro et al (2021) by leveraging household-level information. In

Column (5) we control for a brand's aggregate consumption in the DMA each week, alone and interacted with the household's past purchase frequency. Here we exclude seasonal and geographical dummies because weekly variation of DMA-level consumption of a product is a highly granular proxy for fluctuations in local demand. This is a conservative test: local demand variation is also causally induced by ads themselves, according to our model.

In Column (6) we replace state dummies with household dummies, also interacted with brand dummies, in the baseline specification of Column (4). This test nets away the finest possible cross sectional targeting of ads. Importantly, by using the panel structure of the data, this test also evaluates stable tastes versus past experiences as sources of habits. Empirically separating these drivers, which are conflated in our *Frequency* proxy, is difficult (Heckman 1981). By allowing for highly flexible non-Gaussian heterogeneity and persistent short run fluctuations, Dubé et al. (2010) find that tastes cannot account for substantial positive autocorrelation in brand consumption and show that the data is best explained by state-dependent "brand loyalty", which we micro-found with memory. Here we assess state dependence using the panel structure: controlling for household dummies, *Frequency* captures variation of recent consumption that is independent from tastes. Standard errors are clustered at the year-week-DMA-product level.

Table 3: Robustness tests
Geographical and Time controls, pooled product categories.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	d_τ	d_τ	d_τ	d_τ	d_τ	d_τ
e_τ	0.00757***	0.00564***	0.00543***	0.00196***	0.00252***	0.00304***
	(0.000192)	(0.000210)	(0.000204)	(0.000238)	(0.000183)	(0.000200)
$e_{-\tau}$	-0.00159***	-0.00352***	-0.00373***	0.000158	0.00119***	0.00229***
	(0.000220)	(0.000238)	(0.000229)	(0.000264)	(0.000213)	(0.000216)
$e_\tau \cdot f_\tau$	0.0113***	0.0112***	0.0113***	0.0112***	0.00283***	0.0124***
	(0.000404)	(0.000404)	(0.000406)	(0.000410)	(0.000494)	(0.000365)
$e_{-\tau} \cdot f_\tau$	0.00999***	0.0100***	0.00997***	0.00991***	0.00684***	0.0122***

	(0.000452)	(0.000454)	(0.000454)	(0.000463)	(0.000508)	(0.000393)
Observations	72,116,888	72,116,888	72,116,888	72,116,886	72,116,686	72,116,854
R-squared	0.054	0.055	0.058	0.057	0.068	0.128

Notes: Table 3 estimates the regression coefficients for equation (4) for our pooled dataset adding different geographical and temporal robustness controls. Column (1) represents our baseline regression with month of year, year, state and product dummies, with the first three sets interacted with the last. Column (2) substitutes the month of year with week of year dummies and state dummies with DMA dummies, including month-year dummies too. Column (3) includes further interactions between DMA and week of year, alone and interacted with the product dummies. Column (4) switches from the baseline specification brand from product dummies. Column (5) controls for a brand's aggregate consumption in the DMA each week, alone and interacted with the household's past purchase frequency, excluding the seasonal and geographic dummies. Column (6) builds on column (4), replacing state dummies for household dummies, including their interactions with the brand dummies. Superscripts: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the week-year-DMA-product level.

The results are robust across these tests. The magnitude and statistical significance of the coefficient on the interaction between *Exposure* and *Frequency* are remarkably stable for both direct effects and spillovers. The only exception is the specification in column (5), where direct effects shrink, but is also likely to over-control for changes in local demand. The results in Table 3 confirm both that ads are more effective for habitual buyers of the advertised brand and that they exert positive spillovers on habitual buyers of the competing, non-advertised brand.

The stability of our estimates to finer sources of seasonal, temporal, and geographical variation in demand, and even controlling for household dummies, suggests that our effects arise from the timing of exposure within households, which is consistent with our model. Evidence from column (6) further suggests that stable tastes cannot fully account for the habits-ads interaction. Instead, changes in a household's database are important, suggesting the presence of time varying state dependence, which our model explains through changes in past purchase frequency f_g.[20]

[20] In another version of this test, we proxy tastes by a household's average consumption in the pre-sample period 2012 to 2014. This allows us to see if stable tastes entail a form of state dependence (i.e. whether, in line with the model, they become more predictive of demand when ad exposure is high), if they exhibit spillovers, and if they fail to exhibit concavity. The test confirms these finer model predictions, as shown in Table C.10 and Figure C.2 in the Appendix.

5. Tests of Similarity

Measuring similarity and linking it to demand is key to testing new predictions of memory models. We now do so, revisit spillovers within products, and test Prediction 3 across products.

5.1 Measuring Similarity

Similarity is an input to key cognitive functions such as recall and recognition, and is routinely measured and connected to behavior in psychology (Nosofsky 1992, Bhatia and Aka 2022). Economists are starting to use similarity to test cognitive theories of economic choice (Bordalo, Conlon, et al. 2025, Bordalo, Burro et al 2025). Following this thread, we measure similarity between ads and test the corresponding predictions of our model.

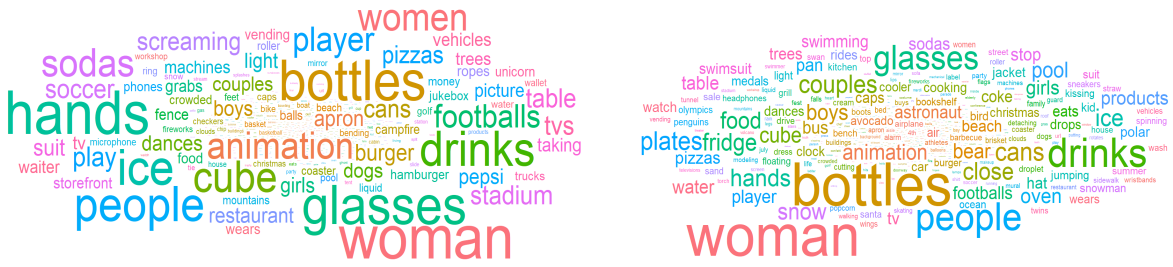
The Ad Intel data reports short (4 to 6 words) descriptions of individual ads. We pool all ads of each brand and then, following Singhal (2001), remove stop words and truncate each word to get its stem. We count, separately for each brand, the number of times each stem is present in any occurrence within our 4 years of interest. This generates i) a joint list of stems for a product category k (across both brands) and ii) word vectors associated with each brand, listing the number of instances $(\#s)_{AD_{kb}}$ of each stem s for each brand b of product k. We compute the cosine similarity score between the brand word vectors as:

$$S(AD_{k1}, AD_{k2}) = \frac{\sum(\#s)_{AD_{k1}} * (\#s)_{AD_{k2}}}{\sqrt{\sum(\#s)^2_{AD_{k1}}} \cdot \sqrt{\sum(\#s)^2_{AD_{k2}}}}$$

If an ad is repeated, each instance is counted separately, so a more frequent ad is more important for similarity. We compute bilateral similarity between the ads of any two products from the 20 we consider. Figure 4 reports the word lists for the 4 products with highest and lowest ad similarity.

Figure 4: Examples of words used in ads for different products and brands.

Panel A. Colas



Panel B. Cheese crackers

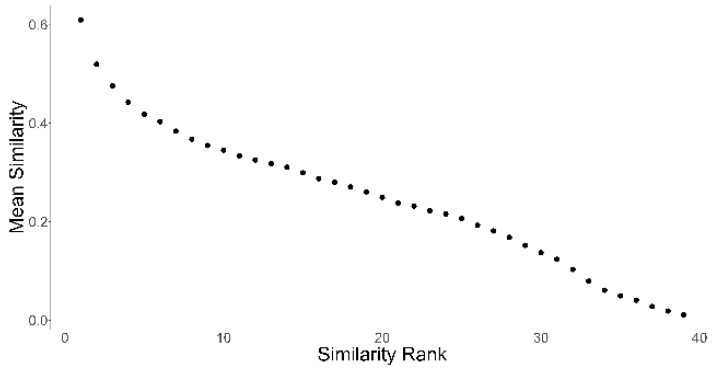




Notes: the two products illustrated are colas (high similarity $S = 0.865$) and cheese crackers (low similarity $S = 0.001$). The more frequently those words are present in each occurrence, the larger they are in the plot.

Figure 5 reports in panel A the average similarity of a brand to its most similar product (rank 1), its second most similar, (rank 2), and so on up to rank 39. There is meaningful variation in measured similarity. Average similarity to a target brand or product declines gradually from 0.6 to 0 as the rank increases.

Figure 5: Similarity across brands



Notes: the figure displays average similarity scores computed across brands from the most similar (rank 1) to the least similar for each observation (rank 39).

Similarity between the ads for brands competing in the same product is on average 0.46, much higher than average similarity across all brands, 0.25. This is consistent with Proposition 2, and hence with positive our evidence on within-product spillovers.

5.2 Similarity and Spillovers Within and Across Product Categories

As in our previous analysis, we study how weekly consumption $d_{j,\tau,t}$ for a "target" brand or category τ is affected by its own ads and frequency but also by the ads and frequency of a "cue" brand or category κ, denoted $e_{j,\kappa,t}$ and $f_{j,\kappa,t}$. Critically, we augment Equation (4) by allowing the spillover effects to be mediated by similarity $S_{\tau\kappa}$ between the ads of the target and the cue:

$$d_{j,\tau,t} = \sum_{s\in\{\tau,\kappa\}} \left[\alpha_s \cdot f_{j,s,t} + \beta_s e_{j,s,t}\right] + \sum_{s,s'\in\{\tau,\kappa\}} \left[\gamma_{s,s'} \cdot e_{j,s,t} \cdot f_{j,s',t} + \delta_{s,s'} \cdot e_{j,s,t} \cdot f^2_{j,s',t}\right]$$

$$+\varphi \cdot e_{j,\kappa,t} \cdot S_{\tau,\kappa} + \varphi_f \cdot e_{j,\kappa,t} \cdot f_{j,\tau,t} \cdot S_{\tau,\kappa} + X'_{j,\tau,t} \cdot \xi + \varepsilon_{j,\tau,t}. \quad (5)$$

The key difference with Equation (4) is the presence of two similarity terms whose impact is measured by coefficients φ and φ_f (the second line). The first coefficient captures the direction and strength of spillovers from exposure to the cue $e_{j,\kappa,t}$ on the demand of the target depending on

their similarity $S_{\tau,\kappa}$. The second coefficient captures the varying strength of this effect based the frequency of the target $f_{j,\tau,t}$. We run three versions of (5), each capturing different spillovers.

Test 1, within product. Here the target is one product, $\tau = kb$, the cue is its competitor in the same market, $\kappa = kb'$. Prediction 3 says that such spillovers should be larger when the cue is more similar to the target, $\varphi > 0$, and more so when the frequency of the target is higher, $\varphi_f > 0$.

The other tests aggregate product-level demand and ads. They thus allow us to detect stronger negative spillovers because: i) similarity falls sharply across product boundaries, and ii) interference "adds up" over a broader group of dissimilar groups.

Test 2, across products, fixed cue. Here, the target is total demand for a product category, so $d_{j,\tau,t} = d_{j,k1,t} + d_{j,k2,t}$. The cue is either a cola (Coke or Pepsi) or a beer (Bud or Miller). These products allow an intuitive categorization of targets into drinks, which are similar to the cue, versus "non-drinks", which are dissimilar. Measured similarity supports this intuition: colas have much higher similarity score to other drinks (0.68 to beer, 0.78 to lemon-colas or other carbonated soft drinks) than to non-drinks (0.34) and the same holds for beer. This allows for an intuitive test of similarity versus interference: "drink" cues should foster recall of and demand for other drinks, which are similar, and interfere with recall of and demand for "non-drinks", which are dissimilar.[21] We estimate Equation (5) by setting similarity to one, $S_{\tau,\kappa} = 1$, if the target product is a drink, and zero otherwise. Prediction 3 then implies that low similarity with the cue (beer or cola) predicts negative spillovers on non-drinks, $\beta_\kappa < 0$ and $\gamma_\kappa < 0$, capturing interference. High similarity with the same cue instead predicts positive spillovers for drinks, $\varphi > 0$ and $\varphi_f > 0$.

[21] Among drinks, we use colas and beer as cues because they are highly advertised and purchased, so they are very significant cues (colas and beer have the highest mean ad exposure among drink categories and are in 2nd and 6th place respectively among all products in our sample).

Test 3, across products, pooled cue. Here the target is again total demand for a product, so $d_{j,\tau,t} = d_{j,k1,t} + d_{j,k2,t}$. The cues for each k are now *all* other 19 product categories, $\kappa = \{k' \in X, k' \neq k\}$. Cue-target similarity is then computed as the exposure-weighted average similarity:

$$S_{\tau\kappa} = \sum\nolimits_{k' \neq k} \frac{e_{k'}}{e_{\chi}} \cdot S_{k,k'},$$

where e_{χ} is total household exposure to the cue in the week, namely to all categories $k' \neq k$ and $S_{k,k'}$ is the cosine similarity between the ads of each product k' and those of the target product k.

In this test, we also expect that low similarity with the cue predicts negative baseline spillovers, $\beta_{\kappa} < 0$ and $\gamma_{\kappa} < 0$, capturing interference. Similarity-driven spillovers instead predicts a tendency for spillovers to become positive when the consumer is exposed to ads in more similar product categories, $\varphi > 0$ and $\varphi_f > 0$.

Table 4 reports the results. Column (1) considers the role of similarity for within product spillovers, test 1. Columns (2) and (3) consider test 2: the role of similarity for spillovers across products when the cue is beer and colas, respectively. Columns (4) implements test 3 on the role of similarity for spillovers from all products into the target.

Table 4: Similarity Regressions

	(1) $d_{\tau=kb}$	(2) $d_{\tau=k}$	(3) $d_{\tau=k}$	(4) $d_{\tau=k}$
f_{τ}	0.223***	0.205***	0.195***	0.212***
	(0.000321)	(0.000370)	(0.000368)	(0.000419)
f_{κ}	-0.0174***	-0.00646***	-0.0144***	-0.00384***
	(0.000116)	(0.000236)	(0.000262)	(4.39e-05)
e_{τ}	0.00747***	0.00400***	0.00372***	0.00488***
	(0.000192)	(0.000265)	(0.000274)	(0.000315)
e_{κ}	-0.00146***	-0.00421***	-0.00344***	-0.00110***
	(0.000222)	(0.000454)	(0.000428)	(0.000101)
$e_{\kappa} \cdot S$	0.000933***	0.0192***	0.0261***	0.00226***
	(0.000155)	(0.000822)	(0.000919)	(0.000275)
$e_{\tau} \cdot f_{\tau}$	0.0109***	0.00763***	0.00914***	0.00798***

	(0.000403)	(0.000320)	(0.000314)	(0.000362)
$e_\kappa \cdot f_\tau$	0.0102***	-0.00508***	-0.00657***	-0.00176***
	(0.000452)	(0.000505)	(0.000460)	(9.16e-05)
$e_\kappa \cdot f_\tau \cdot S$	0.00213***	0.0136***	0.00856***	0.00577***
	(0.000303)	(0.00108)	(0.00143)	(0.000303)
$e_\tau \cdot f_\kappa$	-0.000625***	-0.000324	-0.00144***	1.08e-05
	(0.000182)	(0.000242)	(0.000233)	(3.85e-05)
$e_\kappa \cdot f_\kappa$	-0.00122***	0.000260	-1.44e-05	-9.01e-05
	(0.000206)	(0.000334)	(0.000335)	(6.74e-05)
Observations	72,116,888	34,284,919	32,162,141	28,386,299
R-squared	0.054	0.046	0.042	0.048

Notes: column (1) reports the results of the regression at brand level, with interaction on the cosine similarity measure within the same product (i.e. κ = 'other'). Column (1) is thus exactly like column (2) of Table 2, but with the inclusion of the within product similarity interaction. Column (2) regresses the quantity purchased of every product in our dataset on their own advertising and frequency, and the advertising and frequency of beer (that is, κ = beer). Dummy S takes value 1 only for observations that refer to the purchases of another drink and 0 otherwise. Column (3) does the same for colas (κ= colas). Column (4) represents the 'everything else' as cue regression, where we create an ad Index made of the sum of the other ads, and where the ad * S component is the weighted sum of advertising exposure from other products and the similarity with the respective brand of purchase of a given observation. Standard errors are clustered at the product-week-year-DMA level.

The estimates are consistent with the predictions of our memory model and point to the importance of similarity in determining the sign (positive vs negative) and strength (varying with habits) of spillovers. Consistent with our model's predictions, in Column (1) the positive spillover from a brand on its competitor is stronger when their ads are more similar, as captured by the positive coefficient on $e_j \cdot S$, while the positive coefficient on $e_j \cdot f_k \cdot S$ indicates that the effect of similarity is especially strong for frequent buyers, consistent with our model.

We see both positive and negative spillovers (i.e. similarity vs interference) across product categories. Columns (2) and (3) show these effects when either Colas or Beers act as cues. The positive coefficient on $e_j \cdot S$ in these regressions confirms the prediction that spillovers are positive for other drinks, which are similar, and especially so for frequent buyers of these other drinks, as captured by a positive coefficient on $e_j \cdot f_k \cdot S$. Positive spillovers are consistent with within-product results. We see interference – again both in the baseline and interacted with frequency – for non-drinks, which are dissimilar. Such negative spillovers are captured by the negative

coefficient on e_j and on $e_j \cdot f_k$ in Columns (2) and (3). Memory places a testable structure on the effects of ads that helps explain consumer behavior.[22]

Column (4) confirms the tradeoff between similarity and interference for product k when all other products are treated as cues. Greater exposure to similar products, both baseline and interacted with frequency of the target, are both positive, yielding positive spillovers. This is captured by the positive coefficients in column (4) on the terms $e_j \cdot S$ and $e_j \cdot f_k \cdot S$. Exposure to dissimilar cues instead reduces demand of the target, consistent with interference. This is captured by the negative coefficients on e_j and $e_j \cdot f_k$ in Column (4).

The effects of similarity are quantitatively important. In Column (1), a one standard deviation increase in advertising of the competing brand increases own demand by 0.009 standard deviations for the average household. A one standard deviation higher similarity entails a positive spillover of 0.012 standard deviations (a 35% increase). Such spillovers are correspondingly higher for more frequent consumers.[23] In Columns (2) and (3), a one standard deviation increase in advertising of colas (beer) increases demand of drink categories by 0.025 (0.024). The positive spillover across drinks is thus 198% (186%) higher than the average own-product spillover documented in Table 3, consistent with the fact that drinks are highly similar. Conversely, the same increase in exposure to colas (beer) decreases by 0.010 (0009) standard deviations the

[22] In Table F.4 in the Appendix, we run a similar regression using our Advertising Similarity measure to estimate cross similarity between different products, instead of this subjective measure. We estimate the similarity of each product with Colas and Beer, and check, again using an interaction on the pooled dataset, whether ads have a stronger effect for products more similar to them. The same procedure is repeated for every product on our list, consistently showing how greater similarity implies greater spillovers in cueing past experiences.

[23] For a consumer one standard deviation above the average frequency of the own brand, a one standard deviation increase in exposure to the competitor has a spillover of 0.020 standard deviations at the minimum similarity level, and a one standard deviation higher similarity entails an extra spillover of 0.004, that is, a 20% increase. This is the sum of the coefficients on $e_1 \cdot S$ and on $e_1 \cdot f_1 \cdot S$ which is higher than the coefficient on $e_1 \cdot f_1$.

demand for the less similar non-drink products. This represents a 19% (7%) drop relative to the positive spillovers in the same product category (Table 3).

In sum, ad similarity – which may proxy for both product similarity and similarity in advertising context – has a significant impact on the effect of ads: it entails variation in spillover effects, including changes from positive to negative spillovers, within and across product markets.

6. Positive and Normative Implications of Memory

Our central tenet is that consumers may recall some wants and forget others, driven by well-established features of selective retrieval from memory. Memory-based habits have a strength that varies with context, delivering a new theory in which uninformative ads shape behavior by acting as involuntary "memory spotlights" (Wachter and Kahana 2024). Relative to utility theories BM (1988, 1993) we obtain new predictions, which we test and confirm. We also obtain new positive and normative implications for consumer demand and advertising, which we discuss next.

Positive Implications. We shed new light on substitutability and complementarity of goods. In standard models, substitutability between goods comes from having similar hedonic/functional features so that they (almost) equally satisfy the same want. For such goods, cross-good price elasticity is high because consuming the cheaper product entails little reduction in utility.

Our model suggests, in contrast, that economic substitutes can be vulnerable to habits and price inelasticity when three conditions on the memory mechanism are met. The first, concerning encoding, is that utility from substitutes is stored separately due to spurious "branding", e.g. the consumer experiences pleasure as Coke or Pepsi, not as Cola. The second, concerning the database, is that the consumer has more experiences with one brand, due to social context, supply, or random

factors. The third, concerning cues, is that the brands' ads, intrinsic features, or consumption contexts are similar. Under these conditions, cues lead the consumer to spontaneously represent pleasure with the brand she is loyal to, reducing substitutability and price elasticity.

These conditions are met, to varying degrees, in the product categories in our sample. Our results suggest that across different product categories, where the third condition is more likely to be violated, memory entails a form of interference leading to spurious "substitutability" of goods that satisfy different needs. As an example, Colas and Yogurt are surely encoded and experienced differently by consumers, but their cues are so dissimilar that reminders of one product can distract the consumer from the other, creating substitution "in thought" among goods that are not economic substitutes. Our model also highlights when we should see standard substitutability: it arises when the brands of intrinsically similar products are not salient during consumption. In this case, their utility flows are encoded together (e.g. most consumers do not encode the pleasure of eating a cake with different brands of baking flour), so no particular brand is retrieved based on cues.[24]

Consider next standard complementarity, which arises across different product categories when the marginal utility from a product increases with consumption of another, for instance the marginal utility of bread increases with cheese on it and vice versa. Memory implies that familiar complements cue each other, so consumers retrieve bread when thinking of cheese. At the same time, forgetting can dampen demand for unfamiliar complements, e.g. we may forget to buy or use an unfamiliar ingredient when trying a new recipe for a cake. It also implies that some products

[24] These effects can also shape the working of standard tools such as price discounts. Memory suggests that a salient discount in the price of a good cues the product itself and products similar to it (but with smaller demand spillovers than for purely contextual ads), while interfering with recall of other consumer wants. An estimated price elasticity then combines a consumer's willingness to pay q_g for a good with i) a recall $r_g(\mathbf{f}, \mathbf{e})$ of this good, and ii) recall $r_{g'}(\mathbf{f}, \mathbf{e})$ of competing wants. To recover a consumer's tastes q_g, price variation is no longer sufficient. Non-choice data on similarity and recall/attention are also needed.

may erroneously seem complements in their intrinsic utility just because they are associated with the same context so consuming one cues the other, such as movies and popcorn (Neal et al. 2011).

A second set of implications concerns firms' choices of advertising intensity. A major puzzle in marketing is why firms advertise so much despite apparent low average returns (Lodish 1995). Shapiro et al. (2021) study the link between advertising and demand for more than two hundred packaged products and document low effectiveness of own advertising for many granularly defined brands (e.g. Diet Coke ads have little effect on Diet Coke demand). Our analysis may help understand the intensity of firms' advertising despite these facts.

First, one important return to advertising is not to acquire new customers or boost current sales but to keep consumers loyal by strengthening their memory of the product. Even though current advertising may lead to a small increase in current sales, it can lead to a dynamic accumulation of experience for habitual consumers of the good that boosts their demand also in the future. Advertising does more than cueing the brand, it likely also renders a brand salient during consumption, promoting its separate encoding in memory and subsequent retrieval. This effect creates a low-frequency stability of demand by habitual buyers that may be hard to detect empirically using marginal high frequency variation in ads.

Second, significant benefits of advertising may be connected to spillovers. Some spillovers occur within multi-product firms. For instance, Diet Coke ads may have low effect on Diet Coke sales but boost demand for other varieties of Coke. Regulators are increasingly concerned that ads for non-alcoholic beverages promotes purchases of their alcoholic counterparts (Critchlow et al 2024). These spillovers increase the firm's private incentive to advertise. Other spillovers occur across firms: the ads of Coke may benefit Pepsi and vice-versa. These spillovers reduce a firm's private incentive to advertise, but can sustain explicit or implicit cooperative equilibria in which

substitute brands advertise in concert so as to keep their product category salient in consumers' minds, boosting joint sales.

Consider finally advertising content. Understanding the content of ads has proven elusive due to the difficulty of specifying why it matters. Our model highlights two mechanisms that can guide future research. The first is similarity between ads, which could be measured from perceptions or based on specific features. A firm that enters the market may want to sharply differentiate its product from the market leader in order to avoid positive spillovers on them.[25,26] The second is encoding the brand, both in terms of distinctiveness relative to competitor brands and in terms of the pleasure associated with it. Thus, ads that either cue moments of consumption pleasure such as parties or travel, or co-occur with them, can prompt the consumer to encode more distinctly the advertised brand and recalling it in similar future experiences.

These implications rely on the structure of memory and its measurable components – database, cues, and similarity – which in turn have important implications for market equilibria, to which we return in the conclusion.

Welfare Implications. Under the information or utility views of advertising, ads are good for welfare: they allow the consumer to make more accurate choices or increase utility directly or through complementarities with consumption of advertised goods. Under most behavioral views, instead, ads are bad for welfare. In Galbraith (1963) ads create superfluous wants (because

[25] An example of such a spillover, referred to a "halo effect" in the marketing literature, is the 1980's UK marketing campaign for Cinzano vermouth, whose ads were pitched as a parody of (and thus similar to) the lifestyle ads of the market leader, Martini. The campaign increased overall consumption of vermouth, likely benefitting Martini.

[26] If the positive and negative features of a product can be selectively recalled, a firm may rather cue the greater benefit of its product than its lower cost. Cueing cost may in fact prompt recall of the disadvantages of all products, reducing demand for the entire product category. Beard (2013) reviews several episodes of intense competition, and quotes McDonald's president Michael Quinlan among others: "If you've got good [product], flaunt it, but don't tear down someone else. It's not good for the industry as a whole, and I think we ought to stop."

necessary goods cue themselves with physiological stimuli, e.g. hunger for food), allowing firms to extract resources from consumers. In the coarse thinking model of Mullainathan et al. (2008), ads also cause overestimation of the value of a good. Memory offers a different perspective: ads can act as useful reminders, but because some goods yield higher consumer surplus than others, they can hinder recall of less salient but more valuable goods. To see this, note that in our model consumer welfare is affine in the following index (see Appendix A for the derivation):

$$\mathrm{W}(\mathbf{f}, \mathbf{e}) = \sum_{x \in X} r_x(\mathbf{e}, \mathbf{f}) \cdot \mathrm{W}_x, \tag{6}$$

where $r_x(\mathbf{e}, \mathbf{f})$ is the probability of recalling option x while W_x is the surplus associated with the option conditional on recalling it (W_x increases in the utility q_x of x and decreases in price p_x).

The neoclassical consumer recalls all options, $r_x(\mathbf{e}, \mathbf{f}) = 1$ for all x. Here, welfare depends solely on tastes and prices through surpluses W_x. Under selective memory, welfare is instead shaped also by the database $\mathbf{f}$ and the cues $\mathbf{e}$ a consumer is exposed to, which affects which wants are or are not recalled and satisfied. Let $\overline{\mathrm{W}} \equiv \frac{\sum_{x \in X} \mathrm{W}_x}{|X|}$ be the average surplus across all options. Higher exposure to a cue of g then affects welfare as:

$$\frac{\partial \mathrm{W}(\mathbf{f}, \mathbf{e})}{\partial e_g} = \left[r_g(g, \mathbf{f}) - r_g(\mathbf{e}, \mathbf{f})\right] \cdot \left(\mathrm{W}_g - \overline{\mathrm{W}}\right) + \sum_{x \neq g} \left[r_x(g, \mathbf{f}) - r_x(\mathbf{e}, \mathbf{f})\right] \cdot \left(\mathrm{W}_x - \overline{\mathrm{W}}\right). \tag{7}$$

Cuing g boosts its recall, $r_g(g, \mathbf{f}) - r_g(\mathbf{e}, \mathbf{f}) > 0$, which has a positive welfare impact if g gives high surplus, $\mathrm{W}_g - \overline{\mathrm{W}} > 0$. Crucially, by the second "covariance" term, the cue of g also spills over, promoting recall of options similar to it, $r_x(g, \mathbf{f}) - r_x(\mathbf{e}, \mathbf{f}) > 0$, and inhibiting recall of dissimilar ones, $r_x(g, \mathbf{f}) - r_x(\mathbf{e}, \mathbf{f}) < 0$. Consistent with our evidence, the advertised good draws the consumer's limited attention, it reduces attention to a range of other goods. The ad of g then

boosts welfare if it fosters recall (directly or via spillovers) of high surplus options $W_x - \overline{W} > 0$ and inhibits lower surplus ones $W_x - \overline{W} < 0$, while it reduces welfare otherwise.

When the cued option has high surplus, $W_g - \overline{W} > 0$, ads carry an "information benefit", akin to the benefit of good reminders (Ericson 2017). Crucially, though, the ad also exerts positive and negative spillovers on other goods. The information benefit is then *amplified* if it positively spills over to high surplus goods and inhibits low surplus ones, while it is dampened or even reversed otherwise. Seeing a Coke ad cues the consumer about her favourite drink but also spills over to i) similar drinks she likes such as Lemoncolas, enabling surplus extraction from them, and possibly also to ii) complements such as pretzels, again amplifying welfare effects. There can however be detrimental effects, via inhibition of high value goods.[27] Exposure to a Pepsi ad can interfere with recall of healthier ice tea. Seeing car ads cues the convenience of driving, causing forgetting of the pleasure of biking. The consumer then extracts too little surplus from forgotten goods. A memory cue provides "information" by moving the spotlight toward previously forgotten goods but also casts other goods into oblivion. Critically, ads are spontaneous cues, so they cannot be ignored by the consumer. They influence her choice even if such influence is detrimental.[28]

Memory accommodates Galbraith's (1963) view because necessities cue themselves: total exposure e_g to cues of a g can well depend also on internal stimuli. However, it also implies that many beneficial goods lack strong physiological cues (e.g., healthy food) while less beneficial

[27] In our baseline model, the marginal utility q_x is constant (but see footnote 9) and utility is quasilinear in the numeraire. Quasi-linearity implies that the numeraire clears the budget constraint, so the consumer optimizes by comparing the utility of the activities that come to mind to their associated numeraire cost. With decreasing marginal utility for all goods, the consumer would equate the marginal utility of goods coming to mind, while forgotten goods have a high marginal utility. In this case, cuing good g is particularly powerful in boosting consumption of g, but it is also costly in terms of precluding diversification of goods whose recall is inhibited.

[28] Even if the consumer trains herself not to purchase the advertised brand, it is challenging to control spontaneous associations to similar goods. It is plausibly even harder to counteract the inhibition in thinking about goods that are not mentioned but could have come to mind in the absence of the ad.

ones are often cued in the environment (e.g., junk food), so that they may be over-demanded. Second, if a necessity is strongly cued through ads beyond basic physiological stimuli, it can inhibit recall of goods with higher surplus and reduce welfare. Consumption habits, including for "necessities" can be associated with welfare losses. The consumer can engage excessively in an activity simply because it is top of mind due to repetition or widespread social cues, despite easily available higher value alternatives.[29]

7. Conclusion

We started from a truism: among two goods yielding similarly positive hedonic stimuli, we demand the one for which we recall pleasurable experiences instead of the one we forget about. We easily recall and prepare usual dishes, and forget to prepare dishes we very much like but have less often. We easily recall and wear certain clothes, and forget to wear a nice shirt that fell in the back of a drawer. Memory breaks the neoclassical link between utility and demand.

We model a consumer who optimizes conditional on the goods she is thinking about, and discipline recall with regularities of memory. These regularities prove powerful. They yield established phenomena such as habits and uninformative advertising, but also new predictions on how ads work for consumers who had different experiences and for goods that are more or less similar, within as well as across product categories. Using empirical proxies for memory variables and similarity, we find robust support for these predictions in the data.

[29] Memory can explain higher willingness to pay for a familiar good (e.g. an apple) when the consumer is presented with the good, rather than an image or description of it (Bushong et al. 2010) and when the DM physically holds the good (Knetsch Wong 2009). Seeing the good or holding it are stronger cues for q than mere descriptions.

Our work shows the payoff of introducing cognition, in particular realistic memory and attention, into economic analysis. It allows to complement *in a theoretically structured way* conventional choice and price data with increasingly available non-choice data on past experiences, cues, and similarity. We discuss three research directions that the framework opens up.

First, as described in Section 6, we show how incentives for advertising depend on the habits of consumers and the intrinsic features of goods. These incentives will entail new, testable comparative statics about the equilibrium joint distribution of market share, market power, advertising intensity, and strength of habits (in the sense of strength of spillover effects). For example, advertising may be stronger in branded but (intrinsically) undifferentiated markets, relative to highly differentiated markets, where advertising by a firm can have the more standard negative spillovers on its competitor.

Second, the model stresses the importance of dynamics. As in the Stigler anecdote, the equilibrium level of advertising should balance the long-run benefit of building consumer databases against the short-term cost of advertising.[30] This logic sheds light on entry and competition, rationalizing the marketing dictum "differentiate anyway you can" (Kim and Mauborgne 2005). This captures the firm's effort to be encoded separately from the habitual product, potentially by drawing attention to high contrast features, either intrinsic (as in Bordalo, Gennaioli, and Shleifer 2016) or branding. Yet, memory also produces market segmentation: once a habit foothold has been established, potentially due to idiosyncratic shocks (such as place of birth) or intensive ad campaigns targeted to specific demographics, it is persistent if accompanied by continuous investment in the database. As alluded to earlier, this can take the form of

[30] This is consistent with the advertising industry's focus on metrics that measure the long-term impact of ads. Hartnett et al (2021) documents that brands that stopped advertising for a year suffered substantial drops of demand over that period, which were difficult to reverse with renewed advertising.

advertising, but can also leverage branding and product associations; if Corona is associated with a party at the beach, similar environments will help retrieve Corona but also rehearse the association.

Finally, our results have implications for other markets where persuasion is important. A prime example is politics, where much effort is spent on "getting out the vote." Our model suggests several lines of inquiry. A party affiliation, like a brand, is associated with memories and is thus liable to memory-based habits. In this context, habits may be of the form that voters may care more for issues they hear a lot about. Conversely, habit-cue interactions suggest that a partisan message may have a higher return than a median one: a partisan message (including negative advertising about the opposition) is more effective at increasing the motivation of the party base, both because it retrieves more experiences of a similar action in the past and because it focuses the base on issues it cares about.

In our current analysis, memory influences the consumer's prediction of the utility she will derive from consuming a good, not the utility experienced during the good's consumption. But memory can also influence experienced utility, as in Bordalo et al (2025): engaging in an activity (e.g. working in a specific firm) cues pleasant or unpleasant thoughts (e.g. past workplace disappointments), shaping the current experienced utility from the activity itself. We leave it to future work to study the real-world implication of these forces.

References

Atkin, D. (2013). Trade, tastes, and nutrition in India. *American economic review*, *103*(5), 1629-1663.

Beard, F. (2013). A History of Comparative Advertising in the United States. *Journalism & Communication Monographs* 15(3): 114–216.

Becker, G., and Murphy, K. (1988). A theory of rational addiction. *Journal of Political Economy*, 96(4): 675–700.

Becker, G., and Murphy, K. (1993). A simple theory of advertising as a good or bad. *Quarterly Journal of Economics*, 108(4):941–964.

Bhatia, S., & Aka, A. (2022). Cognitive modeling with representations from large-scale digital data. *Current Directions in Psychological Science*, 31(3): 207-214.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2020). Memory, attention and choice. *Quarterly Journal of Economics*, 135(3): 1399-1442.

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

Bordalo, P., Burro, G., Coffman, K. B., Gennaioli, N., and Shleifer, A. (2025). Imagining the future: Memory, simulation and beliefs. *Review of Economic Studies*, 92(3): 1532-1563.

Bordalo, P., Conlon, J., Gennaioli, N., Kwon, S., and Shleifer, A. (2025). How people use statistics. *Review of Economic Studies*, forthcoming.

Bordalo, P., Gennaioli, N., Schwartzstein, J., and Shleifer, A. (2025). The utility of memories. *Working Paper.*

Bronnenberg, B., and Dubé, J.-P. (2017). The formation of consumer brand preferences. *Annual Review of Economics*, 9(1): 353-382.

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

Bronnenberg, B., Dubé, J.-P. and Syverson, C. (2022). Marketing investment and intangible brand capital. *Journal of Economic Perspectives*, 36(3): 53-74.

Bushong, B., King, L., Camerer, C., and Rangel, A. (2010). Pavlovian processes in consumer choice: The physical presence of a good increases willingness-to-pay. *American Economic Review*, 100(4): 1556-1571.

Cattaneo, M.R., Crump, R.K., Farrell, M.H., and Feng, Y. (2024). On binscatter. *American Economic Review*, 114(5):1488:1514.

Critchlow, N., Holmes, J., & Fitzgerald, N. (2025). Alibi marketing? Surrogate marketing? Brand sharing? What is the correct terminology to discuss marketing for alcohol-free and low-alcohol products which share branding with regular strength alcohol products?. *Addiction*, *120*(1), 4-6.

Dubé, J. P., Hitsch, G. J., & Rossi, P. E. (2010). State dependence and alternative explanations for consumer inertia. *The RAND Journal of Economics*, *41*(3), 417-445.

Ehrenberg, A. S. C. (1974). Repetitive advertising and the consumer. *Journal of Advertising Research*, 14(April):25–34.

Einav, L., Leibtag, E., and Nevo, A. (2010). Recording discrepancies in Nielsen Homescan data: Are they present and do they matter? *QME*, *8*, 207-239.

Ericson, K. (2017). On the interaction of memory and procrastination: Implications for reminders, deadlines, and empirical estimation. *Journal of the European Economic Association*, 15(3):692-719.

Frank, R. (1962). Brand choice as a probability process. *Journal of Business,* 35 (1), 43-56.

Galbraith, J. K. (1963). *The affluent society*. Houghton Mifflin Harcourt.

Hartnett, N., Gelzinis, A., Beal, V., Kennedy, R., & Sharp, B. (2021). When Brands Go Dark: Examining Sales Trends when Brands StopBroad-Reach Advertising for Long Periods. *Journal of Advertising Research*, *61*(3), 247-259.

He, C., and Klein, T. (2023). Advertising as a reminder: Evidence from the Dutch State Lottery. *Marketing Science*, 42 (5): 892-909.

Heckman, J. J. (1981). Heterogeneity and state dependence. In *Studies in labor markets* (pp. 91-140). University of Chicago Press.

Kahana, M. (2012). *Foundations of Human Memory*. Oxford, UK: Oxford University Press.

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

Keller, K. (1993). Conceptualizing, measuring, and managing customer-based brand equity. *Journal of Marketing*, *57*(1), 1-22.

Laibson, D. (2001). A cue-theory of consumption. *Quarterly Journal of Economics*, 116(1):81-119.

Lodish, L. M., Abraham, M., Kalmenson, S., Livelsberger, J., Lubetkin, B., Richardson, B., and Stevens, M. E. (1995). How t.v. advertising works: A meta-analysis of 389 real world split cable t.v. advertising experiments. *Journal of Marketing Research*, 32(2):125–139.

Kim, W. and Mauborgne, R. (2005). Blue ocean strategy: How to create uncontested market space and make the competition irrelevant. Harvard Business Review Press.

McClure, S., Li, J., Tomlin, D., Cypert, K., Montague, L., & Montague, P. (2004). Neural correlates of behavioral preference for culturally familiar drinks. *Neuron*, 44(2): 379-387.

Mullainathan, S., Schwartzstein, J., and Shleifer, A. (2008). Coarse thinking and persuasion. *Quarterly Journal of Economics,* 123 (2): 577-619.

Neal, D., Wood, W., Labrecque, J., Lally, P. How do habits guide behavior? Perceived and actual triggers of habits in daily life. *Journal of Experimental Social Psychology* 48(2): 492-498.

Nelson, P. (1974). Advertising as information. *Journal of Political Economy*, 82(4): 729-754.

Nosofsky, R. (1992). Similarity scaling and cognitive process models. *Annual review of Psychology*, *43*(1), 25-53.

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

Shapiro, J. (2006). A "memory-jamming" theory of advertising. *Working paper*.

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

Shapiro, B., Hitsch, G., and Tuchman, A. (2021). TV advertising effectiveness and profitability: Generalizable results from 288 brands. *Econometrica*, 89(4):1855–1879.

Singhal, A. (2001). Modern information retrieval: A brief overview. *IEEE Data Eng. Bull.*, *24*(4), 35-43.

Stigler, G. (1968). Price and non-price competition. *Journal of Political Economy*, 76(1), 149-154.

Tellis, G. (1988). Advertising exposure, loyalty, and brand purchase: A two-stage model of choice. *Journal of Marketing Research*, 25(2):134–144.

Wachter, J. A., & Kahana, M. J. (2024). A retrieved-context theory of financial decisions. *The Quarterly Journal of Economics*, *139*(2), 1095-1147.

Webb, R., Zhao C., Osborne M., Landry P., and Camerer C. (2022). A neural auto-pilot theory of habit: evidence from canned tuna. Working Paper.

Webb, R., Fong, J., Mazar, A., Levine, J., Wellsjo, A. S., Natan, O. R., ... & Camerer, C. (2025). Integrating neuro-psychological habit research into consumer choice models. Working Paper.

Appendix

A. Proofs

Prediction 1. For a database $\mathbf{f} = (f_{11}, f_{12}, \ldots, f_n)$ we have:

$$r_g(c, \mathbf{f}) = \frac{S(c, g) \cdot f_g}{\sum_{x \in G} S(c, x) \cdot f_x + S(c, n) \cdot f_n}.$$

And recall cued by an exposure vector $\mathbf{e} = (e_{11}, \ldots, e_{K2}, e_n)$ is given by

$$r_g(\mathbf{e}, \mathbf{f}) = \sum\nolimits_{c \in X} r_g(c, \mathbf{f}) \cdot \eta_c,$$

where $\eta_c = \frac{e_c}{e}$ and $e = \sum_{x' \in X} e_{x'}$.

We now obtain a condition ensuring that $r_g(c, \mathbf{f})$ strictly increases in similarity between the cue c and g, formally $S(c, g) > S(c', g) \Rightarrow r_g(c, \mathbf{f}) > r_g(c', \mathbf{f})$. Under Assumption 1, we know that $r_g(c, \mathbf{f}) \geq \frac{S(c,g) \cdot f_g}{\overline{S} \cdot (1 - f_n) + \underline{S} \cdot f_n}$ and that $r_g(c', \mathbf{f}) \leq \frac{S(c', g) \cdot f_g}{\underline{S}}$. A sufficient condition is then that, whenever $S(c, g) > S(c', g)$ we have:

$$\frac{S(c, g) \cdot f_g}{\overline{S} \cdot (1 - f_n) + \underline{S} \cdot f_n} > \frac{S(c', g) \cdot f_g}{\underline{S}},$$

which can be rewritten as $\frac{S(c,g)}{S(c',g)} > \frac{\overline{S}}{\underline{S}} - f_n \cdot \left(\frac{\overline{S} - \underline{S}}{\underline{S}} \right)$, which is in turn equivalent to:

$$f_n > \frac{\overline{S} - \underline{S} \cdot \frac{S(c, g)}{S(c', g)}}{\overline{S} - \underline{S}}.$$

The tightest condition occurs at cues that are adjacent in similarity for g, for instance a cue/good with rank r and the immediate next similar cue/good with rank $r+1$. Define the tightest similarity ratio (across cues and goods g) at the value:

$$\gamma = \min_{g,r=1,\ldots,1+2K} \frac{S(r,g)}{S(r+1,g)},$$

where r is the similarity rank. Such γ exists and is unique, giving the condition:

$$f_n > f_n^* \equiv \frac{\overline{S} - \underline{S} \cdot \gamma}{\overline{S} - \underline{S}} > 0, \qquad (A.1)$$

which ensures that recall increases with similarity for $c \in G$.

Note that condition A.1 implies not only that cueing good g increases recall of g, but it also implies that cueing *any* consumer good g' increases the recall of good g compared to a neutral cue. This will be important later. To see why this holds, write this condition as:

$$\min_{c \in G} r_g(c, \mathbf{f}) \geq \frac{S(2K) \cdot f_g}{\overline{S} \cdot (1 - f_n) + \underline{S} \cdot f_n} > \frac{\underline{S} \cdot f_g}{\underline{S} \cdot (1 - f_n) + S(n,n) \cdot f_n} = r_g(n, \mathbf{f})$$

where $S(2K) = \min_{g,r=1,\ldots,1+2K} S(r,g)$. This can be written as:

$$\left[\frac{S(2K)}{\underline{S}}\right]\left[\frac{S(n,n)}{\underline{S}} - 1\right] \cdot f_n + \left(\overline{S} - \underline{S}\right) \cdot f_n > \left[\overline{S} - \underline{S} \cdot \frac{S(2K)}{\underline{S}}\right],$$

which is implied by (A.1) due to $S(n,n) \geq \underline{S}$.

Equipped with this result, we can characterize the key comparative statics of demand for a consumer with a cue-database combination $(\mathbf{e}, \mathbf{f})$. By deriving Equation (3) it with respect to past purchases f_g of the same good we obtain:

$$\frac{\partial d_g(\mathbf{e},\mathbf{f})}{\partial f_g} = q_g \cdot \frac{\partial r_g(\mathbf{e},\mathbf{f})}{\partial f_g},$$

which, by using (1) and (2), gives:

$$\frac{\partial d_g(\mathbf{e},\mathbf{f})}{\partial f_g} = q_g \cdot \sum\nolimits_{c\in X} \frac{\partial r_g(c,\mathbf{f})}{\partial f_g} \cdot \eta_c$$

$$= q_g \cdot \sum\nolimits_{c\in X} \frac{S(c,g)\cdot f_g \cdot \sum_{x'\neq g} S(c,x')\cdot f_{x'}}{f_g \cdot [\sum_{x'\in X} S(c,x')\cdot f_{x'}]^2} \eta_c$$

$$= q_g \cdot \sum\nolimits_{c\in X} \frac{r_g(c,\mathbf{f})\left[1-r_g(c,\mathbf{f})\right]}{f_g} \cdot \eta_c, \qquad \text{(A. 2)}$$

which is positive if $r_g(c,\mathbf{f}) \in (0,1)$ for at least one $c \in X$, which is guaranteed by Assumption 1 (in fact, condition (A.1) further implies that $r_g(c,\mathbf{f}) \in (0,1)$ for all $c \in X$). Thus, demand of g increases in f_g, as in point i).

To assess concavity in f_g, we can further derive (3) and obtain:

$$\frac{\partial^2 d_g(\mathbf{e},\mathbf{f})}{\partial f_g^2} = q_g \cdot \sum\nolimits_{c\in X} \frac{\partial^2 r_g(c,\mathbf{f})}{\partial f_g^2} \cdot \eta_c$$

$$= -q_g \cdot \sum\nolimits_{c\in X} \frac{2r_g(c,\mathbf{f})^2\left[1-r_g(c,\mathbf{f})\right]}{f_g^2} \cdot \eta_c < 0, \qquad \text{(A. 3)}$$

as implied by our previous argument. Thus, demand of g is concave in past consumption f_g.

Consider now the comparative static relative to exposure e_g to own ads. By deriving Equation (3) with respect e_g we obtain:

$$\frac{\partial d_g(\mathbf{e},\mathbf{f})}{\partial e_g} = q_g \cdot \frac{\partial r_g(\mathbf{e},\mathbf{f})}{\partial e_g},$$

which, by using (1) and (2), gives:

$$\frac{\partial d_g(\mathbf{e},\mathbf{f})}{\partial e_g} = q_g \cdot \sum\nolimits_{c\in X} r_g(c,\mathbf{f}) \cdot \frac{\partial \eta_c}{\partial e_g}$$

$$= q_g \cdot r_g(g,\mathbf{f}) \cdot \frac{\sum_{x'\in X} e_{x'} - e_g}{e^2} - q_g \cdot \sum\nolimits_{c\neq g} r_g(c,\mathbf{f}) \frac{e_c}{e^2}$$

$$= \frac{q_g}{e}\left[r_g(g,\mathbf{f})\left(1-\eta_g\right) - \sum\nolimits_{c\neq g} r_g(c,\mathbf{f})\eta_c\right]$$

$$= q_g \cdot \frac{r_g(g,\mathbf{f}) - r_g(\mathbf{e},\mathbf{f})}{e}, \qquad \text{(A. 4)}$$

The expression is positive for $r_g(g,\mathbf{f}) - r_g(\mathbf{e},\mathbf{f}) > 0$. Again, Assumption 1 and condition (A.1) are sufficient to guarantee this condition.

Finally, by deriving (A.3) with respect to f_g we find:

$$\frac{\partial^2 d_g(\mathbf{e},\mathbf{f})}{\partial e_g \partial f_g} = \frac{q_g}{e} \cdot \left[\frac{\partial r_g(g,\mathbf{f})}{\partial f_g} - \frac{\partial r_g(\mathbf{e},\mathbf{f})}{\partial f_g}\right]$$

$$\propto r_g(g,\mathbf{f})\left[1 - r_g(g,\mathbf{f})\right] - \sum\nolimits_{c\in X} r_g(c,\mathbf{f})\left[1 - r_g(c,\mathbf{f})\right] \cdot \eta_c . \qquad \text{(A. 5)}$$

The derivative is positive provided product $r_g(c,\mathbf{f})\left[1 - r_g(c,\mathbf{f})\right]$, evaluated at $c = g$, is larger than its average value. Since $r_g(g,\mathbf{f}) \geq r_g(c,\mathbf{f})$ (strictly if $c \neq g$), this condition is satisfied provided $r_g(g,\mathbf{f}) \leq \frac{1}{2}$. To characterize this condition, note that:

$$r_g(g,\mathbf{f}) \leq \frac{\overline{S} \cdot f_g}{\underline{S}} < \frac{\overline{S} \cdot (1 - f_n)}{\underline{S}}$$

So a sufficient condition for $r_g(g,\mathbf{f}) < 1/2$ is:

$$\frac{\overline{S}\cdot(1-f_n)}{\underline{S}}<\frac{1}{2},$$

which can be rewritten as:

$$f_n > f_n^{**} \equiv \frac{2\overline{S}-\underline{S}}{2\overline{S}}.$$

The proposition is thus proven for $f_n > \overline{f_n} = \max(f_n^*, f_n^{**})$.

Finally, we consider concavity of the effect of exposure, namely $\frac{\partial^3 d_g(\mathbf{e},\mathbf{f})}{\partial e_g \partial^2 f_g} < 0$. Starting from (A.5) we have:

$$\begin{aligned}
\frac{\partial^3 d_g(\mathbf{e},\mathbf{f})}{\partial e_g \partial f_g^2} &= -\frac{q_g}{e f_g^2}\Big\{r_g(g,\mathbf{f})\big[1-r_g(g,\mathbf{f})\big]-\sum\nolimits_{c\in X} r_g(c,\mathbf{f})\big[1-r_g(c,\mathbf{f})\big]\eta_c\Big\} \\
&\quad+\frac{1}{e\cdot f_g}\Big\{\partial_{f_g} r_g(g,\mathbf{f})\big[1-2r_g(g,\mathbf{f})\big]-\sum\nolimits_{c\in X}\partial_{f_g} r_g\big(c,\bar{\mathbf{f}}\big)\big[1-2r_g\big(c,\bar{\mathbf{f}}\big)\big]\eta_c\Big\} \\
&= \frac{1}{e f_g^2}\Big\{-r_g(g,\mathbf{f})\big[1-r_g(g,\mathbf{f})\big]+\sum\nolimits_{c\in X} r_g(c,\mathbf{f})\big[1-r_g(c,\mathbf{f})\big]\eta_c \\
&\quad+ r_g(g,\mathbf{f})\big[1-r_g(g,\mathbf{f})\big]\big[1-2r_g(g,\mathbf{f})\big] \\
&\quad-\sum\nolimits_{c\in X} r_g(c,\mathbf{f})\big[1-r_g(c,\mathbf{f})\big]\big[1-2r_g(c,\mathbf{f})\big]\eta_c\Big\} \\
&= \frac{1}{e f_g^2}\Big\{-2r_g^2(g,\mathbf{f})\big[1-r_g(g,\mathbf{f})\big]+2\sum\nolimits_{c\in X} r_g^2(c,\mathbf{f})\big[1-r_g(c,\mathbf{f})\big]\eta_c\Big\}
\end{aligned}$$

which is negative provided $r_g^2(g,\mathbf{f})\big[1-r_g(g,\mathbf{f})\big] > \sum_{c\in X} r_g^2(c,\mathbf{f})\big[1-r_g(c,\mathbf{f})\big]\eta_c$, which coincides with the condition that (A.5) is positive and is guaranteed by the condition $f_n > \overline{f_n}$, proving point c).

Proposition 2. Consider similarity-based spillovers, namely the effect on the demand for good $g = kb$ of consumption experiences and ads of the competing brand $g' = kb'$. Higher exposure $e_{g'}$ to cue $c = g'$ impacts retrieval of g as follows:

$$\left.\frac{\partial r_{\mathrm{g}}(\mathbf{e},\mathbf{f})}{\partial e_{g'}}\right|_{\bar{\mathbf{e}},\bar{\mathbf{f}}} = \frac{\left[r_g(g',\mathbf{f}) - r_g(\mathbf{e},\mathbf{f})\right]}{e}. \tag{A.6}$$

This is positive if and only if cue g' prompts above-average recall of g given the baseline cue e. Given the assumption $S(kb, kb') > S(kb, x)$ *for* $x \notin k$, the latter has the upper bound:

$$r_g(\mathbf{e},\mathbf{f}) \leq \eta_g \cdot r_g(g,\mathbf{f}) + \left(1 - \eta_g - \eta_n\right) \cdot r_g(g',\mathbf{f}) + \eta_n \cdot r_g(n,\mathbf{f}).$$

So a sufficient condition for (A.6) to be positive is:

$$r_g(g',\mathbf{f}) \geq \frac{e_g}{e_g + e_n} \cdot r_g(g,\mathbf{f}) + \frac{e_n}{e_g + e_n} \cdot r_g(n,\mathbf{f}),$$

leading to:

$$\frac{e_n}{e_g} > \overline{e} \equiv \frac{r_g(g,\mathbf{f}) - r_g(g',\mathbf{f})}{r_g(g',\mathbf{f}) - r_g(n,\mathbf{f})},$$

that is, the share of the least similar neutral events in the cue must be sufficiently large (since by condition A.1 the right-hand side is positive). This is intuitive: increasing the share of g' in the cue promotes recall of g if it makes the cue closer to g, in particular if e_g is not too large, proving point d), namely $\partial d_g(\mathbf{e},\mathbf{f})/\partial e_{g'} > 0$, with $e^* = \overline{e} \cdot e_g$. Together with Proposition 1, this also proves the first part of point f), namely that demand for product k as a whole increases, $\partial\left[d_g(\mathbf{e},\mathbf{f}) + d_{g'}(\mathbf{e},\mathbf{f})\right]/\partial e_{g'} > 0$.

Consider next the interaction between $c = g'$ and past purchases of f_g. By deriving (A.6) with respect to f_g we find:

$$\frac{\partial^2 r_g(\mathbf{e},\mathbf{f})}{\partial f_g \partial e_{g'}} = \frac{1}{e \cdot f_g}\Big\{r_g(g',\mathbf{f})\big[1 - r_g(g',\mathbf{f})\big] - \sum\nolimits_{c\in X} r_g\big(c,\bar{\mathbf{f}}\big)\big[1 - r_g\big(c,\bar{\mathbf{f}}\big)\big]\eta_c\Big\}. \quad (A.7)$$

Due to concavity of the expectation on the right-hand side of (A.7) satisfies:

$$\sum\nolimits_{c\in X} r_g(c,\mathbf{f})\big[1 - r_g(c,\mathbf{f})\big]\eta_c \le r_g(\mathbf{e},\mathbf{f})\big[1 - r_g(\mathbf{e},\mathbf{f})\big]$$

Then, under the condition $r_g(g,\mathbf{f}) < \frac{1}{2}$ from Proposition 1 – which is satisfied when $f_n > \overline{f_n}$ – it follows that (A.7) is positive whenever (A.6) is positive, i.e. $r_g(g',\mathbf{f}) > r_g(\mathbf{e},\mathbf{f})$ (a sufficient condition for the latter condition being $e_n/e_g > \overline{e}$). This proves the first part of point e), namely $\partial^2 d_g(\mathbf{e},\mathbf{f})/\partial e_{g'}\partial f_g > 0$, as well as the second part of point f), namely $\partial^2\big[d_g(\mathbf{e},\mathbf{f}) + d_{g'}(\mathbf{e},\mathbf{f})\big]/\partial e_{g'}\partial f_x > 0,\ x = g, g'$.

Finally, we consider concavity. Similar to the analysis in Proposition 1c):

$$\frac{\partial^3 d_g(\mathbf{e},\mathbf{f})}{\partial e_{g'}\partial f_g^2} = \frac{1}{e f_g^2}\Big\{-2r_g^2(g',\mathbf{f})\big[1 - r_g(g',\mathbf{f})\big] + 2\sum\nolimits_{c\in X} r_g^2\big(c,\bar{\mathbf{f}}\big)\big[1 - r_g\big(c,\bar{\mathbf{f}}\big)\big]\eta_c\Big\}$$

This is negative provided

$$r_g^2(g',\mathbf{f})\big[1 - r_g(g',\mathbf{f})\big] > \sum\nolimits_{c\in X} r_g^2\big(c,\bar{\mathbf{f}}\big)\big[1 - r_g\big(c,\bar{\mathbf{f}}\big)\big]\eta_c$$

which holds because $r_g(c,\mathbf{f}) < r_g(g',\mathbf{f}) < \frac{1}{2}$.

From Prediction 1 and the previous results in Prediction 2, it follows that under the assumptions of the latter, increasing exposure to one brand in the product category increases

demand for the category as a whole, $\partial\left[d_g(\mathbf{e},\mathbf{f}) + d_{g'}(\mathbf{e},\mathbf{f})\right]/\partial e_{g'} > 0$, and more so for more frequent buyers of either good, $\partial^2\left[d_g(\mathbf{e},\mathbf{f}) + d_{g'}(\mathbf{e},\mathbf{f})\right]/\partial e_g \partial f_g > 0$ and $\partial^2\left[d_g(\mathbf{e},\mathbf{f}) + d_{g'}(\mathbf{e},\mathbf{f})\right]/\partial e_{g'}\partial f_{g'} > 0$, proving the second part of point e).

Proposition 3.

Consider first how recall of good g depends on exposure e_c to cue c:

$$\left.\frac{\partial r_g(\mathbf{e},\mathbf{f})}{\partial e_c}\right|_{\mathbf{e},\mathbf{f}} = \frac{\left[r_g\left(c,\vec{\mathbf{f}}\right) - r_g\left(\bar{\mathbf{e}},\vec{\mathbf{f}}\right)\right]}{e}, \qquad (A.8)$$

Similar to Proposition 2b), this is positive if and only if cue c prompts above average recall of g, that is $r_g(c,\mathbf{f}) > r_g(\mathbf{e},\mathbf{f})$. Denote by $\sigma(g) \in \{1,2,\ldots,|G|\}$ the ranking of similarity of other goods to good g. By Prediction 1, for each good g there is a good $\sigma_1(g) \in \{1,2,\ldots,|G|\}$ such that $(A.9)$ is positive if $c = x_\sigma$ for $\sigma < \sigma_1(g)$ and it is zero or negative otherwise.

In particular, consider how the sensitivity of recall of g to exposure e_c depends on the similarity $S(g,c)$ of g and $c \neq g$. We have:

$$\left.\frac{\partial r_g(\mathbf{e},\mathbf{f})}{\partial S(g,c)}\right|_{\mathbf{f},\mathbf{e}} = \sum\nolimits_x \frac{\partial r_g(x,\mathbf{f})}{\partial S(g,c)}\eta_x = \frac{\partial r_g(g,\mathbf{f})}{\partial S(g,x)}\eta_g + \frac{\partial r_g(c,\mathbf{f})}{\partial S(g,c)}\eta_c$$

$$= -\frac{r_g(g,\mathbf{f})\cdot r_c(g,\mathbf{f})}{S(g,c)}\eta_g + \frac{r_g(g,\mathbf{f})\cdot\left[1 - r_g(g,\mathbf{f})\right]}{S(g,c)}\eta_c. \qquad (A.9)$$

Then:

$$\left.\frac{\partial^2 r_g(\mathbf{e},\mathbf{f})}{\partial e_c \partial S(g,c)}\right|_{\mathbf{f},\mathbf{e}} = \frac{1}{e}\left[\frac{\partial r_g(c,\mathbf{f})}{\partial S(g,c)} - \frac{\partial r_g(\mathbf{e},\mathbf{f})}{\partial S(g,c)}\right]$$

$$= \frac{r_g(g,\mathbf{f})\cdot\left[1-r_g(g,\mathbf{f})\right]}{eS(g,c)}(1-\eta_c) + \frac{r_g(g,\mathbf{f})\cdot r_c(g,\mathbf{f})}{eS(g,c)}\eta_g$$

$$> 0 \qquad (A.10)$$

Consider next how the sensitivity of recall of g to exposure e_c depends on the frequency of g. We have:

$$\left.\frac{\partial^2 r_g(\mathbf{e},\mathbf{f})}{\partial f_g \partial e_c}\right|_{\mathbf{f},\mathbf{e}} = \frac{1}{e\cdot f_g}\left\{r_g(c,\mathbf{f})\left[1-r_g(c,\mathbf{f})\right] - \sum\nolimits_{x\in X} r_g(x,\mathbf{f})\left[1-r_g(x,\mathbf{f})\right]\eta_x\right\}. \quad (A.11)$$

Similar to Proposition 2, point b), if $r_g(g,\mathbf{f}) < \frac{1}{2}$, then the function $r_g(x,\mathbf{f})\left[1-r_g(x,\mathbf{f})\right]$ increases in $r_g(x,\mathbf{f})$. Thus, (A.11) identifies a threshold $\sigma_2(g) \in \{1,2,\ldots,|G|\}$ such that $(A.11)$ is positive if $c = x_\sigma$ for $\sigma < \sigma_2(g)$ and it is zero or negative otherwise. Because $r_g(x,\mathbf{f})\left[1-r_g(x,\mathbf{f})\right]$ is concave, $\sigma_2(g) \geq \sigma_1(g)$.

Finally, consider the triple interaction, between exposure to the cue e_c, the frequency of the target f_g, and similarity between the cue and the target, $S(g,c)$. Starting from point a), we obtain:

$$\left.\frac{\partial^3 r_g(\mathbf{e},\mathbf{f})}{\partial e_c \partial S(g,c)\partial f_g}\right|_{\mathbf{f},\mathbf{e}} = \frac{\partial}{\partial f_g}\frac{r_g(g,\mathbf{f})}{eS(g,c)}\left[\left(1-r_g(g,\mathbf{f})\right)(1-\eta_c) + r_c(g,\mathbf{f})\eta_g\right]$$

The first term is proportional to:

$$\partial_{f_g} r_g(g,\mathbf{f})\cdot\left[1-2r_g(g,\mathbf{f})\right] = \frac{r_g(g,\mathbf{f})\left[1-r_g(g,\mathbf{f})\right]\left[1-2r_g(g,\mathbf{f})\right]}{f_g}$$

and the second term is proportional to:

$$\partial_{f_g} r_g(g,\mathbf{f}) \cdot r_c(g,\mathbf{f}) + r_g(g,\mathbf{f}) \cdot \partial_{f_g} r_c(g,\mathbf{f}) = \frac{r_g(g,\mathbf{f})\left[1 - r_g(g,\mathbf{f})\right] r_c(g,\mathbf{f})}{f_g} - \frac{r_c(g,\mathbf{f}) r_g(g,\mathbf{f})}{f_g}$$

$$= \frac{r_c(g,\mathbf{f}) r_g(g,\mathbf{f})}{f_g}\left[1 - r_g(g,\mathbf{f}) - 1\right] = -\frac{r_c(g,\mathbf{f}) r_g^2(g,\mathbf{f})}{f_g}$$

because

$$\partial_{f_g} r_c(g,\mathbf{f}) = \partial_{f_g} \frac{S(c,g) f_c}{\sum_x S(x,g) f_x} = -\frac{r_c(g,\mathbf{f}) r_g(g,\mathbf{f})}{f_g}$$

In particular, the expression is positive provided:

$$(1 - \eta_c)\left[1 - r_g(g,\mathbf{f})\right]\left[1 - 2r_g(g,\mathbf{f})\right] > \eta_g r_c(g,\mathbf{f}) r_g(g,\mathbf{f})$$

Given the assumption $r_c(g,\mathbf{f}) \leq \frac{1}{2}$ for all c, g, a sufficient condition for this to hold is that current exposure to g is not too high, or current exposure to c is not too low:

$$\frac{1 - \eta_c}{\eta_g} > \frac{r_c(g,\mathbf{f})}{1 - 2r_g(g,\mathbf{f})}.$$

Welfare Formula and Comparative Static. The consumer's overall utility is the probability of engaging in each option x times the surplus (utility minus price) obtained from the option, plus numeraire consumption. The latter can be thought of as the utility of increased savings, for simplicity we assume it is always available to the consumer. Thus, total utility is equal to:

$$\widetilde{\mathrm{W}}(\mathbf{e},\mathbf{f}) = \sum_x \Pr(x|\mathbf{e},\mathbf{f}) \cdot (q_x + \overline{\varepsilon}) + \theta \cdot \left[A - \sum_x \Pr(x|\mathbf{e},\mathbf{f}) \cdot p_x\right],$$

where $\Pr(x|\mathbf{e},\mathbf{f})$ is the probability of choosing option x, $(q_x + \overline{\varepsilon})$ is the average utility from consuming x conditional on choosing it, p_x is the deterministic price of the option, and θ is the marginal utility from the numeraire. Welfare can thus be rewritten as:

$$\widetilde{\mathrm{W}}(\mathbf{e},\mathbf{f}) = \sum_x \Pr(x|\mathbf{e},\mathbf{f}) \cdot (q_x + \overline{\varepsilon} - p_x) + \theta \cdot A.$$

In this expression, both the probability of choosing $\Pr(x|\mathbf{e},\mathbf{f})$ and the average utility conditional on choosing, $(q_x + \overline{\varepsilon} - p_x)$, are implicit functions of recall and the utility shock ε. Specifically, if option x is recalled, which occurs with probability $r_x(\mathbf{e},\mathbf{f})$, the consumer chooses it if and only if $\varepsilon > \theta \cdot p_x - q_x$. Otherwise, the consumer chooses it if $\varepsilon > \theta \cdot p_x$. In both cases, choosing x yields utility $q_x + \varepsilon - \theta \cdot p_x$. We can thus write:

$$\Pr(x|\mathbf{e},\mathbf{f}) \cdot (q_x + \overline{\varepsilon} - p_x) = r_x(\mathbf{e},\mathbf{f}) \cdot \int_{\theta \cdot p_x - q_x}^{\hat{\varepsilon}} (q_x + \varepsilon - \theta \cdot p_x) \cdot \frac{d\varepsilon}{2 \cdot \hat{\varepsilon}} +$$

$$[1 - r_x(\mathbf{e},\mathbf{f})] \cdot \int_{\theta \cdot p_x}^{\hat{\varepsilon}} (q_x + \varepsilon - \theta \cdot p_x) \cdot \frac{d\varepsilon}{2 \cdot \hat{\varepsilon}}$$

where for simplicity we assume the utility shock is iid and uniformly distributed in $[-\hat{\varepsilon}, \hat{\varepsilon}]$. We then have:

$$\Pr(x|\mathbf{e},\mathbf{f}) \cdot (q_x + \overline{\varepsilon} - p_x) = \frac{r_x(\mathbf{e},\mathbf{f})}{2 \cdot \hat{\varepsilon}} \cdot \left[\frac{(q_x - \theta \cdot p_x)^2}{2} + \hat{\varepsilon} \cdot \left[(q_x - \theta \cdot p_x) + \frac{\hat{\varepsilon}}{2}\right]\right] +$$

$$\frac{[1 - r_x(\mathbf{e},\mathbf{f})]}{2 \cdot \hat{\varepsilon}} \cdot \left[\frac{(\theta \cdot p_x)^2}{2} + \hat{\varepsilon} \cdot \left[-\theta \cdot p_x + \frac{\hat{\varepsilon}}{2}\right]\right]$$

$$= \frac{r_x(\mathbf{e},\mathbf{f})}{4 \cdot \hat{\varepsilon}} \cdot \left[\frac{q_x^2 + 2 \cdot q_x \cdot (\hat{\varepsilon} - \theta \cdot p_x)}{2}\right] + \frac{(\hat{\varepsilon} - \theta \cdot p_x)^2}{4 \cdot \hat{\varepsilon}}$$

Let $\mathrm{W}_x \equiv \left[\frac{q_x^2+2\cdot q_x\cdot(\hat{\varepsilon}-\theta\cdot p_x)}{2}\right]$. We can write:

$$\widetilde{\mathrm{W}}(\mathbf{e},\mathbf{f}) = \frac{1}{4\cdot\hat{\varepsilon}}\cdot\sum_x r_x(\mathbf{e},\mathbf{f})\cdot \mathrm{W}_x + \frac{(\hat{\varepsilon}-\theta\cdot p_x)^2}{4\cdot\hat{\varepsilon}} + \theta\cdot A,$$

Welfare is an affine function of the recall-weighted utility of options, $\mathrm{W}(\mathbf{e},\mathbf{f}) \equiv \sum_x r_x(\mathbf{e},\mathbf{f})\cdot \mathrm{W}_x$. The term that does not depend on recall accounts for choices that occur despite no recall, as well as the implications of purchases on the savings component.

The welfare impact of greater exposure e_g to good g is then equal to

$$\frac{\partial\widetilde{\mathrm{W}}(\mathbf{e},\mathbf{f})}{\partial e_g} = \frac{1}{4\cdot\hat{\varepsilon}}\frac{\partial \mathrm{W}(\mathbf{e},\mathbf{f})}{\partial e_g},$$

where:

$$\frac{\partial \mathrm{W}(\mathbf{e},\mathbf{f})}{\partial e_g} = \sum_x \frac{\partial r_x(\mathbf{e},\mathbf{f})}{\partial e_g}\cdot \mathrm{W}_x = \sum_x [r_x(g,\mathbf{f}) - r_x(\mathbf{e},\mathbf{f})]\cdot \mathrm{W}_x$$

$$= \left[r_g(g,\mathbf{f}) - r_g(\mathbf{e},\mathbf{f})\right]\cdot \mathrm{W}_g + \sum_{x\neq g}[r_x(g,\mathbf{f}) - r_x(\mathbf{e},\mathbf{f})]\cdot \mathrm{W}_x =$$

$$= \left[r_g(g,\mathbf{f}) - r_g(\mathbf{e},\mathbf{f})\right]\cdot \left(\mathrm{W}_g - \overline{\mathrm{W}}\right) + \sum_{x\neq g}[r_x(g,\mathbf{f}) - r_x(\mathbf{e},\mathbf{f})]\cdot \left(\mathrm{W}_x - \overline{\mathrm{W}}\right)$$

where the last equality follows from $\sum_x r_x(g,\mathbf{f}) = \sum_x r_x(\mathbf{e},\mathbf{f}) = 1$.

B. Descriptive Statistics

Table B.1 lists the variables used for the 20 product markets in our analysis. Following Shapiro et al (2021), we compute product level statistics and aggregate them across products. Table B.2 presents aggregate descriptive statistics of individual variables (as well as product level statistics on demand), while Table B.3 presents aggregate correlations across variables at the product level. Table B.4 presents these correlations for the colas market. Product level statistics for other markets are available upon request.

Table B.1

Description of the main variables in the model.

Note:

Variable	Description
d_1, d_2	Household's total weekly quantity of purchase of the leading brand 1 and follower brand 2 (ounces or packs)
f_1, f_2	Number of times in the previous year in which the household purchased brand 1 or 2
e_1, e_2	Average weekly expected household exposure to advertising of brand 1 or 2 (seconds)
$e_1 \cdot f_1$ $e_2 \cdot f_2$	Interactions capturing the direct effect of exposure to an ad on demand for that brand
$e_2 \cdot f_1$ $e_1 \cdot f_2$	Interactions capturing the spillover effect of exposure to an ad on demand for the other brand

Table B.2

Summary statistics of the main variables of the model among the 20 products

Notes: Panel A displays the median, mean and other percentiles of the average value among the 20 products of the main variables in our sample. To construct the table, we initially compute the sample mean of each variable for each product, and then on those 20 product estimates construct the statistics above. Panel B presents product level statistics on demand.

Panel A: aggregate statistics

				Percentiles		
	Median	Mean	Min	25%	75%	Max
d_1	3.63	10.72	0.91	2.14	18.14	39.42
d_2	2.34	6.29	0.48	0.94	6.59	27.11

f_1	3.23	3.98	1.75	2.71	5.39	8.08
f_2	2.16	2.41	0.82	1.17	2.58	5.39
e_1	14.26	16.11	2.91	5.69	24.70	37.07
e_2	7.30	9.23	2.44	4.44	11.79	31.00
$e_1 \cdot f_1$	63.32	73.12	5.04	16.56	88.05	269.73
$e_2 \cdot f_1$	20.66	44.81	7.79	12.09	49.24	191.14
$e_1 \cdot f_2$	22.61	40.51	4.04	14.63	51.51	177.86
$e_2 \cdot f_2$	12.34	26.96	2.42	7.26	43.11	126.40

Panel B: product level demand statistics

		$d_1 + d_2$		d_1		d_2	
Product	Unit	mean	std	mean	std	mean	std
beer	OZ	63.14	187.97	37.81	148.69	25.33	119.62
Candy chocolate	OZ	2.66	13.80	2.12	13.00	0.54	3.71
Cereal	OZ	2.92	11.15	2.16	9.30	0.76	5.90
Coffee	OZ	4.19	14.73	2.33	11.37	1.86	9.71
Colas	OZ	66.53	189.14	39.42	142.54	27.11	124.03
Cheese crackers	OZ	1.90	8.30	0.96	5.74	0.94	5.77
Detergent heavy	OZ	8.47	34.96	5.65	27.64	2.47	20.62
Diapers	CT	8.61	38.86	4.68	26.55	3.93	28.15
Dog food dry	OZ	38.73	158.02	33.57	148.31	5.16	56.26
Facial tissue	CT	44.35	222.54	30.24	199.00	14.10	100.59
Frozen pizza	OZ	4.20	16.13	2.62	13.26	1.58	9.27
Lemon colas	OZ	18.40	79.33	10.38	59.08	8.02	53.25
Mouthwash	OZ	3.27	15.68	2.33	13.30	0.94	8.35
Other carbonated Soft drink	OZ	49.08	164.81	25.89	127.02	23.19	102.40
Pain remedy headache	CT	5.69	33.56	3.34	24.93	2.35	22.27

Shampoo	OZ	1.42	7.50	0.94	6.19	0.48	4.21
Soap	OZ	1.86	9.77	0.91	6.29	0.95	7.43
Soup canned	OZ	6.24	24.99	3.92	18.17	2.32	16.39
Toilet tissue	CT	1.69	6.15	1.16	5.32	0.53	3.25
Yogurt	OZ	7.16	20.44	3.97	15.35	3.19	13.57

Table B.3

Average correlations of the main model variables among the 20 products

Notes: the table displays the average correlation of the main variables in our model across the 20 products in the sample defined in the text. We first computed the correlation of each of the 20 products separately, and then computed the average between them.

	d_1	d_2	f_1	f_2	e_1	e_2	$e_1 \cdot f_1$	$e_2 \cdot f_1$	$e_1 \cdot f_2$
d_2	0.00	1.00							
f_1	0.21	-0.04	1.00						
f_2	-0.03	0.24	-0.10	1.00					
e_1	0.01	0.00	0.01	0.00	1.00				
e_2	0.01	0.01	0.01	0.01	0.23	1.00			
$e_1 \cdot f_1$	0.15	-0.02	0.66	-0.07	0.44	0.10	1.00		
$e_2 \cdot f_1$	0.13	-0.02	0.57	-0.05	0.11	0.47	0.52	1.00	
$e_1 \cdot f_2$	-0.02	0.17	-0.07	0.68	0.33	0.08	0.06	0.00	1.00
$e_2 \cdot f_2$	-0.02	0.16	-0.06	0.61	0.08	0.37	-0.01	0.09	0.55

Table B.4

Average correlations of the main variables for colas

Notes: the table displays the average correlation of the main variables in our model for colas in our sample.

	d_1	d_2	f_1	f_2	e_1	e_2	$e_1 \cdot f_1$	$e_2 \cdot f_1$	$e_1 \cdot f_2$
d_2	1.00								
f_1	0.00	1.00							
f_2	0.34	-0.03	1.00						
e_1	-0.03	0.37	-0.02	1.00					

e_2	0.01	0.00	0.00	-0.01	1.00				
$e_1 \cdot f_1$	0.00	0.00	0.00	-0.01	0.24	1.00			
$e_2 \cdot f_1$	0.28	-0.03	0.80	-0.03	0.32	0.07	1.00		
$e_1 \cdot f_2$	0.28	-0.03	0.83	-0.02	0.07	0.29	0.74	1.00	
$e_2 \cdot f_2$	-0.03	0.31	-0.03	0.82	0.25	0.05	0.04	-0.01	1.00

Next, Figure B.1 plots the time variation of ad exposure in the colas market, as well as its cross-section variability. There is substantial time variation, consistent with Table B.2 Panel A, but also substantial variation across DMAs at a given point in time.

Figure B.1: ad exposure to colas brands

Note: The figure shows summary statistics for our measure of expected exposure to advertising, calculated at the household level. For each week of the year, we obtained the average exposure across all the years included in the analysis. The figure shows the median and the interquartile range of this measure.

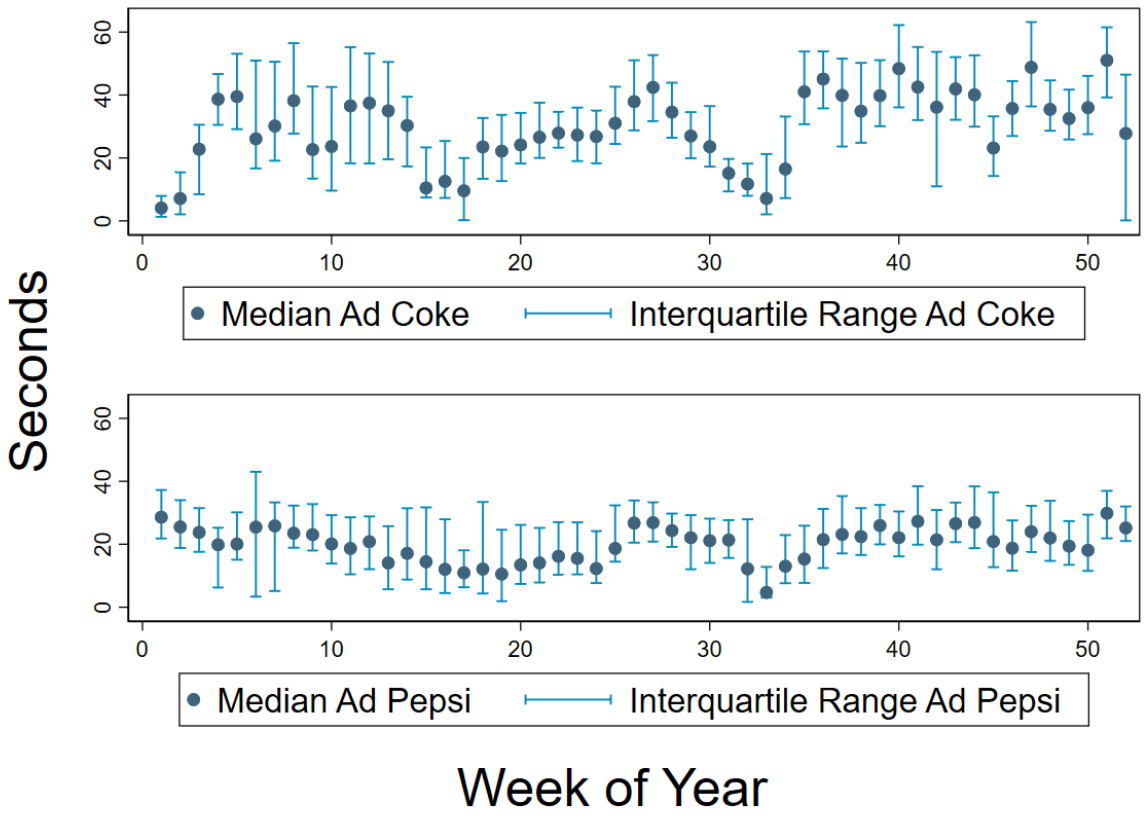
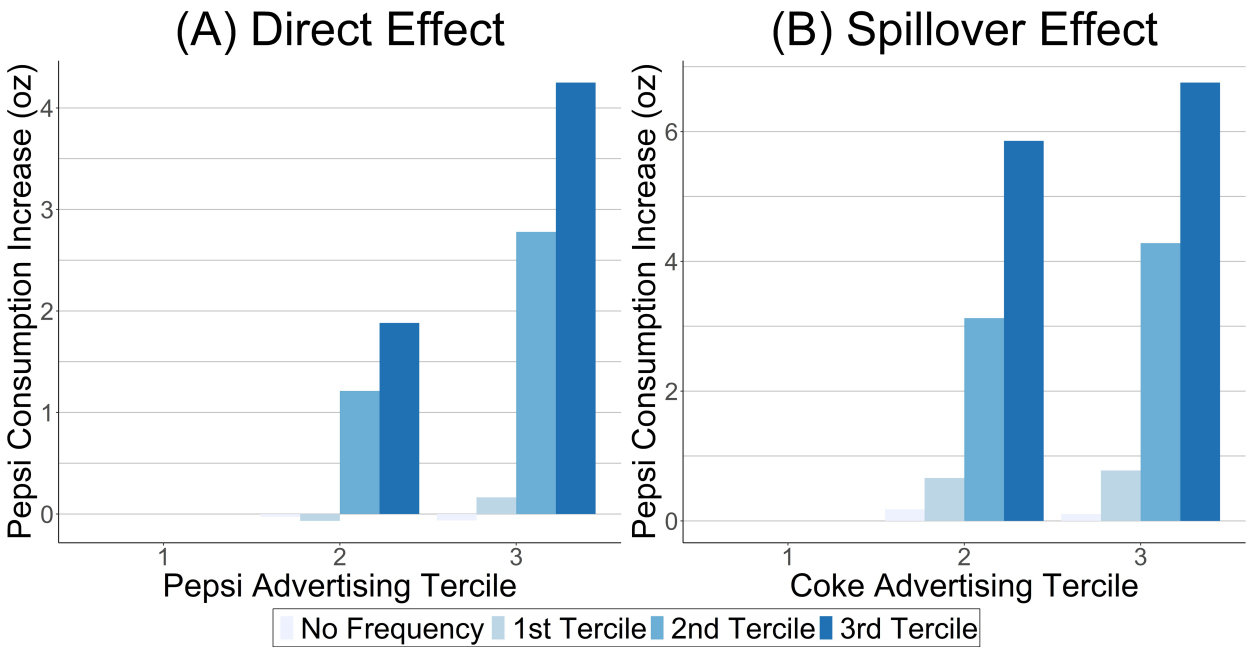


Figure B.2 reports the purchased quantity of Coke (top panels) and Pepsi (bottom panels), as a function of a household's *Exposure* to advertising (left) and past purchase *Frequency* (right) of the brand. The left panels focus on weeks during which advertising of the alternative brand is in its bottom quintile, isolating the association between a brand's consumption and its advertising intensity.

Figure B.2: The role of ads for sporadic and habitual Pepsi consumers.



Notes: The bars display the mean weekly increase in consumption of Pepsi in different terciles of *Exposure* to Pepsi (left panel) or Coke (right panel) ads, compared to in the first *Exposure* tercile, grouped by different levels of Pepsi *Frequency*. For each group, the average consumption of Pepsi at the lowest tercile is set to 0. The Frequency groups include the observations for households in the three terciles of Pepsi, among those who bought Pepsi at least once in the previous year. We include also a column for those that did not purchase Pepsi in the previous year.

Table B.5

Baseline regression at the brand level

Notes: This table estimates the regression coefficients for equation (4) one brand at the time, with ads measured as expected weekly exposure, for all brands in our sample. Dependent variables are standardized. Panel A reports the brand regressions for colas, Panel B reports the 40 brand level coefficients from the corresponding column (2) and (4) as in Panel A for each product, with superscripts *** indicating significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the week-year-DMA level. Panel C presents the percentage of brands for which: the estimated coefficients are significant at 5% or not, in each case split between positive and negative. The specification is that of table 2 in the main text, but for a single brand, and repeated for the 40 brands in our sample.

Panel A: Colas Regression

VARIABLES	(1) d_{Coke}	(2) d_{Coke}	(3) d_{Pepsi}	(4) d_{Pepsi}	(2) $d_{Coke}+d_{Pepsi}$
f_{Coke}	0.0261***	0.0261***		-0.00184***	0.0185***
	(0.000222)	(0.000302)		(0.000109)	(0.000234)
f_{Pepsi}		-0.00208***	0.0350***	0.0346***	0.0211***
		(0.000114)	(0.000319)	(0.000391)	(0.000270)
e_{Coke}	-2.67e-05***	-7.08e-06**		3.11e-06	-3.30e-06
	(2.48e-06)	(2.82e-06)		(2.76e-06)	(2.86e-06)
e_{Pepsi}		-8.21e-06***	-6.86e-06***	-2.23e-06	-7.65e-06***
		(1.67e-06)	(1.25e-06)	(1.49e-06)	(1.66e-06)
$e_{Coke} \cdot f_{Coke}$	7.26e-06***	3.57e-06***		-8.29e-07***	2.15e-06***
	(3.26e-07)	(3.89e-07)		(1.99e-07)	(3.19e-07)
$e_{Pepsi} \cdot f_{Coke}$		2.63e-06***		4.38e-08	2.01e-06***
		(2.60e-07)		(1.18e-07)	(2.19e-07)
$e_{Coke} \cdot f_{Pepsi}$		-4.03e-07***		3.65e-06***	2.09e-06***
		(1.38e-07)		(6.04e-07)	(4.08e-07)
$e_{Pepsi} \cdot f_{Pepsi}$		-4.01e-08	2.78e-06***	1.34e-06***	8.49e-07***
		(8.19e-08)	(2.84e-07)	(3.40e-07)	(2.33e-07)
$e_{Coke} \cdot f_{Coke}^2$	-7.92e-08***	-3.59e-08***		1.47e-08***	-1.74e-08***
	(3.61e-09)	(4.41e-09)		(2.46e-09)	(3.62e-09)
$e_{Pepsi} \cdot f_{Coke}^2$		-3.24e-08***		-2.33e-09	-2.60e-08***
		(3.75e-09)		(1.64e-09)	(3.17e-09)
$e_{Coke} \cdot f_{Pepsi}^2$		4.59e-09***		-4.20e-08***	-2.41e-08***
		(1.16e-09)		(8.21e-09)	(5.41e-09)
$e_{Pepsi} \cdot f_{Pepsi}^2$		2.21e-09***	-4.48e-08***	-2.54e-08***	-1.50e-08***
		(6.95e-10)	(3.55e-09)	(5.04e-09)	(3.30e-09)
Constant	-1.462***	-1.631***	-1.570***	-1.800***	-2.409***
	(0.0904)	(0.110)	(0.0962)	(0.113)	(0.114)
Observations	3,894,746	3,894,108	3,894,215	3,894,108	3,894,108
R-squared	0.123	0.125	0.146	0.147	0.114

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Panel B. Brand level coefficients

Product	f_1	f_2	e_1	e_2	$e_1 \cdot f_1$	$e_2 \cdot f_2$	$e_2 \cdot f_1$	$e_1 \cdot f_2$

bud	2.29E-02***	-9.57E-04***	-1.07E-03***	-2.47E-03***	2.45E-04***	5.39E-04***	-1.24E-05*	8.84E-06
miller	4.02E-02***	-8.19E-04***	-1.63E-03***	-4.86E-04***	5.03E-04***	1.98E-04***	2.68E-05**	-3.53E-07
mnm	1.56E-02***	-1.37E-03***	9.82E-05	-6.03E-04***	4.16E-05***	1.16E-04***	-3.28E-05**	2.49E-05***
reese	3.44E-02***	-3.38E-04	-8.66E-04***	9.13E-05	3.67E-04***	-1.64E-04***	4.43E-05***	-1.41E-05
cheerios	4.36E-02***	-9.51E-05	-2.87E-04***	3.84E-04***	1.04E-04***	-7.28E-05	-3.85E-05**	-1.19E-04***
post	6.78E-02***	-6.90E-04***	-2.06E-04	3.01E-05	7.55E-04***	3.20E-04***	-4.85E-05*	-2.76E-05***
folgers	4.45E-02***	-4.30E-03***	4.46E-04**	1.55E-04	1.62E-04***	6.71E-04***	-2.99E-04***	-5.51E-04***
maxwell	5.22E-02***	-4.55E-03***	-1.77E-04	-7.98E-05	6.62E-04***	2.39E-04***	-3.73E-04***	-1.85E-04***
coke	2.55E-02***	-1.99E-03***	-1.83E-04***	-5.13E-04***	9.16E-05***	1.01E-04***	-1.75E-05***	7.61E-06**
pepsi	3.29E-02***	-1.83E-03***	-2.64E-04***	-1.17E-04**	1.01E-04***	8.49E-05***	-1.08E-05**	-9.26E-06***
pepperidge	4.61E-02***	7.23E-04**	2.78E-04	6.30E-04***	-4.38E-04***	-3.90E-04***	5.46E-05	-9.08E-05***
sunshine	3.86E-02***	-1.43E-03***	-2.76E-04	-1.08E-04	3.59E-04***	-6.51E-06	-2.09E-04***	8.85E-05
tide	4.09E-02***	-1.24E-02***	3.82E-04***	4.48E-04	9.07E-05**	6.00E-05	-3.37E-04***	-2.77E-04**
purex	1.01E-01***	-6.85E-03***	1.24E-04	3.64E-05	-1.35E-04	-8.82E-06	-1.98E-04**	-1.18E-04***
pampers	3.98E-02***	-7.95E-03***	-4.63E-04	4.50E-04	3.99E-04**	3.87E-04	1.50E-04*	-5.60E-04***
huggies	3.13E-02***	-8.69E-03***	-1.95E-03***	6.74E-05	1.12E-03***	-1.17E-04	-2.24E-04	1.17E-04
purina	3.05E-02***	-4.95E-03***	3.65E-04**	4.26E-04	-2.13E-04***	-1.78E-04	-1.34E-04***	-3.45E-04*
iams	9.40E-02***	-3.16E-03***	-4.49E-04	1.78E-04	1.91E-04	-5.71E-05	-6.48E-05	-1.12E-04***
kleenex	3.10E-02***	-6.53E-03***	-3.72E-04**	2.27E-04	5.68E-04***	3.84E-04***	-4.53E-04***	-3.96E-04***
puffs	4.96E-02***	-4.42E-03***	2.44E-04	-7.12E-04***	6.65E-04***	7.65E-04***	-3.15E-04***	-1.56E-04***
digiorno	4.42E-02***	-3.41E-03***	-1.85E-04	8.20E-04***	4.47E-04***	-3.93E-04***	-2.85E-04***	-8.21E-05**
redbaron	5.61E-02***	-2.62E-03***	8.55E-04***	3.62E-04**	-1.26E-04	-1.03E-04	-1.21E-04***	-1.96E-04***
sprite	4.39E-02***	-6.38E-04**	-8.70E-04***	-1.14E-03***	1.04E-03***	5.75E-04***	-1.87E-04***	-7.04E-05*
sevenup	6.95E-02***	-1.11E-03***	1.83E-04	2.36E-03***	-6.94E-06	-1.24E-03***	-7.36E-05**	-1.75E-04***
listerine	4.74E-02***	-9.87E-03***	-1.64E-04	-6.53E-04*	5.98E-06	3.09E-04**	-1.85E-04***	-3.12E-04***
crest	7.73E-02***	-6.21E-03***	-6.72E-04*	6.91E-05	6.33E-04***	3.83E-04***	-1.24E-04*	-1.36E-04***
mountaindew	2.83E-02***	-2.45E-03***	-1.14E-03***	-8.83E-04***	3.46E-04***	2.66E-04***	1.54E-05*	1.47E-05*
drpepper	3.33E-02***	-2.04E-03***	-1.35E-03***	-1.53E-03***	3.71E-04***	4.07E-04***	1.99E-05**	2.01E-05**
aleve	3.90E-02***	-7.77E-03***	-2.95E-04**	2.25E-04	3.40E-04***	-2.45E-04	-1.23E-04***	-1.56E-04*
tylenol	9.06E-02***	-4.88E-04	5.50E-04*	3.52E-04**	-2.28E-04	-1.58E-04	-2.90E-04***	-1.98E-04***
suave	5.63E-02***	-6.51E-03***	-2.87E-04	-7.40E-04***	1.74E-04	7.48E-04***	1.17E-04	-1.80E-04
pantene	8.68E-02***	-1.70E-03**	4.25E-04	3.01E-04	-2.69E-04	-1.07E-04	-1.22E-04	-2.84E-04**
suave	6.08E-02***	-3.97E-03***	3.37E-04	-3.29E-04	4.37E-04	-3.66E-05	-1.04E-04	-7.97E-06
dove	7.05E-02***	-6.04E-03***	-2.25E-07	7.14E-05	-2.31E-05	1.53E-04	-2.38E-04***	-2.66E-04
campbell	2.87E-02***	6.22E-04***	-2.76E-04***	-8.51E-04***	1.79E-04***	2.52E-04***	-6.91E-05***	1.09E-04***
progresso	4.18E-02***	-4.09E-05	-5.51E-04***	-2.98E-04***	1.58E-03***	7.16E-04***	-1.73E-04***	-1.23E-04***
charmin	3.92E-02***	-7.73E-03***	3.37E-04***	2.17E-04	7.94E-05***	2.95E-05	-2.11E-04***	-1.07E-04***
angelsoft	6.35E-02***	-8.62E-03***	3.97E-05	-2.73E-04**	7.68E-05	2.64E-04***	-6.02E-05**	-7.70E-05***

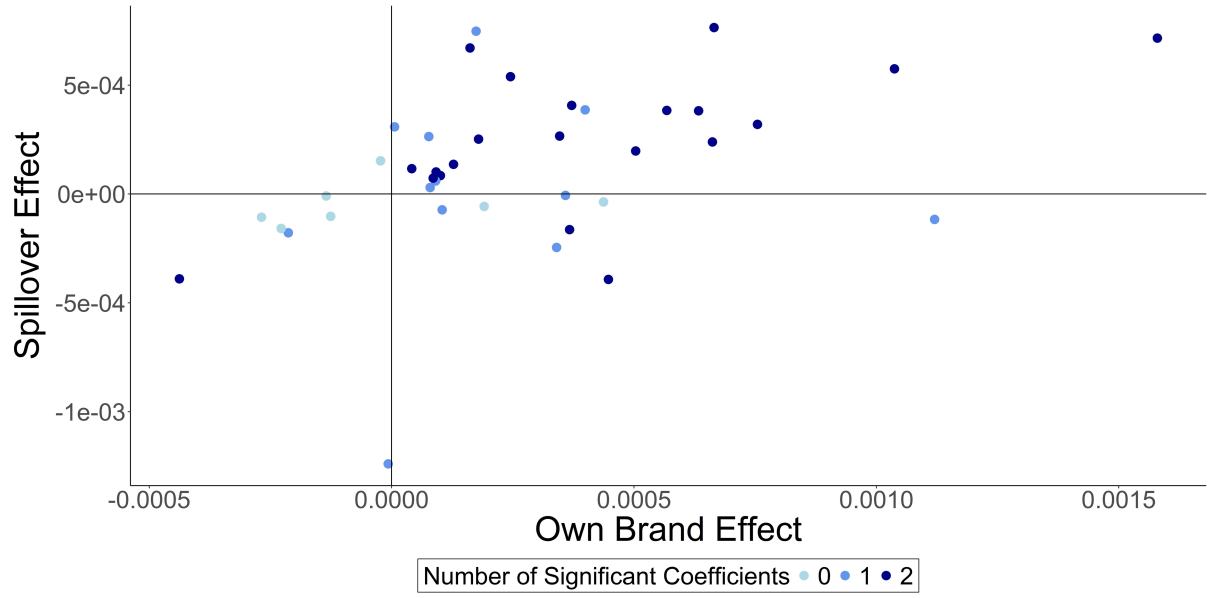
Panel C. Summary of baseline regression for the 40 brands.

Coefficients of 40 brands				
	% $p \geq 0.05$		% $p < 0.05$	
	> 0	**< 0**	**> 0**	**< 0**
f_1	0%	0%	**100%**	0%
f_2	0%	10%	5%	**85%**
e_1	23%	25%	15%	38%
e_2	38%	10%	15%	38%
$e_1 \cdot f_1$	13%	15%	**68%**	5%
$e_2 \cdot f_1$	10%	28%	**53%**	10%
$e_1 \cdot f_2$	10%	18%	8%	**65%**

$e_2 \cdot f_2$	10%	20%	10%	**60%**
$e_1 \cdot f_1^2$	18%	18%	5%	**60%**
$e_2 \cdot f_1^2$	15%	18%	15%	**53%**
$e_1 \cdot f_2^2$	23%	8%	**65%**	5%
$e_2 \cdot f_2^2$	23%	10%	**65%**	3%

Figure B.3: Estimates for the direct and spillover effects in each product

Notes: each of the 40 points in the graph represents the combination of the estimates of the linear interactions associated with the frequency of a different brand within the 40 brand regressions. For example, one point in the plot comes from the coke regression, with (x,y) coordinates given by the coefficients on $e_{Coke} \cdot f_{Coke}$ and $e_{Pepsi} \cdot f_{Coke}$. Significance displayed is at 95% confidence level.



C. Robustness

Here we collect results on the analysis of Section 4.2 and the robustness tests in Section 4.3, and also present further analysis.

C.1 Baseline analysis

Table C.1 presents the results of our baseline regression (1) for colas, based on a different measure of advertising, that is the total weekly count of ad occurrences for each brand in each DMA, as described in section 2.1. This measure does not account for the variability among demographics other than the

different DMAs. Table C.2 summarizes the results of this alternative definition of advertising across the 40 brands.

Table C.1

Colas Regression, Occurrences

Notes: the regression includes the same controls as our baseline specification, with ads measured through occurrences. Superscripts: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the week-year-DMA level.

VARIABLES	(1) d_{Coke}	(2) d_{Coke}	(3) d_{Pepsi}	(4) d_{Pepsi}	(2) $d_{Coke} + d_{Pepsi}$
f_{Coke}	0.0261***	0.0261***		-0.00184***	0.0185***
	(0.000222)	(0.000302)		(0.000109)	(0.000234)
f_{Pepsi}		-0.00208***	0.0350***	0.0346***	0.0211***
		(0.000114)	(0.000319)	(0.000391)	(0.000270)
e_{Coke}	-2.67e-05***	-7.08e-06**		3.11e-06	-3.30e-06
	(2.48e-06)	(2.82e-06)		(2.76e-06)	(2.86e-06)
e_{Pepsi}		-8.21e-06***	-6.86e-06***	-2.23e-06	-7.65e-06***
		(1.67e-06)	(1.25e-06)	(1.49e-06)	(1.66e-06)
$e_{Coke} \cdot f_{Coke}$	7.26e-06***	3.57e-06***		-8.29e-07***	2.15e-06***
	(3.26e-07)	(3.89e-07)		(1.99e-07)	(3.19e-07)
$e_{Pepsi} \cdot f_{Coke}$		2.63e-06***		4.38e-08	2.01e-06***
		(2.60e-07)		(1.18e-07)	(2.19e-07)
$e_{Coke} \cdot f_{Pepsi}$		-4.03e-07***		3.65e-06***	2.09e-06***
		(1.38e-07)		(6.04e-07)	(4.08e-07)
$e_{Pepsi} \cdot f_{Pepsi}$		-4.01e-08	2.78e-06***	1.34e-06***	8.49e-07***
		(8.19e-08)	(2.84e-07)	(3.40e-07)	(2.33e-07)
$e_{Coke} \cdot f_{Coke}^2$	-7.92e-08***	-3.59e-08***		1.47e-08***	-1.74e-08***
	(3.61e-09)	(4.41e-09)		(2.46e-09)	(3.62e-09)
$e_{Pepsi} \cdot f_{Coke}^2$		-3.24e-08***		-2.33e-09	-2.60e-08***
		(3.75e-09)		(1.64e-09)	(3.17e-09)
$e_{Coke} \cdot f_{Pepsi}^2$		4.59e-09***		-4.20e-08***	-2.41e-08***
		(1.16e-09)		(8.21e-09)	(5.41e-09)
$e_{Pepsi} \cdot f_{Pepsi}^2$		2.21e-09***	-4.48e-08***	-2.54e-08***	-1.50e-08***
		(6.95e-10)	(3.55e-09)	(5.04e-09)	(3.30e-09)
Constant	-1.462***	-1.631***	-1.570***	-1.800***	-2.409***
	(0.0904)	(0.110)	(0.0962)	(0.113)	(0.114)
Observations	3,894,746	3,894,108	3,894,215	3,894,108	3,894,108
R-squared	0.123	0.125	0.146	0.147	0.114

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table C.2

Summary of baseline regression for the 40 brands: occurrences

Notes: For each of the main coefficients in our regression we compute the percentage of brands for which: the estimated coefficients are significant at 5% or not, in each case split between positive and negative. The specification is exactly that of table 2 in the main text, except that ads are not measured as exposure but as total weekly occurrences in the dma. The regressions are repeated for the 40 brands in our sample.

Coefficients of 40 brands, occurrences				
	% $p \geq 0.05$		% $p < 0.05$	
	> 0	**< 0**	**> 0**	**< 0**
f_1	0%	0%	**100**%	0%
f_2	5%	10%	3%	**83**%
e_1	23%	23%	15%	40%
e_2	40%	18%	10%	33%
$e_1 \cdot f_1$	15%	10%	**70**%	5%
$e_2 \cdot f_2$	20%	8%	**55**%	18%
$e_2 \cdot f_1$	15%	13%	8%	**65**%
$e_1 \cdot f_2$	8%	23%	10%	**60**%
$e_1 \cdot f_1^2$	20%	13%	8%	**60**%
$e_2 \cdot f_1^2$	20%	10%	13%	**58**%
$e_1 \cdot f_2^2$	18%	8%	**68**%	8%
$e_2 \cdot f_2^2$	18%	5%	**73**%	5%

C.2 Sample of households with recent experiences

We reproduce Table 2 (on the baseline analysis for the pooled sample) and Table B.5 Panels A and B (on colas, and the summary results for our 40 brands) using the sample of households who purchase a product at least once in the previous year. Unlike our baseline sample, this does not condition on uninterrupted consumption of the product from first observed purchased of the product. The resulting sample is larger, including 100,942,238 (4,949,645 for colas) household-week observations. This is 40% (27%) more than those included in our baseline sample from Table 2 (n=72,116,888). Notice that our results are also robust to another sample selection criterion (available upon request) where we only analyze decisions of those households who have a positive consumption of either brand in 2015. This is the sample of those who have some positive memories in the first year of our analysis, independently from the fact that they keep purchasing the products in subsequent years. In this way we address the concern that our sample selection

only includes those on whom advertising works, since they keep buying the products over the entire timespan we analyze.

Table C.3: Pooled regression: recent experiences sample

Notes: This table estimates the regression coefficients for columns (1) and (2) of Table 2 with ads measured as expected weekly exposure, with the recent experience sample. Dependent variables are standardized. Robust standard errors are shown in parentheses. Superscripts: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the week-year-DMA level.

	(1)		(2)
VARIABLES	d_τ	VARIABLES	$d_1 + d_2$
f_τ	0.226***	f_1	0.149***
	(0.000315)		(0.000287)
$f_{-\tau}$	-0.00784***	f_2	0.157***
	(0.000104)		(0.000384)
e_τ	0.00655***	e_1	0.00416***
	(0.000161)		(0.000248)
$e_{-\tau}$	-0.00144***	e_2	0.00270***
	(0.000175)		(0.000287)
$e_\tau \cdot f_\tau$	0.0129***	$e_1 \cdot f_1$	0.0178***
	(0.000393)		(0.000348)
$e_{-\tau} \cdot f_\tau$	0.00776***	$e_2 \cdot f_1$	-0.00674***
	(0.000392)		(0.000393)
$e_\tau \cdot f_{-\tau}$	0.000219	$e_1 \cdot f_2$	0.0173***
	(0.000143)		(0.000449)
$e_{-\tau} \cdot f_{-\tau}$	-0.000110	$e_2 \cdot f_2$	-0.00608***
	(0.000166)		(0.000527)
$e_\tau \cdot f_\tau^2$	-0.000800***	$e_1 \cdot f_1^2$	-0.000889***
	(8.11e-05)		(6.25e-05)
$e_{-\tau} \cdot f_\tau^2$	-0.000520***	$e_2 \cdot f_1^2$	-0.000105**
	(7.20e-05)		(5.28e-05)
$e_\tau \cdot f_{-\tau}^2$	-6.97e-05***	$e_1 \cdot f_2^2$	-0.000197**
	(2.06e-05)		(9.87e-05)
$e_{-\tau} \cdot f_{-\tau}^2$	0.000127***	$e_2 \cdot f_2^2$	0.000174
	(2.90e-05)		(0.000114)
	-0.0453***		-0.0620***
Observations	100,942,238	49,599,747	Observations
R-squared	0.054	0.050	R-squared

Table C.4

Panel A: Colas Regression: recent experience sample

Notes: This table estimates the regression coefficients for columns (2) and (4) and (5) of Table 3 with ads measured as expected weekly exposure, with the recent experience sample. Dependent variables are standardized. Robust standard errors are shown in parentheses. Superscripts: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the week-year-DMA level.

	(1)	(2)	(3)

VARIABLES	d_{Coke}	d_{Pepsi}	$d_{Coke} + d_{Pepsi}$
f_{Coke}	0.0286***	-0.00160***	0.0204***
	(0.000260)	(9.16e-05)	(0.000205)
f_{Pepsi}	-0.00163***	0.0365***	0.0225***
	(9.50e-05)	(0.000383)	(0.000262)
e_{Coke}	-0.000233***	-0.000151***	-0.000273***
	(4.32e-05)	(4.11e-05)	(4.19e-05)
e_{Pepsi}	-0.000477***	-0.000260***	-0.000527***
	(6.87e-05)	(6.08e-05)	(6.60e-05)
$e_{Coke} \cdot f_{Coke}$	0.000103***	-4.90e-06	7.42e-05***
	(9.27e-06)	(3.07e-06)	(7.07e-06)
$e_{Pepsi} \cdot f_{Coke}$	0.000109***	-1.24e-05***	7.41e-05***
	(1.43e-05)	(4.78e-06)	(1.13e-05)
$e_{Coke} \cdot f_{Pepsi}$	-9.80e-06***	9.54e-05***	5.48e-05***
	(2.53e-06)	(1.48e-05)	(9.75e-06)
$e_{Pepsi} \cdot f_{Pepsi}$	6.74e-06*	0.000112***	7.80e-05***
	(3.71e-06)	(1.93e-05)	(1.30e-05)
$e_{Coke} \cdot f_{Coke}^2$	-9.09e-07***	1.41e-07***	-5.91e-07***
	(1.40e-07)	(4.28e-08)	(1.04e-07)
$e_{Pepsi} \cdot f_{Coke}^2$	-1.20e-06***	7.70e-08	-8.52e-07***
	(1.97e-07)	(5.74e-08)	(1.56e-07)
$e_{Coke} \cdot f_{Pepsi}^2$	1.12e-07***	-8.64e-07***	-4.78e-07***
	(2.52e-08)	(2.73e-07)	(1.74e-07)
$e_{Pepsi} \cdot f_{Pepsi}^2$	-5.45e-08	-1.25e-06***	-8.55e-07***
	(3.57e-08)	(3.01e-07)	(2.01e-07)
Constant	-1.448***	-1.610***	-2.135***
	(0.0938)	(0.0973)	(0.0975)
Observations	4,949,645	4,949,645	4,949,645
R-squared	0.131	0.149	0.127

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Panel B: Summary of baseline regression for 40 brands: recent experience sample

Notes: For each of the main coefficients in our regression we compute the percentage of products for which: the estimated coefficients are not significant at 5%, the estimated coefficients are positive and significant at 5%, the estimated coefficients that are negative and significant at 5%. Relative to Table 3, count of ad occurrences is used instead of ad exposure.

Coefficients of 40 brands, Recent Experiences				
	% $p \geq 0.05$		% $p < 0.05$	
	> 0	**< 0**	**> 0**	**< 0**
f_1	0%	0%	**100**%	0%
f_2	8%	10%	23%	**60**%
e_1	15%	23%	13%	50%
e_2	25%	23%	8%	45%
$e_1 \cdot f_1$	13%	10%	**68**%	10%

$e_2 \cdot f_2$	5%	28%	**58**%	10%
$e_2 \cdot f_1$	20%	25%	18%	**38**%
$e_1 \cdot f_2$	13%	40%	18%	**30**%
$e_1 \cdot f_1^2$	18%	18%	8%	**58**%
$e_2 \cdot f_1^2$	18%	20%	13%	**50**%
$e_1 \cdot f_2^2$	20%	28%	**38**%	15%
$e_2 \cdot f_2^2$	20%	23%	**43**%	15%

C.3 Further robustness tests

We start with a test that considers household j's relative demand for the leading brand of product k in week t, $d_1 - d_2$ as the dependant variable, for our baseline specification for colas. The regression specification is:

$$d_1 - d_2 = \hat{\alpha}_1 \cdot f_1 + \hat{\alpha}_1 \cdot f_2 + \hat{\beta}_1 \cdot e_1 + \hat{\beta}_2 e_2 +$$

$$\hat{\gamma}_1 \cdot e_1 \cdot f_1 + \hat{\gamma}_2 \cdot e_2 \cdot f_2 + \hat{\gamma}_3 \cdot e_2 \cdot f_1 + \hat{\gamma}_4 \cdot e_1 \cdot f_2 +$$

$$\hat{\delta}_1 \cdot e_1 \cdot f_1^2 + \hat{\delta}_2 \cdot e_2 \cdot f_2^2 + \hat{\delta}_3 \cdot e_2 \cdot f_1^2 + \hat{\delta}_4 \cdot e_1 \cdot f_2^2 + X'\hat{\beta} + \varepsilon.$$

for brands $b = 1,2$. The signs of the coefficients should match those for our baseline specification (1), according to Propositions 1 through 3 and following the discussion in Section 4.1. Table C.5 confirms these results for colas. The results for other products are similar.

Table C.5
Baseline regression for colas, difference regression.

Notes: the regression includes the same controls as our baseline specification, with ads measured through exposure. Column (1) displays the results with ads measured as expected weekly exposure, whereas column (2) does the same with ads measured as total weekly occurrences. Superscripts: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the week-year-DMA level.

	(1)	(2)
	$d_{Coke} - d_{Pepsi}$	$d_{Coke} - d_{Pepsi}$
f_{Coke}	0.0205***	0.0209***
	(0.000200)	(0.000244)

f_{Pepsi}	-0.0232***	-0.0243***
	(0.000247)	(0.000271)
e_{Coke}	-6.12e-05	-7.40e-06***
	(5.33e-05)	(2.72e-06)
e_{Pepsi}	-0.000214***	-4.74e-06***
	(8.03e-05)	(1.53e-06)
$e_{Coke} \cdot f_{Coke}$	7.53e-05***	3.24e-06***
	(7.31e-06)	(3.23e-07)
$e_{Pepsi} \cdot f_{Coke}$	8.37e-05***	1.96e-06***
	(1.08e-05)	(2.02e-07)
$e_{Coke} \cdot f_{Pepsi}$	-6.90e-05***	-2.70e-06***
	(1.03e-05)	(4.11e-07)
$e_{Pepsi} \cdot f_{Pepsi}$	-6.03e-05***	-9.12e-07***
	(1.29e-05)	(2.31e-07)
$e_{Coke} \cdot f_{Coke}^2$	-7.00e-07***	-3.67e-08***
	(1.04e-07)	(3.77e-09)
$e_{Pepsi} \cdot f_{Coke}^2$	-8.80e-07***	-2.30e-08***
	(1.38e-07)	(2.88e-09)
$e_{Coke} \cdot f_{Pepsi}^2$	6.36e-07***	3.11e-08***
	(1.78e-07)	(5.51e-09)
$e_{Pepsi} \cdot f_{Pepsi}^2$	7.32e-07***	1.84e-08***
	(1.87e-07)	(3.40e-09)
Constant	-0.0443	-0.0485
	(0.108)	(0.108)
Observations	3,894,108	3,894,108
R-squared	0.154	0.154

Table C.6
Baseline regression for pooled, difference regression.

Notes: the regression includes the same controls as our baseline specification, with ads measured through exposure. Column (1) displays the results with ads measured as expected weekly exposure, whereas column (2) does the same with ads measured as total weekly occurrences. Superscripts: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the week-year-DMA-product level.

	(2)	(3)
	$d_{Own} - d_{Other}$	d_{Own}- d_{Other}
f_{Own}	0.177***	0.176***
	(0.000310)	(0.000309)
f_{Other}	-0.150***	-0.150***
	(0.000327)	(0.000327)
e_{Own}	0.00291***	0.00121***
	(0.000273)	(0.000285)
e_{Other}	0.000237	-0.000809***
	(0.000291)	(0.000303)
$e_{Own} \cdot f_{Own}$	0.00823***	0.00734***
	(0.000484)	(0.000515)
$e_{Other} \cdot f_{Own}$	0.00735***	0.00284***
	(0.000498)	(0.000555)
$e_{Own} \cdot f_{Other}$	-0.00524***	-0.00692***
	(0.000513)	(0.000534)
$e_{Other} \cdot f_{Other}$	-0.00549***	-0.00671***
	(0.000528)	(0.000536)
$e_{Own} \cdot f_{Own}^2$	-0.000762***	-0.000692***
	(0.000122)	(0.000132)
$e_{Other} \cdot f_{Own}^2$	-0.00101***	1.21e-05
	(0.000118)	(0.000146)
$e_{Own} \cdot f_{Other}^2$	-0.000172*	0.000210**
	(0.000103)	(0.000103)

$e_{Other} \cdot f_{Other}^2$	-0.000199*	0.000432***
	(0.000104)	(0.000103)
Constant	-0.226***	-0.224***
	(0.00804)	(0.00784)
Observations	36,058,444	36,552,157
R-squared	0.059	0.059

As an additional robustness check, we estimate Equation (4) by controlling for a brand's or a product's aggregate consumption in the DMA each week, alone and interacted with the household's past purchase frequency. Weekly variation of DMA-level consumption of a product is a highly granular proxy for fluctuations in local demand for the brand to which advertisers may respond, allowing us to isolate the role of household-specific ad exposure. These are challenging tests: part of the variation in local demand can be causally induced by ads themselves, according to our model. Table C.7 shows that the estimated coefficients barely change when we add these controls in the colas regression (see also column (5) of Table C.9 for the pooled version), further assuaging the reverse causality concern.

Table C.7: Controlling for DMA-level Demand in Colas

Notes: the regressions include controls for weekly purchases of colas in the DMA, first only at baseline (1) and (3) and then interacted with frequency (2) and (4). The average weekly purchases by DMA, week, and year is computed considering every purchase in the DMA, regardless of whether the household making the purchase is included in our sample or not. Note that ads and Frequency are the same as the previous analysis, as we consider the same sample. Robust standard errors are shown in parentheses. Superscripts: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the week-year-DMA-product level.

	(1)	(2)	(3)	(4)	(5)
	d_{Coke}	d_{Coke}	d_{Pepsi}	d_{Pepsi}	$d_{Coke} + d_{Pepsi}$
q_{Colas}	0.00332***	0.00165***	0.00258***	0.00114***	0.00199***
	(3.51e-05)	(4.86e-05)	(3.51e-05)	(4.35e-05)	(4.69e-05)
$q_{Colas} \cdot f_{Coke}$		0.000217***		-3.02e-05***	0.000144***
		(4.69e-06)		(2.13e-06)	(3.83e-06)

$q_{Colas} \cdot f_{Pepsi}$		-3.77e-05***		0.000298***	0.000167***
		(2.24e-06)		(7.43e-06)	(5.32e-06)
f_{Coke}	0.0253***	0.00820***	-0.00198***	0.000410**	0.00645***
	(0.000249)	(0.000358)	(8.94e-05)	(0.000170)	(0.000300)
f_{Pepsi}	-0.00209***	0.000782***	0.0329***	0.0101***	0.00719***
	(9.26e-05)	(0.000179)	(0.000358)	(0.000589)	(0.000422)
e_{Coke}	-0.000306***	-0.000387***	-0.000213***	-0.000228***	-0.000442***
	(5.33e-05)	(5.46e-05)	(5.20e-05)	(5.41e-05)	(5.28e-05)
e_{Pepsi}	-0.000475***	-0.000682***	-0.000235***	-0.000281***	-0.000698***
	(8.33e-05)	(8.35e-05)	(7.49e-05)	(7.95e-05)	(8.04e-05)
$e_{Coke} \cdot f_{Coke}$	9.27e-05***	0.000109***	-8.42e-06***	-1.13e-05***	7.50e-05***
	(8.87e-06)	(9.22e-06)	(3.06e-06)	(3.14e-06)	(6.97e-06)
$e_{Pepsi} \cdot f_{Coke}$	0.000104***	0.000136***	-9.15e-06**	-1.17e-05**	9.52e-05***
	(1.36e-05)	(1.38e-05)	(4.60e-06)	(4.85e-06)	(1.09e-05)
$e_{Coke} \cdot f_{Pepsi}$	-1.76e-05***	-1.72e-05***	8.48e-05***	9.09e-05***	4.66e-05***
	(2.69e-06)	(2.76e-06)	(1.46e-05)	(1.57e-05)	(9.89e-06)
$e_{Pepsi} \cdot f_{Pepsi}$	8.12e-06**	9.16e-06**	0.000101***	0.000113***	8.07e-05***
	(3.86e-06)	(3.91e-06)	(1.86e-05)	(1.98e-05)	(1.29e-05)
$e_{Coke} \cdot f_{Coke}^2$	-7.91e-07***	-9.72e-07***	1.58e-07***	1.96e-07***	-6.04e-07***
	(1.25e-07)	(1.32e-07)	(3.96e-08)	(4.19e-08)	(9.68e-08)
$e_{Pepsi} \cdot f_{Coke}^2$	-1.09e-06***	-1.26e-06***	8.59e-08*	7.23e-08	-9.01e-07***
	(1.77e-07)	(1.85e-07)	(5.16e-08)	(5.65e-08)	(1.46e-07)
$e_{Coke} \cdot f_{Pepsi}^2$	1.69e-07***	1.89e-07***	-7.73e-07***	-9.20e-07***	-4.61e-07***
	(3.00e-08)	(3.42e-08)	(2.49e-07)	(2.83e-07)	(1.71e-07)
$e_{Pepsi} \cdot f_{Pepsi}^2$	-8.92e-09	-2.94e-08	-1.13e-06***	-1.14e-06***	-7.72e-07***
	(3.36e-08)	(3.71e-08)	(2.75e-07)	(2.97e-07)	(1.92e-07)
Constant	-0.904***	-0.881***	-1.238***	-0.759***	-1.162***
	(0.108)	(0.108)	(0.110)	(0.108)	(0.110)
Observations	3,894,072	3,894,072	3,894,072	3,894,072	3,894,072
	0.129	0.134	0.149	0.155	0.124

Table C.8: Robustness for Geographical and Time controls for Coke.

Notes: The table estimates Columns 1, 2, 3 and 6 of Table 3 for coke. Column (1) represents our baseline regression with month of year and state dummies. Column (2) substitutes the month of year with week of year dummies and state dummies with DMA dummies, including month-year dummies too. Column (3) includes further interactions between DMA and week of year. Superscripts: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the week-year-DMA level.

VARIABLES	(1) d_{Coke}	(2) d_{Coke}	(3) d_{Coke}
f_{Coke}	0.0255***	0.0253***	0.0253***
	(0.000250)	(0.000250)	(0.000250)
f_{Pepsi}	-0.00199***	-0.00188***	-0.00189***
	(9.25e-05)	(9.24e-05)	(9.26e-05)
e_{Coke}	-0.000183***	-0.000374***	-0.000390***
	(5.36e-05)	(6.22e-05)	(6.16e-05)
e_{Pepsi}	-0.000513***	-0.000533***	-0.000503***
	(8.42e-05)	(9.21e-05)	(9.24e-05)
$e_{Coke} \cdot f_{Coke}$	9.16e-05***	9.22e-05***	9.21e-05***
	(8.87e-06)	(8.85e-06)	(8.87e-06)
$e_{Pepsi} \cdot f_{Coke}$	0.000101***	0.000103***	0.000103***
	(1.36e-05)	(1.36e-05)	(1.36e-05)
$e_{Coke} \cdot f_{Pepsi}$	-1.75e-05***	-1.86e-05***	-1.80e-05***
	(2.69e-06)	(2.68e-06)	(2.68e-06)
$e_{Pepsi} \cdot f_{Pepsi}$	7.61e-06**	6.99e-06*	6.69e-06*
	(3.86e-06)	(3.85e-06)	(3.85e-06)
$e_{Coke} \cdot f_{Coke}^2$	-7.89e-07***	-7.86e-07***	-7.86e-07***
	(1.25e-07)	(1.25e-07)	(1.25e-07)
$e_{Pepsi} \cdot f_{Coke}^2$	-1.09e-06***	-1.09e-06***	-1.09e-06***
	(1.77e-07)	(1.77e-07)	(1.77e-07)
$e_{Coke} \cdot f_{Pepsi}^2$	1.70e-07***	1.68e-07***	1.64e-07***
	(3.00e-08)	(2.92e-08)	(2.89e-08)
$e_{Pepsi} \cdot f_{Pepsi}^2$	-1.57e-08	-2.43e-08	-1.96e-08
	(3.37e-08)	(3.34e-08)	(3.33e-08)
Constant	-1.620***	-1.101***	-1.060***
	(0.110)	(0.110)	(0.111)
Observations	3,894,108	3,894,108	3,894,103
R-squared	0.125	0.126	0.129

The exercise above can also be run in a pooled regression with 40 brands. Table C.9 reports the baseline estimation of Equation (4) in Column (1) and then sequentially refines the variation we control for. This step-by-step introduction of dummies shows the sensitivity of our coefficients to different possible

sources of reverse causality. In Column (2) implements the main identification strategy in Shapiro et al. (2021), replacing the month of year with week of year dummies and state dummies with DMA dummies, and including month-year dummies, allowing us to control for finer seasonal and geographical variation. Column (3) further saturates the specification by allowing for DMA specific seasonality: we include interactions between DMA and week of year. Column (4) builds on column (1), saturating individual time invariant variations, replacing state dummies for household dummies.

The results are remarkably robust across these tests. The coefficients on past purchases, exposure, and similarity-based spillovers barely change across specifications. The stability of our coefficients to adding finer sources of seasonal, temporal and geographic variation that is increasingly difficult for firms to track in their supply of ads is comforting.

Table C.9: Robustness tests
Robustness for Geographical and Time controls, 40 products.

Notes: This table is the complete version of Table 3, displaying all our usual coefficients of interest. It estimates the regression coefficients for equation (4) for our pooled dataset adding different geographical and temporal robustness controls. Column (1) represents our baseline regression with month of year, year, state and product dummies, with the first three sets interacted with the last. Column (2) substitutes the month of year with week of year dummies and state dummies with DMA dummies, including month-year dummies too. Column (3) includes further interactions between DMA and week of year, alone and interacted with the product dummies. Column (4) switches from the baseline specification brand from product dummies. Column (5) controls for a brand's aggregate consumption in the DMA each week, alone and interacted with the household's past purchase frequency, excluding the seasonal and geographic dummies. Column (6) builds on column (4), replacing state dummies for household dummies, including their interactions with the brand dummies. Superscripts: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the week-year-DMA-product level.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Q_own	Q_own	Q_own	Q_own	Q_own	Q_own
e_{Own}	0.223***	0.223***	0.223***	0.221***	0.196***	0.0241***
	(0.000321)	(0.000322)	(0.000322)	(0.000333)	(0.000309)	(0.000469)
e_{Other}	-0.0174***	-0.0176***	-0.0176***	-0.0159***	-0.0114***	0.000328
	(0.000116)	(0.000117)	(0.000117)	(0.000115)	(0.000132)	(0.000228)
f_{Own}	0.00757***	0.00564***	0.00543***	0.00196***	0.00252***	0.00304***
	(0.000192)	(0.000210)	(0.000204)	(0.000238)	(0.000183)	(0.000200)
f_{Other}	-0.00159***	-0.00352***	-0.00373***	0.000158	0.00119***	0.00229***
	(0.000220)	(0.000238)	(0.000229)	(0.000264)	(0.000213)	(0.000216)
$e_{Own} \cdot f_{Own}$	0.0113***	0.0112***	0.0113***	0.0112***	0.00283***	0.0124***
	(0.000404)	(0.000404)	(0.000406)	(0.000410)	(0.000494)	(0.000365)

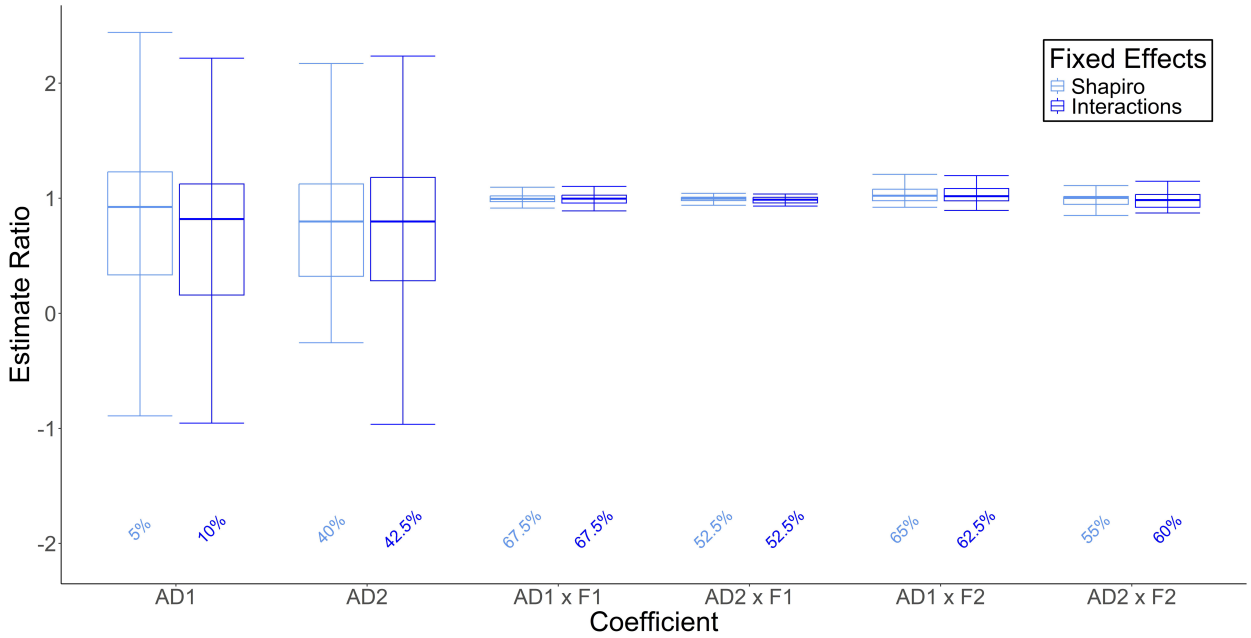
$e_{Other} \cdot f_{Own}$	0.00999***	0.0100***	0.00997***	0.00991***	0.00684***	0.0122***
	(0.000452)	(0.000454)	(0.000454)	(0.000463)	(0.000508)	(0.000393)
$e_{Own} \cdot f_{Other}$	-0.000598***	-0.000578***	-0.000622***	-0.000479***	0.00296***	-0.000784***
	(0.000182)	(0.000182)	(0.000182)	(0.000183)	(0.000187)	(0.000187)
$e_{Other} \cdot f_{Other}$	-0.00115***	-0.00119***	-0.00117***	-0.00114***	-0.00256***	-0.000259
	(0.000206)	(0.000206)	(0.000206)	(0.000209)	(0.000203)	(0.000197)
$e_{Own} \cdot f_{Own}^2$	-0.000567***	-0.000562***	-0.000563***	-0.000554***	-0.000635***	-0.000630***
	(9.28e-05)	(9.28e-05)	(9.29e-05)	(9.29e-05)	(0.000121)	(7.96e-05)
$e_{Other} \cdot f_{Own}^2$	-0.000933***	-0.000936***	-0.000933***	-0.000940***	-0.00109***	-0.000888***
	(0.000103)	(0.000103)	(0.000103)	(0.000103)	(0.000128)	(7.65e-05)
$e_{Own} \cdot f_{Other}^2$	-4.41e-05	-4.68e-05	-4.43e-05	-4.27e-05	-0.000181***	6.90e-05**
	(3.46e-05)	(3.47e-05)	(3.48e-05)	(3.47e-05)	(3.42e-05)	(3.18e-05)
$e_{Other} \cdot f_{Other}^2$	0.000282***	0.000287***	0.000286***	0.000280***	0.000350***	-4.01e-06
	(4.21e-05)	(4.22e-05)	(4.24e-05)	(4.24e-05)	(4.13e-05)	(3.58e-05)
Constant	-0.0550***	-0.0570***	-0.0573***	-0.0570***	-0.0664***	0.00964**
	(0.00130)	(0.00132)	(0.00132)	(0.00132)	(0.00128)	(0.00457)
Observations	72,116,888	72,116,888	72,116,888	72,116,886	72,116,686	72,116,854
R-squared	0.054	0.055	0.058	0.057	0.068	0.128

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The Figure below compares the coefficients estimated, in pooled regressions, across different levels of granularity. Specifically, it shows boxplots summarizing the ratios, relative to the baseline specification, of the coefficients computed in columns 2 through 4 in Table C.8.

Figure C.1: a graphical representation of the fixed effects robustness test

The figure displays boxplots summarizing, for the 40 brands, for each greater granularity specification (columns (2) to (3) of table C.8), the ratio between the estimated coefficients with respect to our baseline specification (the one with Month and State dummies, column (1) of table C.8). These ratios are computed separately for each specification and each coefficient. The percentages below display the proportion of products where each coefficient is significant at the 5% level for each specification.



C.4 Cues for Tastes versus Database

Here we report the results of the exercise, from Section 4.4, that separates these effects household tastes from those of household experiences in our data by developing a proxy for household preferences for a good. We proxy the latter by the average yearly quantity purchased in the pre-sample period 2012 to 2014, denoted $Q_{j,g}$. We estimate Equation (4) by including as regressors $Q_{j,g}$, alone and interacted with current ad exposure. Table C.10 reports the estimates. In column (1) we report our baseline estimation for comparison, in column (2) we report our estimates after controlling for $Q_{j,g}$ and its interactions.

Table C.10: Tastes versus Databases

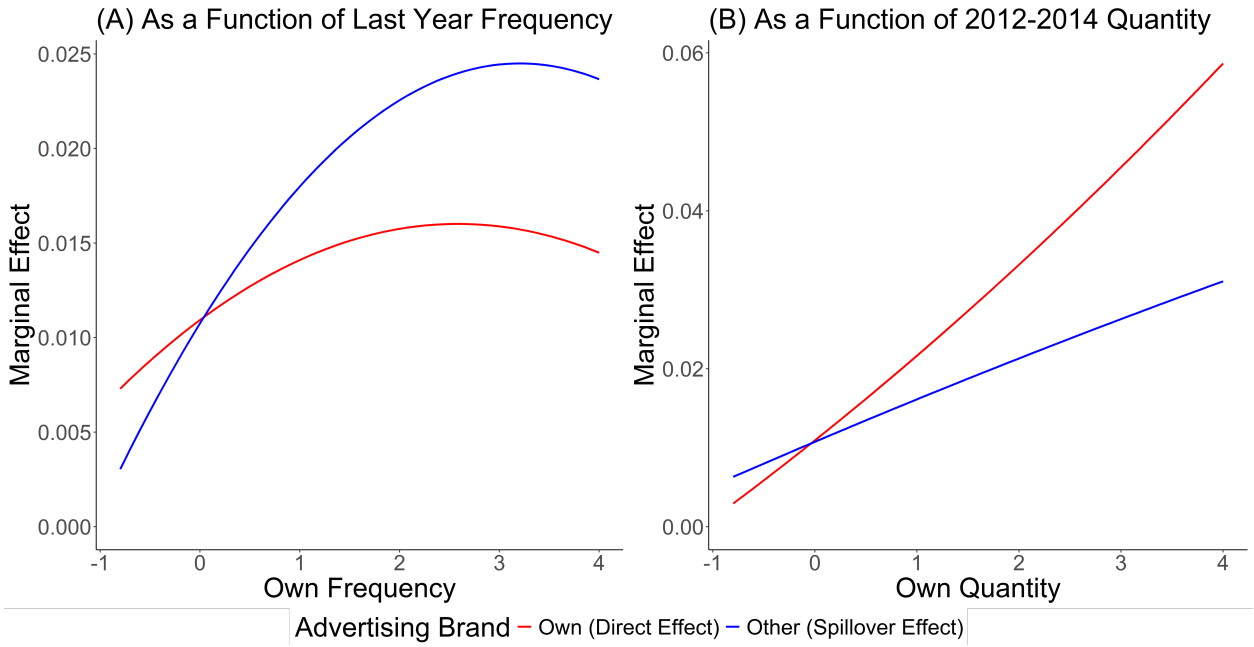
VARIABLES	(1) Q_own	(2) Q_own
e_{Own}	0.0114***	0.0109***
	(0.000161)	(0.000319)
e_{Other}	0.00922***	0.0108***
	(0.000192)	(0.000469)
f_{Own}	0.223***	0.151***
	(0.000260)	(0.000636)
f_{Other}	-0.0169***	-0.0153***
	(0.000106)	(0.000158)
$e_{Own} \cdot f_{Own}$	0.0108***	0.00393***
	(0.000371)	(0.000787)

$e_{Other} \cdot f_{Own}$	0.0106***	0.00857***
	(0.000418)	(0.00108)
$e_{Own} \cdot f_{other}$	-0.000247	-2.08e-05
	(0.000178)	(0.000234)
$e_{Other} \cdot f_{other}$	-0.00130***	-0.00109***
	(0.000202)	(0.000268)
$e_{Own} \cdot f_{Own}^2$	-0.000499***	-0.000760***
	(9.26e-05)	(0.000145)
$e_{Other} \cdot f_{Own}^2$	-0.00102***	-0.00134***
	(0.000102)	(0.000184)
$e_{Own} \cdot f_{Other}^2$	-6.47e-05*	-0.000148***
	(3.48e-05)	(3.05e-05)
$e_{Other} \cdot f_{Other}^2$	0.000302***	9.02e-05***
	(4.22e-05)	(3.42e-05)
$Past_{Own}$		0.157***
		(0.00129)
$Past_{Other}$		0.00549***
		(0.000203)
$e_{Own} \cdot Past_{Own}$		0.0103***
		(0.00184)
$e_{Other} \cdot Past_{Own}$		0.00547**
		(0.00256)
$e_{Own} \cdot Past_{Other}$		0.00229***
		(0.000214)
$e_{Other} \cdot Past_{Other}$		0.00163***
		(0.000233)
$e_{Own} \cdot Past_{Own}^2$		0.000405
		(0.000348)
$e_{Other} \cdot Past_{Own}^2$		-9.78e-05
		(0.000485)
$e_{Own} \cdot Past_{Other}^2$		-9.60e-05***
		(9.39e-06)
$e_{Other} \cdot Past_{Other}^2$		-5.40e-05***
		(7.92e-06)
Constant	0.0303***	0.0601***
	(0.00409)	(0.00602)
Observations	72,648,838	47,508,672
R-squared	0.054	0.073

Notes: Total Purchases, *Frequency* and *Exposure* are standardized at the product level. Total consumption is computed after standardization. Past refers to the average yearly purchased quantity in 2012-2014. For each product, the variables associated with brand 1 (e.g. e_1 and f_1) refer to the market leader. Regressions include the usual controls and product dummies. Column (1) estimates Equation (4) for product demand Column (2) estimates Equation (6). Robust standard errors are shown in parentheses. Superscripts: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the week-year-DMA-product level.

The coefficients associated with the time-varying database, *Frequency*, decrease somewhat, as expected, but they remain significant and comparable in magnitude, especially spillovers.

Figure C.2: a graphical representation of the fixed effects robustness test



Notes: Panel A shows the marginal effect (in standard deviations) of *Exposure* to τ and $-\tau$ for different levels of τ *Frequency*. We estimate the effect of ads by adding the baseline standardized *Exposure* coefficient with its linear and second order interaction with standardized τ *Frequency*, keeping standardized $-\tau$ *Frequency* at 0. For the τ advertising line (in red) in panel A, the plotted equation is: $me_\tau = \hat{\beta}_1 + \hat{\gamma}_1 * f_\tau + \hat{\delta}_1 * f_\tau^2$. The estimates are taken from Column (2) of Table C.10. Panel B shows the marginal effect (in standard deviations) of *Exposure* to τ and $-\tau$ for different levels of τ 2012-2014 Quantity. We estimate the effect of ads by adding the baseline standardized *Exposure* coefficient with its linear and second order interaction with standardized τ 2012-2014 Quantity, keeping standardized $-\tau$ 2012-2014 *Quantity* at 0. For the τ advertising line (in red) in panel A, the plotted equation is: $me_\tau = \hat{\beta}_1 + \hat{\gamma}_1 * q_\tau + \hat{\delta}_1 * q_\tau^2$. The estimates are taken from Column (2) of Table C.10.

D. Alternative mechanisms: Prices and Intertemporal Shifts

Here, we check the robustness of our results for two alternative mechanisms. One possibility is that ads are correlated with, and possibly highlight, price discounts. A second possibility is that ads prompt consumers to upfront their purchase but reduce future ones and have no effect on overall consumption.

D.1 Prices

Consider the price mechanism first. We first check whether ads mention words associated with promotions and sales. For each of the 20 products we count the number of times the following words are

present in our descriptions: price, sale, dollar, bargain, discount, promotion. The result is that only the word 'sale' is mentioned, and even that extremely rarely ('sale' is present in about 1% of descriptions of ads for Coke and Mountain Dew). We next assess the correlation between ad exposure and prices. Table D.1 shows there is correlation, indeed negative but mostly small. To test whether these correlations hide an indirect effect of ads through pricing, we run a regression where we control for the interactions between price and frequency (of each product). If purchases from frequent consumers were only moved by price drops, which are correlated to greater ad exposure, adding these controls would make the ad times frequency interactions coefficients drop. Table D.2 shows that introducing these controls does not cause any substantial shift in our ad estimates.

Table D.1
Correlation between Prices and Ad Exposure across Products.

Notes: The Table presents the correlation coefficients $\rho_{e_b,price_{b'}}$ between average ad exposure to brand b and price of brand b', where b, b' can be either brand. P-values are presented below the corresponding coefficients.

	$\rho_{e_1,price_1}$	$\rho_{e_2,price_2}$	$\rho_{e_1,price_2}$	$\rho_{e_2,price_1}$
Soup canned	-30%	-32%	-32%	-29%
	0.00	0.00	0.00	0.00
Beer	-10%	-6%	-5%	-7%
	0.00	0.00	0.00	0.00
Detergent heavy duty	-10%	-6%	3%	-1%
	0.00	0.00	0.00	0.32
Frozen pizza	-9%	-1%	-1%	2%
	0.00	0.17	0.14	0.06
Soap	9%	4%	-6%	-7%
	0.00	0.00	0.00	0.00
Facial tissue	-8%	4%	-2%	2%
	0.00	0.00	0.07	0.02
Candy chocolate	7%	3%	9%	-10%
	0.00	0.01	0.00	0.00
Colas	-6%	4%	-6%	6%
	0.00	0.00	0.00	0.00
Mouthwash	-6%	3%	0%	-7%
	0.00	0.00	0.96	0.00

Pain remedy headhache	-6%	-3%	2%	-3%
	0.00	0.00	0.05	0.00
Lemon colas	-4%	-4%	-4%	0%
	0.00	0.00	0.00	0.75
Dog food dry	-4%	-4%	2%	-2%
	0.00	0.00	0.14	0.10
Toilet tissue	-3%	-6%	-2%	-4%
	0.00	0.00	0.07	0.00
Yogurt	3%	-4%	-5%	1%
	0.00	0.00	0.00	0.13
Cheese crackers	2%	-3%	2%	-5%
	0.07	0.00	0.12	0.00
Cereal	-2%	0%	2%	3%
	0.10	0.74	0.07	0.00
Shampoo	-1%	-1%	-2%	-3%
	0.50	0.23	0.06	0.01
Diapers	0%	-1%	-1%	-1%
	0.69	0.25	0.23	0.63
Coffee	0%	6%	1%	5%
	0.74	0.00	0.44	0.00
Other carbonated soft drink	0%	-2%	0%	-1%
	0.89	0.10	0.86	0.24

TABLE D.2

Baseline regression with price interactions

Notes: The regression follows the baseline specification, but includes interactions between the log of weekly average dma prices for each product and each household frequency of purchase of each product. Superscripts: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the week-year-DMA level.

VARIABLES	(1) d_τ	(2) d_τ	(3) d_τ	(4) d_τ
f_{Own}	0.221***	0.222***	0.222***	0.223***
	(0.000317)	(0.000320)	(0.000320)	(0.000320)
f_{Other}	-0.0169***	-0.0169***	-0.0169***	-0.0169***
	(0.000115)	(0.000115)	(0.000115)	(0.000115)
e_{Own}		0.00478***	0.00478***	0.00459***
		(0.000196)	(0.000196)	(0.000198)
e_{Other}		0.000573**	0.000573**	0.000539**
		(0.000225)	(0.000225)	(0.000225)

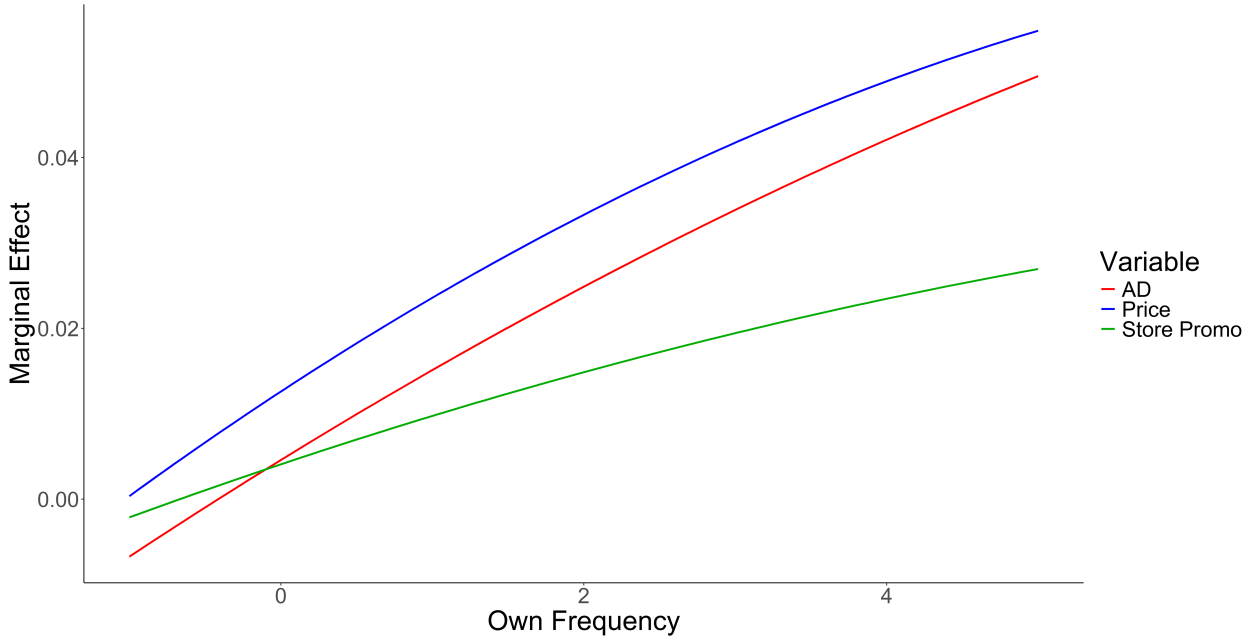
$e_{Own} \cdot f_{Own}$		0.0116***	0.0116***	0.0109***
		(0.000412)	(0.000412)	(0.000413)
$e_{Other} \cdot f_{Own}$		0.00726***	0.00726***	0.00673***
		(0.000456)	(0.000456)	(0.000454)
$e_{Own} \cdot f_{Other}$		-0.000540***	-0.000540***	-0.000554***
		(0.000184)	(0.000184)	(0.000187)
$e_{Other} \cdot f_{Other}$		-0.00113***	-0.00113***	-0.00107***
		(0.000208)	(0.000208)	(0.000203)
$e_{Own} \cdot f_{Own}^2$		-0.000544***	-0.000544***	-0.000385***
		(9.37e-05)	(9.37e-05)	(9.56e-05)
$e_{Other} \cdot f_{Own}^2$		-0.000929***	-0.000929***	-0.000828***
		(0.000103)	(0.000103)	(0.000103)
$e_{Own} \cdot f_{Other}^2$		-3.73e-05	-3.73e-05	-4.62e-05
		(3.44e-05)	(3.44e-05)	(3.55e-05)
$e_{Other} \cdot f_{Other}^2$		0.000284***	0.000284***	0.000260***
		(4.19e-05)	(4.19e-05)	(3.99e-05)
price_own	-0.0107***	-0.0120***	-0.0120***	-0.0126***
	(0.000566)	(0.000572)	(0.000572)	(0.000584)
price_other	-0.00185***	-0.000704	-0.000704	-0.00116*
	(0.000577)	(0.000587)	(0.000587)	(0.000598)
store_promo_own	0.00829***	0.00808***	0.00808***	0.00876***
	(0.000583)	(0.000590)	(0.000590)	(0.000604)
store_promo_other	0.000411	0.000439	0.000439	0.000750
	(0.000613)	(0.000618)	(0.000618)	(0.000627)
$Pri_{Own} \cdot f_{Own}$	-0.00876***	-0.00867***	-0.00867***	-0.0116***
	(0.000306)	(0.000317)	(0.000317)	(0.000435)
$Pri_{Other} \cdot f_{Own}$	-0.00101***	-5.16e-06	-5.16e-06	-0.00251***
	(0.000308)	(0.000318)	(0.000318)	(0.000433)
$Pri_{Own} \cdot f_{Other}$	0.000146	0.000275**	0.000275**	0.000303
	(0.000114)	(0.000120)	(0.000120)	(0.000187)
$Pri_{Other} \cdot f_{Other}$	0.00107***	0.000788***	0.000788***	0.000748***
	(0.000118)	(0.000122)	(0.000122)	(0.000173)
$Pri_{Own} \cdot f_{Own}^2$				0.000632***
				(9.87e-05)
$Pri_{Other} \cdot f_{Own}^2$				0.000524***
				(9.62e-05)
$Pri_{Own} \cdot f_{Other}^2$				-5.57e-05*
				(3.36e-05)
$Pri_{Other} \cdot f_{Other}^2$				-3.41e-05
				(2.63e-05)
$Pro_{Own} \cdot f_{Own}$	0.0113***	0.0103***	0.0103***	0.0127***
	(0.000314)	(0.000319)	(0.000319)	(0.000461)
$Pro_{Other} \cdot f_{Own}$	0.00306***	0.00271***	0.00271***	0.00475***
	(0.000292)	(0.000299)	(0.000299)	(0.000423)
$Pro_{Own} \cdot f_{Other}$	-0.00105***	-0.00102***	-0.00102***	-0.00117***
	(0.000112)	(0.000113)	(0.000113)	(0.000188)
$Pro_{Other} \cdot f_{Other}$	-0.000906***	-0.000868***	-0.000868***	-0.00107***
	(0.000118)	(0.000120)	(0.000120)	(0.000207)
$Pro_{Own} \cdot f_{Own}^2$				-0.000583***
				(0.000108)
$Pro_{Other} \cdot f_{Own}^2$				-0.000478***
				(9.98e-05)
$Pro_{Own} \cdot f_{Other}^2$				8.81e-05**
				(3.49e-05)
$Pro_{Other} \cdot f_{Other}^2$				9.94e-05**

				(4.21e-05)
lny	0.00275***	0.00269***	0.00269***	0.00270***
	(0.000151)	(0.000152)	(0.000152)	(0.000152)
$Pri_{Own} \cdot lny$	0.00181***	0.00189***	0.00189***	0.00190***
	(0.000137)	(0.000138)	(0.000138)	(0.000138)
$Pri_{Other} \cdot lny$	-0.000461***	-0.000498***	-0.000498***	-0.000502***
	(0.000140)	(0.000142)	(0.000142)	(0.000142)
$Pro_{Own} \cdot lny$	0.000981***	0.00101***	0.00101***	0.000998***
	(0.000142)	(0.000144)	(0.000144)	(0.000144)
$Pro_{Other} \cdot lny$	-0.000281*	-0.000252*	-0.000252*	-0.000269*
	(0.000150)	(0.000151)	(0.000151)	(0.000151)
Constant	-0.0573***	-0.0552***	-0.0552***	-0.0553***
	(0.00127)	(0.00129)	(0.00129)	(0.00129)
Observations	74,992,857	73,294,362	73,294,362	73,294,362
R-squared	0.053	0.054	0.054	0.054

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

From column (4), we show the direct marginal effect of a one standard deviation increase in ads and promotions, and decrease in price as a function of past frequency. The effect is computed, for each variable, $\hat{\alpha}_{var,0} + \hat{\alpha}_{var,1} * f_{own} + \hat{\alpha}_{var,2} * f_{own}^2$, where $\hat{\alpha}_{var,n}$ corresponds to the coefficient of column (4) of table D.2 corresponding to each variable and its n-th interaction with Frequency Own. The coefficients for price are swapped in sign, to have the effect of a decrease in price.

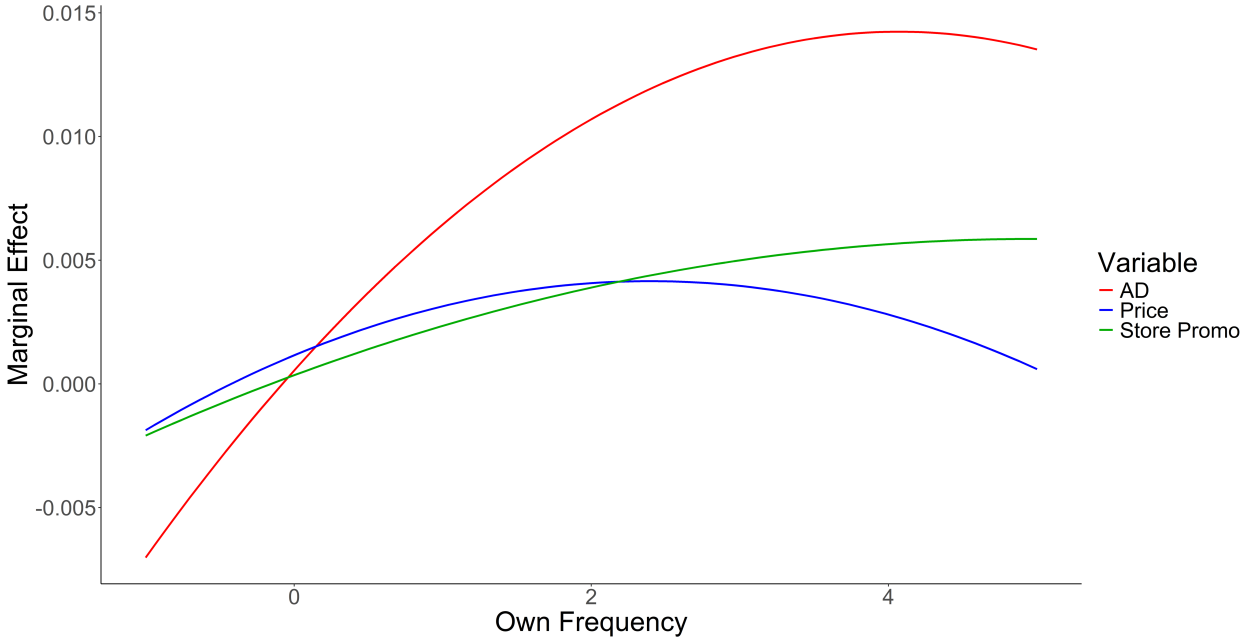
Figure D.1: Marginal direct effects of price, discount and ad exposure.



Notes: The figure shows the marginal effect (in standard deviations) of a one standard deviation increase in Own Advertising and Promotion, and decrease for Own Price, for different levels of Own Frequency. For any variable, the original marginal effect is computed as the sum of the baseline standardized variable coefficient, with its linear and second order interaction with Own Frequency, keeping Other Frequency at 0. In formula, this is $me_{var} = \hat{\alpha}_{var,0} + \hat{\alpha}_{var,1} * f_{own} + \hat{\alpha}_{var,2} * f_{own}^2$, where $\hat{\alpha}_{var,n}$ corresponds to the coefficient of column (4) of table D.2 corresponding to each variable and its n-th interaction with Frequency Own. The coefficients for price are swapped in sign, to have the effect of a decrease in price.

Next, we have the spillovers. Again from Column (4):

Figure D.2: Marginal spillover effects of price, discount and ad exposure.



Notes: The figure shows the marginal spillover effect (in standard deviations) of a one standard deviation increase in Own Advertising and Promotion, and decrease for Own Price, for different levels of Own Frequency. For any variable, the original marginal effect is computed as the sum of the baseline standardized variable coefficient, with its linear and second order interaction with Own Frequency, keeping Other Frequency at 0. In formula, this is $me_{var} = \hat{\alpha}_{var,0} + \hat{\alpha}_{var,1} * f_{own} + \hat{\alpha}_{var,2} * f_{own}^2$, where $\hat{\alpha}_{var,n}$ corresponds to the coefficient of column (4) of table D.2 corresponding to each variable spillover and its n-th interaction with Frequency Own. The coefficients for price are swapped in sign, to have the effect of a decrease in price.

D.2 Intertemporal shifts

Consider next the possibility that ads entail intertemporal shifts in purchases. While consistent with memory, this mechanism would reduce the allocative consequences of advertising. Table D.3 assesses this possibility by estimating the pooled specification of Table 3 at the monthly level, a frequency at which

we do not expect intertemporal substitution to matter. Our results on habits, cues and their interaction, including spillovers, are robust. Table D.4 estimates the impulse response of demand for each product to ads, by controlling for lagged advertising exposure. The influence of ads is again robust, and does not cause a drop in future consumption.

Table D.3
Baseline specification at monthly frequency

Notes: This table estimates the regression coefficients for equations (1, 2). d_b is the monthly demand for brand b, and ads are measured as expected monthly exposure. Total Purchases, *Frequency* and *Exposure* are standardized at the product level. Robust standard errors are shown in parentheses. Superscripts: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the month-year-DMA-product level.

	(1)	(2)	(3)
VARIABLES	d_{Own}	d_{Own}	d_{Own}+ d_{Other}
f_{Own}	0.366***	0.363***	0.239***
	(0.000726)	(0.000729)	(0.000592)
f_{Other}		-0.0269***	0.239***
		(0.000249)	(0.000592)
e_{Own}	0.0256***	0.0245***	0.0263***
	(0.000429)	(0.000433)	(0.000430)
e_{Other}		0.0126***	0.0263***
		(0.000553)	(0.000430)
$e_{Own} \cdot f_{Own}$	0.0348***	0.0299***	0.0195***
	(0.000862)	(0.000939)	(0.000632)
$e_{Other} \cdot f_{Own}$		0.0278***	0.0166***
		(0.00124)	(0.000896)
$e_{Own} \cdot f_{Other}$		-0.00357***	0.0166***
		(0.000410)	(0.000896)
$e_{Other} \cdot f_{Other}$		-0.00272***	0.0195***
		(0.000534)	(0.000632)
$e_{Own} \cdot f_{Own}^2$	-0.00237***	-0.00192***	-0.000890***
	(0.000208)	(0.000231)	(0.000147)
$e_{Other} \cdot f_{Own}^2$		-0.00273***	-0.00178***
		(0.000321)	(0.000229)
$e_{Own} \cdot f_{Other}^2$		0.000202**	-0.00178***
		(8.18e-05)	(0.000229)
$e_{Other} \cdot f_{Other}^2$		0.000648***	-0.000890***
		(0.000118)	(0.000147)
Constant	-0.260***	-0.675***	-0.959***
	(0.00558)	(0.0148)	(0.0208)
Observations	21,508,114	21,477,476	21,477,476
R-squared	0.144	0.146	0.123

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

We next examine directly whether an increase in ad exposure negatively predicts future purchases. To do so, we re-run the specification in Table 1 adding lagged exposure to ads, up to three weeks in the past. Table D.4 shows the results for the pooled dataset. Direct and spillover effects are robust to the inclusion of lags up to 3 weeks, and lagged exposure is generally insignificant. When lagged exposure is significant, it boosts current demand in the same way as contemporaneous exposure. In sum, intratemporal substitution does not seem a substantial effect for coke. Similar results hold for other products.

Table D.5
Impulse Response to Ads

Notes: This table estimates the regression coefficients for equation (1) as in Table 1, adding lagged exposure to ads. $\Delta_t e_b$ represents the expected weekly exposure to ads of brand b at a lag of t weeks before demand is measured. Dependent variables are standardized. Superscripts: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the week-year-DMA-product level.

Dep. Variable: d_{Own}

	(1)	(2)	(3)	(4)	(5)
f_{Own}	0.216***	0.216***	0.216***	0.213***	0.212***
	(0.000260)	(0.000260)	(0.000260)	(0.000268)	(0.000275)
f_{Other}	-0.0158***	-0.0158***	-0.0158***	-0.0152***	-0.0149***
	(0.000109)	(0.000109)	(0.000108)	(0.000112)	(0.000115)
$e_{Own,t}$	0.00985***		0.00996***	0.00893***	0.00865***
	(0.000159)		(0.000155)	(0.000166)	(0.000173)
$e_{Own,t-1}$	0.00766***	0.0131***	0.00747***	0.00578***	0.00503***
	(0.000163)	(0.000127)	(0.000154)	(0.000175)	(0.000185)
$e_{Own,t-2}$				0.00340***	0.00206***
				(0.000164)	(0.000184)
$e_{Own,t-3}$					0.00305***
					(0.000171)
$e_{Other,t}$	0.00449***		0.00439***	0.00405***	0.00394***
	(0.000168)		(0.000173)	(0.000186)	(0.000192)
$e_{Other,t-1}$	0.00269***	0.00542***	0.00281***	0.00213***	0.00187***
	(0.000161)	(0.000145)	(0.000173)	(0.000193)	(0.000207)
$e_{Other,t-2}$				0.00120***	0.000706***
				(0.000182)	(0.000202)

$e_{Other,t-3}$					0.000970***
					(0.000186)
$e_{Own} \cdot f_{Own}$	0.00932***		0.00669***	0.00590***	0.00581***
	(0.000264)		(0.000348)	(0.000377)	(0.000389)
$\Delta_1 e_{Own} \cdot f_{Own}$		0.00834***	0.00411***	0.00259***	0.00199***
		(0.000262)	(0.000346)	(0.000393)	(0.000421)
$\Delta_2 e_{Own} \cdot f_{Own}$				0.00267***	0.00143***
				(0.000372)	(0.000418)
$\Delta_3 e_{Own} \cdot f_{Own}$					0.00267***
					(0.000391)
$e_{Other} \cdot f_{Own}$	0.00665***		0.00468***	0.00398***	0.00391***
	(0.000310)		(0.000388)	(0.000420)	(0.000433)
$\Delta_1 e_{Other} \cdot f_{Own}$		0.00608***	0.00296***	0.00158***	0.000810*
		(0.000308)	(0.000384)	(0.000427)	(0.000453)
$\Delta_2 e_{Other} \cdot f_{Own}$				0.00246***	0.00145***
				(0.000407)	(0.000452)
$\Delta_3 e_{Other} \cdot f_{Own}$					0.00209***
					(0.000413)
$e_{Own} \cdot f_{Other}$	-0.000297**		-0.000554***	-0.000737***	-0.000742***
	(0.000128)		(0.000165)	(0.000179)	(0.000185)
$\Delta_1 e_{Own} \cdot f_{Other}$		0.000154	0.000406**	-9.88e-05	-0.000266
		(0.000127)	(0.000166)	(0.000189)	(0.000200)
$\Delta_2 e_{Own} \cdot f_{Other}$				0.00109***	0.000718***
				(0.000185)	(0.000211)
$\Delta_3 e_{Own} \cdot f_{Other}$					0.000848***
					(0.000185)
$e_{Other} \cdot f_{Other}$	-0.00125***		-0.000927***	-0.000774***	-0.000837***
	(0.000136)		(0.000159)	(0.000174)	(0.000178)
$\Delta_1 e_{Other} \cdot f_{Other}$		-0.00111***	-0.000549***	-0.000363**	-0.000278
		(0.000134)	(0.000158)	(0.000172)	(0.000186)
$\Delta_2 e_{Other} \cdot f_{Other}$				-0.000602***	-0.000448**
				(0.000176)	(0.000197)
$\Delta_3 e_{Other} \cdot f_{Other}$					-0.000414**
					(0.000173)
$e_{Own} \cdot f_{Own}^2$	-2.22e-05***		-1.59e-05***	-1.40e-05***	-1.44e-05***
	(3.69e-06)		(4.94e-06)	(5.15e-06)	(5.15e-06)
$\Delta_1 e_{Own} \cdot f_{Own}^2$		-2.02e-05***	-9.81e-06**	-8.60e-06	-6.29e-06
		(3.58e-06)	(4.87e-06)	(5.40e-06)	(5.67e-06)

$\Delta_2 e_{Own} \cdot f_{Own}^2$				-1.30e-06	2.38e-06
				(5.12e-06)	(5.78e-06)
$\Delta_3 e_{Own} \cdot f_{Own}^2$					-7.55e-06
					(5.09e-06)
$e_{Other} \cdot f_{Own}^2$	-6.24e-06		-4.83e-06	-5.61e-06	-6.43e-06
	(4.51e-06)		(6.04e-06)	(6.24e-06)	(6.27e-06)
$\Delta_1 e_{Other} \cdot f_{Own}^2$		-4.71e-06	-1.54e-06	-1.75e-06	-1.27e-06
		(4.46e-06)	(5.96e-06)	(6.40e-06)	(6.34e-06)
$\Delta_2 e_{Other} \cdot f_{Own}^2$				2.28e-06	9.52e-07
				(5.95e-06)	(6.67e-06)
$\Delta_3 e_{Other} \cdot f_{Own}^2$					2.51e-06
					(6.37e-06)
$e_{Own} \cdot f_{Other}^2$	-1.42e-06		-8.87e-07	-3.28e-07	-2.85e-07
	(1.88e-06)		(2.30e-06)	(2.45e-06)	(2.44e-06)
$\Delta_1 e_{Own} \cdot f_{Other}^2$		-2.15e-06	-1.21e-06	3.17e-07	2.64e-07
		(1.82e-06)	(2.34e-06)	(2.57e-06)	(2.62e-06)
$\Delta_2 e_{Own} \cdot f_{Other}^2$				-3.17e-06	-3.34e-06
				(2.85e-06)	(3.30e-06)
$\Delta_3 e_{Own} \cdot f_{Other}^2$					-2.02e-07
					(2.51e-06)
$e_{Other} \cdot f_{Other}^2$	1.36e-05***		8.10e-06***	6.23e-06**	6.54e-06***
	(2.27e-06)		(2.27e-06)	(2.52e-06)	(2.48e-06)
$\Delta_1 e_{Other} \cdot f_{Other}^2$		1.41e-05***	9.09e-06***	5.60e-06**	4.38e-06*
		(2.28e-06)	(2.41e-06)	(2.29e-06)	(2.54e-06)
$\Delta_2 e_{Other} \cdot f_{Other}^2$				7.81e-06***	5.44e-06*
				(2.87e-06)	(3.11e-06)
$\Delta_3 e_{Other} \cdot f_{Other}^2$					5.24e-06**
					(2.55e-06)
Constant	-0.0405***	-0.0634	-0.0407***	-0.0430***	-0.0438***
	(0.00351)	(0.180)	(0.00351)	(0.00366)	(0.00382)
Observations	64,363,694	64,384,578	64,363,694	58,342,068	53,647,008
R-squared	0.052	0.052	0.052	0.053	0.053

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

E. Extensive versus Intensive Margins.

Here we examine whether ad exposure increases the probability of purchase, the quantity purchased conditional on a purchase, or both. Focusing on the colas market, Table E.1 reports our specification (1) where the dependent variable is a dummy for whether purchases of the given brand is positive. We focus on the colas market, and the results show that ads impact the probability of purchase through the memory mechanism.

Table E.1
Ads and the Extensive Margin

Notes: This table estimates the regression coefficients for equation (4) where the dependent variable is whether purchase of either brand is positive. Ads are measured as expected weekly exposure. Total Purchases, *Frequency* and *Exposure* are standardized at the product level. Superscripts: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the week-year-DMA level.

	(1) $d_{Coke} > 0$	(2) $d_{Pepsi} > 0$
f_{Coke}	0.0111***	-0.000437***
	(8.49e-05)	(2.62e-05)
f_{Pepsi}	-0.000680***	0.0118***
	(3.56e-05)	(0.000107)
e_{Coke}	-7.13e-05***	-7.68e-05***
	(2.05e-05)	(1.90e-05)
e_{Pepsi}	-0.000280***	-0.000253***
	(3.10e-05)	(2.74e-05)
$e_{Coke} \cdot f_{Coke}$	3.20e-05***	-1.36e-06*
	(3.30e-06)	(7.40e-07)
$e_{Pepsi} \cdot f_{Coke}$	5.46e-05***	-9.78e-07
	(4.93e-06)	(1.10e-06)
$e_{Coke} \cdot f_{Pepsi}$	-6.99e-06***	3.27e-05***
	(1.05e-06)	(5.80e-06)
$e_{Pepsi} \cdot f_{Pepsi}$	4.55e-06***	7.24e-05***
	(1.42e-06)	(6.94e-06)
$e_{Coke} \cdot f_{Coke}^2$	-3.63e-07***	3.14e-08***
	(4.96e-08)	(7.57e-09)
$e_{Pepsi} \cdot f_{Coke}^2$	-5.27e-07***	2.81e-08***
	(7.08e-08)	(9.59e-09)
$e_{Coke} \cdot f_{Pepsi}^2$	6.68e-08***	-4.27e-07***
	(1.26e-08)	(1.14e-07)
$e_{Pepsi} \cdot f_{Pepsi}^2$	-1.67e-09	-8.63e-07***
	(1.32e-08)	(1.11e-07)
Constant	-0.135***	-0.205***
	(0.0330)	(0.0297)
Observations	3,894,108	3,894,108
R-squared	0.185	0.216

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

To assess the impact on the quantity purchased, we use the same specification with the dependent variable being the total amount of a given brand bought, conditional on purchase. Table E.2 reports the results, again showing robust direct and spillover effects of ads.

Table E.2
Ads and the Intensive Margin

Notes: This table estimates the regression coefficients for equation (4) where the dependent variable is the amount of either brand bought, conditional on purchase. Ads are measured as expected weekly exposure. Dependent variables are standardized. Robust standard errors are shown in parentheses. Superscripts: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the week-year-DMA level.

	(1) $d_{Coke}\|d_{Coke}>0$	(2) $d_{Pepsi}\|d_{Pepsi}>0$
f_{Coke}	1.762***	-1.253***
	(0.0519)	(0.0841)
f_{Pepsi}	-0.917***	2.300***
	(0.0722)	(0.0798)
e_{Coke}	0.134***	0.0856*
	(0.0322)	(0.0448)
e_{Pepsi}	-0.0445	0.161**
	(0.0504)	(0.0665)
$e_{Coke} \cdot f_{Coke}$	0.00611***	-0.00889***
	(0.00158)	(0.00205)
$e_{Pepsi} \cdot f_{Coke}$	0.00608**	-0.0116***
	(0.00250)	(0.00293)
$e_{Coke} \cdot f_{Pepsi}$	-0.0141***	0.00662**
	(0.00241)	(0.00277)
$e_{Pepsi} \cdot f_{Pepsi}$	-0.00369	-0.00161
	(0.00348)	(0.00366)
$e_{Coke} \cdot f_{Coke}^2$	-5.45e-05***	7.93e-05***
	(1.33e-05)	(1.00e-05)
$e_{Pepsi} \cdot f_{Coke}^2$	-6.74e-05***	4.37e-05***
	(1.87e-05)	(1.38e-05)
$e_{Coke} \cdot f_{Pepsi}^2$	0.000157***	-5.33e-05**
	(2.91e-05)	(2.64e-05)
$e_{Pepsi} \cdot f_{Pepsi}^2$	-6.11e-05*	-1.13e-05
	(3.66e-05)	(3.12e-05)
Constant	-919.8***	-820.6***
	(87.53)	(99.25)
Observations	556,155	372,814

R-squared	0.045	0.059

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table E.3
Ads and the Intensive Margin, Pooled Sample

Notes: This table estimates the regression coefficients for equation (4) where the dependent variable is the amount of either brand bought, conditional on purchase. Ads are measured as expected weekly exposure. Total Purchases, *Frequency* and *Exposure* are standardized at the product level. Robust standard errors are shown in parentheses. Superscripts: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the week-year-DMA-product level.

	(1) $d_{Own} > 0$	(2) $d_{Own} \mid d_{Own} > 0$
f_{Own}	0.0786***	0.0767***
	(8.90e-05)	(0.000732)
f_{Other}	-0.00199***	-0.114***
	(3.18e-05)	(0.00114)
e_{Own}	0.00140***	0.00546***
	(5.26e-05)	(0.00169)
e_{Other}	0.000369***	0.00597***
	(5.78e-05)	(0.00201)
$e_{Own} \cdot f_{Own}$	0.00336***	0.00785***
	(0.000121)	(0.00101)
$e_{Other} \cdot f_{Own}$	0.00247***	-0.00551***
	(0.000134)	(0.00132)
$e_{Own} \cdot f_{Other}$	-0.000128***	-0.00628***
	(4.84e-05)	(0.00152)
$e_{Other} \cdot f_{Other}$	-0.000521***	0.00165
	(4.51e-05)	(0.00183)
$e_{Own} \cdot f_{Own}^2$	-0.000381***	-0.000314***
	(2.71e-05)	(0.000108)
$e_{Other} \cdot f_{Own}^2$	-7.09e-05**	-4.24e-05
	(3.26e-05)	(0.000139)
$e_{Own} \cdot f_{Other}^2$	3.26e-05***	0.000775***
	(8.88e-06)	(0.000255)
$e_{Other} \cdot f_{Other}^2$	3.74e-05***	0.000440
	(8.35e-06)	(0.000311)
Constant	0.0763***	1.401***
	(0.000327)	(0.0136)
Observations	72,116,888	5,117,061
R-squared	0.108	0.097

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

F. Effect of advertising on zero frequency households

Section 4 documents that baseline effects of ad exposure tend to be insignificant across products. In our model, ads impact demand through a memory mechanism whereby they cue retrieval of past experiences, so ads are not predicted to have an effect on their own. Still, it is natural to assume that consumers have some familiarity with the products being advertised. We revisit this issue by examining the sample of households that have not purchased *either* of the two brands in the previous year, and so have $f_1 = f_2 = 0$. We regress differential and total demands on the effect of ads in all 20 products. The regression maintains all the controls present in our baseline regression, but by construction does not include any memory components. Overall, the regressions show no substantial effects, with relatively few instances where ads move consumption in favour of their product significantly at 5%, and the corresponding magnitudes being low, confirming that the impact of ads is manifest essentially only for people with past experiences of the product. Table F.1 reports the results for colas and beer.

Table F.1

Ad effects at zero frequency.

Note: Table F.1 estimates the regression coefficients for equation (4), with ads measured as expected weekly exposure for those consumers who have frequency equal to 0 for both brands. Hence, the coefficients entailing f_i are not estimated for any i. Panel A displays the results for colas, Panel B for beer, and Panel C for the pooled sample. Dependent variables are standardized. Robust standard errors are shown in parentheses. Superscripts: *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Standard errors are clustered at the week-year-DMA level for Panels A and B, and at the week-year-DMA-product level for Panel C.

Panel A: colas

	(1)	(3)	(4)
VARIABLES	d_{Coke}	d_{Pepsi}	$d_{Coke}+ d_{Pepsi}$
e_{Coke}	0.000846	-0.000400	1.48e-05
	(0.00114)	(0.000774)	(4.71e-05)
e_{Pepsi}	0.00609***	0.00153	0.000254***
	(0.00172)	(0.00135)	(7.49e-05)
Constant	-2.165	-5.452**	-0.339**

	(3.344)	(2.403)	(0.137)
Observations	1,550,210	1,550,210	1,550,210
R-squared	0.001	0.001	0.001

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Panel B: beer

	(1)	(2)	(3)
VARIABLES	d_{Bud}	d_{Miller}	$d_{Bud}+ d_{Miller}$
e_{Bud}	0.0162***	0.00426***	0.000588***
	(0.00401)	(0.00128)	(0.000121)
e_{Miller}	0.000651	0.00700**	0.000220
	(0.00480)	(0.00337)	(0.000171)
Constant	-4.571	-2.686	-0.254
	(6.096)	(1.945)	(0.184)
Observations	991,087	991,087	991,087
R-squared	0.001	0.001	0.001

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Panel C: Pooled

(1)		(2)
VARIABLES d_{Own}	VARIABLES	$d_1 + d_2$
e_{Own} 0.00578***	e_1	0.00244***
(0.000168)		(0.000235)
e_{Other} -0.00333***	e_2	0.000572**

	(0.000152)		(0.000271)
Constant	0.00699***	Constant	0.00961***
	(0.00150)		(0.00221)
Observations	57,515,514	991,087	27,731,678
R-squared	0.001	0.001	0.001

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

The table below summarizes the corresponding results, aggregating across the 20 products. The first two rows concern the coefficients on our measures of ad exposure to either brand, e_1, e_2 of the baseline regression of differential demand (Equation 1). Rows 3 and 4 do the same for the regression on total demand (Equation 2). Column 1 shows the percentage of coefficients which have their sign as theory would predict. For example, in the difference regression the coefficient on e_1 is positive and that on e_2 is negative, while for the total demand regression they should both be positive. The second and third column report the share of coefficients that have the expected sign are also significant at 10% and 5% respectively.

Table F.2: ad effects at zero summary across 20 products

No Frequency Regressions				
	% $p \geq 0.05$		**% $p < 0.05$**	
	> 0	**< 0**	**> 0**	**< 0**
e_1	33%	33%	33%	3%
e_2	45%	35%	8%	13%

Table F.3. Within product similarity

Note: This table contains the similarity score obtained by looking at the ads descriptions of the two brands for any product.

	AD Similarity
beer	0.724
candychoc	0.287

cereal	0.616
coffee	0.461
colas	0.778
crackerscheese	0.001
detergentheavy	0.058
diapers	0.865
dogfooddry	0.473
facialtissue	0.232
frozenpizza	0.405
lemoncolas	0.374
mouthwash	0.673
othercarbonatedsoftdrink	0.397
painremedyheadhache	0.371
shampoo	0.579
soap	0.552
soupcanned	0.526
toilettissue	0.053
yogurt	0.829

Table F.4. Cross product similarity.

Note: This table contains the similarity score obtained by comparing the ads descriptions of any couple of products. For each product we take into account ads descriptions from both brands.

product	beer	candy	cerea	coffe	colas	crack	deter	diape	dogfo	facia	froze	lemon	mouth	other	painr	shamp	soap	soupc	toile	yogur
beer	1.00																			
candy	0.10	1.00																		
cerea	0.24	0.09	1.00																	
coffe	0.11	0.16	0.24	1.00																
colas	0.68	0.12	0.32	0.20	1.00															
crack	0.05	0.07	0.13	0.01	0.08	1.00														
deter	0.31	0.09	0.32	0.30	0.48	0.04	1.00													
diape	0.21	0.07	0.23	0.18	0.35	0.03	0.43	1.00												

dogfo	0.10	0.07	0.18	0.14	0.19	0.01	0.20	0.17	1.00											
facia	0.14	0.20	0.27	0.27	0.29	0.22	0.35	0.25	0.34	1.00										
froze	0.28	0.08	0.29	0.18	0.42	0.03	0.35	0.30	0.14	0.15	1.00									
lemon	0.58	0.11	0.41	0.24	0.78	0.08	0.49	0.40	0.20	0.29	0.46	1.00								
mouth	0.27	0.10	0.31	0.28	0.48	0.06	0.49	0.45	0.26	0.38	0.41	0.58	1.00							
other	0.55	0.10	0.24	0.18	0.78	0.07	0.43	0.31	0.15	0.28	0.25	0.65	0.38	1.00						
painr	0.31	0.12	0.38	0.28	0.50	0.09	0.54	0.47	0.33	0.46	0.41	0.61	0.66	0.44	1.00					
shamp	0.31	0.13	0.29	0.23	0.54	0.08	0.55	0.43	0.27	0.52	0.36	0.60	0.66	0.51	0.70	1.00				
soap	0.30	0.11	0.22	0.24	0.45	0.04	0.49	0.41	0.34	0.49	0.29	0.51	0.64	0.43	0.69	0.77	1.00			
soupc	0.24	0.06	0.29	0.14	0.26	0.04	0.21	0.18	0.18	0.13	0.16	0.25	0.23	0.19	0.24	0.20	0.20	1.00		
toile	0.09	0.13	0.16	0.08	0.20	0.20	0.16	0.16	0.08	0.49	0.08	0.17	0.21	0.15	0.23	0.29	0.21	0.09	1.00	
yogur	0.33	0.16	0.41	0.31	0.54	0.03	0.54	0.47	0.33	0.33	0.47	0.61	0.65	0.40	0.64	0.61	0.57	0.32	0.16	1.00

Table F.5 contains results from 20 regressions performed to study the impact of similarity across different goods. We test a variation of specification (3), with the sole difference that S is defined as the cosine similarity in ads descriptions between the product of the observation in that line and the specific product to which all products are compared to. We repeat this for all 20 products and exclude the product for which we calculate the cross-similarity for that regression.

Table F.5: Cross-Similarity Regression Summary

Notes: the table provides the estimates for the cross products ads' similarity regression for the 20 products. For each product, we run a regression where the ad of such product is included in a regression on the pooled dataset, together with its interactions with frequency and our similarity measure to the product corresponding to each observation.

p	$e_p * S_{i,p}$	$e_p * S_{i,p} * f_i$
Beer	0.036	0.025
Candy chocolate	0.016	-0.099
Cereal	0.005	0.021
Coffee	0.014	0.004
Colas	0.016	0.002
Cheese crackers	-0.010	-0.019
Detergent heavy duty	0.006	0.020
Diapers	0.003	0.015
Dog food dry	0.026	0.014
Facial tissue	0.001	-0.051
Frozen pizza	0.008	0.013
Lemon colas	0.006	0.018
mouthwash	0.006	0.024
Other carbonated soft drink	0.020	0.015

Pain remedy headhache	0.004	0.018
Shampoo	0.006	0.017
Soap	0.003	0.007
Soup canned	-0.001	-0.004
Toilet tissue	0.011	-0.020
Yogurt	0.037	0.024