International Journal of Research in Marketing 33 (2016) 508–523

Contents lists available at ScienceDirect

IJRM
International Journal of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Full Length Article

Effects of TV advertising on keyword search☆

Mingyu Joo [a,*], Kenneth C. Wilbur [b], Yi Zhu [c]

[a] The Ohio State University Fisher College of Business, United States
[b] University of California, San Diego, Rady School of Management, United States
[c] University of Minnesota Carlson School of Management, United States

ARTICLE INFO

Article history:
First received on April 30, 2014 and was under review for 6 1/2 months
Available online 17 April 2015

Keywords:
Advertising
Search engine marketing
Television

ABSTRACT

This paper investigates the possibility that television advertising influences online search using the AOL search dataset. It uses a novel keyword mining technique to classify keywords as brand related, category related (generic), or unrelated, distinguishing between category search and consumers' tendency to search a branded keyword. A three-level conditional choice model is estimated to determine whether hourly changes in brands' television advertising expenditures are related to deviations from baseline trends in search behaviors. The results indicate a statistically significant relationship between TV advertising and consumers' tendency to search branded keywords (e.g. "Fidelity") rather than generic category-related keywords (e.g. "stocks") in the dataset. The effect is largest for relatively young brands during standard business hours with an elasticity, .07, comparable to extant measurements of advertising's impact on sales. However, television advertising is not found to influence category search incidence and has limited effects on click-through rates.

1. Introduction

The online search process starts not with the search itself, but with the idea to search. The search idea may be general or very specific. For example, a consumer may want to learn more about a product category or may want to consider a particular brand. Next, the consumer chooses what query to enter into a search engine. A critical distinction is whether the query's keywords are generic (e.g. "Pick-up truck") or brand-related (e.g. "Ford F-150"). The specificity of the search query determines the breadth of competitive information returned by the search engine. It also influences the marketer's cost per click on paid links, which may differ dramatically between generic and branded keywords (Desai, Shin, & Staelin, 2014; Rutz & Bucklin, 2011). After entering the query, the consumer chooses which results to click.

The average American watches about 5 h of television per day (TVB, 2013), and television accounts for 58% of global ad spending (Nielsen, 2013), so one might suspect that television advertising could influence any of these search behaviors. It could prompt a consumer to search when she otherwise may not have done so. It may make a brand name more salient, increasing the chance that the consumer uses a brand-related query rather than a generic query. And it may influence whether the consumer feels satisfied with the search results and clicks a link.

The purpose of this paper is to estimate how online search for financial services responds to television advertising using the AOL search dataset. It relates individual-level behavior at the hourly level to intertemporal variation in brands' television advertising expenditures. A

☆ The authors gratefully acknowledge financial support from MSI/WIMI multichannel research award #4-1645.
* Corresponding author at: 558 Fisher Hall, 2100 Neil Ave., Columbus, OH 43210, USA.
E-mail addresses: joo.85@osu.edu (M. Joo), kennethcwilbur@gmail.com (K.C. Wilbur), yizhu@umn.edu (Y. Zhu).
URL's: http://www.mingyujoo.com (M. Joo), http://kennethcwilbur.com (K.C. Wilbur), http://www.tc.umn.edu/~yizhu/ (Y. Zhu).

http://dx.doi.org/10.1016/j.ijresmar.2014.12.005
0167-8116/

three-level conditional choice model is developed to isolate the impact of television advertising on category search incidence, branded keyword choice and click-through rates, controlling for brand- and time-specific baselines of search tendency, recent movements in the Dow Jones stock index, and advertising content. The model performs very well in both in-sample and hold-out validation tests.

The empirical results indicate that television advertising is positively related to consumers' choice of branded keywords at the expense of generic keywords. This relationship is largest for relatively young brands during standard business hours, with an elasticity of .07, similar to extant findings of the effect of advertising on sales. However, it does not appear that advertising expenditure changes the number of searches in the product category. In other words, television advertising shifts share of people who searched among competing brands' keywords, but it did not increase the proportion of people searching for financial services category. This suggests that financial services brands' advertising encourages substitution between keywords in the category, but does not increase the number of consumers who are interested in the category.

The next section reviews relevant literature and distinguishes the current article from its nearest neighbor (Joo, Wilbur, Cowgill, and Zhu, 2014). Sections 2 and 3 present the data and the econometric model. Section 4 discusses the results and Section 5 discusses their implications for practice and the academic literature.

1.1. Related literature

Online search has recently become consumers' primary source of information (Ratchford, Talukdar, & Lee, 2007), but consumer information search predates search engines. In fact, marketing academics have studied the antecedents and consequences of consumer search for decades. Their work has found a variety of mechanisms whereby advertising can influence consumer search for information.

Advertising has been shown to initiate consumer search for information about a product category. A consumer with little information about the category cannot search efficiently, whereas a consumer with extensive information has little need to search. Advertising can increase consumers' objective and subjective knowledge (Newman & Staelin, 1973) and stimulate information search in new categories with low prior knowledge (Bettman & Park, 1980; Swasy & Rethans, 1986). The effect of advertising on search can also vary across stages in the purchase funnel. For example, Punj and Staelin (1983) showed that consumers with more product-specific knowledge search less after seeing an advertisement, while those with general category knowledge are more likely to search. Klein and Ford (2003) distinguished between online and offline search, finding that consumers' mix of time spent on online vs. offline search activities depends on the relative importance of attributes that can be reliably verified through online search.

Advertising has also been shown to affect the means of consumers' search for information. Consumers may search using broad or focused means. Focused search strategies are more likely when the consumer has greater uncertainty about differences between brands (Moorthy, Ratchford, & Talukdar, 1997) and when the consumer overestimates her current level of knowledge (Moorman, Diehl, Brinberg, & Kidwell, 2004). Whereas this literature has relied primarily on experimental evidence, the current paper estimates similar effects using field data.

A number of recent papers have shown that online search data can help predict market outcomes, and therefore constitute important information that marketing managers need to track. For example, Kulkarni, Kannan, and Moe (2011) presented convincing evidence that online search data can improve forecasts of new product sales in the motion picture industry. Kulkarni, Ratchford, and Kannan (2012) showed that automobile purchasers who used the internet to search for their cars placed greater emphasis on product attribute ratings, while those who did not use the internet placed greater emphasis on summary recommendations. Hu, Du, and Damangir (2014) estimated a marketing mix model incorporating Google Trends data along with standard data like market shares, prices and advertising expenditures, finding that information about search volume enhanced model fit both in-sample and out-of-sample. Therefore, understanding the drivers of online search may help us to understand how search data can be used in forecasting new product sales and understanding purchaser characteristics.

The current article is also related to research on search engine marketing metrics. For example, Abou Nabout and Skiera (2012) showed that the profit implications of improving advertising quality depend on how ad quality affects costs per click; they recommended coupling advertisement quality improvements with lower keyword bids. Abou Nabout, Skiera, Stepanchuk, and Gerstmeier (2012) investigated the profitability of different search advertising agency compensation structures on search advertising profitability. Sayedi, Jerath, and Srinivasan (2014) showed that firms can free ride on a competitor's traditional ads by advertising on the competitor's keywords. These are just a few examples of a voluminous recent literature. Generally speaking, if online search is affected by a firm-controlled activity such as television advertising, then it would be important to adjust ROI metrics to optimize both offline and online marketing expenditures.

This paper also relates to extant research in attribution and path to purchase. Increasingly, researchers seek to understand how firms' marketing communications in multiple channels influence and interact with consumers' information acquisition strategies, expressions in social media, and purchase decisions. Studies have shown differences in online channels' differential effects on conversion (Li & Kannan, 2014), the impact of clicks on conversion (Jerath, Ma, & Park, 2014; Park & Park, 2014), how display ads influence online search (Kireyev, Pauwels, & Gupta, 2013), and how different payment schemes (such as cost-per-action or cost-per-mille) and the process of "last-touch" attribution influence publishers' strategic decisions (Berman, 2013). While this stream of research mostly examines online environments, the present article adds to this by considering the impact of offline marketing activities on the online channel.[1]

[1] A notable exception that explores TV advertising's effect on online shopping is Liaukonyte et al. (in press).

Table 1
Query mining procedure results.

Brand: Charles Schwab		Brand: Fidelity	
Sample brand keywords			
Schwa	Charlesschwab	Fedelity	Fidelityinvestment
Schwab	Charlesschwabb	Fid	Fidelityinvestments
Scwab	Charlesswab	Fidelity	Fidelty
Shwab	Wwwcharlesschwab	Fidelityadvisor	
Sample generic keywords			
401K	Dividends	Investment	Mortgage
401-K	Equity	Investments	Securities
Bank	Finance	Investor	Security
Banking	Financial	Investors	Shares
Banks	Fund	IRA	Stock
Bond	Funds	Liability	Stocks
Bonds	Income	Liquidity	Tax
Broker	Invest	Loan	Trusts
Brokerage	Investing	Loans	UGMA
Brokers	Investment	Money	Yield
Sample unrelated keywords			
About	Careers	Community	First
Advantage	Ceo	Connecticut	For
After	Chains	Corporation	Form
Are	Charts	Dean	Government
Best	City	Death	Group
Buy	Closed	Delta	Hall
Buying	Code	Distribution	Harbor
Calculator	College	DVD	High
Calculators	Comhttp	Education	History
California	Common	Examples	HIV

The current paper is most closely related to Joo et al. (2014), which investigated the relationship between television advertising and online search using aggregated Google data recorded in 2011. They found that advertising was associated with a small increase in product category queries and a larger increase in the share of category search terms containing branded keywords. Although Joo et al. (2014) analyzed data that were newer and broader than the current article, there were a number of limitations of that aggregate analysis that we seek to address here.

The most important difference between Joo et al. (2014) and this work is that we allow for heterogeneity in consumer behavior. The marketing literature has been emphatic in its agreement that ignoring heterogeneity in consumer preferences may bias elasticity estimates (e.g., Abramson, Andrews, Currim, & Jones, 2000; Guadagni & Little, 1983; Kamakura & Russell, 1989; Murthi & Srinivasan, 1998), underestimate the impact of marketing-mix variables (e.g., Chiang, Chib, & Narasimhan, 1999), and overestimate the effect of state dependence variables (e.g., Keane, 1997; Roy, Chintagunta, & Haldar, 1996). Researchers have also shown that models with representations of heterogeneity outperform homogeneous models with better model fit and better parameter recovery (e.g., Andrews, Ainslie, & Currim, 2002) while sources of observed heterogeneity predict consumer decisions better than unobserved heterogeneity (e.g., Horsky, Misra, & Nelson, 2006). Similarly, research in online advertising pointed out the importance of incorporating consumer heterogeneity (Rutz & Trusov, 2011) and showed that marketing mix effects vary across consumer segments (Pauwels, Leeflang, Teerling, & Huizingh, 2011). This paper segments consumers into four groups based on two usage patterns (Lim, Currim, & Andrews, 2005; Neslin, Henderson, & Quelch, 1985; Viswanathan, Kuruzovich, Gosain, & Agarwal, 2007): frequency of search and time elapsed since last search. The segment-level heterogeneity parameters estimated in this paper help to explain the effects of television advertising on online search by specifying *whom* it affects and *how*.

The individual level data in this paper further allow us to measure consumer response to television advertising by counting the number of *people* undertaking a given search activity in each hour, rather than the number of *actions*. This confers two important benefits. First, aggregate data may conflate "search reformulations" with increased category search; therefore the category-search effect of TV advertising may be overestimated using aggregate data. Second, counting consumers gives a better proxy to predict conversions (e.g., Kulkarni et al., 2012; Li & Kannan, 2014). We also include additional data on advertising content (Bertrand, Karlan, Mullainathan, Shafir, & Zinman, 2010; Weinberger & Spotts, 1989; Xu, Wilbur, Siddarth, & Silva-Risso, 2014) to help explain *when* and *why* television advertising affects online search, which were not available to Joo et al. (2014).

2. Data

This study explores data from the financial services category, because i) television advertising and online search are both frequent in this category, ii) brand names in the category are easily distinguished from common unrelated words and names, and iii) there is limited basis to expect simultaneity or unobserved factors that drive both the timing of television advertising and online search activity. This section introduces the datasets and describes how their characteristics influence the modeling choices presented below.

Table 2
Descriptive statistics.

Brand	Mon.–Fri. daily averages			Saturday–Sunday averages			% adv. exp. emphasizing web	% adv. exp. emphasizing phone	Brand age (in years, as of 2006)
	Number of queries	Click-through rate	Adv. exp. ($000)	Number of queries	Click-through rate	Adv. exp. ($000)			
AG Edwards	4	51.6%	13.7	2	47.9%	45.3	16.4%	9.1%	119
Charles Schwab	38	58.0%	160.8	18	48.4%	298.8	0.0%	0.0%	35
CitiGroup	85	60.6%	0.0	68	61.4%	0.0	–	–	194
E-Trade	35	54.7%	67.4	13	52.9%	71.5	5.8%	13.5%	24
Edward Jones	10	60.6%	56.0	7	71.4%	45.5	12.3%	22.8%	84
Fidelity	189	74.1%	122.8	97	72.7%	337.5	7.3%	18.3%	60
Forex	2	51.0%	0.0	3	51.3%	0.0	–	–	7
FXCM	1	71.4%	0.0	1	68.8%	0.0	–	–	7
Legg Mason	1	84.2%	0.0	0	66.7%	0.0	–	–	107
Merrill Lynch	25	36.5%	19.4	18	32.3%	344.5	7.3%	12.6%	92
Morgan Stanley	12	65.1%	0.0	6	66.7%	0.0	–	–	71
Nuveen	1	75.5%	0.0	1	84.6%	0.0	–	–	108
Oppenheimer	4	60.9%	21.8	2	70.5%	55.7	7.2%	11.1%	46
OptionsXpress	2	76.7%	0.0	1	75.8%	0.0	–	–	6
Raymond James	3	68.7%	10.6	2	55.8%	8.7	6.0%	33.2%	44
Scottrade	25	27.6%	14.0	7	17.5%	1.5	16.8%	11.4%	26
ShareBuilder	4	32.3%	0.0	2	14.8%	0.0	–	–	6
T. Rowe Price	3	69.4%	65.9	2	63.6%	153.8	12.5%	22.6%	69
TD Ameritrade	68	56.0%	313.6	32	50.8%	452.6	9.3%	23.2%	35
Vanguard	33	28.5%	0.0	28	32.1%	0.0	–	–	31
Van Kampen	1	78.6%	2.3	0	100.0%	39.9	7.5%	22.6%	79
Wachovia	155	56.7%	20.3	115	60.0%	182.0	3.7%	22.8%	98
All brand keywords	699	58.9%	888.7	424	58.0%	2037.2	7%	16%	–
All generic keywords	2586	61.6%	–	2108	63.4%	–	–	–	–

2.1. Online search data

The online search data include all queries entered by 180,000 randomly selected users of AOL.com between March 1 and May 30, 2006; of these people, approximately 64,000 searched a financial services product category keyword during the sample period. In the following discussion, a "query" is a search term and a "keyword" is a word within a query. Each query consists of one or more keywords and almost all keywords appear in multiple queries. For each query, the data include an anonymous user ID, the constituent keyword(s), the date and time of the search, and if the consumer clicked a result, the link and position of the result clicked.[2] Appendix 1 discusses the legal and ethical implications of using the AOL dataset for academic research.

A critical aspect of the study is to distinguish between advertising's possible effects on *total category searches* and *keyword choice within the focal category*. A naïve approach would look simply at the number of search queries that contain the brand name, since these are more easily countable. However, this might lead to an erroneous conclusion about the rate at which television advertising creates new searches. In fact, consumers might have searched anyway, but are now using branded keywords in favor of generic keywords. This possible mistake has important implications. A customer who searches a branded keyword is cheaper to acquire and more likely to convert than a customer who searches a generic keyword. Branded keywords limit the amount of competitive information the consumer is exposed to and reduce competition in the keyword auction.

To avoid this mistake, it is necessary to identify generic product category-related keywords. However, the full set of brand- and category-related keywords might not be available. For example, consumers may frequently misspell a brand name or may use unanticipated generic keywords.

A query mining algorithm was developed to identify the full set of product category-related keywords. The algorithm is related to the literature on keyword suggestions (Abhishek & Hosanagar, 2007; Chen, Xue, & Yu, 2008; Fuxman, Tsaparas, Achan, & Agrawal, 2008) and is described in Appendix 2. The algorithm serves three purposes: (1) identify unanticipated keywords for each brand, (2) identify generic keywords related to multiple brands in the product category, and (3) filter out unrelated queries. The results indicate that the procedure has high face validity. Table 1 displays some sample brand keywords identified by this procedure, including several unanticipated mis-spellings of branded keywords such as "Charlesswab" and "Fedelity." It also shows some sample generic keywords, including such examples as 401K, Stocks, Bonds, and Dividends. The query mining technique also identified some reasonable generic keywords which were not anticipated, such as Equity, Liquidity, and UGMA, a tax shelter for intergenerational wealth transfers. Finally, the table shows some keywords which were identified as not specifically related to the financial services category, such as calculators, careers, and government.

[2] AOL displayed zero, one or two sponsored links above the organic search results and did not display sponsored links on the side of the page. The sponsored links were typically the same as the highest-ranking organic results and directed consumers to the same landing pages. The paid links appeared in a lightly shaded area of the page but were otherwise similar in appearance to the organic links. The data do not report how many sponsored links were presented on each page of search results.

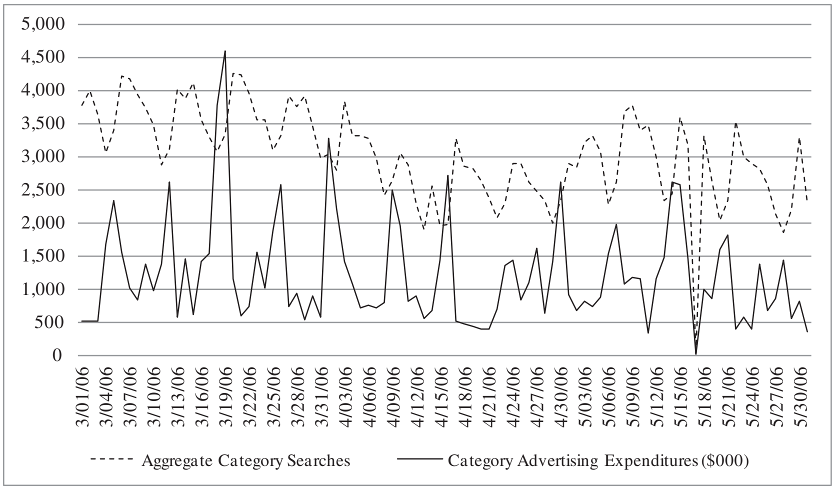

Fig. 1. Average category search and ad spending by date.

In summary, the query mining algorithm identified about 57,500 queries containing financial services brand-related keywords and 227,621 generic product category queries that did not contain any branded keywords.

2.2. Television advertising data

The television advertising data were drawn from Kantar Media's "Stradegy" database. They report the brand, date, exact time and estimated expenditure of all paid advertisements carried by national networks and local television stations in the US. In the absence of audience data, it is necessary to presume that ad prices correlate with program audience sizes (see e.g., Wilbur, 2008).

2.3. Consumer segmentation

A question of particular interest is how advertising's impact on online search behavior varies across consumer segments. The search data do not provide consumer demographics or geographic information, so consumers are segmented according to usage, as in Viswanathan et al. (2007). Discrete consumer segmentation is used so that segment-specific results have clear interpretations.

The segmentation scheme divides consumers along two dimensions: frequency of search and time elapsed since last search. A consumer is categorized as a "Frequent Searcher" if she enters more non-financial-services-related queries during the sample than average; this distinction is often labeled "heavy/light" in the scanner panel data literature and is very similar to the frequent/infrequent

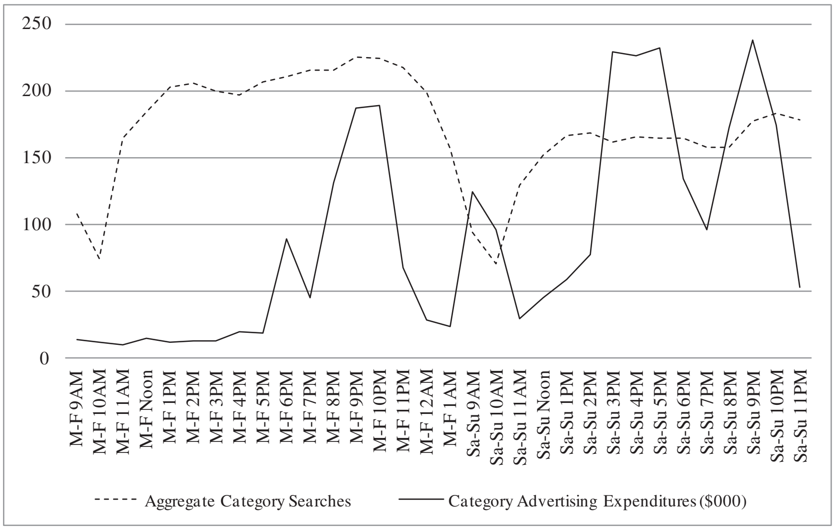

Fig. 2. Aggregated category search and ad spending by day/hour.

website visitor distinction commonly used in analysis of website visitation data (e.g. Pauwels et al., 2011). A consumer is categorized as a “Recent Searcher” if she has entered any query within the previous hour. Otherwise, she is classified as “not recent” (NR). Using non-financial-services-related queries ensures that the segmentation does not rely on the dependent variables in the regression. The four segments (Frequent/NR, Frequent/Recent, Infrequent/NR, Infrequent/Recent) have average sizes of 48%, 13%, 34%, and 5%, respectively.

2.4. Brand characteristics, advertising content, time effects, and time-varying shifters

Several brand characteristics were considered: brand awareness, total assets under management, and asset growth in the previous twelve months. Brand awareness metrics were unavailable for most brands in the category, so brand age is used as a proxy. After gathering data on these three brand characteristics, it was found that age and total assets are highly correlated and that total assets and recent asset growth are almost perfectly correlated. Therefore, brand age is the only brand characteristic included in the analysis, but the results are very similar if total assets or recent asset growth is used in its place.

Similarly, a number of advertising content characteristics were considered. Intuitively, if an advertisement emphasizes the brand's website, it would likely increase web traffic and search activity for the brand. “Web Emphasis” was defined to indicate that the web address was spoken during the course of the commercial while “Phone Emphasis” indicates that the phone number was spoken during the course of the commercial. These variables are very highly correlated with other measures of web and phone prominence such as font size and duration on screen.

Several subjective measures of advertising content were also considered, such as whether the ad contained a celebrity spokesperson or a humor appeal. However, preliminary analysis did not reveal any impact of these effects on consumer search behavior. They were excluded from the analysis as they are more susceptible to coding error.

We used the Dow Jones Industrial Average (DJIA) as an exogenous time-varying shifter of online search behaviors. The DJIA average is reported in many news outlets and therefore may influence consumers' tendencies to search for financial services keywords. The model includes two DJIA variables: one indicates the absolute positive change since the most recent trading day's opening value and the other measures the absolute negative change since the most recent trading day's opening value. This allows for asymmetric effects. For example, a 1% decrease may have a bigger effect on consumer search than a 1% increase if consumers are loss-averse.

Baseline consumer tendencies to search are captured by brand dummies and time fixed effects. Two time controls are used. Week fixed effects allow the baseline to vary across weeks. Two sets of weekday/hour and weekend/hour fixed effects (collectively referred to as “day-hour fixed effects”) allowed the baseline to vary across hours during weekdays and hours during weekends. This approach resulted in a better balance between model fit and complexity (as measured by AIC and BIC) than a full set of 168 weekday/hour dummies, perhaps because access to desktop computers changed consumer search tendencies between weekdays and weekend days, but not much within those two classifications.

The hours between 2:00 AM and 9:00 AM are dropped due to near-zero advertising and search activity. In summary, the dataset contains 1548 h over three months and 22 financial services brands, resulting in 34,056 total observations.

2.5. Descriptive statistics

Table 2 shows that keyword search and result click rates vary substantially across brands. Some brands, such as E-Trade and TD Ameritrade, focused primarily on providing web services and were searched frequently. Others, like Van Kampen, targeted financial advisors and other intermediaries, and were searched less frequently. These differences demonstrate the importance of estimating brand-specific search tendencies to cleanly identify how these actions change with variation in advertising over time. Search tendency is higher on weekdays than on weekend days, helping to explain the structure of the day/hour fixed effects described above.

2.6. Simultaneity

Television advertising cannot directly cause search behavior because ads are normally purchased far in advance of the dates on which they air. However, it could be that brands anticipated when consumers would search and placed ads at points in time most likely to influence search. To explore this possibility, Fig. 1 shows the aggregate category searches and advertising expenditures for each day in the sample. The correlation between the two variables (−.02) is not statistically distinguishable from zero. Fig. 2 shows how average search tendency and advertising expenditures moved over the day/hour periods in the analysis. Again, the correlation (.08) is not statistically significantly different from zero.

It is notable that sample advertisements emphasized telephone contact more than twice as often as web contact. In fact, several frequently-searched brands like Charles Schwab did not even provide their website addresses in their television advertisements. These observations are consistent with typical advertising practice, in which television and search advertising campaigns are run by different agencies and designed to achieve different objectives.

While it is impossible to prove exogeneity, prior domain knowledge and conversations with advertising practitioners in this product category indicated that financial services brands did not use expectations about online search data in planning their advertising campaigns in 2006. Therefore, we attach a causal interpretation to the findings reported below.

3. Model and estimation

The goals of the model are to derive estimating equations that relate online search behavior to TV advertising. Because the search data are recorded at the individual level and do not report individual exposures to television advertising, they are aggregated over individuals within each segment, brand and hour. Thus, the data are analyzed at the level of the brand-hour.

Three search behaviors are of primary interest. The first is the number of searchers in each segment who conducted at least one product category-related search in each hour of the dataset. The second is the number of searchers in each segment who, conditional on having searched at least once, entered at least one branded keyword. The third is the number of searchers in each segment who, conditional on having searched at least one branded keyword in a given hour, clicked at least one result link. All three search behaviors count the number of *searchers* who took an action rather than the number of actions themselves. Counting searchers, rather than search behaviors, gives a truer indication of the potential profit impact of each search behavior.

The second behavioral variable, the choice of at least one branded keyword, was considered very carefully. It was initially thought that searchers might often search multiple brands' keywords within a single hour, for example a search for “Fidelity” followed shortly afterwards by a search for “Charles Schwab.” In fact this is quite rare. 99.2% of the searchers in the dataset never searched keywords for two or more brands within any single hour. The overwhelming prevalence of searching for exactly zero or one brands within any given hour suggested a multinomial choice model as an appropriate representation of searchers' keyword choice behavior.

3.1. Search behavior

The first behavior modeled is the share of consumers who search for any financial services keyword. At the individual level, this is a dichotomous action (does the individual search or not), so a binary response model is used and then aggregated across individuals within each segment. A searcher in segment z searches at time t ($S_z^t = 1$) with probability

$$\Pr_{1t}^z = \Pr(S_t^z = 1) = \frac{\exp(\gamma_1^z + \ln(1+A_t)\beta_{1t}^z + DJIA_t\varphi_1^z + X_t\phi_1^z)}{1 + \exp(\gamma_1^z + \ln(1+A_t)\beta_{1t}^z + DJIA_t\varphi_1^z + X_t\phi_1^z)}, \tag{1}$$

where λ_1^z is a segment-specific baseline tendency to search, A_t is a category advertising stock, $DJIA_t$ is a vector containing the two stock index variables described in Section 2.5, and X_t is a vector containing week fixed effects and day-hour fixed effects.

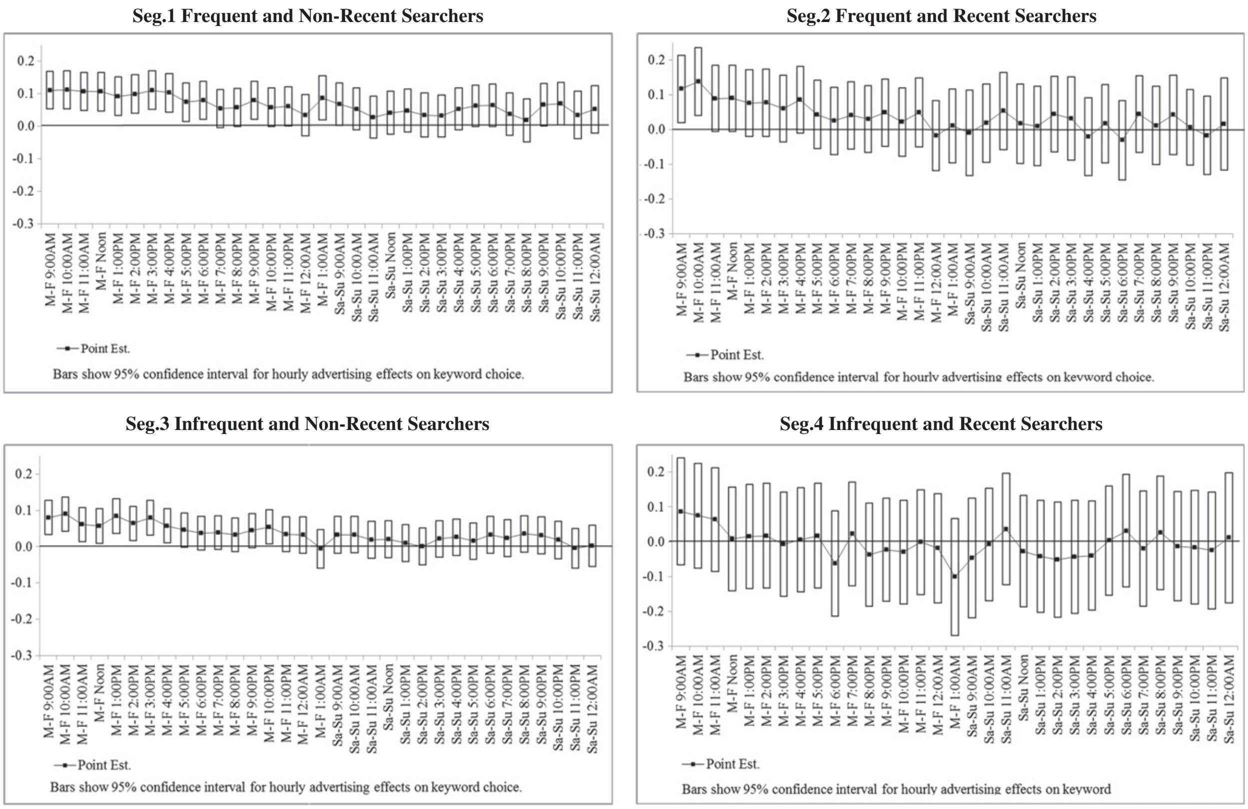

Fig. 3. Keyword choice day/hour advertising effects.

Table 3
Brand specific advertising effects for keyword choice.

	Consumer segment							
	Frequent, NR		Frequent, Recent		Infrequent, NR		Infrequent, Recent	
	Est.	St. Err.	Est.	St. Err.	Est.	St. Err.	Est.	St. Err.
Brand age	−.001**	(.000)	−.001	(.002)	−.001**	(.000)	−.001**	(.000)
Web emphasis	*.009*	(.005)	−.018	(.011)	.000	(.003)	.003	(.012)
Phone emphasis	−.007	(.004)	.000	(.007)	−*.004*	(.002)	−.024**	(.009)

Note: Estimates in italic are significant at the 90% confidence level.
** Significant at the 99% confidence level.

Table 4
Brand specific advertising effects for click behavior.

	Consumer segment							
	Frequent, NR		Frequent, Recent		Infrequent, NR		Infrequent, Recent	
	Est.	St. Err.	Est.	St. Err.	Est.	St. Err.	Est.	St. Err.
Brand age	.004**	(.000)	.004**	(.001)	.001**	(.000)	.002**	(.001)
Web emphasis	.012	(.010)	.015	(.020)	.002	(.006)	.011	(.021)
Phone emphasis	−.004	(.008)	−.010	(.014)	−.001	(.004)	.023	(.016)

** Significant at the 99% confidence level.

Advertising stock is typically modeled using either a distributed-lag specification or an exponential smoothing specification. Both specifications were considered. Appendix 3 shows that the decay of parameter estimates in a distributed-lag specification appears approximately exponential. Therefore, the more parsimonious model is presented. Advertising stock is given by

$$A_t = adv_t + \lambda^z A_{t-1} \tag{2}$$

where adv_t is the cumulative category advertising expenditure at time t, and λ^z is an advertising carryover parameter for segment z measuring persistence of effects in hours.

Advertising responsiveness β_{1t}^z is

$$\beta_{1t}^z = \theta_1^z + \theta_{1t}^z, \tag{3}$$

where θ_1^z represents the baseline effect of category advertising on search probability, and x_t contains week fixed effects and day-hour fixed effects.

Let n_t^z be the number of people in segment z at time t, and let c_{1t}^z be the number of people in segment z who search at least one financial services keyword in time t. The distribution of any single consumer in segment z searching for at least one keyword in time t is Bernoulli with the probability given in Eq. (1), so the probability of observing c_{1t}^z consumers in segment z searching at time t is binomial:

$$L_{1t}^z = \frac{n_t^z!}{c_{1t}^z!(n_t^z - c_{1t}^z)!}\left(\Pr_{1t}^z\right)^{c_{1t}^z}\left(1 - \Pr_{1t}^z\right)^{n_t^z - c_{1t}^z}. \tag{4}$$

3.2. Keyword choice

The second behavior modeled is the choice of a keyword for brand $K = 1, \ldots, K$ or a generic query (without any branded keywords, denoted $K = 0$). The conditional probability that a searcher in segment z chooses a keyword associated with brand k at time t ($K_t^z = k$)

Table 5
Advertising carryover parameter estimates.

	Carryover Est.	St. Err.	90% life (h)
Frequent, Not Recent	.630**	(.060)	5
Frequent, Recent	.830**	(.037)	13
Infrequent, Not Recent	.500**	(.075)	4
Infrequent, Recent	.300**	(.030)	2

** Significant at the 99% confidence level.

Table 6
Influence of advertising on keyword choice during business hours.

	Est.	95% bootstrap CI	Firm age	Adv. exp. ($000)
AG Edwards[a]	−.018	(−.054, .012)	119	2085
Wachovia	.004	(−.025, .029)	98	6073
Merrill Lynch	.011	(−.021, .037)	92	10,236
Edward Jones	.020	(−.013, .046)	84	4877
Van Kampen	.026	(−.008, .053)	79	1188
T. Rowe Price	**.037**	(.001, .066)	69	8349
Fidelity	**.043**	(.007, .073)	60	16,881
Oppenheimer	**.062**	(.019, .101)	46	2888
Raymond James	**.065**	(.021, .104)	44	926
Charles Schwab	**.073**	(.026, .116)	35	18,381
TD Ameritrade	**.072**	(.026, .113)	31	32,465
Scottrade	**.083**	(.032, .129)	26	964
E-Trade	**.084**	(.033, .133)	15	6309
Young Brands[b]	**.074**	(.037, .109)		

Note: Estimates in bold are significantly different from zero.
[a] Only brands with average daily spending in excess of $10,000 are shown.
[b] Young brand stands for brands of which ages are less than 50 years as of 2006, and business hours are from 9 AM to 5 PM weekdays.

Table 7
Goodness of fit.

	Pseudo R-squared				
	Freq., NR	Freq./rec.	Infreq., NR	Infreq./rec.	Overall
Search behavior	.806	.790	.619	.630	.732
Keyword choice	.615	.685	.550	.670	.588
Click behavior	.191	.160	.067	.069	.106
All models	.778	.771	.589	.630	.697

conditional on searching financial services keywords is

$$\Pr{}_{2kt}^{z} = \Pr\left(K_t^z = k \middle| S_t^z = 1\right) = \frac{\exp\left(\gamma_{2k}^z + \ln(1+A_{kt})\beta_{kt}^z + DJIA_t\varphi_2^z + X_t\phi_2^z\right)}{1 + \sum_{k'} \exp\left(\gamma_{2k'}^z + \ln(1+A_{k't})\beta_{2k't}^z + DJIA_t\varphi_2^z + X_t\phi_2^z\right),} \tag{5}$$

where γ_{2k}^z is the baseline brand- and segment-specific search tendency, and $A_{kt} = adv_{kt} + \lambda^z A_{kt-1}$ is advertising stock for brand k at time t. The proportion of searches containing only generic keywords is $\Pr_{20t}^z = 1 - \sum_{k=1}^{K} \Pr_{2kt}^z \cdot \beta_{2kt}^z$ relates advertising impact on brand keyword choice to brand and ad content characteristics:

$$\beta_{2kt}^z = \theta_2^z + web_{kt}\theta_{2,web}^z + phone_{kt}\theta_{2,phone}^z + age_k\theta_{2,age}^z + x_t\theta_{2t}^z \tag{6}$$

where θ_2^z represents the baseline impact of brand advertising on brand keyword choice, web_{kt} is the fraction of advertising spending by brand k at time t that emphasized web contact and $phone_{kt}$ is the fraction of advertising spending by brand k at time t that emphasized telephone contact. age_k is the age of brand k, which proxies for brand recognition and assets under management.

Let c_{2kt}^z be the number of people in segment z at time t who searched at least once for brand k. Out of the c_{1kt}^z people who searched in time t, the probability of observing that each of the $K = 0, \ldots, K$ options is searched by $(c_{20t}^z, \ldots, c_{2kt}^z, \ldots, c_{2Kt}^z)$ people is given by a multinomial density,

$$L_{2t}^z = \frac{c_{1t}^z!}{\prod_{k=0}^{K} c_{2kt}^z!} \prod_{k=0}^{K} \left(\Pr{}_{2kt}^z\right)^{c_{2kt}^z}. \tag{7}$$

3.3. Click behavior

The third behavior modeled is, conditional on searching at least once for brand k at time t, the probability of clicking at least one result. This behavior has two outcomes and therefore is modeled using a binary response model.[3] A searcher in segment z clicks at

[3] Tobit and Poisson models produced qualitatively similar results, so the simplest model is presented.

Table 8
Holdout Validation.

		Root Mean Squared Error			
		Freq., NR	Freq./rec.	Infreq., NR	Infreq./rec.
Search behavior	In-sample	.006	.012	.011	.027
	Out-of-sample	.006	.011	.011	.028
Keyword choice	In-sample	.021	.035	.019	.043
	Out-of-sample	.022	.036	.020	.044
Click behavior	In-sample	.158	.106	.211	.110
	Out-of-sample	.170	.112	.214	.116

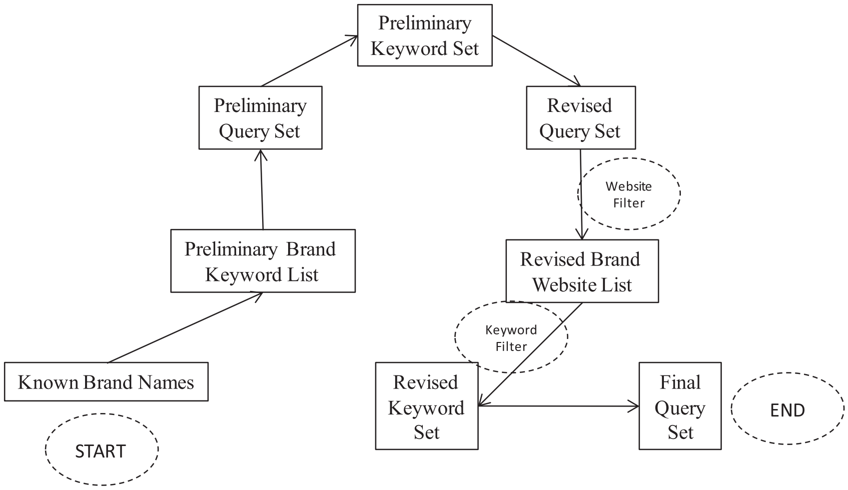

Fig. A1. Query mining logical structure.

least one link related to brand k at time t ($C^z_{kt} = 1$) conditional on obtaining search results for brand k with probability

$$\Pr^z_{3kt} = \Pr(C^z_{kt} = 1 | K^z_t = k) = \frac{exp(\gamma^z_{3k} + \ln(1 + A_{kt})\beta^z_{3kt} + DJIA_t\varphi^z_3 + X_t\phi^z_3)}{1 + exp(\gamma^z_{3k} + \ln(1 + A_{kt})\beta^z_{3kt} + DJIA_t\varphi^z_3 + X_t\phi^z_3)}, \tag{8}$$

where γ^z_{3kt} represents the baseline tendency to click at least one link after searching brand k, and advertising effect β^z_{3kt} has the same structure as β^z_{2kt} given in Eq. (6) above.

Let c^z_{3kt} be the number of those people who clicked at least one result link after searching brand keyword k at time t. Out of the c^z_{2kt} people who searched at least once for brand k at time t, the probability of observing that c^z_{3kt} people clicked at least one link is binomial,

$$L^z_{3kt} = \frac{c^z_{2kt}!}{c^z_{3kt}!(c^z_{2kt} - c^z_{3kt})!} \left(\Pr^z_{3kt}\right)^{c^z_{3kt}} \left(1 - \Pr^z_{3kt}\right)^{c^z_{2kt} - c^z_{3kt}} \tag{9}$$

The parameters of the three-level model are estimated simultaneously by Maximum Likelihood. The joint likelihood function is:

$$LL = \sum_z \log\left(L^z_{1t} L^z_{2t} \prod_k L^z_{3kt}\right). \tag{10}$$

4. Results

This section shows how television advertising is related to search, keyword choice, and click behaviors. It then presents advertising persistence parameters, keyword choice advertising elasticities, model fit and holdout sample results. The non-advertising parameter estimates are generally reasonable, but not the focus of the study, so they are provided in Appendix 4.

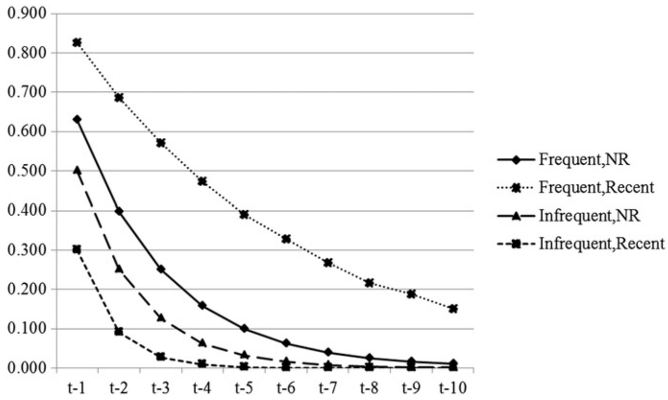

Fig. A2. Keyword choice distributed lag advertising parameter estimates.

4.1. Television advertising and category search

The results show *almost* no evidence that television advertising is related to the number of individuals who search financial product category. The confidence intervals—reported in Fig. OA1 of the Online Appendix—indicate that TV advertising is not found to affect category search incidence in any day-hour combination.

This non-finding makes sense for the financial services category. It suggests that television advertising does not lead searchers to enter the product category as opposed to other categories. This seems likely, as information search for financial services is more likely driven by personal factors—such as a change in family structure or income level—than by television advertising. However, if the set of category-related generic keywords had not been identified, the study might have led us to the erroneous conclusion that television advertising increases category search incidence.

We do not advise generalizing this finding beyond mature product categories. Financial services brands advertise heavily and have high levels of awareness. It seems likely that television advertising would initiate category search for new categories or new brands.

4.2. Television advertising and keyword choice

In contrast to category search incidence, hourly advertising parameter estimates in Fig. 3 show that television advertising is positively related to searchers' tendency to choose branded keywords as opposed to generic category keywords. This effect is largest for searchers who have not searched within the past hour. Within this group, the effects are larger for searchers who search frequently than for infrequent searchers. This may suggest that heavy internet users are more likely to respond to ads by searching than light users. Advertising's effects on branded keyword choice are strongest during business hours on weekdays, especially so for searchers who have not searched within the past hour. Because many financial services consumers are white-collar professionals, it is logical that their online usage is greatest during business hours due to easy access to a computer. As Fig. 2 shows, category search incidence is highest on weekdays between 9 AM and 5 PM. Consumers may be responding to TV advertising seen the previous evening or they may be exposed to TV ads during the day, such as through a television set in a waiting room, lobby, lounge, elevator or retail store.

Advertising effects related to brand-specific characteristics are presented in Table 3. Television advertising for younger brands has larger effects on most segments' keyword choices than advertising for older brands. This stands to reason, as younger firms such as TD Ameritrade and E-Trade started delivering services online earlier than older firms such as Merrill Lynch. Consumers would naturally be more likely to respond to younger firms' advertising by going online. It is also possible that younger firms' advertising targets younger consumers, and that younger consumers are more likely to respond via online search than other means.

The advertising content variables in Table 3 reveal two complementary patterns. First, *web emphasis* increased branded keyword search among frequent searchers who had not searched recently. The effect is positive and significant at the 90% confidence level. Second, *phone emphasis* reduced the frequency of branded keyword choice among infrequent searchers. The effect is significant at the 99% confidence level for infrequent searchers who searched recently, and significant at the 90% level for those who had not searched recently. If frequent searchers are more likely to respond to TV ads online, and if infrequent searchers are more likely to respond to TV ads by phone, then these two effects suggest that choosing a call-to-action that matches an advertisement's target segment is a strategy that can help to maximize advertising effectiveness. In other words, a television advertisement should emphasize a target segment's preferred means of information acquisition in order to increase its impact.

Overall, the finding that television advertising did not increase the number of people who searched in the category, but did increase the number of people who searched branded keywords, suggests that television advertising does not initiate category search but does shift share of searchers among competing brands. This is similar to findings in the advertising effectiveness literature about how advertising influences sales in mature product categories (Tellis, 2004).

4.3. Television advertising and click behavior

The click-through results show limited evidence that television advertising is related to search result click-through rates. Only three of the 192 segment-specific day-hour fixed effects were statistically significant (as shown in Fig. OA2 in the Online Appendix). In consideration of possible Type I error, we interpret the low proportion of significant results as a null effect of advertising timing on click result probability.

Table 4 presents the effects of brand age and advertising content on the probability that a consumer clicks a search result. While the effect of TV advertising on branded keyword choice was smaller for older brands, the probability of a search result click for older brands' keywords *increases* with TV advertising expenditure. In addition, this effect is bigger for frequent searchers. This may be because searches feel more satisfied with search results for older brands' keywords, due to their brand associations, and this effect might be reinforced by advertising. With regard to advertising content, the parameter estimates do not indicate that searchers' tendency to click search results has any relation to television advertising content.

4.4. Advertising carryover

The relationship between television advertising and search behavior persisted for hours, not days, as shown in Table 5. Infrequent searchers who have searched recently have the shortest shelf life, as 90% of the carryover is exhausted within the first two hours, while frequent searchers who have searched recently have the longest memories (13 h). In general, carryover estimates are larger for frequent searchers than for infrequent searchers.

4.5. Keyword choice advertising elasticities

It appears the primary effect of television advertising is on searchers' keyword choice. How big is that effect? A simple way to measure effect magnitudes is to use elasticity calculations, as in the advertising effectiveness literature, e.g., Ataman, van Heerde, and Mela (2010) and Lodish et al. (1995).

The elasticity measures the cumulative increase in the number of people searching branded keywords instead of generic keywords in response to a change in television advertising. The cumulative number of people who search branded keywords in future periods is denoted $Q_{kt} = \sum_{z}\sum_{\tau=t}^{T_z} c^z_{1\tau}\,\Pr^z_{2k\tau}$, where T_z denotes the segment-specific duration for an advertisement aired in time t in Table 5. The elasticity formula is

$$\begin{aligned}\frac{\partial Q_{kt}}{\partial adv_{kt}}\frac{adv_{kt}}{Q_{kt}} &= \frac{adv_{kt}}{Q_{kt}}\sum_{z}\sum_{\tau=t}^{T_z} c^z_{1\tau}\frac{\partial \Pr^z_{2k\tau}}{\partial adv_{kt}} \\ &= \frac{adv_{kt}}{Q_{kt}}\sum_{z}\sum_{\tau=t}^{T_z} c^z_{1\tau}\Pr^z_{2k\tau}\left(1-\Pr^z_{2k\tau}\right)\beta^z_{2k\tau}\frac{\left(\lambda^z\right)^{\tau-t}}{1+A_{k\tau}}.\end{aligned}$$

The interpretation of this measure is the change in the number of searchers using branded keywords given a 1% change in the advertising expenditure of brand k during standard business hours.

Table 6 shows the brand-specific point estimates and confidence intervals for the elasticity of keyword choice during standard business hours (Monday–Friday, 9 AM–5 PM) with respect to television advertising. Consistent with the empirical results presented earlier, television advertising by younger brands has a stronger impact on branded keyword choice. For brands younger than fifty years old, the average elasticity is .07, with a confidence interval ranging from .03 to .10.

The estimated elasticity is larger than the short-term market share elasticity of advertising (.05) and smaller than the long-term market share elasticity of advertising (.10) reported by Lodish et al. (1995). It similarly lies between the short-run (.01) and long-run (.12) advertising elasticities found by Ataman et al. (2010).

4.6. Model validation

Table 7 presents the goodness-of-fit statistics for each model and segment. Pseudo R-squared statistics are based on Likelihood Ratio statistics (McFadden, 1974). Overall, the model explains about 70% of the variation in the dependent variables.

A large number of covariates were included in the model, so the model was re-estimated on a subsample of the data and prediction errors were compared to a holdout sample. The advertising carryover specification prevents a random holdout validation exercise, so holdout validation was done via a nonrandom split (Keane & Wolpin, 2007). The model is re-estimated using data from March and April and fitted values are produced for May.

The results show that the model did not overfit the hold-out period, despite the large number of fixed effects included. Table 8 presents the Root Mean Square Error (RMSE) comparisons between in-sample and out-of-sample observed and predicted choice probabilities, by model and segment. Error magnitudes are comparable across dependent variables and segments; the largest single difference is about 10%, and six of twelve model/segments have differences smaller than 5%.

5. Discussion

This paper has presented how offline advertising expenditures are related to online search using the AOL search dataset. Individual level search data and advertising content data helped to understand *how* advertisements affect *who* searches for product information *when* and *how*. No evidence was found that TV advertising increases the number of people who search financial services-related queries. Instead, advertising expenditures increase searchers' tendency to search brand-related keywords at the expense of generic keywords. The effect is largest for young brands during business hours, with effects comparable to extant measurements of the elasticity of advertising on sales.

The key implication of these results for firms is that television advertising may affect the outcome of their search advertising campaigns. Because television advertising increases the number of searchers using branded keywords, advertisers can expect searches that coincide with television campaigns to i) include a higher share of branded keyword traffic (which tends to be cheaper than ads on generic keywords), and ii) searchers will be exposed to less competitive information than if they were to search generic keywords, which may increase conversion rates. Given these positive spillovers from television to search, ROI metrics should incorporate such interaction effects to optimize both offline and online marketing expenditures.[4] In addition, the results also imply that maximizing television advertising effectiveness requires selecting a call to action that corresponds to the target segment's preferred mode of response.

Future academic work on search advertising could consider that search advertisers may be able to influence consumer search via offline marketing. Perhaps the default assumption about how traditional advertising influences consumer search for established brands in mature product categories should be that it shifts consumers' keyword choices within the product category, rather than their tendencies to search.

Several interesting possibilities could be explored in this area. For example, if television advertising primarily diverts consumers from competitive generic keyword searches to branded keyword searches, then brands' television advertising expenditures are more likely to be strategic complements. On the other hand, if television advertising does more to stimulate category search rather than to alter consumers' keywords, then television advertising expenditures are more likely to be strategic substitutes, as advertisers free-ride on rivals' TV spending. It also may be that the most profitable way to compete for keywords and clicks is to use television advertising to divert consumers into less-competitive branded keyword auctions.

This study has a number of limitations, which suggest future research opportunities. For one, we study a single mature product category over a three-month period, but a new or evolving category might produce different results. For another, the available data limited the scope of the research questions, as the data did not include clicks on internet display ads, direct website traffic, or distinguish among clicks on paid or organic links. Many steps were taken to rule out alternative explanations—allowing baseline search behaviors to vary with brand and time, using exogenous time-varying factors to isolate the effect of advertising, and selecting a category that is not subject to obvious temporal endogeneity and that did not promote its online offerings to a significant extent. We hope to stimulate additional research in this new area as more data become available to researchers in the future.

Appendix 1. Legal and ethical implications of using AOL search data

The search dataset used in this paper was released by AOL for non-commercial research.[5] However, AOL did not seek users' permission prior to releasing the data. Journalists quickly identified a few consumers who had searched personally identifying keywords such as telephone numbers and physical addresses (Barbaro & Zeller, 2006), leading the event to be labeled one of the "101 dumbest moments in business." The executives involved resigned, AOL was sued for violating users' privacy, and an ethical debate ensued. Anderson (2006) summarized the views of Jeffrey Seglin, a syndicated ethics columnist: "'If researchers are using the data to identify individual users, I believe they've crossed an ethical line,' Seglin says. But if they don't try to match up the data with actual people, he believes they are (ethically) in the clear."

Legal and ethical questions are clearly important in the context of using the AOL search data. The "readme.txt" file that AOL distributed with the data claims copyright or the data. According to the US Copyright Office, copyrighted works can be reproduced without the copyright holder's permission when the reproducer's purpose is covered under "fair use," which explicitly includes scholarship and research as a valid purpose.[6] These data have been used by numerous articles published in conference proceedings and scholarly journals. That such purposes are consistent with the "fair use" doctrine might explain why AOL has not initiated any known legal actions against those researchers or outlets.

Ethical concerns are at least as important as legal concerns. On this point the authors ensured compliance with academic institutional review board policies. For example, the Duke University policy on research with human subjects states that "existing data may have been collected for research purposes or non-research purposes, such as driver's license information or school records… if data are provided to researchers without identifiers, the proposed research does not meet the definition of research with human subjects." Further protecting individual privacy, all data preparation steps were performed algorithmically ("scrubbing" the data of any personally identifying information) and the statistical analysis was performed on aggregated data only.

[4] As Swasy and Rethans (1986) suggested, "advertisers should begin to incorporate a measure of curiosity generation and question solicitation in their new-product ad concept and ad-testing procedures."

[5] http://search-logs.com/aol_readme.txt.

[6] http://www.copyright.gov/fls/fl102.html.

Appendix 2. Details of query mining technique to find category keywords

This appendix describes the query mining algorithm. It draws on three papers in particular. Abhishek and Hosanagar (2007) introduced the concept of semantic similarity, using the similarity in the results returned by two different search queries to identify commonalities among the queries. Building on this, Fuxman et al. (2008) found that keyword suggestions are maximally relevant when the analyst uses clickstream data to identify similarities in result clicks to identify similarities in search queries. Chen et al. (2008) added concept hierarchy to semantic similarity, showing that keyword context is an important determinant of keyword meaning, a technique which helps to screen out irrelevant keywords. The algorithm used here was developed independently but is related to ideas in all three of these papers.

The basic idea is to iterate between sets of brand-related websites and the keywords that led to clicks on those sites to identify relevant queries, refining the data at several points. Three examples will facilitate the exposition. Suppose consumer A searches "Information about Fidelity retirement" and clicks Fidelity.com. Suppose consumer B searches "Fidelity NetBenefits 401K" and clicks a Fidelity subsidiary, NetBenefits.com. Assume, for the purposes of this example, that the researcher does not know NetBenefits is affiliated with Fidelity. Consumer C searches the word "retirement homes in Chicago" and clicks brookdaleliving.com. Starting from Fidelity.com, the query mining procedure will classify "Fidelity" and "NetBenefits" as keywords related to the Fidelity brand, "retirement" and "401K" as generic keywords, and "Information," "about," "homes," and "Chicago" as irrelevant keywords. It will identify Fidelity.com and Netbenefits.com as brand-related websites, and brookdaleliving.com as an irrelevant website. The discussion below uses these examples to illustrate how the procedure works.

Fig. A1 displays the logical structure of the query mining technique. The analyst starts with a list of Known Brand Names, such as "Fidelity." From this list, a Preliminary Website List is compiled by searching each of the known brand names, and including the first brand-owned website returned as an organic result to a search on the brand name. From a Known Brand List that only contains "Fidelity," this step would produce a Preliminary Website List that only contains Fidelity.com.

The Preliminary Query Set is identified as all queries in the data that led to result clicks to any website in the Preliminary Website List. In the examples above, the Preliminary Query Set would include only consumer A's query. The Preliminary Keyword Set is identified as all unique keywords contained in this Preliminary Query Set.[7] In the examples above, this set is "information," "about," "Fidelity," and "retirement."

The Revised Query Set is defined as all queries containing any member of the Preliminary Keyword Set. In the examples above, the Revised Query Set would contain all three consumers' queries, since consumer B's query includes the word "Fidelity" and consumer C's query includes the word "retirement."

The Website Filter generates a list of category-related websites. It works in three steps. First, the analyst compiles a list of unique websites consumers clicked after searching queries in the Revised Query Set. In the examples above, this set would include fidelity.com, netbenefits.com, and brookdaleliving.com. Second, for each website, the analyst calculates the aggregate click-through rate from any query in the Revised Query Set to the site in question. Sites with low aggregate click-through rates are dropped. If the above example were replicated on a large scale, this step would drop brookdaleliving.com. Third, the analyst manually visits each remaining site and classifies it as either (i) associated with a particular brand, (ii) a generic site, or (iii) unrelated; unrelated sites are dropped. This completes the Revised Website List.

The Keyword Filter generates a list of category-related keywords from the Revised Website List. Like the Website Filter, it works in 3 steps. First, the analyst compiles the list of unique keywords contained in queries that led to clicks on the Revised Website List. In the examples above, this step includes "Information," "about," "Fidelity," "retirement," "Netbenefits," and "401K." Second, for each keyword, the analyst calculates the aggregate click-through rate from all queries containing that keyword to any site in the Revised Website List and drops those with low aggregate click-through rates. This step removes most unrelated keywords. In the examples above, this would drop "about." Third, the analyst classifies each remaining keyword as either belonging to a particular brand or as a generic keyword. For each keyword w and each brand k, the analyst calculates CTR_{wk} as the aggregate click-through rate from all queries containing keyword w to any website associated with brand k. When the ratio of CTR_{wk} to $CTR_{wk'}$ is larger than the threshold for all brands $k' \neq k$, then keyword w is associated with brand k; otherwise keyword w is classified as a generic keyword. In the examples above, this step would associate "Fidelity" and "Netbenefits" with the brand Fidelity, and would classify "retirement" and "401K" as generic keywords. The output of this step is the Revised Keyword List, which should be reviewed by the analyst for reasonableness.[8]

The Final Query Set is defined as all queries containing any member of the Revised Keyword Set. A branded query is any query that contains at least one brand-related keyword. A generic query is any query that does not contain any brand-related keywords. In the current application, it was virtually never the case that a single query contains keywords associated two or more different brands, but many queries contain multiple brand-related keywords for the same brand, and many queries combine brand-related keywords for a single brand with generic keywords. Note that the Final Query Set will contain many queries that did not lead to any result click.

In the current research, 22 known brands were identified from the advertising database. The Preliminary Website List for these brands produced 37,398 queries in the Preliminary Query Set. This query set included 1843 unique keywords in the Preliminary Keyword Set. The Revised Query Set contained 906,538 queries. These queries produced clicks to 2389 unique URLs, which the Website

[7] Infrequently occurring keywords are excluded from the preliminary keyword set. In the current application, keywords that appeared in fewer than four queries were dropped.

[8] It would be possible to iterate query selection, website filtering and keyword filtering as needed if the analyst suspects that some important category keywords have not been included.

Filter reduced to 53 sites in the Revised Website List. The Keyword Filter resulted in 199 unique keywords in the Revised Keyword List, of which 92 were generic and 107 were branded. The generic keywords produced 245,419 queries in the Final Query Set, and the branded keywords contributed 67,733 queries.

Appendix 3. Distributed lag carryover specifications of advertising effectiveness

This appendix presents an alternative specification of advertising stock variables to see the patterns of decay over time. In place of exponential decay function (e.g., Eq. (2)), distributed lag carryover specification is tested. The advertising stock variables are defined as:

$$A_t = \sum_{\tau=t-L}^{t} \lambda_\tau^z \cdot adv_\tau$$

for category search behavior model, where λ_τ^z are lagged advertising parameters and L is the number of lags, and

$$A_{kt} = \sum_{\tau=t-L}^{t} \lambda_\tau^z \cdot adv_{k\tau}$$

for keyword choice and click decision models. We normalize $\lambda_t^z = 1$ for identification. There is no parametric assumption restricting the λ_τ^z estimates.

Fig. A2 shows the patterns of lagged advertising parameters with 10 lags for keyword choice model. Parameter estimates in all four segments are quite consistent with exponential decay pattern. This pattern holds in all other models — category search incidence and click tendency.

Appendix 4. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ijresmar.2014.12.005.

References

Abhishek, V., & Hosanagar, K. (2007). Keyword generation for search engine advertising using semantic similarity between terms. *Proceedings of the ICEC*.

Abou Nabout, N., & Skiera, B. (2012). Return on quality improvements in search engine marketing. *Journal of Interactive Marketing*, *26*(3), 141–154.

Abou Nabout, N., Skiera, B., Stepanchuk, T., & Gerstmeier, E. (2012). An analysis of profitability of fee-based compensation plans for search engine marketing. *International Journal of Research in Marketing*, *29*(1), 68–80.

Abramson, C., Andrews, R. L., Currim, I. S., & Jones, M. (2000). Parameter bias from unobserved effects in the multinomial logit model of consumer choice. *Journal of Marketing Research*, *37*(4), 410–426.

Anderson, N. (2006). The ethics of using AOL search data. http://arstechnica.com/old/content/2006/08/7578.ars

Andrews, R., Ainslie, A., & Currim, I. (2002). Discrete versus continuous representations of heterogeneity. *Journal of Marketing Research*, *39*(4), 479–487.

Ataman, M. B., van Heerde, H. J., & Mela, C. F. (2010). The long-term effect of marketing strategy on brand performance. *Journal of Marketing Research*, *47*(5), 866–882.

Barbaro, M., & Zeller, T., Jr. (2006, August 9). *A face is exposed for AOL searcher no. 4417749*. New York Times.

Berman, R. (2013). Beyond the last touch: Attribution in online advertising. *Working paper, UC Berkeley*.

Bertrand, M., Karlan, D., Mullainathan, S., Shafir, E., & Zinman, J. (2010). What's advertising content worth? Evidence from a consumer credit marketing field experiment. *The Quarterly Journal of Economics*, *125*(1), 263–306.

Bettman, J., & Park, C. W. (1980). Effects of prior knowledge and experience and phase of the choice process on consumer decision processes: A protocol analysis. *Journal of Consumer Research*, *7*(3), 234–248.

Chen, Y., Xue, G., & Yu, Y. (2008). Advertising keyword suggestion based on concept hierarchy. *Proceedings of the WSWD 2008*.

Chiang, J., Chib, S., & Narasimhan, C. (1999). Markov chain Monte Carlo and models of consideration set and parameter heterogeneity. *Journal of Econometrics*, *89*(1–2), 223–248.

Desai, P. S., Shin, W., & Staelin, R. (2014). The company that you keep: When to buy a competitor''s keyword. *Marketing Science*, *233*(4), 485–508.

Fuxman, A., Tsaparas, P., Achan, K., & Agrawal, R. (2008). Using the wisdom of the crowds for keyword generation. *Proceedings of the WWW 2008*.

Guadagni, P., & Little, J. (1983). A logit model of brand choice. *Marketing Science*, *2*(3), 203–238.

Horsky, D., Misra, S., & Nelson, P. (2006). Observed and unobserved preference heterogeneity in brand-choice models. *Marketing Science*, *25*(4), 322–335.

Hu, Y., Du, R. Y., & Damangir, S. (2014). Decomposing the impact of advertising: Augmenting sales with online search data. *Journal of Marketing Research*, *51*(3), 300–319.

Jerath, K., Ma, L., & Park, Y. (2014). Consumer click behavior at a search engine: The role of keyword popularity. *Journal of Marketing Research*, *51*(4), 480–486.

Joo, M., Wilbur, K. C., Cowgill, B., & Zhu, Y. (2014). Television advertising and online search. *Management Science*, *60*(1), 56–73.

Kamakura, W., & Russell, G. (1989). A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research*, *26*(4), 379–390.

Keane, M. P. (1997). Modeling heterogeneity and state dependence in consumer choice behavior. *Journal of Business & Economic Statistics*, *15*(3), 310–327.

Keane, M. P., & Wolpin, K. I. (2007). Exploring the usefulness of a nonrandom holdout sample for model validation: Welfare effects on female behavior. *International Economic Review*, *48*(4), 1351–1378.

Kireyev, P., Pauwels, K., & Gupta, S. (2013). Do display ads influence search? Attribution and dynamics in online advertising. *Working Paper*. Harvard University.

Klein, L. R., & Ford, G. T. (2003). Consumer search for information in the Digital Age: An empirical study of prepurchase search for automobiles. *Journal of Interactive Marketing*, *17*(3), 29–49.

Kulkarni, G., Kannan, P. K., & Moe, W. (2011). Using online search data to forecast new product sales. *Decision Support Systems*, *52*(3), 604–611.

Kulkarni, G., Ratchford, B. T., & Kannan, P. K. (2012). The impact of online and offline information sources on automobile choice behavior. *Journal of Interactive Marketing*, *26*(3), 167–175.

Li, H., & Kannan, P. K. (2014). Attributing conversions in a multichannel online marketing environment: An empirical model and a field experiment. *Journal of Marketing Research*, *51*(1), 40–56.

Liaukonyte, J., Teixeira, T., & Wilbur, K. C. (2014). Television advertising and online shopping. *Marketing Science* (in press).

Lim, J., Currim, I. S., & Andrews, R. L. (2005). Consumer heterogeneity in the longer-term effects of price promotions. *International Journal of Research in Marketing*, *22*(4), 441–457.
Lodish, L. M., Abraham, M., Kalmenson, S., Livelsberger, J., Lubetkin, B., Richardson, B., & Stevens, M. E. (1995). How TV advertising works: A meta-analysis of 389 real world split cable TV advertising experiments. *Journal of Marketing Research*, *32*(2), 125–139.
McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, *3*(4), 303–328.
Moorman, C., Diehl, K., Brinberg, D., & Kidwell, B. (2004). Subjective knowledge, search locations, and consumer choice. *Journal of Consumer Research*, *31*(3), 673–680.
Moorthy, S., Ratchford, B. T., & Talukdar, D. (1997). Consumer information search revisited: Theory and empirical analysis. *Journal of Consumer Research*, *23*(4), 263–277.
Murthi, B. P. S., & Srinivasan, K. (1998). Performance of the integrated random coefficients covariance probit model: Implications for brand choice. *International Journal of Research in Marketing*, *15*(2), 137–156.
Neslin, S. A., Henderson, C., & Quelch, J. (1985). Consumer promotions and the acceleration of product purchases. *Marketing Science*, *4*(2), 147–165.
Newman, J. W., & Staelin, R. (1973). Information sources for durable goods. *Journal of Advertising Research*, *13*(2), 19–29.
Nielsen (2013). Global ad spend: With double digit gains, display ads drive in-your-face results. http://www.nielsen.com/us/en/newswire/2013/global-ad-spend-with-double-digit-gains-display-ads-drive-in-y.html (Accessed October 2013)
Park, C. H., & Park, Y. (2014). Investigating purchase conversion by uncovering online visit patterns. *Working paper*. Binghamton University.
Pauwels, K., Leeflang, P. S. H., Teerling, M. L., & Huizingh, K. R. E. (2011). Does online information drive offline revenues? Only for specific products and consumers segments! *Journal of Retailing*, *87*(1), 1–17.
Punj, G. N., & Staelin, R. (1983). A model of consumer information search behavior for new automobiles. *Journal of Consumer Research*, *9*(4), 366–380.
Ratchford, B. T., Talukdar, D., & Lee, M. (2007). The impact of the internet on consumers' use of information sources for automobiles: A re-inquiry. *Journal of Consumer Research*, *34*(1), 111–119.
Roy, R., Chintagunta, P. K., & Haldar, S. (1996). A framework for investigating habits, “the hand of the past”, and heterogeneity in dynamic brand choice. *Marketing Science*, *15*(3), 280–299.
Rutz, O. J., & Bucklin, R. E. (2011). From generic to branded: A model of spillover in paid search advertising. *Journal of Marketing Research*, *48*(1), 87–102.
Rutz, O., & Trusov, M. (2011). Zooming in on paid search ads — an individual-level model calibrated on aggregated data. *Marketing Science*, *30*(5), 789–800.
Sayedi, A., Jerath, K., & Srinivasan, K. (2014). Competitive poaching in sponsored search advertising and its strategic impact on traditional advertising. *Marketing Science*, *33*(4), 586–608.
Swasy, J. L., & Rethans, A. J. (1986). Knowledge effects on curiosity and new product advertising. *Journal of Advertising*, *15*(4), 28–34.
Television Bureau of Advertising (TVB) (2013). *TV basics*. (Accessed October 2013).
Tellis, G. J. (2004). *Effective advertising: Understanding when, how and why advertising works*. Thousand Oaks: Sage Publications.
Viswanathan, S., Kuruzovich, J., Gosain, S., & Agarwal, R. (2007). Online informediaries and price discrimination: Evidence from the automotive retailing sector. *Journal of Marketing*, *71*(3), 89–107.
Weinberger, M. G., & Spotts, H. E. (1989). A situational view of information content in TV advertising in the U. S. and U. K. *Journal of Marketing*, *53*(1), 89–94.
Wilbur, K. C. (2008). A two-sided, empirical model of television advertising and viewing markets. *Marketing Science*, *27*(3), 356–378.
Xu, L., Wilbur, K. C., Siddarth, S., & Silva-Risso, J. (2014). Price advertising by manufacturers and dealers. *Management Science*, *60*(11), 2816–2834.