

Brand Attitudes and Search Engine Queries

Jeffrey P. Dotson
Brigham Young University
jeff.dotson@byu.edu

Ruixue Rachel Fan
Google
rachelfan@google.com

Elea McDonnell Feit
Drexel University
efeit@drexel.edu

Jeffrey D. Oldham
Google
oldham@google.com

Yi-Hsin Yeh
Google
yyeh@google.com

30 September 2016

Brand Attitudes and Search Engine Queries

Abstract

Search engines record the queries that users submit, including a large number of queries that include brand names. This data holds promise for assessing brand health. However, before adopting brand search volume as a brand metric, marketers should understand how brand search relates to traditional survey-based measures of brand attitudes, which have been shown to be predictive of sales. We investigate the relationship between brand attitudes and search engine queries using a unique micro-level data set collected from a panel of Google users who agreed to allow us to track their individual brand search behavior over eight weeks and link this search history to their responses to a brand attitude survey. Focusing on the smartphone and automotive markets, we find that users who are actively shopping in a category are more likely to search for any brand. Further, as users move from being aware of a brand to intending to purchase a brand, they are increasingly more likely to search for that brand, with the greatest gains as customers go from recognition to familiarity and from familiarity to consideration. Additionally, users that own and use a particular automotive or smartphone brand are much more likely to search for that brand, even when they are not in market suggesting that a substantial volume of brand search in these categories is not related to shopping or product search. We discuss the implications of these findings for assessing brand health from search data.

Keywords: search engines, brand search, brand metrics, recall, recognition, familiarity, consideration, purchase intent

1 Introduction

Survey-based measures of consumers' brand attitudes have been widely adopted by marketers to monitor brand health relative to competitors, assess the performance of advertising and other marketing tactics, and provide an early indicator of future sales (Keller, 1993; Aaker, 1996). Many large marketing research firms invest heavily in conducting brand tracking studies — periodic consumer surveys designed to gauge customers attitudes towards a brand and its competitors (e.g., Milward Brown BrandExpress, Y&R BrandAsset™ Valuator, YouGov BrandIndex). Brand attitude surveys typically ask respondents to answer a variety of attitude questions about brands including brand awareness (both recall and recognition), familiarity, purchase consideration and purchase intent. Marketers often summarize these attitudes across consumers to produce brand metrics which can be used to assess how the brand is performing in the minds of consumers relative to competitors. (For example, in the brand attitude survey we report, iPhone was recognized by 93.7% of consumers versus 75.0% for HTC; see Table 2.) Past research has shown that these consumer mindset metrics predict company stock returns (Aaker and Jacobson, 2001), as well as future sales (Srinivasan et al., 2010; Hanssens et al., 2014). Further, since these metrics assess intermediate stages in the purchase funnel, it has been suggested that they provide a more nuanced measure of consumer response to advertising (Percy and Rossiter, 1992; Srinivasan et al., 2010; Hanssens et al., 2014) than sales. Thus, brand tracking surveys that assess consumers' brand attitudes have become a standard and widely-used approach to tracking brand health.

Major brands invest a great deal of resources to field these brand tracking studies, an effort that has become increasingly more difficult as fewer consumers are willing to spend the time to answer surveys (Pew Research Center, 2012). At the same time, digital platforms have begun to track consumer behavior in real time, presenting a tantalizing opportunity to measure brand health based on consumer behaviors at a much lower cost. For example, major search engines receive a large volume of search queries that include brand names in the query text. Google Trends (<http://www.google.com/trends>), a free tool that reports an index of the volume of queries submitted to Google, reports that in 2014 ten times more searches that included the word “android” were submitted by U.S. users than searches that included the generic term “smartphone” (Google, 2015). Brand search volume could potentially be used as an ongoing, “dashboard” measure of brand health, the way survey-based metrics like the “percent of

consumers who recognize the brand” have been used in the past to monitor brand health.

While there is little theoretical or empirical literature on why a search engine user would submit a query for a brand, a literature has emerged that implicitly assumes that branded search queries represent an intermediate stage in the path-to-purchase. For example, Joo et al. (2014) and Lewis and Reiley (2013) use branded search query volume as an aggregate measure of advertising response and both find that search queries for the advertised brand increase substantially in the minutes after a television ad is aired. Brand search has also been incorporated in time series models relating advertising to sales and has been shown to improve sales prediction over models that do not include brand search metrics (Chandukala et al., 2014; Hu et al., 2014). Thus, brand search shows promise as a potential brand metric that, similar to survey-based brand attitude metrics, measures some intermediate stage in the purchase process.

Yet, while both brand attitudes and brand search have been proposed as measures of some intermediate stage in the purchase process, it is not clear whether they measure the same thing. Before marketers adopt brand search as a standard metric of brand health, they should understand how search relates to the traditional brand attitude metrics, which have been well-tested as measures of consumer interest in a brand (Lehmann et al., 2008; Srinivasan et al., 2010; Stahl et al., 2012; Hanssens et al., 2014). In this paper, we answer the question:

Are customers who hold positive attitudes toward a brand more likely to search for the brand?

If so, which brand attitudes are most closely associated with brand search?

Unlike past research, which has focused on correlating aggregate time-series on the volume of brand search with aggregate time-series of sales, we conceptualize and measure brand search as a behavior that *individual* consumers engage in. To investigate the relationship between attitudes and search at the consumer level, we assembled a panel of Google users who answered a typical brand attitude survey for two categories: smartphones and vehicles. In addition to answering the survey, these users also consented to allow us to observe their Google Search query counts for brands in those two categories. This unique data set allows us to relate the number of times an individual user searches for a particular brand during the study period to the attitudes s/he holds towards that brand.

Our analysis of this data shows that users who hold positive attitudes towards a brand are much more likely to search for that brand. We find that the likelihood of searching for a brand grows higher as

the brand attitudes go deeper from recognition to purchase intent, with the largest increases for customers who hold positive “mid-funnel” attitudes: familiarity and purchase consideration. We also confirm that users who are actively shopping for a product category are more likely to search for brands in that category. Thus, an increase in search for a brand may be due to an increase in the number of consumers who are shopping for the brand and hold positive attitudes towards the brand.

We also find that consumers who already own a particular brand are also much more likely to search for a brand than non-owners and that customers who are engaged with a category, i.e., have high enduring product involvement (Bloch and Richins, 1983) are more likely to search, regardless of whether they are actively shopping for a new smartphone or vehicle. This suggests that a substantial proportion of the total search volume is not related to shopping and that categories and brands with a higher user base and more “enthusiasts” will have more search volume, all else equal.

We should point out that our analysis is correlational; we show that users with positive brand attitudes are more likely to search for a brand. Following theory, our analysis assumes that brand attitudes lead to product search (cf. Keller, 1993) and specifically brand search queries in a search engine. Because we only measured brand attitudes at one point in time, we did not investigate the possibility that there may be dual causality, i.e. searching for a brand in a search engine changes brand attitudes. We also do not investigate the relationship between branded and generic search as customers move through the purchase process (Rutz and Bucklin, 2011).

The data we report suggests that there are many reasons a user might submit a brand search query. Users who are shopping in a category are more likely to search for any brand in the category; users are more likely to search for brands for which they hold positive attitudes; users who own a brand are more likely to search for the brand; and users who are category enthusiasts are more likely to search for all brands in the category. While we do not expect that these results will generalize to all categories, they shed light on how managers in high-involvement categories should interpret brand search volume from tools like Google Trends. Our data strongly suggests that overall search volumes reported by tools like Google Trends are a composite of different types of searches — shopping, product troubleshooting, keeping up with news and trends in the category — each associated with different consumer objectives and attitudes. In the conclusion, we will discuss some potential strategies for decomposing aggregate search volumes to

predict specific brand metrics such as familiarity or purchase consideration from search engine query data.

In the remainder of the introduction we briefly review the recent literature that has used brand search volume either to assess advertising or predict sales. Then in Section 2, we describe our data collection approach, including details of how we enrolled panelists in the study, linked panelists' survey responses to their search histories and defined a brand search. In Section 3, we describe our approach to modeling users' brand search counts as a function of their attitudes towards the brands and report the model estimates which show a strong relationship between brand attitudes and search. Finally, in Section 4, we discuss the implications of our findings and suggest directions for future research.

1.1 Related Literature

Shortly after Google Trends was launched in 2006, it was shown that search volume for particular words can be a good near-term predictor of disease outbreaks (Ginsberg et al., 2009). Subsequent research showed that search query volumes for particular words are predictive of economic indicators related to consumer behavior: Wu and Brynjolfsson (2014) find that an index of Google Search queries is predictive of housing sales and prices; Choi and Varian (2012) find that Google Search queries for particular words can predict near-term motor vehicles and parts sales, initial claims for unemployment benefits, travel and consumer confidence. A major focus of this literature is how to identify which search terms (from the many possibilities) are most closely associated with a given economic indicator (Scott and Varian, 2014).

More recently, marketers have focused on *brand* search volume, i.e., the volume of search queries that include specific brand names, and have used this metric in a variety of ways. As mentioned above, several authors have used brand search volume to gauge consumer response to advertising. Joo et al. (2014) use Google Search volume for financial services brands as a measure of response to television advertising, finding that search for the brand and the category increases just after a television ad for the brand is aired. Lewis and Reiley (2013) do a similar analysis using Yahoo! search volume for brands that were advertised during the 2011 Super Bowl and find that search volume for a brand increases dramatically the minute the ad is aired, particularly for movies, automotive brands, internet service brands, but less so for food brands (e.g., "Pepsi" and "Snickers"). So, there is some evidence that advertising causes an increase in search engine queries for a brand, at least for products where consumers regularly gather information or purchase

via the internet.

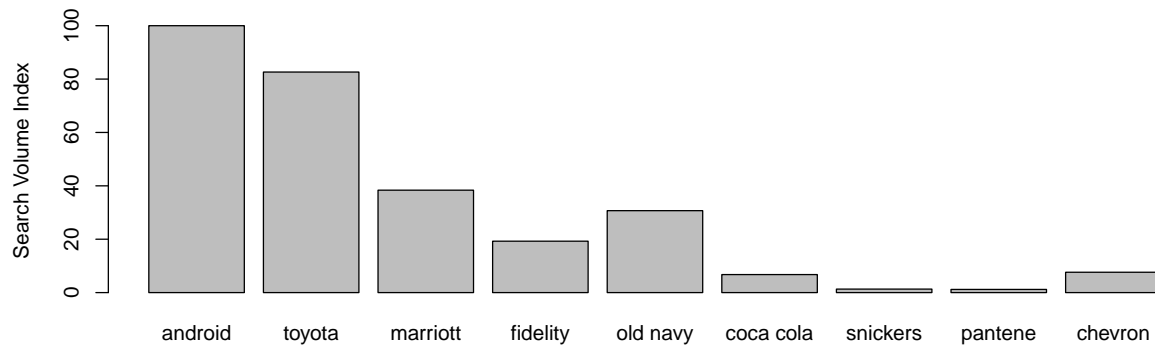
Other researchers have included aggregate measures of brand search volume in marketing mix models intended to predict sales as a function of advertising. For example, Hu et al. (2014) relate Google Trends data for automotive brands to both advertising spending and aggregate sales over time in a vector autoregressive framework. Similarly, Chandukala et al. (2014) relate the count of searches for automotive brands on an automotive shopping site to both advertising spend and sales data. Both of these papers find that brand search volume improves prediction of sales and, since search is a more immediate response metric than sales, it is useful for detecting advertising response. We note that our focus here is on the queries that users submit to a search engine; there are also several papers that investigate the role of paid search advertising in the marketing mix (cf. Rutz and Bucklin, 2011; Dinner et al., 2014; Blake et al., 2015; Pauwels et al., 2016).

Our work stands apart from the prior literature on brand search in two important ways. First, we are focused on understanding whether traditional survey-based measures of brand health relate to brand search, helping to put brand search in the context of other commonly-used brand metrics. The only work we are aware of that brings together data on mind-set metrics with data on brand search is Pauwels and van Ewijk (2013) who use aggregate data and find that both are useful in predicting sales, but do not specifically focus on or report the time-series relationship between aggregate search volume and average brand attitudes. Second, unlike all prior research on brand search, our analysis focuses on understanding how an *individual consumer's* brand attitudes relate to his or her propensity to search for the brand in a search engine. Thus, unlike prior work which can only answer questions of the form, “When the volume of search goes up, what else goes up?,” we are able to directly answer the question, “Are users who hold positive brand attitudes more likely to search for a brand?” This user-level focus allows us to provide a deeper understanding of how user-level attitudes lead to individual-level behavior that is then aggregated up to the brand search volume metric that has been used in prior research.

2 Data Collection

For users who have a Google Account, Google records the full text of each search query the user makes while the user is signed in to their Google Account. This means that for regular signed-in users of Google

Figure 1: 2014 U.S. search volume index for example brands (from Google Trends).



tools, the majority of their search queries can be tracked to the user account.¹ To link Google users' search history to their brand attitudes, we recruited a panel of 1,511 users who agreed to answer a brand attitude survey and to have their Google search history for specific product categories monitored over an 8-week period. Because the panelists agreed to have their search and attitude data linked, the resulting data allows us to explore the relationship between holding a particular brand attitude and searching for that brand in a search engine.

We collected attitudes and search data for the smartphone and vehicle categories. These categories were selected primarily because there is substantial search volume for brands in both of those categories. For example, Figure 1 shows the Google Trends 2014 U.S. search volume index for several well-known, popular brands in a variety of categories. Among these brands, Google Trends reports the greatest search volume for “android” and “toyota”, while search volume for popular fast-moving consumer goods brands like “coca-cola” and “snickers” is much lower. While our data suggest that brand attitudes are associated with brand search for these two categories, we can not say whether these results will generalize to other categories; we would speculate that they would not generalize to categories where product search is low and therefore brand search queries in search engines are also likely to be low. We discuss generalizability to other categories further in the discussion.

Prior to fielding the main survey, we conducted a small pilot study on a panel of 57 U.S.-based

¹In accordance with Google privacy policies, users may review and delete their history at any time at <http://history.google.com>.

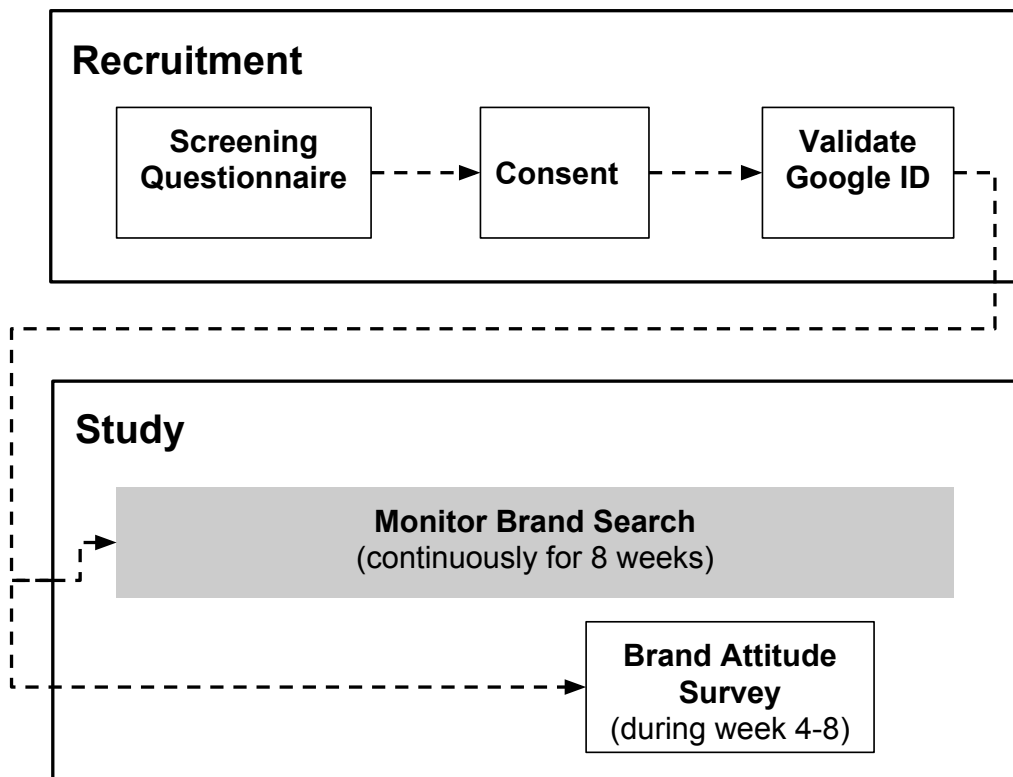
Google users who we recruited from a university-based survey platform. We monitored brand search in the smartphone category for these users over 9 months and during that time the panelists completed three identical quarterly surveys where they were asked about their smartphone brand attitudes. The pilot study findings are consistent with those from the main study and are reported in the online supplement. The pilot study also informed the design of the main study.

The main study was conducted in Fall 2014 in collaboration with the market research firm GfK. The flow of the respondents through the study is depicted in Figure 2. As is typical for commercial brand attitude surveys, GfK recruited panelists from a number of opt-in online survey panels. Respondents were recruited on a rolling basis over a 2-month period. Focusing the study on users who regularly use Google Search, potential panelists were first screened to ensure that they lived in the U.S., had a Google Account and periodically searched while signed into their Google Account. After the initial screening, respondents were sent a more detailed consent form, which explained what data would be collected and analyzed, what was required of panelists to complete the study, the incentives for completing the study, and whom to contact with questions or to withdraw from the study. Upon completing the consent, users validated their Google Account, which provided the mechanism to link their Google search history to their survey responses collected by GfK. Users then entered the main phase of the study where we monitored their brand searches for the two categories over 8 weeks. To complete the study and be counted in the final sample, respondents were required to: 1) submit at least one Google search while signed into their Google Account every two weeks during the observation period and 2) complete the 10-minute online brand attitude survey, which was emailed to respondents in good standing as of day 28 of the study.

2.1 Recruitment

The recruitment for the study was typical of other online panel studies. Of the 15,977 individuals invited to complete the screener, 3,910 (24.5%) passed the screening and consented to participate in the study. Of the 3,910 individuals who agreed to participate in the study at the screening stage, 3,257 (83.3%) verified their Google Account and entered the main phase of the study. Of those 3,257 panelists, 2,302 (70.7%) remained in compliance with the requirement to complete one search while signed in to their Google Account at least once every two weeks. For users who regularly use Google services like Gmail or Google

Figure 2: Panelist flow.



Drive, this requirement is easily met without any special effort on the part of the user. Panelists who had not submitted search queries for a period of 14 days received an email or phone reminder to stay logged into Google. Of the 2,302 search-compliant panelists, 1,511 (65.6%) completed the survey before the end of their individualized 8-week study period.

2.2 Brand Attitude Survey

There is substantial variation across the literature in which specific mind-set metrics should be included in a brand tracking study. Keller (1993) and Aaker (1996) both suggest marketers should measure awareness (recall and recognition) as key measures of brand health. Srinivasan et al. (2010) and Hanssens et al. (2014) focus on advertising awareness, consideration and “liking” in their time-series models linking brand metrics to sales and advertising. Pauwels and van Ewijk (2013) incorporate brand awareness, consideration, and preference into a similar model. A commonly cited problem among those that study aggregate brand metrics is that there is insufficient variation over time in the level of brand awareness (i.e. % of consumers who are aware of a brand) to make it a useful metric in time-series models, a problem which we avoid by focusing on user-level data. Even for the most-popular brands, there are still substantial numbers of panelists who indicate that they do not recognize a brand.

In our brand attitude survey, we included five of the most-common brand attitude questions, which span the range from the upper-funnel metrics of awareness (which can be further broken out into unaided recall and aided recognition) to the mid-funnel metrics of familiarity and consideration to the lower-funnel metric, purchase intent. While some proprietary brand tracking studies include other brand attitude questions in addition to those listed in Table 1, we selected these five questions to represent the typical range of attitude questions included in brand tracking studies (cf. Survey Monkey, 2015). The exact wording of the questions is summarized in Table 1.

For the recall question, which was the first question in the survey, users were asked to type in the names of three brands, and we coded this as 1 if the user typed the name of the brand (or a common misspelling) and 0 otherwise. For the recognition, familiarity, consideration and purchase intent questions, respondents were presented with a list of the target brands (see Appendix A) and asked to check as many brands as they like, resulting in a binary indicator for each metric and each brand. Respondents were quite

Metric	Survey Question
Recognition (Aided Awareness)	Thinking about brands related to [category], which of the following are you aware of?
Recall (Unaided Awareness)	When you think of [category], which three brands come to mind first?
Familiarity	Which of the following brands would you say you are familiar with?
Consideration	Would you consider purchasing this brand when you buy your next [category]?
Purchase Intent	If you were making the purchase today, which of the following [category] brands would you be most likely to purchase?

Table 1: Survey-based brand metrics. Note the Recall question was asked before Recognition, but falls lower in the purchase funnel conceptually.

fast at answering these questions, as they only required the respondent to scan the list of brands and check those that apply. In the smartphone category, we measured attitudes for 9 popular brands, resulting in a total of $9 \times 5 = 45$ binary indicators for each respondent. In the automotive category, we measured attitude for 28 popular brands. For the smartphone category, we included both operating system and device brands, recognizing that many users do not distinguish device brands like Samsung from operating system brands like Android. The survey was organized by category, with the order of categories rotated across respondents.

While our goal is to relate user-level brand attitudes to users' brand search, the brand attitude data can also be summarized across users to compute typical brand metrics such as the percentage of users who can recall the brand. Table 2 summarizes the brand attitudes for several example brands in each category. We find ample variation in brand attitudes both across brands, e.g., iPhone recognition is 93.7% versus 75.0% for HTC, and across users, e.g., even for the least-popular brands a reasonable fraction of users intend to purchase the brand on the next purchase occasion. The model specification we use to analyze the relationship between attitudes and search exploits variation between users (rather than variation between brands or across time) to understand the association between brand attitudes and search.

In addition to collecting data on brand attitudes, we also asked panelists several questions about their relationship to the category as a whole including whether they had made a purchase in the past month,

	Example Smartphone Brands			Example Vehicle Brands			
	iPhone	Samsung	HTC	Toyota	Ford	Kia	BMW
Recognition	93.7	93.2	75.0	92.1	93.8	86.6	89.7
Recall	79.9	79.5	23.0	43.0	62.3	5.8	8.5
Familiarity	60.8	63.6	30.5	57.5	63.4	26.9	29.7
Purchase consideration	46.8	57.4	27.5	43.2	42.2	16.9	18.4
Purchase intent	33.4	22.8	3.8	13.7	12.6	3.2	2.6
Ownership	39.7	27.8	4.4	12.9	12.3	2.6	2.4
Problems (among owners)	15.6	20.1	19.0	8.9	19.3	14.7	18.8

Table 2: Brand attitude metrics for example brands. Numbers indicate the percentage of respondents who hold a particular attitude toward a brand.

	Smartphones	Vehicles
Engaged	57.8	41.3
Recent purchase	24.4	6.4
Actively shopping	31.0	27.3

Table 3: Percentage of respondents who indicate that they are engaged in the category, made a purchase in the past month or were actively shopping, used as controls for overall incidence and volume of brand search expected for each user.

whether they intended to make a purchase in the next month, and whether they “paid attention to the category and watched for announcements or news about the latest product releases,” which measures something similar to Bloch and Richin’s notion of enduring product importance (Bloch and Richins, 1983). We include these binary variables in our model as controls for the overall volume of brand search we expect for each user within each category, hypothesizing that users who are actively shopping or are more engaged in the category are more likely to search for any brand in the category. Table 3 summarizes this fraction of panelists who indicated in the survey that they were engaged, recently made a purchase or were actively shopping. A large proportion of respondents (57.8% and 41.3%) in both categories claim to be generally engaged with the category, while a much smaller fraction say they have made a purchase or plan to make a purchase in each category.

2.3 Search data

The key element of our research design was to link each panelist’s responses to the brand tracking study to the panelist’s brand search at the user level. Our primary metric describing each user’s brand search is the

count of brand search queries for the smartphone and vehicle brands listed in Appendix A. We define a brand query as one that contains one of a set of brand-related keywords. To identify the keywords used for each brand, we began with the brand name, e.g., “iphone” and then used a query clustering method to yield spelling variants. For example, for “iphone,” we identified many variations including “iphone 5”, “iphone 5 s”, “i-phone”, and “iphone”, similar to what one would find if using the Google Trends “Related searches” feature or the Google Adwords Keyword Tool. We supplemented this list of variants with others obtained from the panelist-provided answers to the unaided brand recall questions in the survey. We then obtained counts of the number of queries that included a brand keyword for each user over the 8-week observation period. We also obtained the total number of queries submitted by each user in the observation period.

We find a substantial volume of search for the smartphone brands we study; a majority of panelists made a search related to one of the smartphone brands during the 8-week observation period. We find slightly less search for the automotive brands; a large minority of users made searches related to at least one of the 28 of automotive brands.

As one would expect, the distribution of the brand search count across users is quite skewed and contains a large number of zeros (i.e., instances where a customer does not search for a brand in the two-month observation period). The distribution is consistent with an over-dispersed count distribution such as Zipf or negative binomial. In modeling, we take care to accommodate both excess zeros and over-dispersion in the model specification, as we discuss in Section 3.

2.4 Protecting Panelists’ Privacy

While Google does record the full text of search queries that users submit while logged into their Google Account, Google recognizes that the full text of search queries can contain highly personal information and this data is tightly controlled within Google. While Google does publicly release indexes of the volume of search queries that include specific keywords through Google Trends, the data is aggregated to avoid breaching any user’s privacy. Furthermore, keywords with low search volume are not reported.

Since this study connects individual users’ search data to their survey responses, we obtained specific permission from the panelists to have their search counts for specific categories monitored. While we asked panelists to remain logged in as much as possible, they could log out or use a browser’s privacy

mode, such as Chrome’s incognito mode, to exclude specific search queries from their search history. Like all Google users, they could review their recorded search history using <http://history.google.com>, removing any queries they desire. Since the users knew which categories (but not which brands) were being monitored, there may be some demand effects that increase the overall volume of search, but this should not affect comparisons between brands or the relationship between attitudes and search.

In accordance with the permissions given by the panelists, we used only the brand search counts and total search counts in our analysis, and did not have access to the full text of search queries for each user or information on which links the user clicked on in the search results. Consistent with Google’s privacy policies, the brand search counts and total search counts did not leave Google’s secure computing environment.

3 Relationship between Brand Attitudes and Search

3.1 Analysis Approach

The goal of the analysis is to measure the relationship between users’ brand attitudes and their brand search counts. We do this by regressing a user’s vector of attitudes towards a particular brand on the user’s search count for that brand. To accommodate the empirical distribution of the search counts, we use a hurdle model (cf. Cameron and Trivedi, 1998, Chapter 4) which allows us to specify a binary process for whether a user searches for a given brand at all and a separate count process for the number of times a user searches for the brand in the 8-week observation period. The binary process allows us to accommodate distributions of brand search counts that are zero-inflated (i.e., fewer users search for the brand at all than would be expected from a standard count model like the Poisson) and the count model we used allows for counts that are over-dispersed (i.e., there are users who search for a brand much more than would be expected if the search counts were Poisson). It also allows us to estimate the effect of holding a particular brand attitude on the *incidence* of search separately from the effect of those attitudes on the *volume* of search.

We specify the user’s likelihood of submitting a search for a particular brand to follow a binary

logistic regression.² If y_{ij} is the number of search queries submitted by user j for brand i , then

$$p(y_{ij} > 0) = \frac{\exp(\alpha_i + x_{ij}\beta + z_{ij}\gamma)}{1 + \exp(\alpha_i + x_{ij}\beta + z_{ij}\gamma)} \quad (1)$$

where α_i is an intercept for brand i and β is a vector of parameters that multiplies x_{ij} , a vector of user j 's attitudes toward brand i . The vector z_{ij} includes several additional control variables, which we discuss below. The fixed effects for each brand, α_i , allow for differences in the overall level of search for each brand. The brand fixed effects also ensure that the parameters we estimate for the effects of brand attitudes are informed by differences between users within a brand (e.g., users who are familiar with iPhone are more likely to search for iPhone than users who are not familiar with iPhone) rather than differences between brands.

For those users who exceed the ‘‘hurdle’’ defined by Equation 1, we assume that the user’s count of searches follows a negative binomial distribution truncated below $y_{ij} = 1$. That is, the hurdle model ignores the prediction for the number of zeros from the negative binomial and normalizes the distribution to account for this left truncation. The probability mass function for the truncated negative binomial distribution is given by:

$$p(y_{ij}|y_{ij} > 0) = \frac{\Gamma(y_{ij} + \theta)}{\Gamma(\theta)y_{ij}!} \frac{\mu_{ij}^{y_{ij}} \theta^\theta}{(\mu_{ij} + \theta)^{y_{ij}} [(\mu_{ij} + \theta)^\theta - \theta^\theta]} \quad (2)$$

$$\log(\mu_{ij}) = \tilde{\alpha}_i + x_{ij}\tilde{\beta} + z_{ij}\tilde{\gamma}$$

where θ is an over-dispersion parameter and x_{ij} is the vector of brand attitudes and z_{ij} is the vector of controls.³ In our analysis, we use the same set of covariates to predict whether a user will search (Equation 1) and how much a user will search (Equation 2), although it is possible to allow the covariate vectors to differ between the zero model and the count model. The estimated parameters, however, are allowed to be different between the two models.

The key feature distinguishing the hurdle model from the zero-inflated negative binomial is that the

²We also explored the alternative Cauchit specification, which is more robust to outliers in the search data, but found that the model fit statistics favored the logit and the substantive results were similar. We report the logistic regression here, as we can report odds ratios which are more readily interpretable.

³We explored alternative models for the count conditional on incidence including the truncated (> 0) Poisson and found that the model fits favored the over-dispersed negative binomial.

hurdle model only allows for a single process to create zeros, rather than defining the zeros as arising from a mixture between the zero process and a count process that allows for zeros. Thus, the α_i , β and γ coefficients from Equation 1 can be interpreted independently from the $\tilde{\alpha}_i$, $\tilde{\beta}$ and $\tilde{\gamma}$ parameters in Equation 2. This approach allows us to focus on incidence as a separate process. In fact, the parameter estimates for Equation 1 are the same as what one would obtain using a logistic regression. We estimated the hurdle model by maximum likelihood using the ‘pscl’ package in R (Zeileis et al., 2008).

In addition to the brand attitudes listed in Table 1, we include several control variables in z_{ij} , which affect search for all brands. If Google search is part of the shopping process, then we expect that customers who are actively shopping in a category are more likely to search for any brand in that category. Thus we include indicators for whether the user said s/he had made a purchase or planned to make a purchase in the category as covariates to search for all brands. In the pilot study, we had observed that users who had indicated that they were interested in the category were more likely to search any brand in the category, so, we included a similar indicator for whether the user indicated that s/he was actively engaged in the category as a covariate to search for all brands in the category.

In the survey, we also asked respondents which brand they owned and whether they were having problems with that brand, hypothesizing that owners might submit a search with the brand name seeking information about how to use the product, particularly when they are having problems. We include indicators for which brand the customer owns and whether they are having problems in z_{ij} .

Finally, to control for the fact that some users search a lot more than others (and so are more likely to search for anything), we included the logarithm of the total number of all searches submitted by the user during the entire observation period as a covariate in the hurdle model, allowing for the possibility that users who submit more search queries to Google overall are more likely to search for any brand. We also included the log of the total number of searches as an offset in the count model, effectively modeling the count of searches for a brand as a fraction of the total count of searches for the user.

To summarize, the set of predictors we use in x_{ij} to predict how many times user j will search for brand i includes the 5 brand attitude measures (recognition, recall, familiarity, purchase consideration, and purchase intent), plus indicators for ownership, problems, being in-market and interest in the category. Of the 1511 panelists who completed the study, 1498 completed all the relevant questions for the smartphone

category and 1500 completed all the relevant questions for the vehicle category. We estimated the model using these complete cases.

Given that brand attitudes are often interrelated, we checked for correlations between the brand attitude measures and find that no two metrics have a correlation higher than 0.5 in either category. We report the full covariance matrices for each category in Appendix B. These correlations reflect the variation across consumers in brand attitudes, and so are smaller than one would expect for correlations between aggregated brand-level metrics. For example, while brands that have high familiarity also tend to have high purchase intent, there is more variation among users, with users who are both familiar and intend to purchase the brand, or are familiar and don't intend to purchase the brand or neither.

3.2 Findings

Table 4 shows the estimated parameters for the model defined by Equations 1 and 2. Parameters that appear in bold are significantly different than zero at the 95% confidence level. The upper panel shows the estimated parameters for the logit hurdle equation (Equation 1) and the lower panel shows the estimated parameters for the negative binomial count model (Equation 2). We should note that these parameter estimates are pooled across all brands, that is, we assume that the effect of being familiar with iPhone on iPhone searches is the same as the effect of being familiar with Android on Android searches.⁴

The hurdle portion of the model captures brand search incidence, and so the data are more informative about this portion of the model. Thus we find more significant associations in the upper panel in Table 4. The parameter estimates for the hurdle equation are remarkably similar across the two categories. In the smartphone category, the estimates indicate that all five of the brand attitudes are positively associated with brand search incidence for the category. Thus, the data confirms that customers who hold positive attitudes towards a brand are more likely to search for that brand. For example, in the smartphone category the odds of searching for a brand is 7.0 times higher for a user who holds all five positive brand attitudes, versus a user who holds no positive brand attitudes (from the logit model, we can

⁴We did explore model specifications for incidence which allowed for random effects across brands. When we allowed for random effects of brand attitudes on search across brands, the estimation routine did not reliably converge suggesting that there was insufficient data to estimate them. For a more limited model which only included brand attitudes as predictors with random effects across brands, and excluded the control variables, the population average association between attitudes and search is similar to those reported here.

		Smartphones		Vehicles	
		$N = 1498 \times 8$		$N = 1500 \times 28$	
		est	se	est	se
Hurdle Equation	Recognition	0.20	0.09	0.38	0.09
	Recall	0.25	0.12	0.03	0.13
	Familiarity	0.67	0.08	0.37	0.08
	Consideration	0.44	0.08	0.52	0.08
	Purchase Intent	0.39	0.10	0.34	0.11
	Owns Brand	0.91	0.10	1.26	0.11
	Problems with Brand	0.17	0.14	0.23	0.20
	In-Market*	0.31	0.07	0.32	0.07
	Interested in Category*	0.31	0.07	0.22	0.07
	log(Total Search)*	0.77	0.03	0.78	0.03
Count Equation	Recognition	0.02	0.20	0.22	0.18
	Recall	-0.25	0.24	0.08	0.32
	Familiarity	0.41	0.15	-0.11	0.19
	Consideration	0.57	0.14	-0.08	0.18
	Purchase Intent	0.17	0.18	0.40	0.22
	Owns Brand	0.53	0.17	0.40	0.20
	Problems with Brand	-0.18	0.22	0.57	0.37
	In-Market*	0.56	0.12	-0.06	0.16
	Interested in Category*	0.07	0.13	0.17	0.14
	Dispersion ($\log(\theta)$)	-10.20	35.98	-20.60	2092.17

Table 4: Hurdle model estimates (“est”) and standard errors (“se”) relating brand attitudes to search at the user-brand level. Values in boldface are significant at 95% confidence. * Indicates a category-level control variable that affects search for all brands.

compute the odds ratio = $\exp(0.20 + 0.25 + 0.67 + 0.44 + 0.39) = 7.0$). Similarly, for the automotive category, the odds of searching for a brand for a user who holds all five positive brand attitudes is 5.2 times higher than for a user who doesn't hold any positive attitudes (odds ratio = $\exp(0.38 + 0.03 + 0.37 + 0.52 + 0.34) = 5.3$). Thus we find strong evidence that users who hold positive attitudes towards a brand are more likely to search for that brand.

For both categories, we find the biggest increases in propensity to search for customers who are familiar and would consider purchasing the brand (Familiarity and Consideration coefficients in the upper panel of Table 4). Customers who only recognize a smartphone brand and hold no other positive attitudes, are only 1.22 times more likely to search for a brand than those who don't recognize the brand (odds ratio = $\exp(0.2) = 1.22$). Customers who hold stronger mid- and lower-funnel brand attitudes are much more likely to search for a brand than those who are merely aware of a brand.

We also find evidence that that customers who are actively shopping are more likely to search for any brand in the category. Customers who indicated that they made a purchase or intended to make a purchase ("In-Market") during the observation period were significantly more likely to search for any brand (1.4 times more likely for both categories). Similarly, customers who indicated that they "always pay attention to the category so that they know when to buy" are more likely to search for all brands in the category (1.4 times more likely for smartphones and 1.2 for vehicles.) This suggests that a substantial portion of the brand search queries that are submitted to Google are associated with users who are shopping for the product.

We also find that owning a particular smartphone or vehicle brand is a very strong predictor of brand search, with the odds of searching being 2.5 times greater for brand owners versus non-owners in the smartphone category and 3.5 times greater in the vehicle category. This large increase in brand search among owners (regardless of whether that user is actively shopping), could be partially due to owners searching for information about how to use the product. Marketers who are interpreting total brand search volume (e.g. Google Trends data), should expect that brand search will be higher for brands with more owners, i.e., a larger installed base, irrespective of consumers' attitudes toward the brand. Somewhat surprisingly, we did not find that owners who are having problems with the brand are significantly more likely to search for the brand than other owners, suggesting that information search is not simply associated

with problems or product recalls, but that there is a steady volume of search produced by brand owners, at least for complex durables, where consumers are likely to have questions about usage and maintenance of the product they use every day.

The lower panel of Table 4 reports the estimates for the negative binomial model in Equation 2, which predicts the volume of searches that a user will make for a particular brand, conditional on the user searching at least once for the brand. The model estimates indicate whether a particular attitude is associated with the user searching more frequently (as a fraction of total searches), conditional on searching at least once. As can be seen from the larger standard errors, this portion of the model is less well-identified, due to the relatively low incidence of users searching for brands, particularly in the automotive category. We do, however, find some significant effects.

Among those who search for a smartphone brand, users who are familiar with a brand or would consider purchasing the brand tend to submit more search queries for that brand. That is, users who are familiar with a brand and would consider purchasing that brand are not just more likely to search at least once, but they are also likely to submit more queries over the 8-week period (as a fraction of their total search volume.) In both categories, users who own a brand also tend to submit more queries for that brand. And, in the smartphone category, users who are in-market tend to submit more search queries. Overall, the count equation estimates suggest that at least for smartphones positive mid-funnel brand attitudes—familiarity and consideration—are not only associated with a greater incidence of search, but also shift the distribution of search counts to the right, resulting in more search among those who do search.

4 Discussion

The data we have presented shows that users who hold positive attitudes towards a brand are more likely to search for that brand, with the greatest increases in search propensity for those who hold positive “mid-funnel” attitudes like *familiarity* and *consideration*. We are the first to show direct evidence of a positive association between individual users’ brand attitudes and their brand search and this represents an important step forward in understanding why users search for brands and how search behavior is related to traditional survey-based brand tracking measures. In the remainder of the paper, we discuss implications for marketers and our suggestions for future research on brand search as metric for brand health.

4.1 Limitations

We note several boundaries of our findings. First, in our analysis, we estimated common effects across brands and consumers. While this analysis is sufficient to show the relationship between brand attitudes and brand search, a more complex model with random effects, would allow us to explore heterogeneity among brands and users. Albeit, such a model may require more data to estimate than we have here.

As our focus here was on the association between brand attitudes and brand search, we did not include in the model other covariates that are likely to be predictive of search including advertising (Joo et al., 2014; Lewis and Reiley, 2013; Hu et al., 2014) and product recalls. However, because the model estimates are informed by brand attitude differences between users (and not differences between brands or differences over time), the omission is unlikely to produce an omitted variables bias in our estimates of the association between attitudes and search.

We should also note that at the extremes of brand attitude, the positive relationship between search and attitudes may not hold. For instance, if a consumer is extremely loyal to a particular brand, we may find very little search associated with shopping, as customers who are extremely loyal don't need to do any research through a search engine prior to purchase. However, we found a positive association between purchase intent and brand search, suggesting that our purchase intent question was not a strong enough measure of loyalty to identify those customers whose loyalty is so strong that they wouldn't search at all when shopping.⁵

Finally, our data describes users' search behavior today. In the future, search behavior is likely to evolve as technology and search engines evolve. Search engine providers are constantly innovating to make search results more useful and this could lead to major shifts in brand-search volume that have nothing to do with how consumers perceive those brands. For instance, as mobile search results become more tailored to a user's location and more useful, users may begin to use search more frequently as part of the shopping process, even for goods that are not researched or purchased online today. Similarly, as retailers evolve, shoppers may forgo search engines in favor of brand search at retailer websites. While future changes in

⁵We attempted to identify users with even stronger positive attitudes toward the brand by selecting those who only choose one brand that they would consider for their next purchase. The incidence of brand search was similar for these users, suggesting that this is not a strong enough measure of loyalty to identify users who have such strong positive attitudes that they don't engage in product search. It is also possible that highly-loyal customers submit fewer shopping-related brand queries but submit more non-shopping related queries.

technology may alter how consumers use search engines, we believe this work represents a useful step in helping marketers understand the relationship between traditional survey-based measures of brand health and behavioral measures of the shopping process like search.

4.2 Generalizability

We have shown that there are at least two important product categories where users who hold positive brand attitudes are more likely to submit a web search for those brands. However, we know that the use of search engines in the shopping process varies substantially across categories (cf. Lewis and Reiley, 2013). For example, users seldom search for “Coke” or “Coca-Cola” despite the fact that Coca-Cola generally has high survey-based brand metrics (see Figure 1), so we expect that these findings will likely not hold for all categories. While an extensive, multi-category study is beyond our scope, we provide some speculation on the generalizability of these findings to other categories.

We expect that our findings would extend to categories where users engage in substantial product search during the shopping process resulting in a substantial volume of brand search. In categories where consumers engage in online research prior to purchase such as appliances, furniture, travel, entertainment, financial services, online retailers, and online services, we would expect a similar association between brand attitudes and search as we found for smartphones and automobiles. For these categories, an improvement in brand attitudes may be associated with more search queries for the brand, although there are other potential causes for an increase in search as we discuss below.

Our study also finds that customers who are highly engaged in a category are more likely to search for brands in that category, even when they are not in market. The two categories that we studied, smartphones and automobiles, are categories where a high percentage of the population has enduring product importance and would be expected to continue to engage in product search even when they are not actively shopping (Bloch and Richins, 1983). For these types of categories, brand managers should keep in mind that a substantial volume of search is coming from these “enthusiasts,” whose interest in brands many not reflect the larger community of potential shoppers. So, if an event occurs that encourages enthusiasts to seek out brand information, such as a major new product release, we would expect an increase in brand search for these categories.

However, there are other categories that have high situational importance and low enduring importance, such as refrigerators or furniture or even some subcategories of automobiles like minivans, where consumers engage in a lot of product search when they are shopping, but little or none when they are not. Shoppers in these categories may also engage in generic category search first and then evolve to branded search as they learn more about the category and get closer to purchase (as described by Rutz and Bucklin, 2011). For these categories, where more of the total brand search volume is associated with shopping, we would expect a stronger association between brand attitudes and search.

Similarly, we focused on two categories where *owners* are more likely to search for the brand; we speculate that these owners are seeking out information on how to use the product. For less complex, easy-to-use products, again like a refrigerator or furniture, we would expect less of the total search volume would come from owners and, consequently, there would also be a stronger association between brand search volume and brand attitudes.

While there are some categories where we believe the association between brand attitudes and brand search will be stronger, there are many other categories where customers are far less likely to use a search engine as part of the shopping process, such as fast food or package goods. For these categories, brand search queries will be very sparse and we expect that brand search will be less closely associated with brand attitudes. We encourage future research exploring a broad array of categories to confirm our hypotheses.

4.3 Implications for Marketers

Despite these limitations, our data shows a relationship between brand attitudes and brand search volume for these two categories. This, along with the prior evidence that search is positively associated with both advertising (Joo et al., 2014) and sales (Hu et al., 2014), strongly suggests that brand search volume is usually a positive indicator of brand health and has clearly earned a place among the metrics that belong on the modern marketer's dashboard.

However, our data also suggests a note of caution, as we found that search users who own a brand are also very likely to search for that brand, even when they are not in-market, suggesting that there is a substantial volume of search that is not directly related to shopping. In this study we only observe the total

count of brand search queries for each user, but if it were possible to decompose that total into brand search counts for shopping-related queries versus other types of queries, then we believe that users would submit those shopping queries primarily when they were in market for the product and held positive attitudes towards the brand. This implies that marketers who want to precisely assess advertising response and forecast sales should focus on measuring and reporting shopping-related brand search as a brand metric rather than the total brand search count.

4.4 Future Research

A critical next step in developing brand search as a metric is to distinguish shopping-related brand search from the other sources of brand search such as troubleshooting or searching for homonyms to the brand (e.g., galaxy, Lincoln, apple). Hu et al. (2014) recommend some ad hoc ways to narrow Google Trends data to queries that are shopping-related, including discounting queries with terms that are clearly related to product use or troubleshooting (e.g., “recall”) and using the Google Trends “Categories” filters. Yet, to make brand search a relevant metric across many categories, we need a systematic approach to inferring which queries are shopping-related. Such an approach might infer which queries are shopping-related based on other words included in the query, which links the users clicked on in the search results and what searches they made before and after the target query. (Although this will have to be done with a lot of care toward the privacy of Google users; privacy considerations prevented us from exploring these options here.) We hypothesize that shopping-related brand query volume will be even more closely associated with brand attitudes than the overall brand query volume, making it possible to predict what attitudes users hold based on their shopping-related search history.

Those who seek to incorporate total brand search volume into time-series models today are faced with the problem that their estimates of brand search are potentially “contaminated” by non-shopping related queries, adding noise to their models. Our analysis shows that owners are much more likely to search for a brand suggesting that the number of consumers who own the brand, i.e., the installed base, would be an important (and readily available) control which could account for some of the variation between brands in search volume that is not shopping-related. Similarly, it might be helpful to incorporate indicators for important events that might lead owners to search for the brand that they own, such as

product recalls or major updates to software. An index of the number of customers who are in-market (which could be obtained through a survey), may also be useful in interpreting brand search volume and predicting brand attitudes from brand search volume.

Finally, while we have identified three categories of brand search (shopping, “enthusiast” product search and search related to owning the brand, such as troubleshooting), we note that there could be other reasons we have not considered for why a user would search for a brand. In the spirit of fully understanding brand search, we encourage further exploratory research into why people search for brands, perhaps through an open-ended survey targeted to search users who submit a query for a brand.

Acknowledgements

The authors contributed equally and are listed in alphabetical order. We thank GfK for their assistance with collecting the data described in this paper. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Aaker, D. A. (1996). Measuring brand equity across products and markets. *California Management Review*, 38(3):102–120.

Aaker, D. A. and Jacobson, R. (2001). The value relevance of brand attitude in high-technology markets. *Journal of Marketing Research*, 38(4):485–493.

Blake, T., Nosko, C., and Tadelis, S. (2015). Consumer heterogeneity and paid search effectiveness: A large-scale field experiment. *Econometrica*, 83(1):155–174.

Bloch, P. H. and Richins, M. L. (1983). A theoretical model for the study of product importance perceptions. *The Journal of Marketing*, 47(2):69–81.

Cameron, A. C. and Trivedi, P. K. (1998). *Regression Analysis of Count Data*. Cambridge University Press, New York.

- Chandukala, S. R., Dotson, J. P., Liu, Q., and Conrady, S. (2014). Exploring the relationship between online search and offline sales for better “nowcasting”. *Customer Needs and Solutions*, 1(3):1–12.
- Choi, H. and Varian, H. (2012). Predicting the present with Google Trends. *Economic Record*, 88:2–9.
- Day, G. S. and Pratt Jr., R. W. (1971). Stability of appliance brand awareness. *Journal of Marketing Research*, 8:85–89.
- Dinner, I. M., Van Heerde, H. J., and Neslin, S. A. (2014). Driving online and offline sales: The cross-channel effects of traditional, online display, and paid search advertising. *Journal of Marketing Research*, 51(5):527–545.
- Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., and Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, 457:1012–1015.
- Google (2015). Google Trends for “android” and “smartphone” for U.S. in 2014. <http://trends.google.com>. Accessed August 20, 2015.
- Hanssens, D. M., Pauwels, K. H., Srinivasan, S., Vanhuele, M., and Yildirim, G. (2014). Consumer attitude metrics for guiding marketing mix decisions. *Marketing Science*, 33(4):534–550.
- Hu, Y., Du, R. Y., and Damangir, S. (2014). Decomposing the impact of advertising: Augmenting sales with online search data. *Journal of Marketing Research*, 51(3):300–310.
- Joo, M., Wilbur, K. C., Cowgill, B., and Zhu, Y. (2014). Television advertising and online search. *Management Science*, 60(1):56–73.
- Keller, K. L. (1993). Conceptualizing, measuring and managing customer-based brand equity. *Journal of Marketing*, 57(1):1–22.
- Lehmann, D. R., Keller, K. L., and Farley, J. U. (2008). The structure of survey-based brand metrics. *Journal of International Marketing*, 16(4):29–56.
- Lewis, R. A. and Reiley, D. H. (2013). Down-to-the-minute effects of Super Bowl advertising on online search behavior. In *14th ACM Conference on Electronic Commerce*, volume 9.

- Pauwels, K., Demirci, C., Yildirim, G., and Srinivasan, S. (2016). The impact of brand familiarity on online and offline media synergy. *International Journal of Research in Marketing*.
- Pauwels, K. and van Ewijk, B. (2013). Do online behavior tracking or attitude survey metrics drive brand sales? An integrative model of attitudes and actions on the consumer boulevard. *MSI Working Paper 13-118*.
- Percy, L. and Rossiter, J. R. (1992). A model of brand awareness and brand attitude advertising strategies. *Psychology & Marketing*, 9(4):263–274.
- Pew Research Center (2012). Assessing the representativeness of public opinion surveys. Technical report, Pew Research Center.
- Rutz, O. J. and Bucklin, R. E. (2011). From generic to branded: A model of spillover in paid search advertising. *Journal of Marketing Research*, 48(1):87–102.
- Scott, S. L. and Varian, H. (2014). Bayesian variable selection for nowcasting economic time series. In *Economic Analysis of the Digital Economy*. University of Chicago Press.
- Srinivasan, S., Vanhuele, M., and Pauwels, K. (2010). Mind-set metrics in market response models: An integrative approach. *Journal of Marketing Research*, 47(4):672–684.
- Stahl, F., Heitmann, M., Lehmann, D. R., and Neslin, S. A. (2012). The impact of brand equity on customer acquisition, retention, and profit margin. *Journal of Marketing*, 76(4):44–63.
- Survey Monkey (2015). Brand tracking surveys to measure marketing efforts. <https://www.surveymonkey.com/mp/brand-tracking-surveys-better-brand-advertising/>. Accessed August 20, 2015.
- Wu, L. and Brynjolfsson, E. (2014). The future of prediction: How google searches foreshadow housing prices and sales. In *Economic Analysis of the Digital Economy*. University of Chicago Press.
- Zeileis, A., Kleiber, C., and Jackman, S. (2008). Regression models for count data in R. *Journal of Statistical Software*, 27(8).

Appendices

Appendix A Brands Included in Study

A.1 Smartphone Brands

- Android
- BlackBerry
- Galaxy
- HTC
- iPhone
- LG
- Moto
- Nexus
- Samsung

A.2 Vehicle Brands

- Acura
- Audi
- BMW
- Buick
- Cadillac
- Chevrolet
- Chrysler
- Dodge
- Fiat
- Ford
- GMC
- Honda
- Hyundai
- Infiniti

- Jeep
- Kia
- Land Rover
- Lexus
- Mazda
- Mercedes-Benz
- Mitsubishi
- Nissan
- Porsche
- Scion
- Subaru
- Toyota
- Volkswagen
- Volvo

Appendix B Correlation Matrix for Predictors

	Recall	Recognition	Familiarity	Consideration	Purchase Intent	Owens Brand	Problems with Brand	Interest in Category	In-Market
Recall		0.26	0.33	0.32	0.34	0.22	0.09	0.04	0.01
Recognition			0.42	0.29	0.16	0.17	0.07	0.13	0.05
Familiarity				0.45	0.32	0.37	0.16	0.21	0.12
Consideration					0.49	0.34	0.13	0.08	0.03
Purchase Intent						0.41	0.15	0.01	0.01
Owens Brand							0.36	0.05	0.04
Problems with Brand								0.03	0.06
Interest in Category									0.36
In-Market									

Table 5: Correlation matrix for brand attitudes and control variables for the smartphone category.

	Recall	Recognition	Familiarity	Consideration	Purchase Intent	Owens Brand	Problems with Brand	Interest in Category	In-Market
Recall		0.09	0.32	0.37	0.37	0.37	0.13	0.00	0.00
Recognition			0.25	0.12	0.06	0.05	0.01	-0.07	-0.05
Familiarity				0.38	0.22	0.24	0.09	0.16	0.10
Consideration					0.42	0.29	0.10	0.06	0.04
Purchase Intent						0.45	0.14	0.01	0.01
Ownership							0.40	0.02	0.01
Problem								0.01	0.02
Engage									0.38
In-Market									

Table 6: Correlation matrix for brand attitudes and control variables for the vehicle category.

Online Supplement to “Brand Attitudes and Search Engine Queries”

Pilot Study

The pilot study was conducted over 9 months beginning in October 2013 in conjunction with a university-based consumer panel, using a similar strategy as the main study. As part of the recruitment, the users were required to review their personal Google Search History (see <http://history.google.com>) and complete an online authorization to grant permission for Google to link counts of their subsequent smartphone-related searches to their survey responses, giving us access to the daily counts of search for each brand over the observation period.⁶ Users’ brand searches were monitored over 9 months and during that period, users answered three identical brand attitude surveys in November 2013, February 2014, and May 2014. The resulting data consisted of daily brand search counts for each of 8 popular smartphone brands (Android, Blackberry, HTC, iPhone, LG, Motorola, Nokia and Samsung) for each user, along with that user’s brand attitudes at three points in time.

Brand Attitude Data

Table 7 summarizes the key measures that were collected in the three waves of the survey. The first three columns of the table show the percentage of respondents giving a particular response in any given quarter. For example, 15.9% of respondents indicated that they were “actively shopping for a new smartphone” in November 2013. The last two columns show the percentage of respondents who changed their response between November 2013 and February 2014 and between February 2014 and May 2014. For example, only 5.7% of respondents changed their response for *purchase intent* for the iPhone between November 2013 and February 2014.

The first panel at the top of Table 7 shows the incidence of active shopping and purchasing among the survey respondents. About 12% to 16% of respondents are actively shopping at any given time and between 3% and 8% have made a purchase in the past month. We also found a substantial fraction of these users indicated that “I am not actively looking for a new smartphone, but I’m always paying attention to

⁶The research protocol was approved by both Google’s privacy team and a university Institutional Review Board. Consistent with Google policy, the counts of search queries for each user did not leave Google’s computing environment. See Section 2.4 for more information.

	% of Panelists Responding Positively			% of Panelists Who Changed Their Answer	
	Nov 2013	Feb 2014	May 2014	Nov–Feb	Feb–May
Shopping/Purchasing					
Actively shopping	15.9	12.5	14.8	19.3	13.6
Purchased past month	5.7	8.0	3.4	11.4	11.4
Intend to purchase	10.2	6.8	5.7	10.2	12.5
Pay attention to the market	46.6	44.3	37.5	25.0	27.3
Brand Attitudes (2 examples out of 8 brands)					
iPhone					
Recall	77.3	72.7	77.3	18.2	15.9
Recognition	94.3	90.9	88.6	10.2	15.9
Familiarity	59.1	52.3	53.4	20.5	21.6
Purchase consideration	53.4	47.7	46.6	10.2	17.0
Purchase intent	30.7	29.5	28.4	5.7	8.0
Samsung					
Recall	75.0	85.2	79.5	17.0	19.3
Recognition	97.7	93.2	96.6	9.1	8.0
Familiarity	55.7	55.7	52.3	13.6	19.3
Purchase consideration	68.2	64.8	67.0	14.8	20.5
Purchase intent	29.5	22.7	29.5	13.6	18.2
Ownership					
Own a smartphone	86.4	84.1	86.4	6.8	9.1
Own an iPhone	25.0	25.0	23.9	2.3	5.7
Own a Samsung	27.3	31.8	31.8	6.8	4.5
Problems					
Problems with device	11.4	10.2	10.2	14.8	11.4
Problems with OS	34.1	39.8	34.1	23.9	28.4

Table 7: Summary of brand attitude data in the pilot study.

what is available, so that I know when to upgrade.” The number of respondents who were “paying attention” in this way was 46.6% during the smartphone announcement season in November 2013.

The second panel in Table 7 shows the brand attitude data for iPhone and Samsung. (We omit similar summaries for other brands in the interest of space.) We find minimal variation over time in the aggregate brand metrics. For example, the percentage of respondents who indicated that they would intend to purchase an iPhone is very stable at about 30% over the three quarterly surveys.

The third panel shows the fraction of respondents who own a smartphone, which was holding at around 85% during our observation period in 2013–14. The installed base for the two largest hardware brands is also holding steady with about 30% of respondents indicating they own a Samsung and about 25% indicating that they own an iPhone. The last panel of the table shows that about 10% of respondents indicate that they are having problems with the hardware device that they own (e.g., a cracked case), while nearly 35% of respondents indicate having had some type of problem with their smartphone’s operating system.

The final two columns of Table 7 show the percentage of customers who changed their survey response to the particular question between successive waves of the survey. We find that while most of the aggregate brand metrics are quite stable, the micro-level responses to the survey questions do vary much more over time, consistent with previous research (Day and Pratt Jr., 1971). For example, while the percentage of people who are actively shopping (Table 7’s first row) is quite steady ranging from 12.5 to 15.9%, we see from the fourth column that 19.3% of respondents changed their response to that question between the November 2013 and February 2014 surveys. This means that the group of people who is actively shopping varies from quarter to quarter, even though the percentage who are shopping is relatively stable. (Notably, purchase intent for iPhone is remarkably stable with only 5.7% and 8.0% of respondents changing their responses; this suggests that iPhone has a stable loyal customer base.)

Brand Search Data

The search data spans 7 months and covers the same period during which we collected the survey data: October 2014 to May 2014. During this period, for each respondent, we observe the daily counts of Google Search queries that the user submitted while logged into Google services, as well as the daily counts of

queries that included each of eight popular smartphone brands. To determine the daily counts of each panelist's smartphone queries, a program modified each query to have only lowercase letters, replaced punctuation characters with spaces, and compressed all whitespace to a single space. Then it counted the number of queries that matched each of the strings below using word match. (Note that in the main study, we improved our approach to identifying brand-related terms.)

- Android: "android"
- iPhone: "iphone"
- BlackBerry: "blackberry"
- HTC: "htc"
- LG: "lg"
- Motorola: "moto", "motorola"
- Nokia: "nokia"
- Samsung: "samsung"

Even among our small group of 57 users, there is a steady volume of search over the observation window for the most frequently-searched brands such as "android" and "iphone". But despite this steady volume of search, there was no substantial change in the aggregate search rate over time.

As in the main study, the distribution across users for total search queries and queries for specific brands is quite skewed with more than half of the users making no smartphone-related searches. The counts can also range quite high, as might be consistent with an over-dispersed count distribution such as Zipf or negative binomial.

Relationship Between Brand Attitudes and Brand Search

To produce the estimates reported in Table 7, we regressed the *daily* search counts on the brand attitudes reported in the survey that was closest to that day using the same logit-negative binomial hurdle model described in Section 3. The parameter estimates are remarkably similar to those reported in the main study.

		Smartphones	
		est	se
Hurdle Equation	Recognition	-0.46	0.21
	Recall	0.02	0.14
	Familiarity	0.56	0.13
	Consideration	0.35	0.14
	Purchase Intent	0.34	0.15
	Owns Brand	0.97	0.14
	Problems with Brand	0.54	0.11
	In-Market*	0.40	0.14
	Interested in Category*	0.52	0.11
	News	0.06	0.05
	log(Total Search)*	0.76	0.05
Count Equation	Recognition	1.39	0.48
	Recall	-0.33	0.35
	Familiarity	0.22	0.27
	Consideration	0.24	0.27
	Purchase Intent	-0.26	0.32
	Owns Brand	-0.32	0.28
	Problems with Brand	-0.21	0.42
	In-Market*	-0.32	0.31
	Interested in Category*	0.02	0.23
		log(Total Search)	0.24
	Dispersion (log(θ))	-5.2	2.2

Table 8: Hurdle model estimates (“est”) and standard errors (“se”) relating brand attitudes to search at the user-brand-day level. Values in boldface are significant at 95% confidence. * Indicates a category-level control variable that was not specific to a brand.

Despite the fact that the pilot was conducted with a separate group of users recruited by an entirely different panel, the pilot also found that users who are familiar with or would consider purchasing a smartphone brand are more likely to search for that brand. The only exception is the negative sign for the estimate of the effect of brand recognition on search, which is significant and negative, a finding which we believe is spurious.