

8-1-2012

Is Your Brand Going Out of Fashion? A Quantitative, Causal Study Designed to Harness the Web for Early Indicators of Brand Value

Maureen S. Cole
Georgia State University

Follow this and additional works at: http://scholarworks.gsu.edu/bus_admin_diss

Recommended Citation

Cole, Maureen S., "Is Your Brand Going Out of Fashion? A Quantitative, Causal Study Designed to Harness the Web for Early Indicators of Brand Value." Dissertation, Georgia State University, 2012.
http://scholarworks.gsu.edu/bus_admin_diss/13

This Dissertation is brought to you for free and open access by the Programs in Business Administration at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Business Administration Dissertations by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

PERMISSION TO BORROW

In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Georgia State University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote from, to copy from, or publish this dissertation may be granted by the author or, in his/her absence, the professor under whose direction it was written or, in his absence, by the Dean of the Robinson College of Business. Such quoting, copying, or publishing must be solely for the scholarly purposes and does not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential gain will not be allowed without written permission of the author.

Maureen Schumacher Cole

NOTICE TO BORROWERS

All dissertations deposited in the Georgia State University Library must be used only in accordance with the stipulations prescribed by the author in the preceding statement.

The author of this dissertation is:

Maureen Schumacher Cole
Google, Inc
Atlanta, GA 30309
mschumacher@google.com
Maureen.schumacher@gmail.com

The director of this dissertation is:

V Kumar, Ph.D.
Richard and Susan Lenny Distinguished Chair Professor of Marketing,
Executive Director, Center for Excellence in Brand & Customer Management, and
Director of the Ph.D. Program in Marketing
J. Mack Robinson College of Business
Georgia State University
35 Broad Street, Suite 400
Atlanta, GA 30303

**Is Your Brand Going Out of Fashion?
A Quantitative, Causal Study Designed to Harness the Web for Early
Indicators of Brand Value**

BY

Maureen Schumacher Cole

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree
Of
Doctor of Philosophy
In the Robinson College of Business
Of
Georgia State University

GEORGIA STATE UNIVERSITY
ROBINSON COLLEGE OF BUSINESS
2012

Copyright by
Maureen Schumacher Cole
2012

ACCEPTANCE

This dissertation was prepared under the direction of the Maureen Schumacher's Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctoral of Philosophy in Business Administration in the J. Mack Robinson College of Business of Georgia State University.

H. Fenwick Huss, Dean

DISSERTATION COMMITTEE

V. Kumar, Ph.D. (Chair)

Nita Umashankar, Ph.D.

J. Andrew Petersen, Ph.D. (UNC at Chapel Hill)

Table of Contents

ABSTRACT	1
Chapter 1: INTRODUCTION	2
Chapter 2: LITERATURE REVIEW	10
Research Domain	12
Chapter 3: THEORETICAL FRAMEWORK	15
Customer Information-Processing Model	15
Interaction With a Search Engine	18
Chapter 4: CONCEPTUAL MODEL	20
Brand Value	22
Customer-Based Variables	23
1. CUSTOMER ACTION	24
2. CUSTOMER VALUE	24
Financial Variables	29
1. MARKETING ACTION	29
2. BRAND HISTORY	29
3. FIRMOGRAPHICS	30
4. CONTROL VARIABLES	31
Chapter 5: METHODOLOGY	34
Database Overview	34
Predictor Variables	34
Chapter 6: ANALYSIS	39
Chapter 7: CONTRIBUTION, LIMITATIONS AND FUTURE RESEARCH	49
Contribution	49
Limitations and Future Research Direction	50
Appendix	52
Bibliography	53

Figures and Tables

FIGURE 1: APPLE, COCA COLA, AND GE	5
FIGURE 2: MOTOROLA, RAZR, AND KRISPY KREME.....	8
FIGURE 3: CONSUMER INFORMATION-PROCESSING	16
FIGURE 4: INTERACTION WITH A SEARCH ENGINE	19
FIGURE 5: CONCEPTUAL MODEL.....	22
FIGURE 6: DECISION-MAKING FLOW CHART	39
FIGURE 7: REVISED CONCEPTUAL MODEL.....	44

FIGURE 1: APPLE, COCA COLA, AND GE	5
FIGURE 2: MOTOROLA, RAZR, AND KRISPY KREME.....	8
FIGURE 3: CONSUMER INFORMATION-PROCESSING	16
FIGURE 4: INTERACTION WITH A SEARCH ENGINE	19
FIGURE 5: CONCEPTUAL MODEL.....	22
FIGURE 6: DECISION-MAKING FLOW CHART	39
FIGURE 7: REVISED CONCEPTUAL MODEL.....	44

ABSTRACT

Can Internet search query data be a relevant predictor of financial measures of brand value? Can Internet search query data enrich existing financial measures of brand valuation tools and provide more timely insights to brand managers? Along with the financial based motivation to estimate the value of a brand for accounting purposes, marketers desire to show “accountability” of marketing activity and respond to the customer’s perception of the brand quickly to maintain their competitive advantage and value. The usefulness of the “consumer information processing” framework for brand, consumer and firm forecasting is examined. To develop our hypotheses, we draw from the growing body of work relating web searches to real world outcomes, to determine if a search query for a brand is causal to, and potentially predictive of brand, consumer and firm value. The contribution to current literature is that search queries can predict perception, whereas previous research in this nascent area predicted behavior and events. In this direction, we propose arguments underpinning this research as follows: the theoretical background relative to brand valuation and the theoretical frame based on an in-depth review of how scholars have used search query data as a predictive measure across several disciplines including economics and the health sciences. From a practitioner perspective, unlike traditional valuation methods search query data for brands is more timely, actionable, and inclusive.

Chapter 1: INTRODUCTION

Can Internet search query data be a relevant predictor of financial measures of brand value? Can Internet search query data enrich existing financial measures of brand valuation tools and provide more timely insights to brand managers?

The concept of brand value first emerged in the marketing literature in the late 1980s. The use of a financial term for what was a consumer-based construct was a highly effective technique to communicate the idea that brands are long-lived business assets that can have significant financial value. As a result, there was increased interest in the boardroom in the value of brands; leading to the disclosure of the value of acquired brands on the balance sheet (Hull, 2008).

Brand valuation plays an important role in business practice for three distinct reasons. First, firms are financially motivated to estimate the value of a brand for accounting purposes (Keller, 1993). Second, marketers can show “accountability” for marketing activity and report results using the language of finance (Salinas and Ambler, 2009). Third, the value of a brand name is closely associated with the customer’s satisfaction, awareness, and perception of the quality of the product (Aaker, 1991).

Brand value has typically been measured from either a financial market analysis or a consumer-demand-driver analysis of the brand. Methodologies, like suppliers, have proliferated. In 2009, Salinas and Ambler studied 52 such suppliers of brand valuation methods, admitting that there

were many others.

In general, Salinas and Ambler (2009) categorized suppliers of brand valuation methods as technical or brand management providers. Technical providers define the financial market value of the brand at arm's length between the buyer and seller. Brand Finance is an example of a technical provider using royalty relief, a method that tends to be the choice among this sector. Brand Finance describes royalty relief as the application of an appropriate royalty rate to an estimate of likely future sales which leads to the income attributable to brand royalties in future years. The stream of brand royalty is discounted back to a net present value—the brand value (Brandfinance, 2011). The variables to arrive at the income attributable to brand royalties take into account market data, including quantitative consumer research, insights into competitors, and forecasted brand earnings. Brand Finance is well known for its annual release of a league table of the most valuable brands based on their financial strength (Brand Finance, 2011). The release of the annual league table is in March of the following year.

Brand management providers prefer demand-driver analysis designed to improve marketing effectiveness and accountability. Demand-driver analysis applies an economic use approach to brand valuation using valuation methodology similar to that employed by analysts and accountants. This approach is focused on the intrinsic value of the brand determined by its ability to generate demand. The actual value of the brand is the sum of the future earnings that brand is forecasted to generate, discounted to a present-day value. Millward Brown Optimer is an example of a brand management provider. Millward Brown is known for its highly publicized

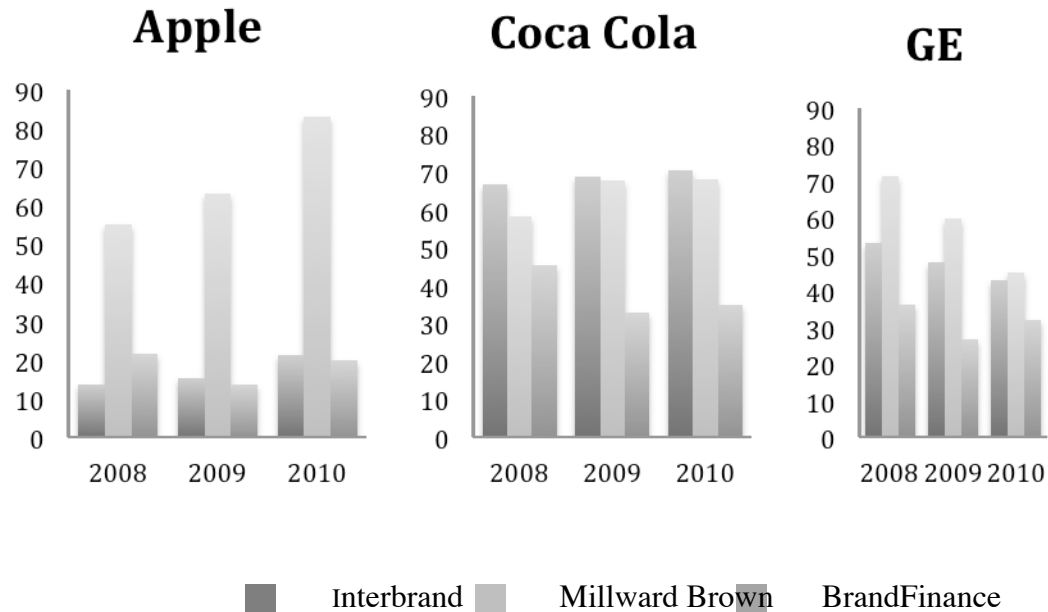
annual league table of the top 100 most valuable global brands. This ranking focuses on consumer-facing brands and, like other rankings, includes customer brand perception metrics in its valuation methodologies. Consumer perceptions are typically gathered over the course of the year through interviews and surveys and incorporated into the final value. The release of the actual valuation or annual league table is in April or May of the following year.

To illustrate the variations between providers, Salinas and Ambler (2009) compared brand valuations published in 2005 for Toyota, Samsung, and Apple across three brand management providers: Interbrand, Millward Brown Optimer, and Vivaldi Partners. The findings varied by a factor of two or more. Likewise, Knowles (2008) compared valuations for eight well-known brands across three providers, Interbrand, Millward Brown, and Brand Finance, again with wide variations in their findings. We conducted our own comparison of Interbrand, Millward Brown Optimer, and Brand Finance valuations for Apple, Coca Cola, and GE during the years 2008, 2009, and 2010, See Figure 1.

Figure 1: Apple, Coca Cola, and GE

Brand Values 2008, 2009, and 2010

As defined by Interbrand, Millward Brown Optimer, and Brand Finance



Despite each provider using legitimate valuation approaches, the disparity in the results is evident and undermines the credibility of current valuation processes. Even if their measurements differ, all tend to be backward looking. Thus, although interesting from a news perspective, practitioners cannot build strategies or predict turning points in the value of their brands.

Two inherent challenges to valuing brands are evident: (1) Brand valuations and rankings are made available after a lag period from the time the data is collected, compiled, and analyzed to the time that it is delivered, and (2) the consumer's perception of the brand is rarely factored into the valuation, and in cases where it is included, the methodology is proprietary. This leaves CEOs, investors, and brand managers to make decisions based on data that is months to a year

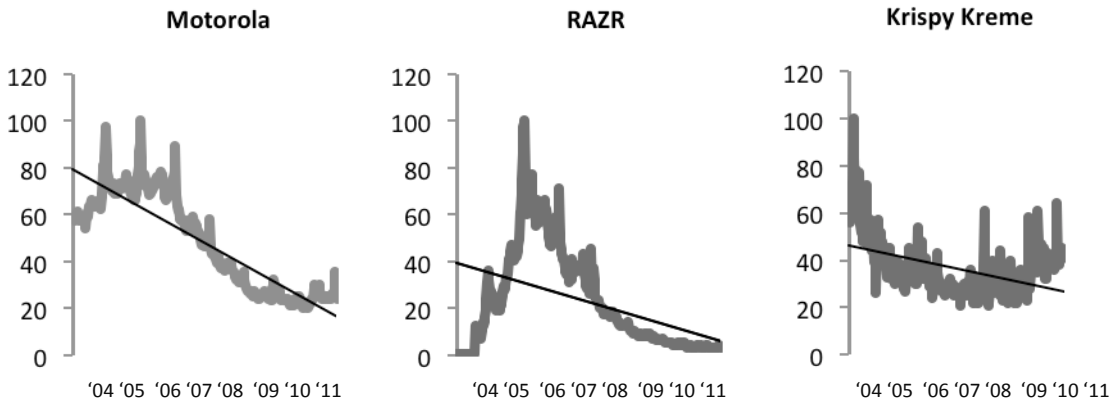
old and that do not adequately represent a key asset, the consumer.

This raises the question as to whether or not brand values can be predicted; i.e., are there any forward-looking metrics that could help us explain brand value? Specifically, can Internet search query data be a relevant, timely predictor of a brand's financial value? In this research, we posit that the firm is able to predict brand value via a consumer's search query on the Internet for brand terms. We believe that branded search queries are the forward-looking metric that explains brand value and could be used to make reliable predictions about a brand's value in the marketplace literally months before that change is validated through traditional brand valuation methods.

The downside of not being able to predict brand value is that this can have a great impact on the practitioner. Motorola is an example of a brand that customers lost interest in, while brand managers were slow to adapt to changing signals in the marketplace for feature-rich phones and 3G. Having successfully created demand for the RAZR, which launched in 2004, Motorola released a series of product extensions including a limited edition version given to guests at the Academy Awards ceremony in 2005. This cemented the RAZR's reputation as an iconic brand and distinctive phone. As sales of the RAZR boomed, Motorola seems to have rested on its laurels, failing to develop a pipeline of exciting replacement products to maintain upward momentum (Frampton, 2007). The RAZR, highly successful at launch and significantly raising the profile of the brand, was not followed by any "blockbuster" handsets before the RAZR approached the end of its life cycle.

Consumer insight is at the heart of successful branding, according to experts. Motorola would have been well advised to gain a deep understanding of its brand value, what drives its consumers, and to keep on top of emerging trends (Frampton, 2007). From the search queries in Figure 2, we see a noticeable rise in queries worldwide for the brand Motorola and RAZR as anticipated during the height of the RAZR releases in 2005. Both queries reached their highest levels the week ending December 25, 2005. Foreshadowing the decline of Motorola, we see a dramatic drop in the number of queries for RAZR, down 30% in one week and 50% within eight months. By early 2007, the number of people searching online for RAZR and Motorola had dropped significantly. Clearly RAZR was no longer top of mind, and without a product pipeline, neither was Motorola.

Figure 2: Motorola, RAZR, and Krispy Kreme
 Google Insight Search Queries
 Worldwide, January 1, 2004–August 31, 2011
 All categories; collected on September 12 and 17, 2011



In 2003, FORTUNE magazine called Krispy Kreme, the doughnut maker, “America’s hottest brand.” What followed was a period of aggressive expansion in New England, including a well-publicized Krispy Kreme store at the Prudential Center in Boston, Massachusetts, which opened on April 15, 2004. During this time, search queries for the Krispy Kreme brand reached their height the week ending March 7, 2004, followed by a sharp decline. Within four months, the number of searches for the branded term “Krispy Kreme” was cut in half. By the end of the year, Krispy Kreme was forced to close all but one store following its recent expansion, and the well-hyped Krispy Kreme at the Prudential Center in Boston was closed sixteen months after it opened (Surowiecki, 2004). By 2005, the company faced defaulting on its \$150M credit line (Anderson, 2008). Would the knowledge of the dramatic decrease in searches for the Krispy Kreme brand have prompted executives to reconsider their expansion plans?

In what follows, we will describe the current state of the relationship between brand value and search queries, as well as other variables of interest including, customer actions, customer values, marketing actions, brand history, and firmographics. The data we will use is available over time, and for multiple brands. Typically this type of research would be a descriptive analysis as it specifies the relationship between brand value and search queries. However, we will use a hierarchical cross-sectional and time-series analysis framework to explain the relationship between our predictive variables and future brand value, along with controlling for numerous other factors that can impact brand value.

Looking ahead, in Chapter 2, we explore the literature concerning the practice of gathering forward-looking metrics to help researchers explain lag data. We will understand how that practice has carried over to digital data, and describe the studies' contributions to the literature by using search queries to predict the present in terms of a consumer's real-time perception of a brand. In addition, we extend the domain of past studies beyond the fields of epidemiology and health to the social sciences specific to marketing. In Chapter 3, we will explain the consumer information-processing model and search engine behavioral model as theoretical underpinnings of our research. In Chapter 4, we describe our conceptual model and hypotheses followed by Chapter 5, where we outline the methodology and data-gathering process. In Chapter 6, we complete our analysis to determine if the number of search queries for a brand is positively associated with brand value. In Chapter 7 we conclude by summarizing the contribution to practitioners and existing literature, we further outline the limitations and provide guidance on future research.

Chapter 2: LITERATURE REVIEW

The practice of gathering forward-looking metrics to help researchers explain lag data is well known. Health practitioners collect information concerning school absenteeism, pharmaceutical sales, and medical diagnoses in an effort to provide early detection of disease outbreaks (Hulth, Rydevik and Linde, 2009). Companies gather sales data, and governments gather unemployment and consumption information. The benefits are that early detection allows for interventions to lower morbidity, allows companies to modify strategies, or governments to enact policy.

Conversely, the use of gathering digital data as a forward-looking metric to explain lag variables is a fairly nascent area. With billions of people online, the digital footprint and the number of search queries on the web is vast. In our literature review, we explore early studies focused on the relationship of digital metrics, and web searches, to lag variables. We will examine the evolution of this body of work that began in health sciences to identify epidemics and outbreaks, through the adoption within economics to identify unemployment and consumption trends. This research is similar in many ways to previous studies as it associates the number of search queries with an outcome. In previous studies, the outcome was a behavior or an event. This research varies in that it is the first of its kind used in marketing sciences, and the first of its kind to associate search queries with perception, specifically the customer's perception or efficacy toward a brand.

The practice of gathering digital data has been challenging in some respects due to a lack of access. Early studies arose from the fields of economics and health sciences with the data utilized in their digital surveillance coming from varied sources. For example, Johnson et al. (2004) used the counts from articles viewed on a U.S. medical site with the Centers for Disease Control and Prevention for surveillance of incidences of influenza. Ettredge, Gerdes and Karuga (2005) used unemployment-related queries captured by Wordtracker, a keyword research tool, to predict the number of unemployed in the United States.

Even with a limited supply of data, researchers continued to study the phenomena of the statistical linkage between online behavior and offline occurrences. Cooper, Mallon, Ledbetter, Pollack, and Peipins (2005) used Yahoo Buzz and LexisNexis news reports to correlate search activity for cancer with their estimated incidence over the period from 2001 to 2003. Some researchers employed creative methods such as launching ad campaigns related to flu symptoms as a means to gather counts of user clicks to correlate with epidemiological data during the 2004–2005 Canadian flu season (Eysenbach, 2006). Others partnered with website operators like a Swedish medical website for query counts to facilitate influenza monitoring (Hulth, Rydevik and Linde, 2009) or with yahoo.com to show that search volume for select influenza queries correlated with caseloads from 2004 to 2008 (Polgreen, Chen, Pennock and Nelson, 2008).

Research Domain

While search engines maintain a record of the searches entered into their websites, it was not until the advent of the Google Trends and Insights for search applications in 2007 that this information was made available to researchers and the public at large. Including counts for all queries on a weekly basis since 2004, the data amounts to a vast and rich database. With this information at their disposal, researchers embraced the opportunity to correlate search queries to real-world outcomes across a variety of disciplines. At the outset, the data was documented and used to describe the evolution of particular queries before the U.S. presidential elections (Constant and Zimmerman, 2008). Using regression models, what soon evolved was the discovery of a high correlation between search queries and unemployment in Germany and the United States. (Askatas and Zimmermann, 2009; D'Amuir and Marucci, 2010).

Ultimately, Ginsberg et al. (2009) used search engine query data to develop a simple model that would estimate the probability that a random physician in a region of the United States would see a patient with influenza. This well-cited, seminal work led to more extensive research that pushed the boundaries within epidemiological studies and econometrics. An early example prior to the release of traditional indicators was Choi and Varian's (2009) research using a simple forecasting model to "predict the present" in terms of travel, as well as auto, retail, and home sales.

Relying on regression analysis, the focus expanded from predicting events to predicting behavior. Goel, Hofman, Lahaie, Pennock and Watts (2010) used search query volume to

determine the relationship of search queries to the opening weekend box office revenue for feature films. Goel et al. (2010) also used search to predict first-month sales of video games, and the rank of songs on the Billboard Hot 100 chart, finding in all cases that search counts were predictive of future outcomes.

Kissan, Wintoki & Zhang (2011, p. 2) believe “the availability of measures of consumer search behavior is only going to increase as we move further into the digital age”. Consistent with this trend, scholars are coming to recognize that what individuals are searching for leaves a trail of “what we collectively think” and “what might happen in the future” (Rangaswamy, Giles and Seres, 2009, p.58). Across the empirical cases that we have considered, we observe that search counts are predictive of an event prior to the release of traditional indicators, or a behavior that will occur within days or weeks in the future. The predictions have typically been epidemiological or economic events including unemployment, consumption, travel, or game, and real estate sales. Our intent is to contribute to the literature by using search queries to predict the present in terms of perception, specifically in terms of a consumer’s real-time perception of a brand. In addition, we will extend the domain of past studies from epidemiological and health fields to the social sciences specific to marketing. See Table 1 for a summary of this study’s contribution with regard to other studies that have focused on relating web searches to real-world outcomes.

Table 1: Summary of Related Studies
 Relating Web Searches to Real-World Outcomes
 Delineated by health, macroeconomics, and marketing sciences

Studies	Contributions to Health Sciences	Studies	Contributions to Macroeconomics
Johnson, Wagner, Hogan, Chapman, Olszewski, Dowling, and Barnas (2004)	Measures the correlation between article access counts and influenza activity.	Ettredge and Karuga (2005)	Internet search data to predict number of unemployed in the United States.
Cooper, Mallon, Leadbetter, Pollack, and Peipins (2005)	Cancer-related searches and volume of news coverage with cancer incidence and cancer mortality.	Constant and Zimmerman (2008)	Evolution of keyword searches before U.S. presidential election; indicator of the traction of the candidates.
Eysenbach (2006)	Tracking flu-related searches on the web for syndromic surveillance.	Askitas and Zimmermann (2009)	Internet search data as a predictor of unemployment in Germany.
Hulth (2007)	Web queries for influenza monitoring.	Choi and Varian (2009)	Search data to predict initial claims for unemployment benefits.
Polgreen, Chen, Pennock, and Nelson (2008)	Internet searches for influenza surveillance; showed that search volume for handpicked influenza-related queries was correlated with caseloads over the period 2004–2008.	Choi and Varian (2009)	Search data to predict travel as well as auto and real estate sales in the United States.
Ginsberg, Mohebbi, Patel, Brammer, Smolinski, and Brilliant (2009)	A simple model to estimate the probability that a random physician in a specific region of the United States would see a patient with influenza.	Suhoy (2010)	Framework for "nowcasting" private consumption; enables early estimates of the monthly index of private consumption in Israel.
Brownstein, Freifeld, and Madoff (2009)	Using web searches for digital disease detection; public health surveillance.	Bughin and Jacques (2011)	Use of online buzz and searches as indicators of telecom sales and company performance.
Wilson and Brownstein (2009)	Early detection of disease outbreaks in Canada.	McLaren and Shanhogue (2011)	Online searches as indicators of labor and housing markets in the United Kingdom.
Wu and Brynjolfsson (2009)	Nanoeconomic data to predict housing prices and sales.	Choi and Lui (2011)	Use of web searches to find turning points—in this case, how the Gulf oil spill negatively impacted parts of Florida while other areas were positively impacted.
Althouse and Cummings (2011)	Prediction of Dengue Incidence Using Search Query Surveillance	Schmidt and Vosen (2011)	Used multiple consumer search categories to build an index of retail searches. Using factor analysis, demonstrated that their index was outperforming key survey-based indicator standards.
Ayers, Ribisl, and Brownstein (2011)	Tracking the rise in popularity of electronic nicotine delivery systems; search query surveillance as a public method to evaluate the diffusion of new health products.	Goel, Hofman, Lahaie Pennock and Watts (2010)	Predicting consumer behavior with web searches; forecasting opening weekend box office revenue, first-month sales of video games, and the rank of songs on the Billboard Hot 100 chart.
Summary of Research Focused on Relating Web Searches to Real-World Perceptions and Contribution to Social Sciences			

The proposed study is designed to determine if brand search query data is positively associated with brand value and if search query data can offer greater empirical evidence as to the brand's efficacy. This study furthers existing science in that it is the first of its kind to relate web searches to real-world perceptions, breaking from the historic disciplines of health sciences and economics.

Chapter 3: THEORETICAL FRAMEWORK

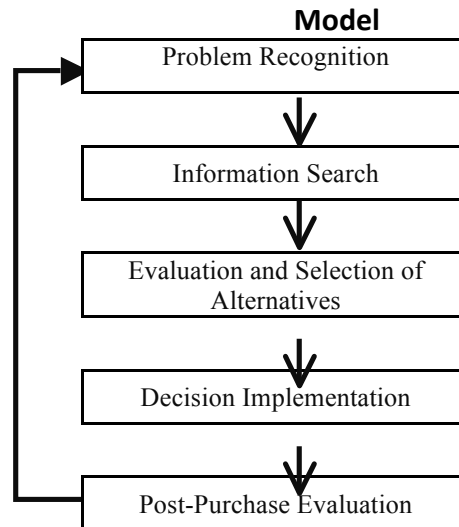
Two models provide the theoretical underpinnings of our research. First, the consumer information-processing model adopted from Kotler (1997), Schiffman and Kanuk (1997), and Solomon (1996) introduces the concept that a consumer has a fundamental need to search for information prior to making a purchase. The perspective being that information search begins after a consumer recognizes a problem, or need for a product. Later theorists, like Bloch, Sherrell and Ridgeway (1986), postulate that a consumer is always searching, either to make a purchase or to gather information for use at a later time. This model lays the foundation that a customer's search query for a brand is a measure of their intent, be it positive or negative.

Second foundational model is the interaction with a search engine (Holscher, 2000). This model details a consumer's searching behavior online as a series of querying, clicking, and browsing. Pentland (2008) describes this behavior as a more "honest signal" of actual interests and preferences. Wu and Brynjolfsson (2009) supports this belief by describing a consumer's search query as a digital trace left by consumers that can be compiled to reveal comprehensive pictures of the true underlying intentions and activities. This model solidifies the foundation that a customer's search query for a brand is a measure of their intent and that search queries in their aggregate are theorized to be a more honest signal of a consumer's value of a brand.

Customer Information-Processing Model

Marketing theorists have long argued that consumers seek information from a variety of sources (Cox, 1967). This concept has evolved into the dominant school of thought among consumer researchers known as the consumer information-processing model depicted in Figure 3 (Matsuno, 1997).

Figure 3: Consumer Information-Processing



Source: Adopted from Kotler (1997), Schiffman and Kanuk (1997), and Solomon (1996)

In this model, the consumer passes through five stages: (1) problem recognition, (2) information search, (3) evaluation and selection of alternatives, (4) decision implementation, and (5) post-purchase evaluation. A principal determinant used in the information process is information search. Though the consumer information-processing model identifies information search as a prominent aspect of the pre-purchase decision process, Bloch et al. (1986) postulates that information search is conducted throughout the entire information processing cycle as consumers are driven by an immediate purchase decision, as well as building a bank of information for future use.

Traditional sources of information that consumers rely on to reduce the uncertainty related to purchase decisions have been advertisements (radio, TV, and print), packaging, store displays, or brochures. Information is also indirectly conveyed by the price and type of store in which the product is sold. Consumers may also draw upon their own prior experience. He or she may obtain information from friends, family, or salespeople or read about products in product-rating publications or specialty magazines. In addition, the consumer may observe products being used by others (Bettman, Johnson and Payne, 1991).

With the advent of the Internet, the consumer can gather near-perfect information (Reibstein, 2002). Historically, the amount of information delivered to the consumer was constrained by the size of the package, the space available to deliver a message, the time available in a TV spot, or the space available in a print ad (Reibstein, 2002). Alternatively, in the digital world the information is not limited by the physical space (Alba and Hutchinson, 1987; Johnson, Lohse and Mandel, 1997). Web sites, natural search links, and ads provide detailed information about their merchandise, and comparison sites offer decision aids for making product price comparisons. Discussion forums, chat rooms, and online consumer reviews provide consumers with detailed experiences, opinions, and knowledge of others, both positive and negative, about a product or a seller.

Another perspective is that the Internet allows consumers to become more efficient in their buying processes. This efficiency results primarily because the cost of information, as well as the cost and time needed to acquire information, is so low (Bakos, 1997). With information being exceedingly easy to access, consumers become fully informed as to their options (Bakos, 1997;

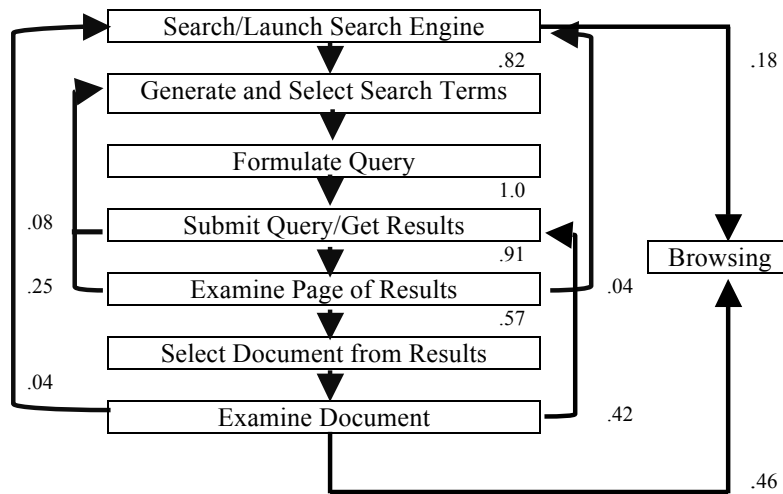
Brynjolfsson and Kahin, 2000). The consumer can benefit from multiple points of view to make informed assessments with less effort required to inform their purchase decisions.

Interaction With a Search Engine

It is hard for many to believe the important role that search engines play in today's world. Be it for academics, work or leisure, search engines can be used to find every conceivable kind of information. Some of the searches conducted are directional, meaning a customer is looking for a specific site. Some searches are fun and frivolous, meaning a customer is exploring hobbies, favorite past times, or entertainment. However, according to Rangaswamy et al. (2009), many searches are imbued with purpose, and search results can influence important decisions about someone's life, health, or a major purchase.

The online search process does not start with the search query itself, but with the idea to search (Joo, Mingyu, Wilbur and Zhu, 2010). The search idea may be general or very specific; a consumer may want to learn more about an entire product category or may want to consider a particular brand. Once the idea to search is formed, the consumer selects a search engine and the query. The query, or set of words known as keywords, can be one word or up to a full sentence. The search query determines the breadth of information returned by the search engine. After entering the query, the consumer examines the results and chooses the one of interest. The consumer makes this choice by clicking on the result, either an ad or a link to the information deemed most relevant. Holscher (2000) describes this process as one where an individual identifies a question or need and seeks information online by switching back and forth between querying, clicking, and browsing. Figure 4.

Figure 4: Interaction With a Search Engine
 Values represent transition probabilities to the next unit



Source: Holscher (2000)

The majority of the querying, clicking, and browsing behavior is conducted on Google: Google accounts for 65.5% of the search volume on the Internet resulting in billions of daily queries that are retained by Google in a database (Comscore, 2011). Pentland (2008) describes search as a more “honest signal” of actual interests and preferences since there is no bargaining, gaming, or strategic signaling involved in contrast to many market-based transactions. As a result, these digital traces left by consumers can be compiled to reveal comprehensive pictures of the true underlying economic intentions and activities (Wu and Brynjolfsson, 2009).

Chapter 4: CONCEPTUAL MODEL

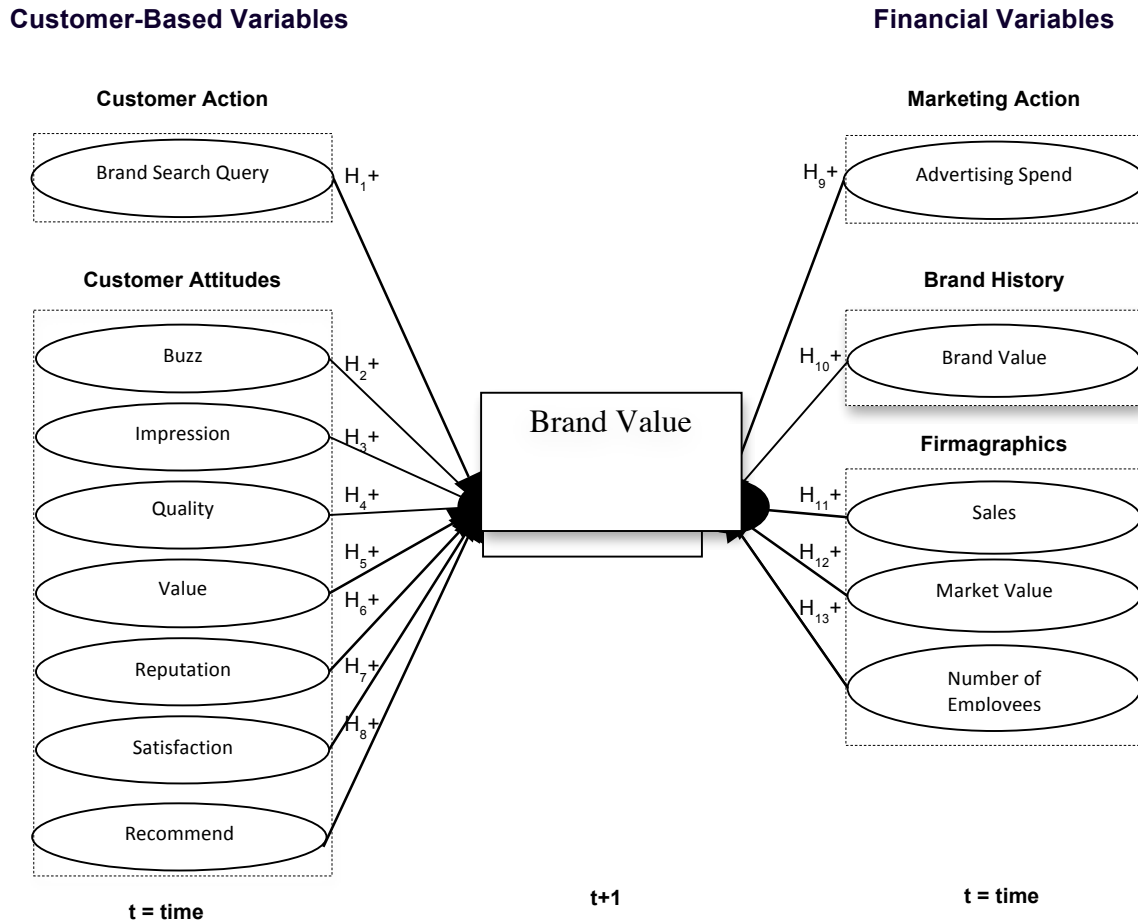
This chapter presents the conceptual model for this study and states the hypotheses tested. The central objective of this study is to determine if the number of search queries is positively associated with brand value. Throughout the literature, researchers identify numerous variables relative to brand value. Some of these variables are connected to financial models like sales and market value (Farquhar, 1989; Simon and Sullivan, 1993; Haigh, 1999). Research has also indicated customer-based variables relative to brand value like satisfaction and awareness (Aaker, 1991; Keller, 1993; Yoo and Donthu, 2000; Vazquez, Del Rio and Iglesias 2002; de Chernatony, Harris and Christodoulides 2004; Pappu, Quester and Cooksey, 2005; Christodoulides, Chernatony, Furrer and Abimbola, 2006). In developing the conceptual framework, we incorporate both financial and customer-based variables into our model for the purpose of explaining the relationship between search queries, consumer and financial variables to brand value. Figure 5.

The selected predictor variables from a financial perspective are those derived from firmographics, the brand's history and related marketing actions. The customer-based perspective is derived from customer values and actions. The conceptual model shows the temporal relationship of the predictor variables represented at time t (year) and the dependent variable, brand value at time $t+1$ (year+1), while controlling for economic factors that can impact brand value.

To address causality, the proposed framework specifies a time-based relationship among the focal variables. Specifically, a one-year lag relationship is defined between branded search queries, other predictor variables and brand value. Although such lags have not been defined in the marketing sciences for search query data, a year lag is largely supported for a change in level of measured communication, followed in the next year by a change in reputation and in the year following by a change in financial performance. A one-year lag is common when investigating the relationship between R&D activities and firm performance, and one-year lag effects are used when accounting for the impact of advertising and R&D expenditures (Peterson and Jeong, 2010).

The model is simple in terms of the number of variables. The rationale, as described by Balasubramanian and Kumar (1990), is that as more variables are added, the model becomes more complex and less interpretable, and understanding the interaction effects becomes more difficult. Also, several predictor variables usually belong to the same domain, thus increasing the potential for high correlations between the model variables.

Figure 5: Conceptual Model



Brand Value

In light of the inconsistencies of the brand valuation methods described earlier, the International Standards Organization (ISO) created brand valuation standards known as ISO 10668. ISO 10668 is the international norm that sets minimum standard requirements for the procedures and methods used to determine the monetary value of brands. It defines a coherent and reliable approach for brand valuation that takes into consideration financial, legal and behavioral science aspects (International Organization for Standardization, 2010). A certification program on the basis of ISO 10668 was developed to attest that a provider's valuation method conforms with

ISO 10668 standards (ISO, 2010). Interbrand and BrandFinance are two such valuation providers whose methodologies have been certified since ISO 10688's inception in October 2010. For purposes of our research, we use BrandFinance valuation characterized by Salinas and Ambler (2009) as a technical provider, with an emphasis on financial market value to minimize collinearity.

Brand Finance first issued its top 250 brand value ranking table in 2007 (reflecting 2006 values). In 2008, the study was extended to report the top 500 brands worldwide. (Brandfinance, 2011). Our data sets consists of 74 brands, each accorded a brand rating with values for each year from 2007 through 2011 (released in years 2008 to 2012). The brands in the dataset represent 15 industries (consumer banking, consumer package goods, department discount and grocery stores, fast food dining, insurance, travel, apparel, auto, consumer electronics, beverages, financial advisors, retail gasoline, retail specialty and telecom) and are headquartered across 15 countries (US, Japan, Switzerland, Sweden, Germany, Australia, Hong Kong, Spain, France, Great Britain, India, Korea, Netherlands, Finland and Italy). Brand values range from a low of \$1.4 trillion for Sprint in 2008, to a high of \$70 trillion for Apple in 2011. In addition, our data set represents brands that have declined and increased in value over the five-year period.

Customer-Based Variables

In most customer-based brand value research, surveys are used to determine the cognitive and behavioral value of a brand at the individual consumer level (Yoo et al., 2001). For our research, we use (1) customer action or the search queries for a brand conducted on Google's search engine, and (2) customer values gathered through online surveys as a measure of the cognitive and behavioral value of a brand at the individual level.

1. CUSTOMER ACTION

Brand Search Query. Chaudhuri and Holbrook (2001) support that brand value includes some degree of predisposition toward a brand by the consumer (Aaker, 1991; Assael, 1998; Beatty and Kahle, 1988; Jacoby and Chestnut, 1978). Historically, brand search queries have not been modeled as a predictive measure of brand value. The hypothesis is that when consumers are browsing, querying and searching on the Internet, this activity represents a predisposition to the brand, and the brand search queries are reliable predictors of a brand's value literally months before that change is validated through traditional brand valuation methods. As described in our literature review, the use of search queries is a growing body of work relating web searches to real world outcome, specifically healthcare and macroeconomics. If we extrapolate the theoretical underpinnings to the social sciences specifically marketing, we would expect increases in branded search queries for a particular brand to reflect in the brand's value released the following year. Therefore,

H₁: increases in brand search queries at t, increased brand value at t+1.

2. CUSTOMER VALUE

Researchers of brand value evaluate customer value indirectly by asking certain questions and drawing particular conclusions about the customer's response (Banyte, Joksaite and Virvilaite, 2007). Consumer surveys are often used to capture customer values, attitudes, beliefs, and behaviors toward a brand (Lehmann, Keller and Farley, 2008; Srinivasan and Hanssens, 1979; Rangaswamy, 1993; Keller, 1993; Berg, Matthews and O'Hare, 2007). A key challenge in developing survey-based brand values and metrics is the wide range of possible measures that could be employed. As Lehmann (1993) observed, no single measure fully captures the richness of brand value, multiple sets of measures and factors must be employed.

For this reason, multidimensional measures are used in this research to capture a customer values toward the brand. A traditional customer value metric used for marketing sciences has been the American Customer Satisfaction Index (ASCI), (Anderson, Fornell and Mazvancheryl, 2004; Fornell, Mithas, Morgeson and Krishnan 2006; Gupta, 2006). As the data is not consistently available at the brand level (Srinivasan and hanssens, 2009), we selected YouGov's Brandindex as our measure of customer value as it is available at the brand level, can be regionally segmented, and based on the segmentation it is available typically since 2007. YouGov's information is interval data, collected daily, on hundreds of brands, rather than just a few predetermined brands. The data is derived from surveys conducted over the Internet using an opt-in panel of the general public. YouGov's survey metrics have proven to be both meaningful and versatile in a variety of research areas including Omnicom's BrandScience projects. In particular, BrandIndex metrics have proven valuable as both upper and lower funnel analytics (Collins et al., 2010).

The customer value metrics from YouGov used in this research include buzz, impression, quality, value, reputation, satisfaction and recommendation. Each metric is taken independently – in other words, in any one survey, any individual respondent is asked about only one measure for the sector, not all seven. Therefore, none of the ratings influence each other within the survey (Brandindex, 2011)

Buzz. Buzz is the “top of mind” accessibility of the brand as perceived by the consumer. Like brand awareness, buzz measures the accessibility of the brand in memory (Chandon, 2003). The

buzz measure reflects what the consumer has heard in the media, news or advertising, or in conversations among friends and family about the brand. Buzz indicates whether a consumer noticed good or bad news, advertising or PR campaigns, product launches or whether there is any 'word on the street' emerging. As buzz reflects the recent brand sentiment and the direction of recent awareness, we hypothesize that an increase in positive and neutral buzz be associated with an increase in the brand value.

H₂: Increase in the customer's positive and neutral buzz at t, increased brand value at t+1

Impression. Impression is the strength, favorability, and uniqueness of perceived attributes and benefits of the brand. The impression variable reflects positive or negative feelings toward the brand. Leuthesser, Kohli and Harich (1995) determined that a brand represents the value of a product's impression, or the global effect, above that which would result for an otherwise identical product without brands name. Thus, we hypothesize that an increase in positive and neutral impressions of the brand, will be associated with an increase in the brand value.

H₃: Increase in the customer's positive and neutral impression of the brand at t, increased brand value at t+1

Quality. Quality reflects whether the brand is perceived as good or poor quality, irrespective of price. Many researchers have identified perceived quality as a dimension of brand value (Aaker, 1991; Kapferer, 1992; Kamakura and Russel, 1993; Martin and Brown, 1991; Feldwick, 1996).

We hypothesizes that an increase in perceived quality of a brand will be associated with an increase in the brand value.

H₄: Increase in the customer's positive and neutral perception of the quality of the brand at t, increased brand value at t+1

Value. Value measures the perception of a brand's price-point, its value offering. Value does not mean "cheap" or "expensive", but what the brand offers a customer in return for the price paid. Value is defined by Lasser, Mittal and Sharma (1995) as the perceived brand utility relative to its costs. From a customer perspective this means the customer weighs and measures what is received, with what must be given up to receive it. This gets at the heart of the value versus the price paid, and the utility the product offers. Therefore, we hypothesize that an increase in positive and neutral value will be positively associated with an increase in the brand value.

H₁₀: Increase in the customer's positive and neutral perception of the value of a brand at t, increased brand value at t+1

Reputation. Reputation informs us about what products to buy, what companies to work for or what stocks to invest in (Fombrun, 1996). Reputation is the image associated with owning or using a brand, the measure of a brand's "prestige" explained via consumers desire to work for a brand. Positive responses indicate that one would be proud to work for the brand, or advise a friend to work for the brand, or one would want to own a product with that brand. Fombrun (1996) demonstrated there are economic returns associated with a brand's reputation informing

our hypothesis that an increase in positive and neutral reputation will lead to an increase in the brand value.

H₅: Increase in the customer's positive and neutral perception of a brand's reputation at t, increased brand value at t+1

Satisfaction. Satisfaction is a measure of performance against customer expectations, it is centrally related to all elements and has some of the strongest correlations with brand health (Berg et al., 2007). It answers the question: Are customers satisfied with their experience with the brand? It measures the extent to which consumers purchase and use the brand, talk to others about the brand, and seek out brand information, promotions, and events. Satisfaction is a measure of the number of people feeling generally 'positive' or 'generally negative' toward a brand. Overall, we postulate the following:

H₆: Increase in the customer's positive and neutral satisfaction measures of the value of a brand at t, will lead to increased brand value at t+1

Recommend. Recommendation, like a referral reflects a positive promotion by individuals, increasing the likelihood of adding new customers or sales, Berg, Matthews and O'Hare (2007). Recommend measures the impact of the ambassadors and the equity destroyers of a brand. Recommend measures if your brand will be recommended for consumption or use, or recommended as a brand to avoid. We hypothesize that an increase in positive and neutral recommendations of the brand will lead to an increase in the brand value.

H₇: Increase in the customers' positive and neutral recommendations at t, increased brand value at t+1

Financial Variables

In contrast to customer-based variables, for financial-based measures researchers collect financial market, accounting, firm level data without contacting consumers; these then identify dollar-metric and financial brand value at the firm or brand level (Yoo, 2000). For our research we use (1) marketing action, (2) brand history and (3) firmographics as a proxy for financial-based measures to predict the brand value at the brand level.

1. MARKETING ACTION

Advertising spend. In the marketing literature there is extant studies on the value creation of a firms advertising spend on brand value (Mizik and Jacobson, 2003; Chaudhuri, 2002; Shah and Akbar, 2008; Peterson and Jeong, 2010; Abdel-khalik, 1975; Hirschey and Weygandt, 1985).

Consistent with the literature, it is our hypothesis that increased advertising spend for a particular brand is reflected in the brands value released the following year.

H₈: Increase in advertising spend by the firm at t, increased brand value at t+1.

2. BRAND HISTORY

Brand Value. Brand loyalties influence brand value (Aaker, 1991). Increased loyalty results in more sales and revenue for the firm. The actions of the consumer become more predicatable resulting in a revenue-stream that can become considerable over time (Gremer, 1999). For this

reason, it is our hypothesis that increases in loyalty reflected in a brand's value in year t will have a spillover effect in a brand's value in year t+1. Thus,

H₉: Increase in brand value at t, increased brand value at t+1

3. FIRMOGRAPHICS

A number of structural variables at the firm level have been used in empirical studies (Proceedings of the International Conference on Information Systems, 1987).

In selecting our variables we indicate net sales (Kumar, Venkatesan, Bohling, and Beckmann, 2008) and the number of employees (Kaen and Baumann, 2003; Evans, 1987) as a measure of growth. We use and market value (Simon and Sullivan, 1993) as a measure of shareholder return.

Sales. Sales reflect the revenue generated by the firm from the sale of a good or service. It is defined in each company's annual report after discounts and allowances. Berg et al., (2007), observed statistically significant correlations between brand value and sales, the healthiest brands having twice the amount of customers increasing spending than the worst-performing brands. These findings allow us to incorporate sales as a key element of brand value. It is our hypothesis that favorable consumer brand value will be reflected in purchase behavior and sales. Therefore,

H₁₀: Increase in firm sales at t, increased brand value at t+1

Market Value. Simon and Sullivan, 1993 outline a technique for estimating a firm's brand value, based on a firms' market value. Market value is the ultimate metric of shareholder value calculated as the value of the firm based on the current price per share multiplied by the number of shares in issue. Total Market Value for Industrials is: Market Capitalization + Preferred

Equity + Short-Term and Long-Term Debt + Other Long-Term Liabilities + Minority Interest – Cash & Equivalents. For Banks and Financials it is calculated at: Market Capitalization + Preferred Equity + Short-Term and Long-Term Debt + Other Long-Term Liabilities + Minority Interest + Total Deposits.

It is our hypothesis, similar to Simon and Sullivan's findings that firms with commonly known brand names have higher estimates of brand value.

H₁₁: Increase in a firm's market value at t, increased brand value at t+1.

Number of Employees. As in previous research by Kaen and Baumann (2003) and Evans (1987), our measure of firm size is the natural log of the number of employees. It is our hypothesis that an increase in the number of employees is indicative of the need to increase production to meet increased demand. Therefore, we expect the following:

H₁₂: Increase in a firm's number of employees at t, increased brand value at t+1

4. CONTROL VARIABLES

Although they are not of primary theoretical interest to our study, we include in our model control variables that have been found in prior research to affect brand outcomes. Beyond whatever substantive interest these control variables possess in their own right, similar to Chaudhuri and Holbrook (2001), their main purpose here is to help remove statistical noise due to omitted-variables bias in a case in which we can capture effects that have been shown elsewhere to make a difference. The control variables include the following macroeconomic factors: gross domestic product, unemployment, and government debt.

A summary of the variables, a detailed description, frequency and characteristic of the data and data source is summarized in Table 2.

Table 2: Variable Description and Source

Dependent Variables	Description	Frequency/ Characteristics	Data Source	Type	Reference
Brand Value	Dollar-based marketing effects or outcomes that accrue to a product or service due to its brand name, as compared with the effects or outcomes that would accrue if the product or service did not have that brand name	Interval-Annual	Brand Finance	\$ Billion	Aaker (1991), Gremer (1999), Keller (2003)
Advertising Spend	Measured media including TV, Magazines, Print, Radio, Outdoor, Internet	Interval-monthly Disaggregate data Brand level	Kantar Media	\$ Million	Mizik (2003), Ho (2005), Chaudhuri (2002), (Shah (2008), Peterson (2010), Abdelkhalik (1975), Hirschey (1985)
Buzz	Measures recent brand sentiment & the direction of recent awareness	Daily	YouGov BrandIndex	Ratio: negative to positive and neutral	Chandon, (2003)
Impression	Strength, favorability, and uniqueness of perceived attributes and benefits of the brand	Daily	YouGov BrandIndex	Ratio: negative to positive and neutral	Leuthesser (1995)
Quality	Whether the brand is perceived as good/poor quality, irrespective of price.	Daily	YouGov Brandindex	Ratio: negative to positive and neutral	Aaker (1991), Kapferer (1991), Kamakura (1991), Martin (1991), Feldwick (1996)
Value	What the brand offers a customer in return for the price paid	Daily	YouGov Brandindex	Ratio: negative to positive and neutral	Lasser, (1995)
Reputation	Measures brand “prestige” via a customers desire to work for a brand	Daily	YouGov Brandindex	Ratio: negative to positive and neutral	Fombrun (1996).
Satisfaction	The extent to which customers purchase and use the brand, talk to others about the brand, and seek out brand information, promotions, and events.	Daily	YouGov Brandindex	Ratio: negative to positive and neutral	Berg (2007)
Recommend	Measures if your brand will be recommended for consumption or use, or recommended as a brand to avoid.	Daily	YouGov Brandindex	Ratio: negative to positive and neutral	Berg (2007)
Sales	Net sales is the revenue generated by the firm from the sale of a good or service	Annual, Lag	Bloomberg and annual financial reports	\$ Million	Berg (2007)
Market Value	Estimation of the value of the firm based on the current price per share multiplied by the number of shares in issue	Annual, Lag	Bloomberg and annual financial reports	\$ Million	Simon and Sullivan (1993)
Number of Employees	Number of full time employees	Annual, Lag	Bloomberg and annual financial reports	Integer > 1	Kaen (2003), Evans (1987)

Chapter 5: METHODOLOGY

Database Overview

To evaluate our conceptual model, we compiled a cross sectional data set of seventy-four brands for five consecutive years from 2007 through 2011. The data set consists of each brands' value, the associated customer search queries for the brand term conducted on Google's search engine, along with the brand's history, the firm's marketing actions, customer attitudes and associated firmographics.

Predictor Variables

The predictor variable is a customer's query for a brand that has been searched for on the Google engine; these queries are stored on Google in what is known as records. Each record stored includes the search query itself, the IP address, the date it was entered, and the type of device it was entered from (mobile phone, desktop computer). These records collectively are referred to as search logs. Most websites store records of visits to their site in a similar way. For more on search logs and privacy see Appendix, Disclosure 1.

To access and analyze search query data, we use two tools, Google Insights found at <http://www.google.com/insights/search/> and Google's internal tool for categorization and counts. Google Insights provides an index of the volume of Google queries by geographic location and category measured against total query volume. Google Insight data are automatically filtered for spam and porn, and the researcher can apply filters by query string, query category and time. The number represents the normalized query share by each filter. For a given query of interest, Google Insights uses a "broad match" algorithm such that any query string containing the query

of interest as a substring is included in the statistics. Google Insights has been used in many studies including capital markets (Vlastakis and Markellos, 2010; Da and Gao, 2010), entertainment (Choi and Varian, 2009; 2010), labor markets (Askatas and Zimmerman; 2009; Suhoy, 2009; D'Amuri and Marcucci, 2010), real estate markets (Wu and Brynjolfsson, 2009; healthcare (Polgreen et al., 2008; Brownstein, Freidfield, and Madoff, 2009; Corley, Mikler, Singh and Cook, 2009; Hulth et al., 2009; Pelat, Turbelin, Bartten, Flahault and Valleron, 2009; Valdivia and Monge-Corella, 2010), economic indices (Ettredge and Karuga, 2005; Schmidt and Vosen, 2009) and travel (Choi and Liu, 2011).

Google's internal tool for categorization and counts is similar to Insights in that it filters for spam and one can apply filters for geography and time. However, Google's internal tool uses an "exact" match algorithm such that any query string containing the exact query will be included in the statistics, queries in a substring will not. Google's internal tool does not index the data and no other filters are automatically applied.

To extract and compile the data we invoke a three stage process involving (1) building query libraries for each brand, (2) quantifying and detrending queries, (3) coding and collating.

Step (1), building query libraries for each brand involves the identification, extraction and storage of brand and associated brand queries. Many queries for a brand are simply the brand itself, like "apple", however many users will input brand association such as apple computer, apple accessories, etc. To develop a library of queries reflecting a brand and its associations, we access Google server logs via Google Insights. The rationale behind using Google Insights is that

it uses the “broadmatch” algorithm that provides us with the query substrings. Researchers typically use Google Insights to obtain the indexed volume of search queries by geographic location and category (Ginsberg et al., 2011; Polgren et al., 2008; Goel et al., 2010; Schmidt and Vosen, 2011; Choi and Lui, 2011 and 2009; McLaren and Shanbhogue, 2011; Askitas and Zimmermann, 2009). For our purposes, we use Insights to gather the query substrings that are displayed as top and rising queries associated with the brand. This process allows us to capture the queries associated with the brand with the most significant level of interest and those that experienced significant growth in a given time period. For example, when you input a specific query into Insights, for a specific period of time, you will see a list of the top 10 rising queries related to that term. The Insights tool determines relativity by examining searches that have been conducted by a large group of users preceding the search term entered, as well as after. For example, if a user inputs "apple," the list of rising searches will include apple, but it may also include a product (apple ipod) or a navigational query (apple location). This process is repeated for each brand, for each year 2007 through 2011, with the associated queries captured in a file called a library.

For purposes of our research we apply filters for the US, to coincide with our dependent longitudinal variables of advertising spend, customer value and brand value, we include all categories and specify the year. If a brand query could mean different things (example, Visa means a credit card association as well as travel and residency documentation) Insights takes a look at broad search patterns among people who search for the credit cards versus those who search for travel and residency documentation. Specifically, people looking for the credit card brand may have also looked for Mastercard in the few searches immediately before and after, while people looking for travel and residency documentation may have conducted a search

related to immigration or customs immediately before or after. In the event the queries significantly cross into other categories (VISA is 10 to 25 percent in the financial category and 25 to 50 percent in law and government) we specify the desired category, in the case of VISA we select financial services. For each brand, the size of the library is $n > 20$.

Step 2 of our process involves obtaining a count of the number of times each query from the newly created brand library was entered into Google as a search. As the brand variables of interest are reported on a calendar basis, we aggregate the search query counts into an annual amount, for a given year. We access Google server logs through the internal tool to obtain the count by reading the number of times the query was entered by desktop, laptop, mobile or tablets, the related year and country. We utilize the internal tool due to the “exact” match algorithm that counts exact queries strings, not substrings which would introduce confounding variables into our process. We repeat this process for each brand and for each year 2007 through 2011. Once we have the counts we normalize the data to 2007 levels, accounting for Internet adoption and Google’s share of core search.

In Step 3 we analyze the emotional content of the queries using SentiStrength, a publically available list of 890 positive and negative sentiment terms (Thelwall, Buckley, Paltoglou, Cai and Kappas, 2010). SentiStrength detects positive and negative sentiment strength in short informal text and assigns an integer from 5 to -5 to reflect sentence-level subjectivity, examples include: dislike = -3, hate=-4, coolest = 3, love = 4. Sentistrength has been used to analyze the emotional content of tweets, blogs and other short text (Bollen, Mao and Zeng, 2009; Paltoglou and Thelwall, in press). For purposes of our analysis, the queries are multiplied by their SentiStrength value and categorized as positive, neutral or negative. This results in a database consisting of 21.2 billion positive and neutral queries, and 32 million negative queries.

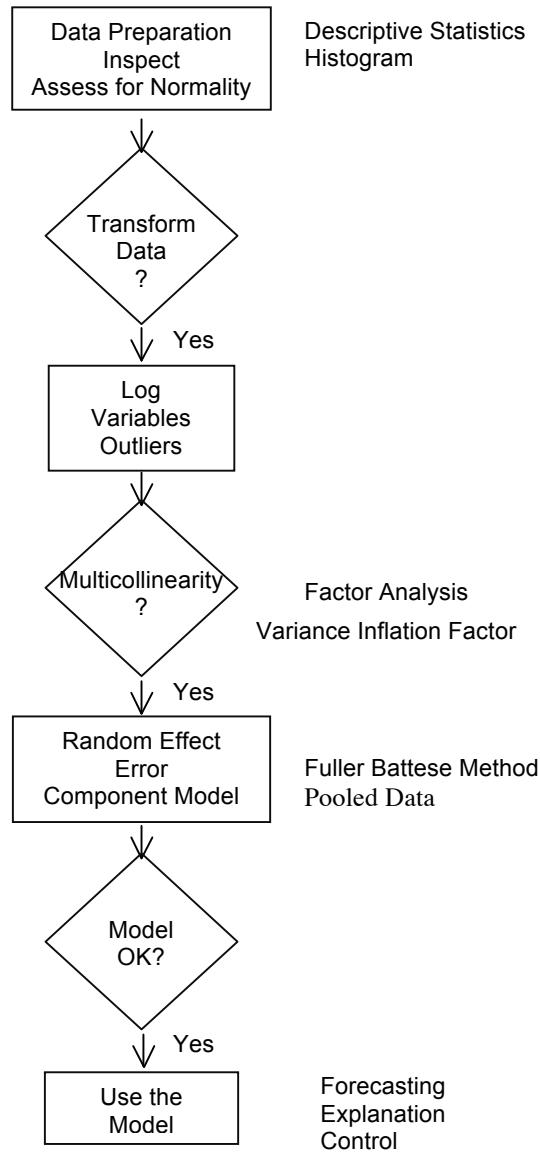
To finalize, we append the values for the remaining predictive variables including brand, customer value and firmographics as well as the control variables.

The methodology described employs qualitative approaches to the large volumes of textual search query data. The amount of data, in the order of billions of queries per day, poses considerable challenges to identifying and manual coding relevant brand queries and counts over the five-year period. In this research we utilize the fully automatic identification and categorization of queries by leveraging the output of Google's machine language algorithms as made available through their Google Insights tool and internal query categorization and access tool. SentiStrength is a semi-automatic coding on textual data (specifically, short text). The use of machine language processing is preferred due to the size of the database, and to eliminate any researcher bias (Crowston, Liu, Allen and Heckman, 2010).

Chapter 6: ANALYSIS

To enable us to assess the proposed hypotheses, we use a modeling framework that accounts for the various sources of errors, outliers, multicollinearity and brand specific effects of search queries. Figure 6 is a flow chart depicting the decision-making steps and selected analysis.

Figure 6: Decision-Making Flow Chart



To begin, standard diagnostic checks were conducted on the data to assess for normalcy.

Descriptive statistics (Table 3) indicate a need to logarithmically transform brand, search and firmographic variables. Outlying search query variable BP is removed, as it is explainable, an anomaly and a negative value that cannot be logged.

Table 3: Descriptive Statistics

	Min	Max	Mean	St. Dev.	Skewness	Kurtosis
GDP	12,703	13,288	13,089	204	-1.14 0.13	-0.20 0.25

Unemployment %	5	10	8	2	-0.69	0.13	-0.87	0.25
Debt % GDP	43	73	59	11	-0.17	0.13	-1.56	0.25
Brand Value \$B	1,424	70,605	13,008	10,785	1.43	0.13	2.43	0.25
Search Query M	-12	1,542	46	115	8.13	0.13	87.46	0.25
Advertising \$M	2	2,900	385	511	2.45	0.13	6.66	0.25
Employees	1,403	2,200,000	156,173	250,222	6.25	0.13	45.91	0.25
Sales \$M	1,605	447,884	59,643	67,957	2.97	0.13	11.17	0.25
Market Value \$M	4,318	2,203,717	261,086	468,745	2.77	0.13	6.83	0.25
Buzz	0.56	0.98	0.88	0.10	-1.38	0.14	1.50	0.28
Impression	0.56	0.98	0.87	0.08	-1.38	0.14	1.87	0.28
Quality	0.61	0.98	0.90	0.07	-1.56	0.14	2.26	0.28
Value	0.52	0.97	0.87	0.08	-1.40	0.14	2.98	0.28
Reputation	0.57	0.98	0.88	0.08	-1.61	0.14	2.76	0.28
Satisfaction	0.73	0.98	0.92	0.05	-1.25	0.14	0.99	0.28
Recommendation	0.61	0.98	0.89	0.07	-1.32	0.14	1.51	0.28

We provide the correlation matrix of the customer value measures in Table 4 to check the presence of collinearity. There seems to be strong linear associations among some variables (corr >0.80) suggesting that multicollinearity could be a problem in the model estimation. The Variance Inflation Factors (VIFs) of the customer value measures also corroborate the possibility of multicollinearity (VIF > 5).

Table 4: Correlation Matrix of Customer Value Measures

	Buzz	Impression	Quality	Value	Reputation	Satisfaction	Recommendation
Buzz	1.00						
Impression	0.63	1.00					
Quality	0.63	0.94	1.00				
Value	0.49	0.85	0.76	1.00			
Reputation	0.57	0.92	0.90	0.81	1.00		
Satisfaction	0.46	0.77	0.76	0.64	0.72	1.00	
Recommendation	0.59	0.95	0.93	0.85	0.90	0.77	1.00

Therefore, in order to control for this multicollinearity we performed a factor analysis on the customer value measures. Kaiser-Meyer-Olin test and Bartlett's test results reported in Table 5 verify the need for factor analysis. Kaiser-Meyer-Olin measure of sampling adequacy was at

.93, which is greater than .5 to precede factor analysis (Hutcheson and Safroniou, 1999).

Bartlett's test of sphericity was $\chi^2 (28) = 3179$ ($p < .001$), rejecting the null hypothesis that the correlation is an identity matrix. This result indicates the appropriateness of factor analysis.

Table 5: Kaiser-Meyer-Olin and Bartlett's Test

Kaiser-Meyer-Olin Measure of Sampling Adequacy		.931
Bartlett's Test of Sphericity	Approx. Chi-Square	3179.061
	df	28
	Sig.	.000

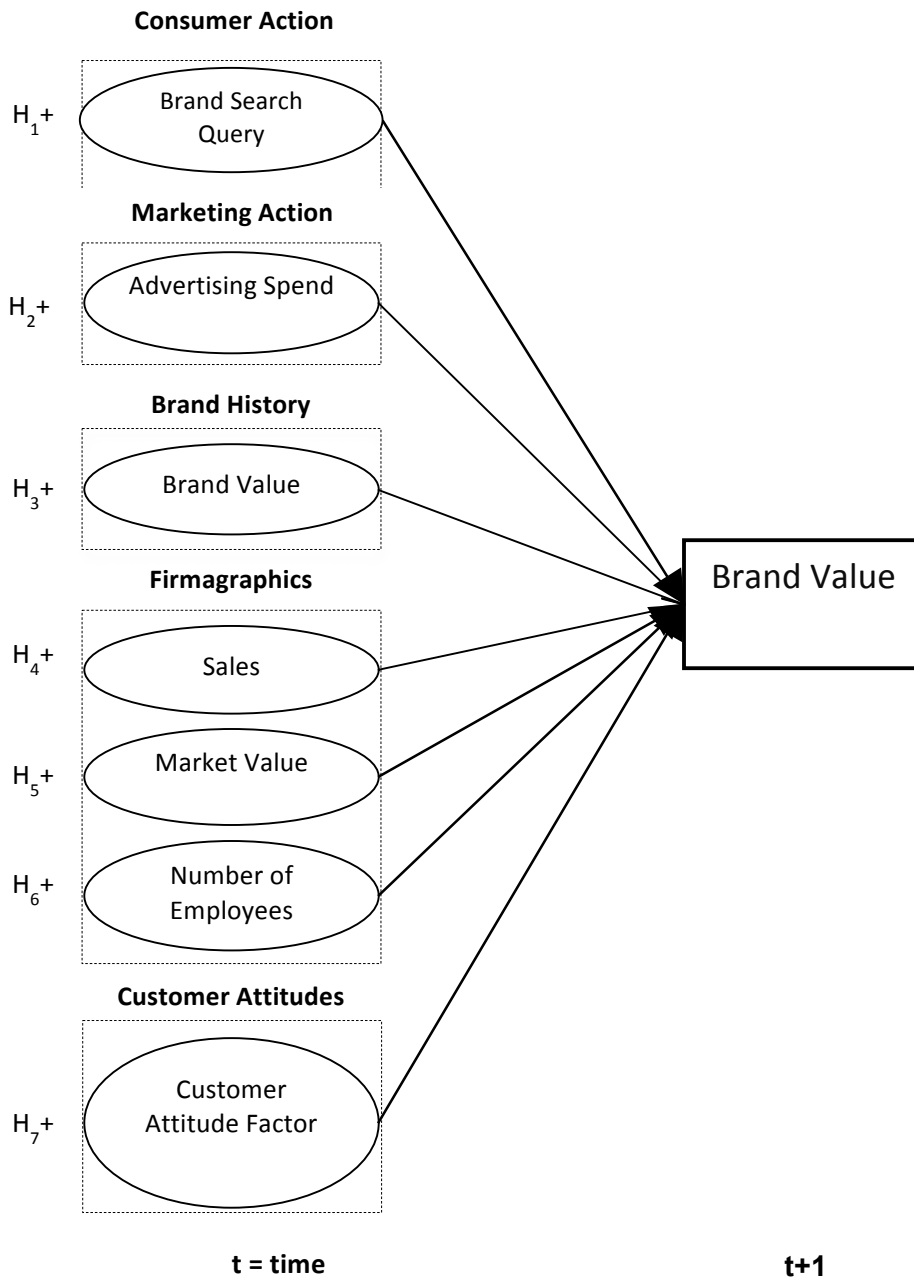
Factor analysis was conducted on the seven measures (see Table 6). One component had an eigenvalue above Kaiser's criterion of 1 and explained 79% of the variance. Given the large sample size, and the Guttman-Kaiser's Rule that the factors with eigenvalue higher than 1 should be retained, we loaded all into one factor named customer's overall attitude using PROC FACTOR in SAS. Factor loadings of all variables exhibit scores higher than .6.

Table 6: Factor Pattern

	Factor
Buzz Ratio	.676
Impression Ratio	.979
Quality Ratio	.951
Value Ratio	.860
Reputation Ratio	.938
Satisfaction Ratio	.825
Recommend Ratio	.966
Eigenvalues	6.329
% of Variance	79.108
Cronbach's Alpha	.946

This resulted in a revised conceptual model shown in Figure 7.

Figure 7: Revised Conceptual Model



The following model was used to investigate the predicative variables for brand value at year t.

$$(1) \quad \text{Log}(BV_t) = \alpha_0 + \alpha_1 \text{Log}(SQ_t) + \alpha_2 \text{Log}(BV_{t-1}) + \alpha_3 MV_t + \alpha_4 SLS_t + \alpha_5 EMP_t + \alpha_6 Debt_t$$

$$(2) \quad \alpha_1 = \beta_0 + \beta_1 ADV_t + \beta_2 \text{Customer Attitude}_t$$

where

SQ_t denotes the value of search queries for brand value in year t,

BV_{t-1} denotes brand value at year t-1,

MV_t denotes the market value for the brands at year t,

SLS_t denotes the sales for the brands at year t,

EMP_t denotes the number of employees for the brands at year t,

$DEBT_t$ denotes the government debt at year t,

ADV_t denotes the advertising spend at year t,

$\text{Customer Attitudes}_t$ denotes the customer attitudes at year t.

By substituting Equation 2 into Equation 1, we estimate the interaction effects of search queries-advertising and search queries-customer attitude as well as the main effects of search queries, lagged brand value, and firmographics.

$$(3) \quad \text{Log}(BV_t) = \alpha_0 + \beta_0 \text{Log}(SQ_t) + \beta_1 ADV_t * \text{Log}(SQ_t) + \beta_2 \text{Customer Attitude}_t * \text{Log}(SQ_t) + \alpha_2 (BV_{t-1}) + \alpha_3 MV_t + \alpha_4 SLS_t + \alpha_5 EMP_t + \alpha_6 Debt_t + \text{error}$$

Equation 1 can be estimated by using ordinary least squares (OLS) regression at the brand level.

The data encompasses four time periods for 74 cross-sections. Estimating a random effects model incorporating heterogeneity may lead to inefficient estimates of the coefficients because of the number of parameters estimated in relation to the sample size.

Given the cross sectional and time series nature of the data and the lack of an *a priori* theoretical reason for the regression coefficients to differ across brands, we pool the data across 74 brands which reduces the number of parameters to 9 by making use of 295 observations. Moreover, pooling provides not only more degrees of freedom, but also a defense against misspecification bias caused by using only time-series or only cross-sectional data (Brobst and Gates, 1977).

According to Kumar and Leone (1988), most similar studies that try to explain variance have analyzed cross-sectional data by stacking the cross-sections and estimating the parameters using OLS regression. However, pooled models can be estimated by procedures that allow both cross sectional and time series variation in the data through the specifications of the error structure (Fuller and Battese 1974, Maddala 1971, Nerlove 1971). The Fuller-Battese procedure was used to estimate the pooled models, (Kumar and Leone, 1988). Variance estimates for each of the cross-sectional, time-specific, and random error components are shown in Table 7.

Table 7: Fuller and Battese Variance Components

Model Description	
Estimation Method	Fuller
Number of Cross Sections	74
Time Series Length	4
Variance Component Estimates	
Variance Component for Cross Sections	0.040119
Variance Component for Time Series	0.003352
Variance Component for Error	0.009805

The benefits of the Fuller-Battese procedure according to Balasubramanian and Kumar (1990) are as follows. First, using cross sectional, time series data allows generalizability of results over time. In contrast, cross-sectional data provides "snapshot" picture specific to a given time period and inferences drawn from such data could be biased by idiosyncrasies associated with that time period. Second, brand value does change (1) across time for a firm (Farris 1979) and (2) across time for an industry (West 1985). Analysis of these two important components of variation can be accomplished by using cross sectional, time series data. For the above reasons, like Brobst and Gates (1977), we estimate the model hierarchically using cross-sectional and time series data and present the results in Table 8.

Consistent with our conceptual framework, Table 8 shows that increases in our customer-based variables are positively associated with increases in brand value. Of primary interest, search queries are positively associated with increases in brand value ($\beta_0=.1874$, $p<.0001$). The first interaction effect reported in Table 8 is between search queries and advertising spend, where a positive association is evident with brand value ($\beta_1=.000015$, $p<.0001$). Next we observe that the

interaction between search queries and customer attitudes are positively associated with brand value ($\beta_2=0.0084$, $p<.0001$).

In addition to the customer-based variables and the effects of search queries and related interaction effects, we observed that the financial variables, market value, sales and number of employees are positively associated with increases in brand value. Also, prior brand values are positively associated with future brand value ($\alpha_2=0.5127$, $p<.0001$). The results also suggest that debt is marginally associated with increased brand value which is contrary to our hypothesis, implying that the increased government debt along with an aggressive economic stimulus was positively associated with brand value ($\alpha_6=.0063$, p-value .0079).

Table 8: Impact of Financial and Customer Based Variables on Brand Value

	Parameters	t-value	p-value		
Intercept	0.0347			Number of Observations	295
log_Search Queries	0.1874	6.68	<.0001	R-square	0.9966
log_Search Queries x Advertising	0.000015	3.12	0.002	Adjusted R-square	0.9965
log_Search Queries x customer attitudes	0.0084	4.37	<.0001		
log_lag_Brand Value	0.5127	13.74	<.0001		
Market Value	0.00000004	4.6	<.0001		
Sales	0.00000101	3.84	0.0002		
Employees	0.0000004	5.48	<.0001		
Debt	0.0063	2.67	0.0079		

Overall, the variance explained by the search query and its interaction is 58.9%, which is quite substantive and significant to our hypotheses. The remaining variance of 40.8% is explained by the lag brand value, and the control variables. See table 9 for a summary of the outcome.

Table 9: Summary of Analysis

	Statement	Results
H ₁	Increase in <i>brand search queries</i> at t, are associated with <i>increased</i> BV at t+1	supported
H ₂	Increase in <i>advertising spend</i> at t, will lead to <i>increased</i> BV at t+1	supported
H ₃	Increase in <i>brand value</i> at t, will lead to <i>increased</i> BV at t+1	supported
H ₄	Increase in <i>firm sales</i> at t, will lead to <i>increased</i> BV at t+1	supported
H ₅	Increase in <i>a firm's market value</i> at t, will lead to <i>increased</i> BV at t+1	supported
H ₆	Increase in <i>a firm's employees</i> at t, will lead to <i>increased</i> BV at t+1	supported
H ₇	Increase in <i>customer's positive or neutral attitude factor</i> at t, will lead to <i>increased</i> BV at t+1	supported

Chapter 7: CONTRIBUTION, LIMITATIONS AND FUTURE RESEARCH

Contribution

Based on our analyses of cross-sectional, time series data spanning several industries, and past studies providing broad based insights into the role of key independent variables in determining brand value, we determine that search queries for a brand term is a forward-looking metric that explains brand value. For the practitioner, the implication is great in that they no longer need to wait for brand valuations and rankings to be made available after a lag period, from the time the data is collected, compiled, and analyzed to the time that it is delivered. With search queries for brand terms, this information is available almost real time so that it can be used to inform marketing plans in terms of forward-looking brand strategies and promotions/tactics to increase customer value and grow overall profitability.

For the industry, since a relationship exists between brand search query data and brand value, then search query data can be used to increase the accuracy of brand valuation tools by adding a robust customer measure to current methodologies. It also can be used to predict subsequent releases, minimizing the lag effects, so brand managers can have early indicators as to how their brands are tracking, and brand values are more timely and actionable.

This study contributes to the nascent literature using search queries to predict real world events and behavior in that it now shows the ability to predict perception, specifically in terms of a consumer's real time perception of a brand. This research extends the domain of past studies from health & macroeconomics to the social sciences specifically, marketing.

Limitations and Future Research Direction

Although the research offers several important implications, some limitations provide the basis for future research. First, as we are working with firm level data, our sample is restricted to larger firms. This was primarily a result of our focus on metrics of financial, customer and brand value as dependent variables, for which longitudinal measures are available only for relative large firms (Krasnikov, 2009). Further research might include smaller firms with newly established brands to investigate whether branded search queries are a valuable early indicator of brand value or if they differ from what we observed in our sample.

Second, our sample data for customer value and advertising is limited to the US. Like firm value, customer value and advertising measures are available only for relative large firms, and within

the US. Further research might include global firms once that information is made available either by the firms self-reporting, or the globalization of the data gathering tools.

Third, branded search queries as our independent variable are derived from the process of developing keyword libraries of the brand and brand associations. Although extensive rigor was established to rule out or minimize endogeneity complete randomization is not possible and causal interpretation could be confounded. To minimize endogeneity, Difference estimation should be used to rule out lagged dependent variables.

Fourth, extend research by introducing moderator variables like firms with single brands versus multiple brands (i.e. Proctor and Gamble), brand duration, creation of new age versus old age brands, or competitive intensity. Last, from a practitioner perspective, Identifying turning points in a customer's perception of brand value will allow brand manager's to implement remediation efforts quickly in terms of changing media mix, pricing, messaging or distribution.

Appendix

Disclosure 1

Search Engine - Definitions and Privacy

Google Privacy Center, Your Data on Google

<http://www.google.com/goodtoknow/data-on-google/search-logs/>

Data Retained

IP address:

123.45.67.89 is the IP address assigned to the user's computer by his or her service provider. Just like other websites, when you ask Google for a page (a search results page, for example), Google uses your computer's IP address to ensure that Google gets the right results back to the right computer. It's important to remember that IP addresses don't say exactly where an individual user is, or who they are. In fact, some Internet Service Providers (ISPs) give users a different IP address every time they log onto the web. The most Google can tell about a user from his computer's IP address is that user's general location (for example, Boston) and possibly the ISP they use to connect to the Internet. Only the ISP (who actually controls the user's account) can match an individual with an IP address.

Time and date:

25/Aug/2011 10:15:32 is the date and time the user typed the query into Google.

Search query:

<http://www.google.com/search?q=cars> is the search query, in this case "cars."

Browsers and operating systems:

Chrome 2.0.0.7; Windows NT 5.1 is the browser and operating system being used.

Cookie:

740674ce2123a969 is the unique cookie ID assigned to a browser the first time a user visits Google. Like an IP address, a cookie doesn't tell Google who a user actually is or where they live, it only identifies a computer. You can delete these cookies at any time.

Time limits on data retention

Google anonymizes IP addresses after 9 months and alter the cookie numbers in our logs permanently after 18 months. This breaks the link between the search query and the computer it was entered from and is similar to the way in which credit card receipts replace digits with hash marks to improve customer security. Here is what an IP address could look like in our logs after 9 months: 123.45.67.XXX. After 18 months, the cookie will be replaced by a newly-generated cookie number.

Google was the first major search engine to announce time limits on the retention of logs data, and are pleased that others in the industry have followed their lead. Online cookies don't last forever, Google cookies expire after two years. Additionally, Google has always allowed people to use its services without cookies (though this may mean losing the use of some features or functions of particular products).

Disclosure 2

Views presented in the publication are those of the researchers, not Google or other data providers.

Bibliography

- Aaker D. (1991). *Managing Brand Equity*. New York: The Free Press, 1991.
- Abdel-khalik, A.R. (1975). Advertising effectiveness and accounting policy, *The Accounting Review*, 50:657-670.
- Assael, H. (1998). *Consumer Behavior and Marketing Actions*, 6th ed., South-Western College Publishing, Mason, OH.
- Anderson, E. W., Fornell, C., and Mazvancheryl, S. (2004). Customer Satisfaction and Shareholder Value,”*Journal of Marketing*, 68 (October), 172–85.
- Anderson, J. (2008). Lessons from Krispy Kreme. *Journal of Business Case Studies*, April, 2008. Volume 4, Number 4. Article Stable URL: <http://www.jstor.org/stable/30162024>.
- Askatas, N., and Zimmermann, K.F. (2009). Google econometrics and unemployment forecasting. IZA Discussion Papers 4201, Institute for the Study of Labor (IZA). Athens University.
- Ayers, J., Ribisl, K., and Brownstein, J. (2011). Tracking the Rise in Popularity of Electronic Nicotine Delivery Systems (Electronic Cigarettes) Using Search Query Surveillance, *American Journal of Preventive Medicine*, Volume 40, Issue 4, April 2011, Pages 448-453.
- Bakos, Y. (1997). Reducing Buyer Search Costs: Implications for Electronic Marketplaces, *Management Science*, Volume 43, Number 12, December.
- Balasubramania, S., and Kumar, V. (1990). Analyzing Variations in Advertising and Promotional Expenditures: Key Correlates in Consumer, Industrial, and Service Markets. *Journal of Marketing* Vol 54, (April 1990), 57 – 68.
- Banyte, J., Joksaite, E., and Virvilaite, R. (2007). Relationship of Consumer Attitude and Brand: Emotional Aspect. *Engineering Economics*. 2007. No 2 (52).
- Beatty, S.E., and Kahle, L.R. (1988). “Alternative Hierarchies of the Attitude-Behavior Relationship: The Impact of Brand Commitment and Habit.” *Journal of Academic Marketing Science* 16 (Summer): 1-10.
- Berg, J. D., Matthews, J. M., & O'Hare, C. M. (2007). Measuring Brand Health to Improve Top-Line Growth. (cover story). *MIT Sloan Management Review*, 49(1): 61-68.

Bettman, J.E., Johnson, and Payne, J. (1991). "Consumer Decision Making." In Handbook of Consumer Behavior. Eds. T. S. Robertson and H. H. Kassarian. Englewood Cliffs, NJ: Prentice Hall, 50-84.

Bloch, P., Sherrell, D. and Ridgeway, N. (1986). Consumer Search: An Extended Framework. *Journal of Consumer Research* 13(1), 119-126.

Bollen, J., Mao, H., and Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, In press, 2011.

Boulding, W.I, and Staelin, R. (1995). Identifying generalizable effects of strategic actions on firm performance: the case of demand-side returns to R&D spending. *Marketing Science*, 14(3), G222-G236.

Brandfinance. BrandFinance 250 and 500: The annual report on the world's most valuable brands, 2007, 2008, 2009, 2010, 2011.

Brandfinance (2011), League Table Brand Valuation Method, How We Measure Brand Value. [http://www.adformatic.nl/images/uploads/League_Table_Methodology_\(Short\).pdf](http://www.adformatic.nl/images/uploads/League_Table_Methodology_(Short).pdf).

BrandIndex. The User Guide: April 2011.

Brobst, R., and Gates, R. (1977). "Comments on Pooling Issues and Methods in Regression Analysis," *Journal of Marketing Research* , Vol. 14, No. 4 (Nov., 1977), pp. 598- 600 Published by: American Marketing Association.

Brownstein, J., Freifeld, C., and Madoff, L. (2009). "Digital Disease Detection? Harnessing the Web for Public Health Surveillance" *The New England Journal of Medicine*.

Brynjolfsson, E., and Kahin, B. (2000). *Understanding the Digital Economy*, MIT Press, 2000.

Chandon, J. L., Chtourou, M., and Fortin, D.R. (2003). Examining the Effects of Advertising Configuration and Exposure Levels on Behavioral Responses to Web-based Advertisements. *Journal of Advertising Research*, 43(2), 1-13.

Chaudhuri, A., and Holbrook, M. (2001). The Chain of Effects from Brand Trust and Brand Affect to Brand Performance: The Role of Brand Loyalty. *Journal of Marketing* Vol. 65 (April 2001), 81-93.

Chaudhuri, A. (2002). How brand reputation affects the advertising-brand equity link. *Journal of Advertising Research* 42 (3), 33-43.

Choi, H. & Shideler, M. (2010) "Predicting Game Sales with Google Trends" Working Paper.

Choi, H., & Slotwiner, D. (2009). Predicting TV Rating with Google Trends, Google Working Paper, Google, 2009.

Choi, H., & Varian, H. (2009). "Predicting initial claims for unemployment benefits" Google Technical Report, <http://research.google.com/archive/papers/initialclaimsUS.pdf>.

Choi, H., & Varian, H. (2009). "Predicting the present with google trends" Google Technical Report, http://google.com/googleblogs/pdfs/google_predicting_the_present.pdf.

Choi, H., and Liu, P. (2011). Reading Tea Leaves in the Tourism Industry: A Case Study in the Gulf Oil Spill (March 24, 2011). Available at SSRN: <http://ssrn.com/abstract=1893078>.

Christodoulides, G., de Chernatony, L., Furrer, O. and Abimbola, T. (2006). Conceptualizing and measuring the equity of online brands. *Journal of Marketing Management*, 22, 7/8, pp. 799–825.

Collins, J., Dixon, D., Eadie, W., Reggimenti, M., Shiffman, D., Soukhareva, J., Vogel, J., and Ware, B. (2010). Commentary - How More Precise Magazine Inputs Can Improve Media Mix Modeling, The Impact of More Balanced Metrics on ROI, March 2010, *JOURNAL OF ADVERTISING RESEARCH*.

Comscore, June 2011. Global leader in measuring the digital world and preferred source of digital marketing intelligence. Search query data for June 2011. Press release retrieved from http://www.comscore.com/Press_Events/Press_Releases/2011/7/comScore_Releases_June_2011_U.S._Search_Engine_Rankings.

Constant, A., and Zimmermann K. (2008). Face to Face with the Financial Crisis: The U.S. Presidential Election from a Transnational Perspective. Im Angesicht der Krise: US Präsidentschaftswahlen in transnationaler Sicht, <http://ideas.repec.org/a/diw/diwwrp/wr4-16.html>.

Cooper, C.P., Mallon, K.P., Leadbetter, S., Pollack, L.A., and Peipins, L.A. (2005). Cancer Internet search activity on a major search engine, United States 2001-2003. *J Med Internet Res* 2005;7:e36.

Corley, C., Mikler, A., Singh, K., and Cook, D. (2009). "Monitoring Influenza Trends through Mining Social Media" *Proceedings of the 2009 International*.

Cox, D. (1967). Risk taking and information handling In consumer behavior (pp. 34-81). Graduate School of Business Administration, Harvard University, Boston.

Crowston, K., Liu, X., Allen, E., and Heckman, R. (2010). Machine Learning and Rule-Based Automated Coding of Qualitative Data, ASIST 2010, October 22–27, 2010, Pittsburgh, PA, USA.

D'Amuri, F. (2009). Predicting unemployment in short samples with Internet job search query data. MPRAPaper 18403, University Library of Munich, Germany15.

- D'Amuri, F. and Marcucci, J. (2010). "Google It! Forecasting the US Unemployment Rate with A Google Job Search Index", Feem Working Paper No. 31-2010.
- Da, Z., Engelberg, J., and Gao, P. (2010). "In Search of Attention" AFA 2010 Atlanta Meetings Paper.
- Davis, J., *Competitive Success: How Branding Adds Value*, Publisher: Wiley, 2010. Hoboken, NH, 2010.
- De Chernatony, L., Harris, F.J., and Christodoulides, G., (2004). Developing a brand performance measure for financial services brands. *Services Industries Journal*, 24, 2, pp. 15–33.
- Erickson, G., and Jacobson, R. (1992). Gaining comparative advantage through discretionary expenditures: The return to R&D and advertising. *Management Sci.* 38(September) 1264–1279.
- Ettredge, M., Gerdes, J., and Karuga, G. (2005). Using Web-based search data to predict macroeconomic statistics. *Communications of the ACM* 48:87–92.
- Evans, D. S. (1987). "The Relationship between Firm Size, Growth and Age: Estimates for 100 Manufacturing Industries", *Journal of Industrial Economics*, 35, 567-581.
- Eysenbach, G. (2006). Infodemiology: Tracking flu-related searches on the Web for syndromic surveillance. *AMI: Symposium Proceedings* 244–248.
- Farquhar, P., (1989). Managing brand equity. *Journal of Advertising Research*, 30, 4, pp. RC7–RC12.
- Farris P. W., and Buzzell R. D. (1979). Why Advertising and Promotional Costs Vary: Some Cross-Sectional Analyses. *Journal of Marketing* , Vol. 43, No. 4 (Autumn, 1979), pp. 112-122.
- Feldwick, P. (1996). What is brand equity anyway and how do you measure it. *Journal of the Marketing Research Society*. 38: 85-104.
- Fombrun, C. J. (1996). *Reputation: Realizing value from the corporate image*. Boston: Harvard Business School Press.
- Fornell, C., Mithas, S., Morgeson, F., and Krishnan. M.S. (2006). "Customer Satisfaction and Stock Prices: High Returns, Low Risk," *Journal of Marketing*, 70 (January), 3–14.
- Frampton, J., All Brands are Not Created Equal, *Best Global Brands 2007*, Interbrand, 34–38.
- Fuller, Wayne A. and George E. Battese (1974), "Estimation of Linear Models With Crossed-Error Structure," *Journal of Econometrics*, 2 (May), 67-78.
- Ginsberg, J., Mohebbi, M.H., Patel, R. S., Brammer, L., Smolinski, M.S., Brilliant, L. (2009). Detecting Influenza Epidemics using Search Engine Query Data, *Nature* 457, 1012 – 1014.

Goel, S., Hofman, J., Lahaie S., Pennock D., and Watts, D. (2010). Predicting Consumer Behavior with Web Search PNAS 2010 107 (41) 17486-17490; published ahead of print September 27, 2010, doi:10.1073/pnas.1005962107.

Google Privacy Center, Your Data on Google, Good to Know:
<http://www.google.com/goodtoknow/data-on-google/search-logs/>.

Gremer, D. D., and Brown, S.W. (1999). "The Loyalty Ripple Effect: Appreciating the Full Value of Customers", *International Journal of Service Industry Management*, 10, 3, 271-293.

Gupta, S., and Zeithaml, V. (2006). "Customer Metrics and their Impact on Financial Performance," *Marketing Science*, 25(6), 687-717.

Haigh, D. (1999). *Understanding the Financial Value of Brands*. Brussels: European Association of Advertising Agencies.

Hirschey, M. and Weygandt, J.J. (1985). 'Amortization policy for advertising and research and development expenditures', *Journal of Accounting Research*, 23: 326-35.

Ho, Y. K., Keh, T., & Ong, J. M. (2005). The effects of R&D and advertising on firm value: an examination of manufacturing and nonmanufacturing firms. *IEEE Transactions on Engineering Management*, 52, 3-14.

Hölscher, C. (2000). Web search behavior of Internet experts and newbies. *Computer networks*, 33(1-6): 337.

Hull, J. (2008). How marketers should use brand valuation. *Market Leader*(42): 51-55.

Hulth, A., Rydevik, G., Linde, A. (2009). Web Queries as a Source for Syndromic Surveillance. *PLoS ONE* 4(2): e4378. doi:10.1371/journal.pone.0004378.

Hutcheson, G. D., and Sofroniou, N. (1999). *The Multivariate Social Scientist: an introduction to generalized linear models*. Sage Publications.

International Organization for Standardization, (2010). "How Strong is Your Company's Brand? An ISO Standard Can Help Tell You, Press Release, Retrieved from <http://www.iso.org/iso/pressrelease.htm?refid=Ref1364>. October 20, 2010.

Jacoby, J. and Chestnut, R. (1978). *Brand Loyalty: Measurement and Management*, John Wiley & Sons, New York, NY.

Johnson, E., Lohse, J., and Mandel M. (1997). *Designing Marketplaces of the Artificial: Four Approaches to Understanding Consumer Behavior in Electronic Environments*. Working Paper. Wharton School, University of Pennsylvania.

Johnson, H. A., Wagner, M. M., Hogan, W. R., Chapman, W., Olszewski, R. T., Dowling, J. and Barnas, G. (2004). "Analysis of Web access logs for surveillance of influenza", *Medinfo*, 11, Pt 2: 1202-6.

Joo, M., Kenneth W., and Yi Z. (2010). *Television Advertising and Online Search*. Available at SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1720713.

Kaen, F.R., and Baumann, H.D. (2003). "Firm Size, Employees and Profitability in U.S. Manufacturing Industries," *Working Paper*, 2003.

Kamakura, A. W., and Russell G. J. (1993). Measuring brand value with scanner data. *International Journal of Research in Marketing* 10(March): 9-22.

Kapferer, J.N. (1992). *Strategic Brand Management*. New York and London: Kogan Page.

Keller, K. L. (2008). *Strategic Brand Management*. Pearson Press.

Keller, K. L. (2009). Editorial: Brand research imperatives. *Journal of Brand Management*, 17(9).

Keller, K. L., & Lehmann, D. R. (2009). Assessing long-term brand potential. *Journal of Brand Management*, 17(1): 6-17.

Keller, K. L., (1993). Conceptualizing, Measuring, and Managing Customer-Based Brand Equity. *Journal of Marketing*: Jan, 1993; vol. 57, 1; ABI/INFORM Global. Page 1 – 22.

Keller, K., and Lehmann, D., (2006). Brands and Branding: Research Findings and Future Priorities. *Marketing Science* 25, 6 (November 2006), 740-759. DOI=10.1287/mksc.1050.0153 <http://dx.doi.org/10.1287/mksc.1050.0153>.

Kissan, J., Wintoki, M., and Zhang, Z. (2011). Forecasting Abnormal Stock Returns and Trading Volume Using Investor Sentiment: Evidence from Online Search. Submitted to *International Journal of Forecasting*, January 23, 2011.

Kmenta, J. (1986). *Elements of Econometrics*, 2nd ed. New York: Macmillan Publishing Company.

Knowles, J. (2008). Varying Perspectives on Brand Equity. (cover story). *Marketing Management*, 17(4): 20-26.

Kotler, P. (1997). *Marketing Management*, 9th Ed., Prentice Hall, New Jersey.

Krasnikov, A., Mishra, S., Orozco, D. (2009). Evaluating the Financial Impact of Branding Using Trademarks: A Framework and Empirical Evidence, *American Marketing Association*, Vol. 73, *Journal of Marketing*, Vol. 73, November 2009, p. 154–166.

- Kumar, V. and Leone, R.P. (1988). "Measuring the Effect of Retail Store Promotions on Brand and Store Substitution," *Journal of Marketing Research*, 25 (May), 178- 85.
- Kumar, V., Venkatesan, R., Bohling, T., Beckmann, D. (2008). Practice Prize Report—The power of CLV: Managing customer lifetime value at IBM. *Marketing Sci.* 27(4) 585–599.
- Kumar, V., Jones, E., Venkatesan, R., and Leone, R. (2011). Is Market Orientation a Source of Sustainable Competitive Advantage or Simply the Cost of Competing? *Journal of Marketing* , 75, 16-30.
- Lassar, W., Mittal, B., Sharma, A. (1995). Measuring customer-based brand equity. *J Consum Mark* 1995;12(4):11 – 9.
- Lehmann, D. R., Keller, K. L., & Farley, J. U. (2008). The Structure of Survey-Based Brand Metrics. *Journal of International Marketing*, 16(4): 29-56.
- Leuthesser, L., Kohli, C.S. and Harich, K.R. (1995). "Brand equity: the halo effect measure", *European Journal of Marketing*, Vol. 29 Iss: 4, pp.57 – 66.
- Maddala, G. S. (1971). "The Use of Variance Components Models in Pooling Cross Section and Time Series Data," *Econometrica*, 39 (March), 341-58.
- Martin, G.S., and Brown, T. J. (1991). In Search of Brand Equity: The Conceptualization and Operationalization of the Brand Impression Construct. Winter Educator's Conference Proceedings, 1991, Terry L. Childers et al. eds. Chicago: American Marketing Association: 431-438.
- Matsuno, K. (1997). Note on Consumer Decision Making Process, Consumers as Problem Solvers, Babson College, 1997.
- McLaren, N., and Shanbhogue, R., (2011). Using Internet Search Data as Economic Indicators (June 13, 2011). Bank of England Quarterly Bulletin No. 2011 Q2.
- Mela, C.F., Gupta, S and Lehmann, D.R. (1997). "The Long-Term Impact of Promotion and Advertising on Consumer Brand Choice," *Journal of Marketing Research*, 34 (May), 248-61.
- Mizik, N., & Jacobson, R. (2003). Trading off between value creation and value appropriation: the financial implications of shifts in strategic emphasis. *Journal of Marketing*, 67, 63–76.
- Nerlove, M. (1971). "Further Evidence on the Estimation of Dynamic Economic Relations from a Time Series of Cross Sections," *Econometrica*, 39 (March), 359-82.
- Paltoglou, G., & Thelwall, M. (in press). Twitter, MySpace, Digg: Unsupervised sentiment analysis in social media. *ACM Transactions on Intelligent Systems and Technology*.

Pappu, R., Quester, P.G. and Cooksey, R.W. (2005). Consumer-based brand equity: improving the measurement – empirical evidence. *Journal of Product and Brand Management*, 14, 3, pp. 143–154.

Pelat, C., Turbelin, C., Bar-Hen, A., Flahault, A., and Valleron A. (2009). “More diseases tracked by using Google Trends.” *Emerging Infectious Diseases*.

Pentland, A. (2008). *Honest Signals, How They Shape the World*, MIT Press 2008.

Peterson, R., and Jeong, J. (2010). Exploring the impact of advertising and R&D expenditures on corporate brand value and firm-level financial performance. *Journal of the Academy of Marketing Science*, 38(6), 677–690.

Polgreen, P.M., Chen, Y., Pennock, D.M., and Nelson, F.D. (2008). Using internet searches for influenza surveillance. *Clin Infect Dis* 47: 1443–1448.

Proceedings of the International Conference on Information Systems (1987). Society for Management Information Systems (U.S.), Research and Education in Information Systems, Association for Computing Machinery, Institute of Management Sciences. College of Information Systems. International Conference on Information Systems, 1987 Volume 8.

Rangaswamy, A., Giles L., and Seres S. (2009). A Strategic Perspective on Search Engines: Thought Candies for Practitioners and Researchers. *Journal of Interactive Marketing*, Volume 23, Issue 1, February 2009, Pages 49-60.

Reibstein, D. (2002). What attracts customers to online stores, and what keeps them coming back? *Journal of the Academy of Marketing Science*, Volume 30, Issue 4, September 1, 2002, Pages 465 – 473.

Salinas, G., & Ambler, T. (2009). A taxonomy of brand valuation practice: Methodologies and purposes. *Journal of Brand Management*, 17(1): 39-61.

Schiffman, Leon G. and Leslie Lazar Kanuk (1997). *Consumer Behavior*, Upper Saddle River, New Jersey: Prentice Hall.

Schiffman, L.G. and Kanuk, L.L. (2004). *Consumer Behavior*, Englewood Cliffs, N.J.: Prentice Hall.

Schmidt, T., and Vosen, S. (2009). “Forecasting Private Consumption: Survey-based Indicators vs. Google Trends” *Ruhr Economic Papers Series*, No 0155.

Shah, S. Z. A., and Akbar, S. (2008). Value relevance of advertising expenditure: a review of the literature. *International Journal of Management Reviews*, 10(4), 301–325.

Simon, C. J., and Sullivan, M.W. (1993). “The Measurement and Determinants of Brand Equity: A Financial Approach.” *Marketing Science* 12, 28–52.

- Solomon, M.R. (1996). *Consumer Behavior: Buying, Having, and Being*, Englewood Cliffs, New Jersey: Prentice Hall.
- Solomon, M., Bamossy, G. and Askegaard, S. (2002). *Consumer Behaviour: a European Perspective* 2nd edition.
- Srinivasan, S., and Hanssens, D. (2009). *Marketing and Firm Value: Metrics, Methods, Findings, and Future Directions*, 2009, American Marketing Association.
- Suhoy, T. (2009). "Query Indices and a 2008 Downturn: Israeli Data" Bank of Israel Discussion Paper, No. 2009.06.
- Surowlecki, J. (2004). *The Decline of Brands*, Wired Magazine, November 2004.
- Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., and Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12), 2544-2558. <http://sentistrength.wlv.ac.uk/>.
- Valdivia, A. & Monge-Corella, S. (2010). "Diseases tracked by using Google Trends, Spain" *Emerging Infectious Diseases*.
- Varian, H., and Choi, H. (2009). *Predicting the Present with Google Trends*. 1 - 20.
- Vázquez, R., Del Rio, A. and Iglesias, V. (2002). Consumer-based brand equity: development and validation of a measurement instrument. *Journal of Marketing Management*, 18, 1/2, pp. 27–48.
- Vlastakis, N., and Markellos, R. (2010). "Information Demand and Stock Market Volatility" Athens University.
- Vosen, S., and Torsten, S.A. (2011). Monthly consumption indicator for Germany based on Internet search query data, *Applied Economic Letters*, Vol 19 Iss 7, 2011.
- West, D.C. (1985). "Advertising Expenditures by Industries - A Long-Run Perspective," *International Journal of Advertising*, 4, 327-39.
- Wu, L., and Brynjolfsson, E. (2009). "The Future of Prediction: How Google Searches Foreshadow Housing Prices and Quantities" (2009). ICIS 2009 Proceedings. Paper 147. <http://aisel.aisnet.org/icis2009/147>.
- Yoo, B., Donthu, N. and Lee, S. (2000) An examination of selected marketing mix elements and brand equity. *Journal of the Academy of Marketing Science*, 28, 2, pp. 195–211.

Maureen Schumacher Cole leads Google's media sales and account management team in the Southeast providing Fortune 100 Financial Service customers with innovative marketing solutions as well as centralized sales and service.

Formerly a GE marketing veteran, Maureen brings 20-years of experience including strategy, planning, and advertising. She has worked in the US and abroad to expand GE's financial products globally. Prior to her career with GE, Maureen worked for Shell Oil Company and SunTrust.

Maureen is an active member of Georgia State University's CMO Roundtable, and is member of Google Grants, a program that supports non-profit organizations with in-kind advertising.

Maureen is a Pennsylvania State University graduate; she holds a BA and MBA and studied management and economics in Lima, Peru. Currently she is pursuing her Executive Doctorate degree from Georgia State University.