Hyoryung Nam & P.K. Kannan

The Informational Value of Social Tagging Networks

Social tagging is a new way to share and categorize online content that enables users to express their thoughts, perceptions, and feelings with respect to diverse concepts. In social tagging, content is connected through user-generated keywords—“tags”—and is readily searchable through these tags. The rich associative information that social tagging provides marketers new opportunities to infer brand associative networks. This article investigates how the information contained in social tags can act as a proxy measure for brand performance and can predict the financial valuation of a firm. Using data collected from a social tagging and bookmarking website, Delicious, the authors examine social tagging data for 44 firms across 14 markets. After controlling for accounting metrics, media citations, and other user-generated content, they find that social tag–based brand management metrics capturing brand familiarity, favorability of associations, and competitive overlaps of brand associations can explain unanticipated stock returns. In addition, they find that in managing brand equity, it is more important for strong brands to enhance category dominance, whereas it is more critical for weak brands to enhance connectedness. These findings suggest a new way for practitioners to track, measure, and manage intangible brand equity; proactively improve brand performance; and influence a firm’s financial performance.

Keywords: user-generated content, social tags, social media, brand equity, firm valuation

Online Supplement: http://dx.doi.org/10.1509/jm.12.0151

With the advent of social media, customers have become active content creators by expressing and sharing their opinions, thoughts, and perceptions toward brands, products, and firms through online reviews, blogs, Tweets, and Facebook posts, collectively referred to as “user-generated content” (UGC). User-generated content is a significant indicator of customers’ brand knowledge and perceptions that helps marketing managers predict product sales (e.g., Liu 2006), understand firm valuation (Tirunillai and Tellis 2012), and infer competitive market structure (e.g., Netzer et al. 2012). Through various forms of UGC such as user reviews, blog posts, Tweets, and social tags, practitioners can now readily uncover and access customers’ thoughts and perceptions about a brand.

Despite the availability of rich semantic information in UGC, most previous studies (e.g., Liu 2006; Tirunillai and Tellis 2012) have focused on the volume and valence of UGC, leaving the associative nature inherent in UGC underexplored. For example, a recent study by Netzer et al. (2012) shows that UGC reveals competitive market dynamics yet leaves the relationship between firms’ performance and the associative structure reflected in UGC for further study. In this article, we investigate the informational value of the associative structure of user-generated keywords linked to a corporate brand. We focus on social tags, a specific form of UGC; infer how consumers conceptualize and create associative structures of brands; and demonstrate the informational value of social tagging networks in the context of stock market valuation.

We infer brand associative structure by examining social tagging networks created by online users. Social tagging is a way for online users to categorize and share web content. Within a social tagging system, users describe and categorize web content with a set of their own keywords, called “tags,” and diverse content is searched and shared using these tags. Because tags generated by individual users also help other users search and organize content, the collections of individually generated tags are called “social tags.” More formally, we define “social tagging” as a user behavior that results in user-defined keywords being associated with any content, with the keywords reflecting the user’s interpretation of the content internally and/or externally through a social process. Social tagging systems have also been known as “folksonomy” (a portmanteau of “folk” and “taxonomy”), which signifies a “grass-roots categorization system” (e.g., Pink 2005). Many social media platforms employ a social tagging system: for example, web links are tagged on Delicious, images and photos are tagged and shared on Pinterest and Facebook, videos are tagged on YouTube, and Tweets are tagged (using hashtags) on Twitter.

One of the primary motivations for creating social tags is to describe and efficiently categorize the vast amount of content that users encounter online. Tagging systems enable users to employ their own knowledge structure and inter-

Hyoryung Nam is Assistant Professor of Marketing, School of Economics, Erasmus University Rotterdam (e-mail: nam@ese.eur.nl). P.K. Kannan is Ralph J. Tyser Professor of Marketing Science and chair, Department of Marketing, R.H. Smith School of Business, University of Maryland, College Park (e-mail: pkannan@rhsmith.umd.edu). John Hulland served as area editor of this article.

ISSN: 0022-2429 (print), 1547-7185 (electronic)

pretations as they read and process content. Tagging systems also enable users to collaborate on developing structures that facilitate the content discovery process (Körner et al. 2010). As an illustration, a user could tag a *Wall Street Journal* article related to an automobile recall with tags such as “Brand X,” “quality issue,” and “recall.” Other users could access this content using these tags and associated links and use their own keywords such as “lemon” and “inferior quality” to associate with “Brand X” in interpreting that article. Through social interactions, tags organically develop dense associative structures around issues, concepts, or content by enabling users to filter brand-related content from news articles, blog posts, and corporate web page information through their own beliefs and biases. Social tags generated from users’ interpretations of brand-related content thereby become a valuable source for inferring the associative networks of users’ thoughts and perceptions about brands, products, and firms.

The typology of UGC in Table 1 highlights how social tags are different from other forms of UGC such as online reviews, blogs, and microblogs. Compared with online product reviews (the most popular UGC format studied in extant research), social tags can have a much broader focus (brand, product, and firm levels) and are not necessarily confined to product purchase or experience specific information. The collaborative interaction among users is also the highest in social tagging because users influence each other in sharing, retrieving, and tagging content. At the same time, self-presentation opportunities in social tagging tend to be the lowest (see also Kaplan and Haenlein 2010). This leads to less individual-specific, idiosyncratic noise in the data while capturing the overall sentiment in user groups rather well.

For our research purposes, the most significant advantages of social tags over other forms of UGC are (1) the high semantic gist of social tags, which enables the harvesting of user-defined, high-level semantic data, and (2) the associative structure of social tag data. The high semantic gist characteristic of social tags provides the essence of how users perceive and think about the concepts in as few words/phrases as possible. Compared with a much longer blog post, tags exist at the highest level of semantics. The network of interconnected tags across users also provides information on how these tags and concepts are related. This associative network of semantically rich tags and concepts can easily be used to form brand associations and obtain volume and valence of tags associated with a brand. Furthermore, networks of social tags can be used to reveal competitive market structure characterizing a set of brands and track dynamics in the perception of the competitive overlap of brand associations.

Such data are not readily available in other forms of UGC. Although similar data can still be obtained through text-mining of other forms of UGC, the keywords deemed salient for such associations must be selected by the researcher or by a text-mining model (for more details, see Web Appendix A1). In creating social tags, users tend to abstract away details at the word level to derive the semantic representations of words spontaneously. In contrast, text-mining machines tend to process at the word level to derive such associations between concepts, which limits the extraction of semantic gist. Finally, in a social tagging system, users can process a diverse set of content, including other forms of UGC such as blogs, microblogs, and reviews, all tagged using user-defined salient keywords. In this sense, information contained in social tagging can be inclusive of other UGC in addition to providing independent additional information. (For example, in Web Appendix A2, we find a positive correlation between the content in social tags with the content in blogs.)

In this article, we present brand management and performance metrics capturing brand familiarity, favorability of brand associations, and competitive overlaps of brand associations by examining social tags created on a social tagging and bookmarking site, Delicious. To test our propositions regarding the impact of these social tag metrics on firm value, we collected the social tags of 44 firms across 14 product markets (based on the North American Industry Classification System [NAICS]) created on Delicious from January 2006 to May 2010. We focus on “monobrand” firms—that is, firms with a single brand that represents the primary business (e.g., Bharadwaj, Tuli, and Bonfrer 2011; Mizik and Jacobson 2008)—to guarantee that the changes in stock returns can be fully related to the changes in brand management metrics, controlling for other effects. We find that an unexpected change in the proposed brand management metrics derived from social tags (hereinafter “social tag metrics”) explains variations in unanticipated stock returns even after controlling for the firm’s accounting metrics (unanticipated sales growth and unanticipated earn-

TABLE 1
Typology of UGC

	Online Reviews	Blogs	Social Tags	Microblogs
Focus	Narrow (product focused)	Broad (brand/firm/product)	Broad (brand/firm/product)	Broad (brand/firm/product)
Self-presentation	Low	High	Low	High
Collaborative interaction	Moderate	Low	High	Low
Social interaction richness	Moderate	High	Moderate	High
Semantic gist	Moderate	Low	High	Moderate
Ease of inferring associative structure	Low	Low	High	Low

Notes: Although tags are widely used in Twitter and Facebook, in the current research, we focus on tags used in social bookmarking platforms (e.g., Delicious, Digg).

ings per share [EPS]); selling, general, and administrative (SG&A) expenses; media citations; and volume of other forms of UGC. We find brand familiarity, favorability of associations, and competitive overlaps of brand associations captured by social tag metrics to be significantly associated with firm value.

Table 2 presents the primary distinction of our research relative to existing research. First, this article focuses on the associative structure of concepts and attributes linked to a brand from all types of web content (e.g., reviews, articles, advertisements, news). Although research has suggested that associative relationships in UGC can reveal competitive market surveillance (e.g., Netzer et al. 2012), no previous study has empirically demonstrated the relationship between a brand associative structure and firm/brand performance such as sales or stock market valuation. We believe that our findings expand the current understanding of the relationship between UGC and firm valuation, which thus far has been mostly confined to the volume and valence of online product reviews.

Second, we present a conceptual model that shows how firm value can be inferred from the associative information in social tags. More specifically, we propose a method to obtain a proxy measure of brand familiarity, association favorability, and competitive overlap of brand associations and evaluate the informational value of each social tag metric. To the best of our knowledge, our study is the first to examine the informational value of social tags in the context of firm valuation.

Third, we investigate the differential informational value of social tag metrics depending on brand strength. Despite its significance, most previous literature has not elaborated on the cross-sectional heterogeneity of UGC metrics and brand performance metrics (Luo 2009; Tirunillai and Tellis 2012). In this article, we determine and test which types of competitive overlap of brand associations can indicate higher firm value for strong brands (i.e., brands with strong brand equity) versus weak brands (i.e., brands with weak brand equity). Our findings suggest that for strong brands, being strongly related to primary associations in the category (category dominance) is positively related to firm value. However, for weak brands, being more connected to competitors' associations is positively related to firm value. Thus, managers should strategically position their brand as reflected in the competitive brand association overlap according to the strength of the brand.

Background

Conceptual Foundation of Social Tags

It is well known in psychology that people reason about new concepts by placing those concepts into preformed categories and extrapolating from the attributes of other concepts belonging to the same categories. Although various theories have been proposed (e.g., the prototype theory [Rosch 1973], the exemplar theory [Medin and Schaffer 1978]) to explain how people mentally represent different concepts, the central premise of all the theories is the existence of a knowledge structure or schema as the basis for memory organization. This knowledge structure is formed and shaped by a person's interactions and experiences with stimuli over time. As new categories of concepts are encountered, they interact with the knowledge structure and change the organization of existing categories (e.g., Buchanan, Simmons, and Bickart 1999; John, Loken, and Joiner 1998). The cognitive bases for individual beliefs and biases also arise from the same knowledge structure. Extant research has also shown that categorization is not the only process by which consumer learning takes place; rather, consumers may also learn from analogies (Gregan-Paxton and John 1997). This would imply that such analogical learning can facilitate knowledge structure changes and memory reorganization.

Literature in natural language processing (e.g., Fountain and Lapata 2010; Griffiths, Steyvers, and Tenenbaum 2007) has highlighted that a person's understanding of sentences, conversations, and documents necessitates retrieval of appropriate concepts from memory. The retrieval of concepts from a stream of information is aided by using the semantic context or the "gist" of the material to predict and

TABLE 2
The Impact of UGC on Firm Performance

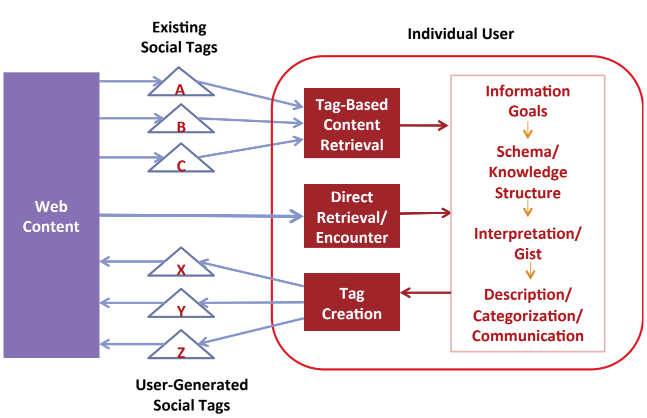
	This Article	Tirunillai and Tellis (2012)	Luo (2009)
Focus	Brand associative networks	Online word of mouth	Negative word of mouth
Data	Social tags	Product reviews	Consumer complaints filed with U.S. Department of Transportation
Dependent metric	Stock return	Stock return, risk, and trading volume	Cash flow, stock return, and stock volatility
Model	Stock response model	Vector autoregression	Vector autoregression
Time window	Quarterly	Daily	Monthly
Sample	44 firms in 14 markets	16 firms in 6 markets	10 firms in airline industry
UGC volume	Yes	Yes	Yes
UGC valence	Yes	Yes	Negativity only
Associative structure	Yes	Not investigated	Not investigated
Brand strength	Yes	Not investigated	Not investigated

identify related concepts from memory to be associated with the content being processed. As people encounter, retrieve, read, and process documents on the Web, the same process plays a key role in the usage and creation of social tags.

When people create social tags, their motivations can be classified into two categories: content organization (categorization and description of the content) and social communication (information sharing and opinion expression regarding the content). Users employ a tagging system to organize the resource (e.g., articles, photos, images). When users intend to categorize content, they more commonly use high-level attributes as tags; yet when users intend to describe content, they use more semantic, contextual attributes as tags (Strohmaier, Körner, and Kern 2010). It has been shown that tags created for describing content might be more useful for understanding rich interpretations of a document or article than those for categorizing content. Social communication is another significant driver of social tagging behavior. In an in-depth study of users on web-based photo-sharing systems, Ames and Naaman (2007) identify four distinct motivations: self-oriented organization (e.g., "I like order," p. 976), self-oriented communication (e.g., "reconstruct what I was thinking," p. 976), social organization (e.g., "wanted to tell people what it was," p. 977), and social communication (e.g., "I can give my friends the basic story," p. 978). The authors find that self-oriented organization and social organization for the public are more commonly observed than self-oriented communication and social communication for the public; however, social communication for friends and family is more commonly observed than social organization for friends and family.

Regardless of the motivation, the process of social tag creation is as follows (see Figure 1): The information goals or motivations of individual users while browsing or searching online content might lead them to use social tags created by other users to retrieve appropriate content matching their goals. However, individual users can also directly retrieve or encounter such content through other means. All content is filtered through the user's schema or knowledge structure that enables the interpretation and creation of the gist of the content. This gist relates the content to the concepts organized in memory, which in turn are

FIGURE 1
Conceptual Foundation of Social Tag Creation



used in the description, categorization, or communication process that underlies the motivation for creating social tags. Therefore, the tags that an individual user creates reflect not only the content being tagged but also the user's knowledge structure—that is, his or her mental representation of related concepts. In addition, the tags could also reflect the social interpretation of the content, if the user retrieved the content using others' social tags (Fu et al. 2010). Thus, social tags can be viewed as the categorization or description of content filtered through the lens of an individual user's knowledge structure as well as through the lens of others' social tags. Taken together, these processes provide a social interpretation of the content. We argue, therefore, that social tags provide insights into a person's beliefs and biases in context of the content (natural language) processed in creating the tags.

In the context of a brand, an individual user's tags should be reflective of his or her knowledge structure specific to the brand as well as the brand-related social content with which he or she chooses to interact. Similarly, the differences in tags that different people use for the same brand should provide the extent of heterogeneity in their knowledge structures and social interactions with regard to the brand. This renders tags ideal for exploring how users think and feel about brands and their relationships with attributes, features, and other concepts over time. Thus, tags can be useful input for constructing brand associative networks dynamically over time.

Social Tagging System in Delicious

An exemplary model of a social tagging system is the bookmarking site Delicious, which allows users to manage a personalized collection of web links, including corporate websites, product reviews, articles, and blog posts.[1] Users create their personal collection of online content by generating a "bookmark" that connects a web link to the set of social tags. By adding or deleting bookmarks, users can create or remove associations between social tags and web links. Bookmarked content on Delicious comprises various topics, such as technology, politics, media, business, and entertainment. In 2006, the site acquired its one-millionth registered user. In 2008, the number of registered users was more than 5 million (eBizMBA.com 2010).[2] Considering

[1]Delicious is the most widely used social bookmarking site. Its market share in social widgets was 10.5% in April–June 2011, while that for similar social bookmarking sites was 7.8% (StumbleUpon; www.stumbleupon.com), 5.2% (Reddit; www.reddit.com), and 9.7% (Digg; www.dig.com) (Web3 Technology 2012). One of the most significant differences between Delicious and these social bookmarking sites is its functionality. Whereas the aforementioned sites focus on finding interesting web content (e.g., links, news) by social voting (e.g., thumbs-up or -down function in StumbleUpon and Reddit), Delicious provides more functions, such as sorting/categorizing web content and searching the content by the user-generated keywords. Thus, user-generated keywords on Delicious are more appropriate to our research purposes to infer the brand associative networks.

[2]According to the July 2013 report by Alexa.com, 18% of site visitors are from the United States/Canada and 13% are from Europe. In our analysis, we focus on user-generated keywords in English, which covers more than 90% of the keywords in the raw data.

that the site had more than 5.5 million unique average monthly visitors in September 2010 (eBizMBA.com 2010), at such scale, the emergent semantic information is well established (Mika 2007).

The social tags on Delicious provide rich associative networks of users' thoughts and perceptions about various topics (e.g., brands, products, firms). Figure 2, Panel A, presents a snapshot of bookmarks associated with the brand name "Google" created on Delicious. Each bookmark is linked to the chain of user-specified social tags. For example, Bookmark 3 shows that a review on Nexus S published by Engadget is tagged with a set of keywords such as "review," "innovative," "iPhone," and "Samsung." These tags capture not only the descriptive summary of the article but also the user's interpretations of this article. As such, the collective distribution of tags connected to this article can reveal the collective perceptions about it. On Delicious, users share various forms of web content, including corporate web pages (Bookmark 1), blog posts (Bookmarks 2 and 4), news articles (Bookmarks 3 and 5), and product reviews (Bookmark 3). In addition, marketers can observe the extent to which a brand is connected to descriptive tags (e.g., "search," "trends"), positive tags (e.g., "cool," "innovative"), negative tags (e.g., "fails"), its own sub-brands (e.g., "nexuss," "android"), its competing brands (e.g., "Apple"), and product category–related tags (e.g., "mobile").

FIGURE 2
Illustration of Bookmarks and Social Tags in Delicious

A: Snapshot of Bookmarks About Google in Delicious

1) Google
http://www.google.com/
search | google | web | reference | tools | news | useful

2) Official Google Blog: The future of mobile
http://googleblog.blogspot.com/2008/09/future-of-mobile.html
mobile | google | future | trends | technology

3) Nexus S review -- Engadget
http://www.engadget.com/2010/12/10/nexus-s-review/
google | android | nexuss | review | iPhone | samsung | innovative

4) Google Operating System: Google Voice Local Search
http://googlesystem.blogspot.com/2007/04/google-voice-local-search.html
google | free | phone | Google | cool

5) The 10 biggest tech 'fails' of 2010 - CNN.com
http://edition.cnn.com/2010/TECH/innovation/12/28/tech.fails.year/index.html
fails | Google | Apple | ten | top

B: Visual Illustration of Social Tags Linked to Google

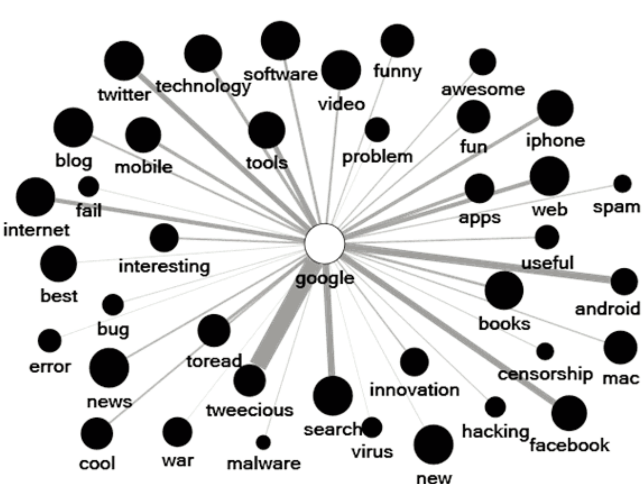


Figure 2, Panel B, illustrates how a collection of social tags obtained from bookmarks can be visualized. The node in the center of the network is the focal brand, Google, and the surrounding nodes are social tags related to Google. The size of each node represents the popularity of each tag on Delicious, captured as the volume of bookmarks linked to the tag. The width of the link between two nodes represents the associative strength between two nodes.[3] Note that we consider only the associations between Google and tags linked to Google and ignore intertag relationships. If we were to include intertag relationships, Figure 2, Panel B, would be more complex.

Useful Characteristics of Social Tagging Networks

Social tags linked to a corporate brand can provide useful insights into brand associations. Table 3 shows the classification of 7,019 sampled tags linked to 44 firms across industries and how they correspond to brand image components.[4] Although the largest group of tags (41.6%) is related to some neutral descriptive information, 4.5% of tags are related to brand attitude or brand personality, and 5.1% of tags are related to product attributes such as reliability and compatibility or nonproduct attributes such as price, promotion, or tutorials. Social tags also contain information about how a firm-corporate brand is connected to product categories, sub-brands, and competitors. We found that 6.9% of tags are about product category associations and 12.2% of tags are related to brand names and sub-brand names of the focal brand and other brands.

In addition, social tagging networks can reveal the competitive market structure. The extent to which tags are shared with other brands can reflect competitive and complementary relationships between brands. Figure 3 shows an example of interrelationships among three corporate brands: Apple, Microsoft, and Google. For example, the set of shared tags between Google and Apple—"mobile," "android," "iphone," "windows," and "apps"—reflects their rivalry in a mobile device market.

Finally, social tagging networks evolve dynamically, reflecting the change in social attention directed toward a

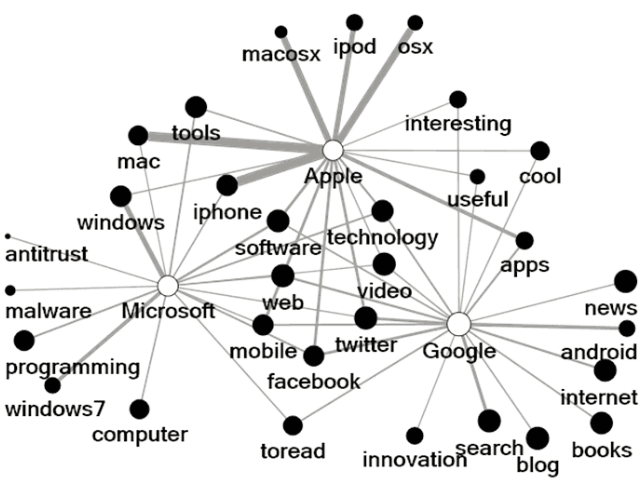
[3]The associative strength between a tag and a brand is defined as the volume of bookmarks tagged with both the brand name and the tag, which is the co-occurrence volume of the tag with the brand on Delicious.

[4]We employed a framework from extant research on consumer knowledge typology (e.g., Brucks 1985; Cohen and Basu 1987) and UGC (e.g., Liu 2006; Tirunillai and Tellis 2012).

TABLE 3
Types of User-Generated Tags Linked to Brands

Types of Associations	Example	Volume	%
Attitude/evaluation	Cool, humorous, innovative, creative	317	4.5%
Product attributes	Accessibility, reliability, stability	90	1.2%
Nonproduct attributes	Price, promotion, tutorial, service	277	3.9%
Product category	MP3, television, toys, hotel	485	6.9%
Brands	Apple, Microsoft, iPod, Zune, Google	857	12.2%
Unique place or name	Michigan, China, Benjamin, Ann	1,633	23.2%
Descriptive words	Article, behavior, business, bus	2,920	41.6%
Other		440	6.2%

FIGURE 3
Visual Illustration of Social Tags Linked to Google and Its Competitors



brand and the change in perceptions regarding a brand over time. Figure 4 shows the trend in the number of bookmarks generated for six corporate brands. The volume of bookmarks associated with a brand represents the extent to which users share online content related to the brand, indicating the level of social attention the brand displays in Delicious. For example, the volume of bookmarks for Apple was smaller than that for Microsoft until 2007. Then, in 2008 and 2009, the volume of bookmarks for the two corporate brands is almost tied, reflecting the growth in popularity and interest in Apple compared with Microsoft (possibly due to the successful introduction of new products such as the iPod, iPhone, and MacBook).

Conceptual Framework

Marketing literature has shown that the financial market reacts to information about brand value revealed through brand surveys, expert reviews, or user-generated product reviews, largely because investors view such information as an indicator of firm value (e.g., Mizik and Jacobson 2008; Tellis and Johnson 2007; Tirunillai and Tellis 2012). In this article, we employ social tags as an informational source of a brand's customer-based brand equity (CBBE) (Keller 1993). Considering that CBBE is based on attributes and thoughts associated with a brand, social tags can serve as a primary source to capture these elements of CBBE. We argue that tagging activity can be a proxy for underlying changes in sentiment and perceptions about a brand. Note that this occurs not because the market incorporates the tagging activity per se into the firm valuation process but rather because the brand perceptions and brand image reflected in the social tags drive the relationship between social tag metrics and firm valuation.

Social Tag Metrics and Hypotheses

In this section, we present the social tag metrics that reflect the elements of CBBE and present a model that relates these elements to a firm's stock return. Figure 5 presents the conceptual framework. At the conceptual level (top of Figure 5), when evaluating firm value, we expect the financial market to incorporate all the available information about a firm's profitability, including (1) the firm's marketing actions (e.g., advertising, new product innovations, public relations), (2) information in traditional/social media, and (3) the firm's accounting performance (e.g., sales, earnings, return on assets [ROA]). We further contend that the firm's marketing actions have an impact on two key elements of CBBE—brand awareness and brand image. As such, we expect that the CBBE elements at least partially mediate the impact of these influences on firm value. As the bottom of Figure 5 shows, we test the conceptual model—specifically, the role of CBBE—using a reduced form model that uses social tag metrics as proxies for elements of CBBE. We employ a stock response model to evaluate whether information contained in a social tag metric is associated with stock returns. The model assesses the relationship between unexpected change in social tag metrics and stock return, thereby evaluating whether the metric is truly indicative of the information used by financial markets. Because the price of a stock reflects all the available information to investors about the profitability of the firm, only "unanticipated" shock in a metric can explain abnormal returns of a firm (e.g., Mizik and Jacobson 2004). Thus, we focus on an unanticipated shock in each social tag metric and its informational value to the financial market. In the discussion that follows, we show the correspondence between key CBBE elements and our proposed social tag metrics and hypothesize how an unexpected shock in social tag metric relates to firm value.

Brand familiarity. One of the key elements of CBBE is brand awareness, defined as "the likelihood a brand comes into the mind and the ease which it does" (Keller 1993, p.

FIGURE 4
Trend of the Number of Bookmarks for Six Selected Brands

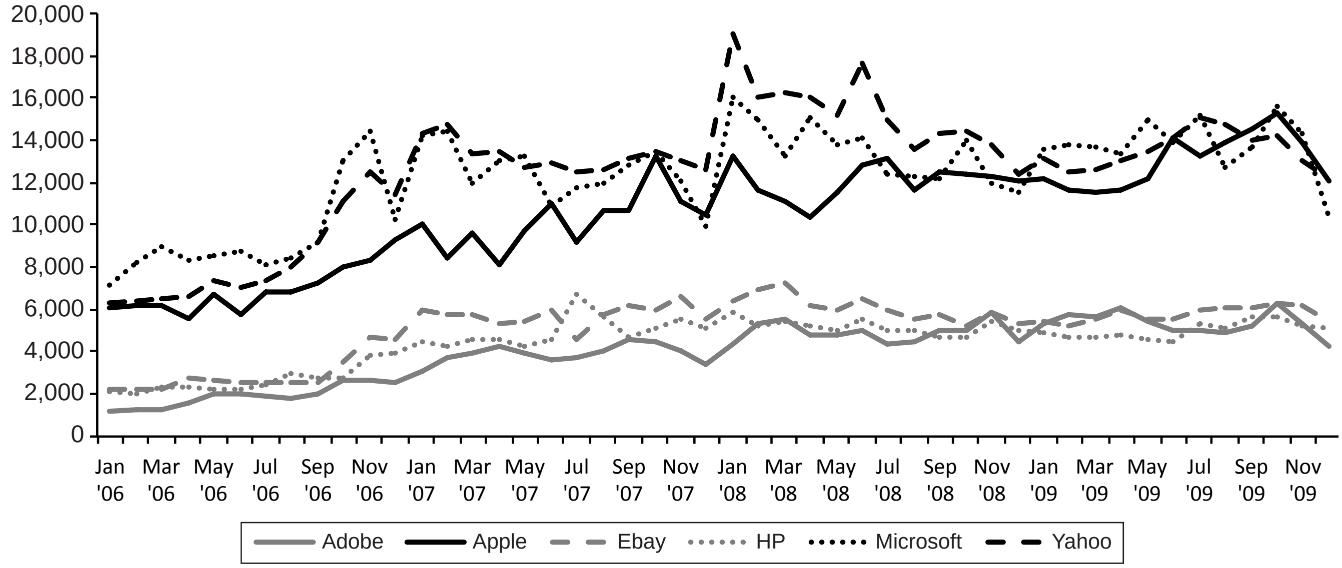
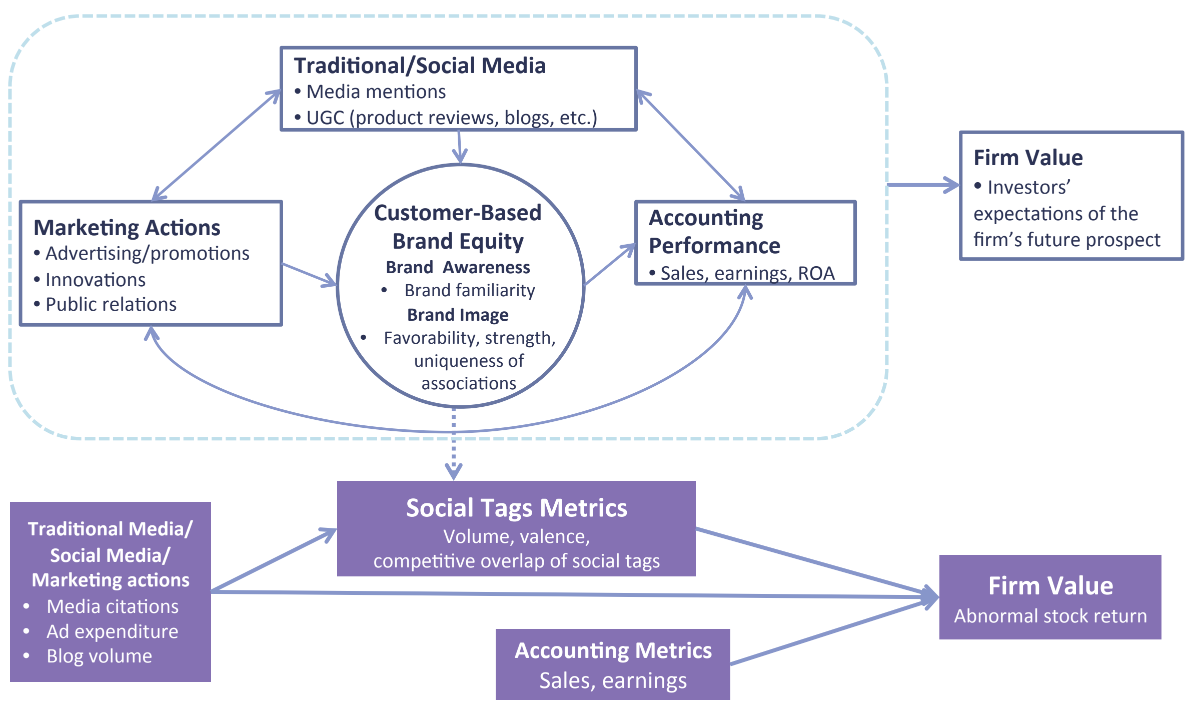


FIGURE 5
Conceptual Framework



3). Brand awareness helps increase the likelihood of a brand being included in the consideration set (e.g., Nedungadi 1990) as well as the likelihood of being selected in competing options (e.g., Jacoby, Syzabillo, and Busato-Schach 1977). In addition, brand awareness is a prerequisite for creating brand image (Keller 1993). As Keller (1993) argues, brand awareness is related to brand familiarity. Brand familiarity can be defined as the number of product-related experiences the consumer has accumulated (Alba and Hutchinson 1987). Familiarity increases through product usage, advertising, word of mouth, buzz from social media, and so on. It is through such repeated exposures that consumers' ability to recognize and recall a brand increases. As Keller (1993, p. 10) states, "Frequent and prominent mentions in advertising and promotion vehicles can intrusively increase consumer exposure to the brand, as can event or sports sponsorship, publicity, and other activities."

We argue that such mentions, promotions, and word of mouth can lead to an increased frequency of the brand name being used as a tag, and they can thereby serve as proxy measures for brand familiarity. That is, we focus on measuring brand familiarity rather than awareness per se. The social tagging creation process is similar to the brand recognition or brand recall test that evaluates how easily consumers recall or retrieve a brand when given the brand or the category as a cue. Given a cue (brand-related content such as articles or blog posts), users create keywords that come up in their mind. Thus, the volume of content tagged with a brand name captures not only the volume of brand-related content available on the Internet but also the extent to which users recognize and interpret brand-related content. We posit that marketers can use this metric as a complement to the classic brand familiarity metric because it can be related to firm valuation. For example, the volume of UGC has been used as a measure of consumer interest and attention and has been found to be significantly related to future product demand (e.g., Liu 2006) and stock returns (Tirunillai and Tellis 2012). This leads to our first hypothesis.

> H_1: The unanticipated increase (decrease) in the volume of content tagged with a brand—a proxy for brand familiarity—is positively (negatively) related to stock returns.

We expect that the extent to which a brand is related to socially popular keywords/concepts can serve as an additional proxy measure for brand familiarity. A brand that is closely connected to social trends and issues that are creating buzz may acquire additional attention and become more familiar to customers. That is, we expect that the close associations a brand has with social trends/issues may result in transference of the popularity of the tags to the brand through these associations. Over time, this leads to greater brand recognition and familiarity by enhancing the relevance of the brand to its potential customers. As such, we expect that a brand can position itself to be more easily recalled or remembered by associating the brand with socially popular concepts, issues, or events (e.g., "Super Bowl," "Grammys," "World Cup," "Olympics"). Correspondingly, we expect that an increase in the volume of socially popular tags linked to a corporate brand indicates an increase in social attention for the brand that can be positively related to the firm's stock returns.

> H_2: The unanticipated increase (decrease) in the volume of socially popular social tags linked to a brand—a proxy for brand familiarity—is positively (negatively) related to stock returns.

Favorability of brand associations. Favorable, strong brand associations are significant sources of CBBE (Keller 1993). Previous literature has shown that brand personalities and brand evaluations are indicative of significant brand value and, in turn, are related to firm valuation. Aaker and Jacobson (2001) show that brand attitude can be a good proxy of firm value by finding that changes in brand attitude are significantly related to changes in return on equity. Mizik and Jacobson (2008) find that brand metrics based on five central favorable brand attributes (differentiation, relevance, esteem, knowledge, and energy) provide incremental information content to accounting performance measures in explaining stock returns.

Here, we propose that the favorability of brand associations can be captured through the valence of social tags. As we discussed in the previous section, social tags contain rich textual information regarding brand personalities and brand evaluations that evolve over time. Broadly, they can fall into three categories: positive associations (e.g., cool, creative, innovative, excellent, premium), neutral associations (e.g., computer, technology), and negative associations (e.g., bad, dead, dirty, hypocrisy). By revealing customers' thoughts and perceptions, social tags serve as a good source for obtaining brand personalities and brand evaluations. This is in line with previous research that shows how the valence of UGC reflects customers' evaluations and, thus, predicts product sales and stock returns. For example, the valence of online user reviews can lead to an increase in book sales (Chevalier and Mayzlin 2006). Studies have found that negative word of mouth has a negative impact on cash flows and stock returns (Luo 2009) and that the average daily product review ratings are significantly related to daily stock returns (Tirunillai and Tellis 2012). Thus, we posit that the dynamics in positive and negative associations are related to investors' expectations about firm value.

> H_{3a}: The unanticipated increase (decrease) in the volume of the positive social tags linked to a brand—a proxy for favorability of brand associations—is positively (negatively) related to stock returns.
>
> H_{3b}: The unanticipated increase (decrease) in the volume of the negative social tags linked to a brand—a proxy for favorability of brand associations—is negatively (positively) related to stock returns.

Competitive overlap of brand associations. A competitive overlap of brand associations is defined as "the extent to which brand associations are linked to the product category (i.e., identification) and are, or are not, shared with other brands (i.e., uniqueness)" (Keller 1993, p. 13). Given that almost all the brands share part of the brand associations with direct or indirect competitors, it is important to build brand equity by understanding the extent to which brand associations should be shared or not shared with a brand's competitors. A strong brand should have unique, differentiated brand associations (Aaker 1982; Keller 1993) yet also maintain primary category associations that help strengthen the category membership (Keller 1993).

In this article, we posit that social tagging networks are exceptionally good sources of information about the competitive overlap of brand associations because they allow researchers to automatically observe which keywords are shared among competing brands. Our approach is in line with previous studies that examine the competitive market structure of brands by employing associative structure of textual description. For example, Netzer et al. (2012) show that the associative network structure of keywords such as brands, product attributes, or evaluations retrieved from online user forums can indicate competitive market structure. In addition, Hoberg and Philips (2010) show that asset complementarities derived from the similarity based on the textual description of a firm's assets can be a measure of potential synergies between firms.

We propose two measures to capture the competitive overlap of brand associations. The first measure is "the extent to which brand associations are linked to the product category" (Keller 1993, p. 13). Being strongly linked to primary associations in the category can help a brand be easily identified as a category member (e.g., MacInnis and Nakamoto 1992; Sujan and Bettman 1989). An especially strong "prototypical" association in the category often makes a brand representative of the category (e.g., Cohen and Basu 1987). Thus, we argue that a stronger association with the primary associations in the category indicates a stronger position or emerging dominance of a brand in that category and, consequently, is a positive signal for future performance and stock returns.

> H_4: The unanticipated increase (decrease) in primary associations in the category—a proxy for category dominance—is positively (negatively) related to stock returns.

However, the overlap of brand associations with competitors can cause brand image dilution or consumer confusion. Comparable attributes or references between competitors, which were previously deemed differential points for the focal brands, often threaten the positioning of the focal brand (Keller, Sternthal, and Tybout 2002). In addition, the competitors' communication campaign may have a negative influence on the focal brand's communication effectiveness by confusing consumers about brand identity or attributes (Burke and Srull 1988; Keller 1987). We capture the extent to which a brand's image is diluted by its competitors by measuring the competitors' connectedness to the focal brand's associations. We expect that the increase in competitors' connectedness to a focal brand's associations can be viewed as interference by competitors or brand image dilution and, thus, is negatively related to firm value.

> H_5: The unanticipated increase (decrease) in competitors' connectedness to the focal brand's associations—a proxy for brand image dilution—is negatively (positively) related to stock returns.

Role of brand strength on the competitive overlap of brand associations. We allow for the possibility that the hypothesized relationships regarding the competitive overlap of brand associations (H_4 and H_5) vary according to brand strength. We define brand strength mainly in line with brand equity.[5] Brand equity is defined as "the differential effect of brand knowledge on consumer response to the marketing of the brand" (Keller 1993, p. 1). Marketing researchers suggest that customers' reactions to marketing activities of strong brands are systematically distinct from their reactions to marketing activities of weak brands (i.e., brands without strong brand equity). A recent study by Ho-Dac, Carson, and Moore (2013) shows that the impact of product reviews on sales is distinct across brand strength. Specifically, researchers note double-jeopardy effects, suggesting that strong, large brands tend to gain more loyalty and attraction from customers than small brands (e.g., Ehrenberg, Goodhardt, and Barwise 1990). In addition, strong brands are likely to experience positive feedback on the profitability of dominant market share, which signals higher product quality (e.g., Smallwood and Conlisk 1979). Furthermore, strong brands are known to have greater advertising and marketing communications effectiveness (e.g., Hoeffler and Keller 2003; Srivastava and Shocker 1991). Thus, we expect that the positive feedback from stronger links to primary associations in the product category, which indicates category leadership or dominance, is stronger for strong brands than for weak brands.

> H_6: The relationship between an unanticipated increase in primary associations in the category and stock returns is more positive for strong brands than for weak brands.

In addition, we expect differential relationships between the increase in competitors' connectedness to the focal brand's associations and firm value. Considering that differentiation strategy can drive the future sales of strong brands (Ehrenberg, Goodhardt, and Barwise 1990), when a strong brand's competitors equate their brands to the focal brand through a greater number of points of parity (i.e., increase in competitors' connectedness to the focal brand), this higher level of competitor interference can result in a lower firm value for the focal brand. In contrast, for a weak brand, having more associations connected to its competitors can result in assimilated associations that benefit the weak brand by positioning it more closely to the strong brands in the market (Ehrenberg, Goodhardt, and Barwise 1990). Thus, creating more points of parity to other competing brands in the category could be a viable strategy for weak brands (Keller, Sternthal, and Tybout 2002). We expect that the increase in competitors' connectedness to the focal brand's associations, which should be deemed a negative signal to strong brands, may not necessarily negatively affect the stock returns of weak brands, because they benefit from the shared associations with strong brands.

> H_7: The relationship between an unanticipated increase in competitors' connectedness to the focal brand's associations and stock returns is more negative for strong brands than for weak brands.

Research Design

Data

Sampling firms. Our selection of firms and markets for the analysis was based on the following criteria. First, we selected monobrand firms that serve in consumer goods industries (e.g., retail, consumer electronics, Internet service companies). To obtain sufficient data to reliably derive brand associative networks, we further limited these firms to those with more than 100 average monthly social bookmarks on Delicious. Next, we selected firms that (1) are U.S. based and listed on the U.S. stock exchange, (2) contain at least ten quarterly data points for financial and accounting metrics from the first quarter of 2006 to the first quarter of 2010, and (3) had not experienced serious mergers and acquisitions from the first quarter of 2006 to the

[5]Brand strength could also be related to the concept of brand prominence (e.g., Han, Nunes, and Drèze 2010), which captures conspicuousness or "loudness" of a brand. In this article, we focus on brand strength derived from brand equity.

first quarter of 2010. Finally, we carefully investigated the content in the sample of 100 bookmarks linked to each corporate brand and excluded corporate brand names from the sample if more than 5% of the bookmarks were considered unrelated. Because a corporate brand name often may have other connotations, this procedure was critical in producing valid data. Our criteria yielded 683,822 bookmarks generated for 44 firms in 14 markets for the analysis.[6]

Social tagging data collection. We employed the following procedures for collecting social tags from Delicious. First, we identified the set of social tags for the data collection. To this end, for each firm in our sample, we collected 2,000 bookmarks and obtained all of the social tags linked to each bookmark. From the collected tags we constructed a dictionary of tags. Among 60,377 tags in the tag dictionary, we selected 7,019 key tags that had more than five bookmarks linked to corporate brands in our sample. This set of tags was refined for each brand. We retained tags that could cumulatively explain more than 99% of tag–brand co-occurrence volume for the final analysis. Then, using Delicious search algorithms, we obtained the historical monthly trends of bookmarks tagged with each key tag and each firm-brand name. We excluded the bookmark data before 2006 because Delicious was launched in 2003, and the number of users did not stabilize until approximately 2006.

We obtained firm financial performance, including quarterly sales, ROA, and stock returns from Center for Research in Security Prices (CRSP) and Compustat databases. We matched the financial data from CRSP and Compustat to social tagging data based on a calendar quarter window. Table 4 shows the measures we employed in our analysis.

[6]The firms in our sample are characterized by the following list of the NAICS codes: manufacturing (311930, 33411, 334112, 334119, 334220, 339932, and 339931), retail (452910, 452990, and 453210), information (519130, 511210, 515120, and 515210), and food service (722211).

TABLE 4
Measures

Conceptual Variable	Notation	Measured Variable	Data Source	
Independent Variables				
Brand familiarity	$VolBK_{it}$	Log of volume of bookmarks linked to brand i during quarter t	Delicious	H_1
	$SocialPop_{it}$	Number of socially popular tags linked to brand i during quarter t scaled by the number of all tags linked to brand i during quarter t	Delicious	H_2
Favorability of brand associations	$Positive_{it}$	Volume of positive tags linked to brand i during t scaled by volume of bookmarks	Delicious	H_{3a}
	$Negative_{it}$	Volume of negative tags linked to brand i during t scaled by volume of bookmarks	Delicious	H_{3b}
Competitive overlap of brand associations	$CatDominance_{it}$	Associative strength of primary associations in the product category of brand i during quarter t	Delicious, NAICS	H_4, H_6
	$Connectedness_{it}$	Average number of competitors associated with each association of brand i during quarter t	Delicious, NAICS	H_5, H_7
Dependent Variables				
Stock returns	$FFRET_{it}$	Abnormal stock returns obtained from Fama–French–Carhart model	CRSP, Kenneth French's website	
Control Variables				
Firm characteristics	$BkMk_{it}$	Log of book-to-market ratio of firm i at quarter t	Compustat	
	$MktValue_{it}$	Log of market value of firm i at quarter t	Compustat, CRSP	
Firm performance	$Sales_{it}$	Log of sales of firm i at quarter t scaled by total assets of firm i at quarter t	Compustat	
	ROA_{it}	ROA of firm i at quarter t	Compustat	
	EPS_{it}	EPS of firm i at quarter t (excluding extraordinary items)	Compustat	
Analyst forecast	$EPSForecast_{it}$	The mean of analysts' forecast on EPS of firm i for quarter t	I/B/E/S	
Media citations	$Media_{it}$	Log of the number of articles on firm/brand i during quarter t retrieved by Dow Jones Factiva Database	Factiva	
UGC volume	$Blogs_{it}$	Log of the number of blog posts on firm/brand i during quarter t identified by Google search engine	Google Blog Search	
SG&A expenses	SGA_{it}	SG&A expenses of firm i at quarter t scaled by total assets of firm i at quarter t	Compustat	

Social Tag Metrics

Networks of firms and social tags. We borrow the framework of affiliation networks (Faust 1997) to construct the bipartite networks of firms/brands and social tags. Our approach is in line with previous literature suggesting that brand image can be represented with an associative network of a brand node and a variety of associations linked to a brand node (e.g., John et al. 2006; Keller 1993). We denote the set of firms/brands as $F = \{f_1, f_2, \ldots, f_{NF}\}$ and the set of tags as $K = \{k_1, k_2, \ldots, k_{NT}\}$, where NF is the number of firms/brands and NT is the number of tags in this network. The affiliation network matrix of firms and keywords for time window t, A_t, consists of element a_{ijt} as defined for each keyword–brand name pair as presented in Equation 1a:

$$\text{(1a)} \qquad A_t = \begin{pmatrix} a_{ijt} & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & a_{NF,NT,t} \end{pmatrix},$$

$$a_{ijt} = N_t(F_i, K_j),$$

where $N_t(F_i, K_j)$ is the number of bookmarks linked to both firm/brand i and keyword j during time window t, which captures how frequently a firm is shown together with a keyword.

Alternatively, the affiliation network can be constructed using the following cosine distance measure, a'_{ijt}, which was shown to effectively capture the associative relationship between tags (Robu, Halpin, and Shepherd 2009):

$$\text{(1b)} \qquad a'_{ijt} = \frac{N_t(F_i, K_j)}{\sqrt{N_t(F_i) N_t(K_j)}},$$

where $N_t(F_i)$ is the number of bookmarks linked to firm/brand i and $N_t(K_j)$ is the number of bookmarks linked to keyword j during time window t.

Brand familiarity. We measured brand familiarity with two metrics: (1) the volume of bookmarks linked to a brand ($VolBK_{it}$) and (2) the number of socially popular tags linked to the brand scaled by the number of all tags linked to the brand ($SocialPop_{it}$). For the second metric, we identified socially popular tags on the basis of the growth in the volume of tags mentioned in the community. Note that we considered tagging activities by the entire community and did not restrict our consideration to tags related to the brands in the sample. This metric captures the indirect social attention attached to a brand through connections to social fads, trends, and interests. When determining the set of socially popular tags, we focused on nonnegative social tags and excluded negative socially popular tags from the list of socially popular tags. We obtained socially popular tags by employing multiple criteria: top 100, 200, 500, 1,000, and 1,500 tags in terms of the volume growth between t and t – 1 identified. In this article, we report the results based on the top 500 socially popular tags.

Favorability of brand associations. We measured valence of brand evaluations is measured as the volume of positive tags ($Positive_{it}$) and negative tags ($Negative_{it}$) scaled by the volume of bookmarks linked to a brand ($VolBK_{it}$), as specified in Equation 2. These metrics capture the fraction of positive (negative) tags among all tags linked to a brand. To obtain this measure, we manually classified 7,019 tags into three categories: positive descriptions, negative descriptions, and neutral descriptions. Three raters participated in this classification process. Fleiss Kappa index (Fleiss 1971) for the reliability across three raters was .904 (z = 44.7, $p < .001$), indicating a reasonable level of agreement. For keywords that raters disagreed on, we took the majority of opinion.

$$\text{(2)} \qquad Positive_{it} = \frac{\sum_{j \in POSTAG} a_{ijt}}{VolBK_{it}}, \quad Negative_{it} = \frac{\sum_{j \in NEGTAG} a_{ijt}}{VolBK_{it}},$$

where POSTAG is the set of predefined positive tags by the three raters and NEGTAG is the set of predefined negative tags by the three raters.

Competitive overlap of brand associations. We measure the competitive overlap of brand associations by employing two metrics capturing category dominance and competitors' connectedness. (For a more detailed explanation of the two metrics, see Web Appendix A3.) The first metric, category dominance ($CatDominance_{it}$) is assessed by how strongly a brand name is connected to primary associations in the category. As presented in Equation 3a, this term is measured as the sum of the product of the associative similarity between the focal brand and the competitors. The more strongly a focal brand is connected to a category's core associations (i.e., an association that is strongly connected to competitors in the market), the higher the value of category dominance. We define the set of competitors as the set of firms with the same NAICS code.[7] For firms without the same NAICS code, we used the first four or two digits of the NAICS code.

$$\text{(3a)} \qquad CatDominance_{it} = \sum_{j=1}^{NT} \sum_{i' \in CF_i} a'_{ijt} a'_{i'jt},$$

where CF_i is the set of competitors for firm/brand i identified by NAICS code.

We define the second metric, competitors' connectedness to the focal brand's associations ($Connectedness_{it}$), as the mean number of competitors linked to each association of the focal brand, as presented in Equation 3b. This metric captures the extent to which competitors share an association with the focal brand.

$$\text{(3b)} \qquad Connectedness_{it} = \frac{\sum_{j=1}^{NT} \sum_{i' \in CF_i} I(a_{ijt}) I(a_{i'jt})}{\sum_{j=1}^{NT} I(a_{ijt})},$$

where $I(a_{ijt}) = 1$ if $a_{ijt} > 0$ and 0 otherwise.

[7]We also identified the set of competitors from the extent to which a brand shares tags with other brands. The competitive metrics derived from this method were highly correlated to the metrics based on the competitor set identified by NAICS code (Pearson's correlation coefficient ranged from .63 to .88).

Measure of Firm Value

We estimated the abnormal stock returns of a firm using the Fama–French–Carhart asset pricing model (Carhart 1997; Fama and French 1993). The model posits that abnormal returns are a function of the overall market return, the difference between returns of small-firm and large-firm stocks, the difference between returns of high and low book-to-market stocks, and the momentum as specified by the following equation[8]:

$$(4) \qquad R_{it} - R_{ft} = \alpha_i + \beta_i(R_{mt} - R_{ft}) + s_i SMB_t + h_i HML_t + m_i MOM_t + \varepsilon_{it},$$

where R_{it} is firm i's stock market return at quarter t, R_{ft} is the risk-free rate at quarter t. R_{mt} is market return, SMB_t is the difference between returns of small-firm and big-firm stocks, HML_t is the difference between returns of high and low book-to-market stocks, MOM_t is the momentum factor that accounts for the tendency for increasing asset prices to increase further, and abnormal return for a firm is calculated as $(R_{it} - R_{ft}) - (\hat{R}_{it} - R_{ft})$.

Measures of Control Variables

We employ four main control variables in the stock response model: analysts' forecasts, media citations, blog posts, and SG&A expenses. We expand on these control variables next.

Analysts' forecasts. We used EPS forecasts from the Thomson Financial I/B/E/S database analyst to obtain the measure of unanticipated earnings. Unanticipated earnings surprise is calculated as the difference between actual quarterly EPS and the median of the analyst earnings forecasts from the I/B/E/S.

Media citations. We used the number of articles retrieved from the Dow Jones Factiva database for each month as the measure of media citations. We used the company tag in Factiva and considered the articles tagged with a company name published in a major newspapers and journals. Subsequently, we matched monthly media citations to social tagging data and stock return data based on a calendar quarter window and aggregated to a quarterly level.

Blog posts. Using Google Blog Search engine, we tracked the monthly volume of blog posts containing each brand name. Subsequently, we matched the number of blog posts to social tagging data and stock return data based on a calendar quarter window.

SG&A expenditure. We used quarterly SG&A expenses in the Compustat database as a proxy measure to capture a brand's promotional activities including advertising expenditures.

Brand Strength

To explain the different responses to social tag metrics across brands, we classified 44 brands into two categories: strong brands versus weak brands. To construct the brand strength variable, we employed diverse sources of brand information. First, we employed external brand equity rankings from Best Global Brands 2006–2009 (Interbrand 2009). We classified strong brands as brands that appeared in the top 100 from 2006 to 2009. There were 15 brands in the sample classified as strong brands. Second, we classified the brands on the basis of the volume of blog posts from January 2006 to May 2010 obtained through the Google Blog Search engine. We used a median split to classify brands into strong versus weak brands. Third, we classified the brands on the basis of the volume of media citations from January 2006 to May 2010 obtained through the Factiva search engine. Again, we used a median split to classify brands into strong versus weak brands. To obtain a complete measure of brand strength, we coded a brand as a strong brand if it satisfied at least two of these criteria. We classified 18 brands as strong brands.

Model Formulation

Stock response model. Our objective is to test whether the social tag metrics defined in the previous section have incremental informational value relative to accounting performance measures in explaining stock returns. To this end, we employ the stock return response modeling approach. Stock response modeling is an analytical tool to evaluate whether information contained in a metric is associated with the changes in stock returns (e.g., Mizik and Jacobson 2004). The underlying assumption of stock return response modeling is the financial market efficiency that the price of a stock reflects all available information related to the profitability of the firm (LeRoy 1989). We believe that for our quarterly data, stock market efficiency holds.[9] Note that the stock response model evaluates whether the information in social tag metrics is associated with abnormal stock returns yet does not establish the causality between social tag metrics and stock return. With the model, we can confirm the informative value of social tag metrics for the financial market (i.e., whether the market incorporates the information reflected in social tag metrics).

We model abnormal returns as a function of unanticipated changes in accounting performance and unanticipated changes in social tag metrics between t and t – 1, as presented in Equation 5. We include control variables such as book-to-market ratio, market value, and calendar year dummies. These variables stem from the existing literature and capture the impacts of economy-wide factors and firm-specific characteristics on the stock returns (Fama 1998; Fama and French 1993).

$$(5) \qquad StkRet_{it} = Eret_{it} + Uret_{it}, \text{ and}$$

[8]Information regarding these benchmarks as well as the actual data used for this analysis are available on Kenneth French's website (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

[9]Although some empirical evidence indicates that the stock market is not always efficient and needs some time to incorporate the available market information (e.g., Luo 2009; Tirunillai and Tellis 2012), in most cases, this market anomaly has been found to resolve in several days or weeks (Mizik and Jacobson 2009).

$$\text{StkRet}_{it} - \text{Eret}_{it} = \sum_{j=1}^{J} \gamma_j \text{U}\Delta\text{Acc}_{jit} + \sum_{k=1}^{K} \beta_k \text{U}\Delta\text{SocTag}_{kit} + \sum_{\ell=1}^{L} \rho_{i\ell}\text{Control}_{\ell it} + \varepsilon_{it},$$

where Eret_{it} is the expected rate of return for firm i at quarter t, Uret_{it} is the unexpected rate of return for firm i at quarter t obtained from the Fama–French–Carhart model (see Equation 4), $\text{U}\Delta\text{Acc}_{jit}$ is the unanticipated change in accounting variable j for firm i between t and t – 1, $\text{U}\Delta\text{SocTag}_{kit}$ is the unanticipated change in social tag metric k for firm i between t and t – 1, and $\text{Control}_{\ell it}$ is control variable ℓ for firm i at t.

The full model of Equation 5 includes social tag metrics, accounting metrics, control variables, and brand strength as a moderator of the impact of the competitive overlap of brand associations on stock returns. We employed a fixed-effects time-series panel model to account for time-invariant unobservables that could be correlated with social tag metrics and accounting metrics (Wooldridge 2009). The model includes year-specific dummy variables to control for time-variant, economy-wide unobserved factors. To account for heteroskedasticity, we adjust standard errors of the model according to White's (1980) method.

Unanticipated change in social tag metric. To obtain the unanticipated change in each social tag metric, we built a random-mixed model (see Equation 6). This model accounts for the expectation on the basis of t – 1 levels of a social tag metric, the correlation of a social tag metric with a firm's lagged financial performance (sales and stock return), unobserved random time effects, and unobserved random firm effects. We used the standardized residual associated with each equation as the measure of unanticipated change of each social tag metric ($\text{U}\Delta\text{SocTag}_{kit}$).[10] By focusing on the unanticipated change in social tag metrics (which is not correlated to a firm's lagged performance), we address the potential endogeneity concerns that a firm's financial performance affects brand perceptions and thus social tag metrics can only be a reflection of a firm's past financial performance.

$$\text{SocTag}_{kit} = \beta_0 + \beta_1\text{SocTag}_{kit-1} + \beta_2\text{Sales}_{it-1} + \beta_3\text{FFRET}_{it-1} + \theta_i + \delta_t + \varepsilon_{it}, \quad (6)$$

where SocTag_{kit} is the kth social tag metric of firm i at quarter t, Sales_{it-1} is the log of the sales of firm i at quarter t, FFRET_{it-1} is the abnormal returns of firm i at quarter t, θ_i is unobserved random firm effects, and δ_t is unobserved random time effects. We similarly obtain the measure of the unanticipated change in media citations ($\text{U}\Delta\text{Media}_{it}$) and blog mentions ($\text{U}\Delta\text{Blogs}_{it}$).

[10]Note that social tag metrics can be constructed in a more granular time window. The primary reasons we employ quarterly data are because (1) our focus is on more persistent dynamics in brand image rather than transient change captured by daily or weekly trends and (2) we plan to control for the impact of the current-term sales, earnings, and SG&A expenditures (available on a quarterly basis).

Unanticipated change in accounting metric. We obtain the measure of the unanticipated change in quarterly sales, ROA, and SG&A expenditures with the standardized residuals from Equation 7, which addresses a seasonal pattern in quarterly accounting data (Foster 1977). We obtain the earnings surprise by subtracting the actual realized EPS from the consensus forecast retrieved from I/B/E/S.

$$\text{Acc}_{jit} = \beta_0 + \beta_1\text{Acc}_{jit-4} + \beta_2(\text{Acc}_{jit-1} - \text{Acc}_{jit-5}) + \theta_i + \delta_t + \varepsilon_{it}, \quad (7)$$

where Acc_{jit} is the jth accounting metric of firm i at quarter t, Acc_{jit-4} is the seasonal lag of jth accounting metric of firm i at quarter t, $(\text{Acc}_{jit-1} - \text{Acc}_{jit-5})$ captures the fourth differences in a quarterly accounting metric, θ_i is unobserved random firm effects, and δ_t is unobserved random time effects.

Results

Descriptive Statistics

Table 5 presents descriptive statistics of the measures reported in Table 4. There is substantial variation in stock return measures and accounting metrics as well as social tag metrics. On average, the log of volume of bookmarks published on each brand during each quarter is 7.26 (SD = 1.78). The mean proportion of socially popular tags linked to each brand during each quarter is 19% (SD = 7%). On average, .26 positive tags are associated with a brand per each bookmark (SD = .13), and .1 negative tags are associated with a brand per each bookmark (SD = .1). The category dominance metric ranges from .006 to 1.045 (M = .24, SD = .23) and, on average, 1.86 competitors are connected to a focal brand's associations (SD = 1.46). The mean of the abnormal returns of the firms in the sample obtained from the Fama–French–Carhart model is –.023 (SD = .156). The mean ROA of the firms is .17 (SD = .06), and the mean log of sales is 8.04 (SD = 1.28). On average for each quarter, the log volume of blog posts on each brand is 11.11 (SD = 1.92), and the log of media citations on each brand is 4.12 (SD = 1.28).

Table 6 presents the correlation of the unexpected change in social tag metrics and accounting metrics, as well as control variables. The abnormal stock return (FFRET) is positively correlated with a firm's accounting performance such as unanticipated earnings surprise (UΔEPS) (r = .104, $p < .05$) and unanticipated sales growth (UΔSale) (r = .0792, $p < .05$). In addition, the unanticipated growth in socially popular keywords (UΔSocialPop) is positively correlated with the abnormal returns (r = .102, $p < .05$), which is consistent with our hypothesis that the financial market evaluates the increase in associations with socially popular keywords as a positive signal for a firm's prospect. Yet contrary to our expectation, connectedness to competitors' associations (UΔConnectedness) is positively related to abnormal returns. We suspect that such a relationship is related to differential informational value of connectedness to competitors' associations (UΔConnectedness) for weak brands versus strong brands. We further confirmed that

TABLE 5
Descriptive Statistics

	Levels[a]					Unexpected Change[b]				
	N[c]	M[d]	SD[d]	Min[e]	Max[e]	N[c]	M[d]	SD[d]	Min[e]	Max[e]
FFRET						710	–.023	.156	–.500	.620
VolBK	710	7.263	1.783	3.784	12.267	710	–.006	.986	–4.098	6.942
SocialPop	710	.192	.070	.096	.419	710	–.009	.957	–3.221	3.099
Positive	710	.264	.131	.045	1.036	710	.011	.962	–2.958	8.316
Negative	710	.104	.100	.000	.805	710	.002	.949	–4.055	7.901
CatDominance	710	.246	.234	.006	1.045	710	.005	.965	–3.559	4.463
Connectedness	710	1.867	1.469	.144	5.724	710	–.007	.963	–4.657	4.522
Media	710	4.125	1.284	.000	7.725	710	–.003	.954	–5.252	5.061
Blogs	710	11.110	1.927	6.669	16.744	710	.000	.988	–4.131	5.792
Sales	710	8.043	1.283	5.273	11.641	710	.004	.968	–8.912	3.977
EPS	710	.042	.161	–1.116	1.627	710	.042	.161	–1.116	1.627
ROA	694	.174	.069	–.159	.401	660	.004	.956	–6.125	7.000
SGA	710	.075	.036	.014	.211	710	.004	.959	–5.261	5.503
MktValue	710	9.938	1.261	7.121	12.713					
BkMk	710	–.497	.751	–2.471	2.325					

[a]This column reports the descriptive statistics of each variable in levels.
[b]This column reports the descriptive statistics of the unexpected change in each metric obtained from Equations 6 and 7.
[c]The number of observations.
[d]Mean and standard deviation of quarterly firm-level values.
[e]Based on all quarterly firm-level values.

unanticipated change in connectedness to competitors' associations (UΔConnectedness) is positively related to abnormal returns for weak brands (r = .101, $p < .05$) but not significantly related to abnormal returns for strong brands (r = –.03, $p > .10$). We find that some of the social tag metrics are intercorrelated (for more discussion, see Web Appendix A4.)

Informational Value of Social Tag Metrics

Stock response model. Table 7 presents the results from the stock response model specified in Equation 5. Column 1 presents the results of stock response model with control variables capturing firm characteristics (market value and book-to-market ratio); Column 2 presents the results with accounting metrics,[11] SG&A expenses, media citations, blog posts, and all the variables in Column 1; Column 3 presents the results with social tag metrics and all the variables in Column 2; Column 4 presents the results with brand strength as a moderator for the impact of competitive overlaps of brand associations on stock returns, including all the variables in Column 3. In Column 4, we code brand strength variable as 1 for a strong brand and –1 for a less strong (weak) brand. The following discussion is based primarily on Column 4. Note that the results in Table 7 indicate that the information in social tag metrics is associated with abnormal stock returns, yet no inference can be drawn regarding the causality between social tag metrics and stock return. We believe that the market reaction to social tag metrics is not to social tags and tagging activities per se but rather to the brand information reflected in social tag metrics. We do not test whether investors investigate social tags and employ the information from tagging for firm valuation. Rather, we contend that the market considers the brand associative structure and brand performance for which social tagging metrics serve as proxies.

We find that the growth in quarterly volume of blog posts (UΔBlogs) and quarterly volume of media citations (UΔMedia) are not reflective as informative signals with regard to a firm's prospect. This indicates that a shock in these volume metrics may not successfully explain the quarterly stock returns (long-term valuation). Also, contrary to H_1, the growth in quarterly volume of social bookmarks tagged with a brand name (UΔVolBK) (a proxy of brand familiarity) is not significant. Note that our finding differs from Tirunillai and Tellis's (2012) study, which shows the volume of online chatter affects the market return. One of the primary reasons that the volume of content tagged with a brand name is not reflected in stock returns (yet the volume of product reviews significantly explains the stock returns) is that the nature of social tagging and product reviews is different (see our discussion on Table 1). The volume of product reviews is more directly related to previous daily sales, which investors directly react to, and captures a focused attention on the products of a firm. Yet the volume of content tagged with a brand is more related to the brand itself and thus is a less direct signal of previous daily sales.

However, as predicted in H_2, we find that the change in associations with socially popular keywords (UΔSocialPop) as a proxy of brand familiarity is positively related to stock returns ($\beta = .0134$, $p < .05$). Thus, stronger association with socially popular keywords, concepts, or themes is indicative that the financial market views a brand in a positive light. Together with the results of H_1, we contend that the financial market's evaluation of a brand's prospects is essentially an evaluation of the brand's associations and the volume of

[11]Note that to save the degree of freedom of our model, Table 7 reports the results with EPS and sales. We report the model with EPS, sales, and ROA in Web Appendix A5-5 and find that the results significance of social tag metrics remain consistent across the models.

TABLE 6
Correlation of Measures

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. FFRET														
2. UΔVolBK	–.0304													
	(710)													
3. UΔSocialPop	.1026*	.0099												
	(710)	(710)												
4. UΔPositive	–.0111	–.0471	–.1391*											
	(710)	(710)	(710)											
5. UΔNegative	–.0758*	–.0355	–.0361	.3724*										
	(710)	(710)	(710)	(710)										
6. UΔCatDominance	.0196	.1907*	.0148	.0952*	.0398									
	(710)	(710)	(710)	(710)	(710)									
7. UΔConnectedness	.0744*	–.0624	.1023*	–.2112*	–.0868*	–.0039								
	(710)	(710)	(710)	(710)	(710)	(710)								
8. UΔMedia	–.0272	.0810*	.0647	–.0846*	.1152*	.0381	.0448							
	(710)	(710)	(710)	(710)	(710)	(710)	(710)							
9. UΔBlogs	–.0213	.1670*	–.1516*	–.1084*	.0131	.0742*	.0429	.0286						
	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(710)						
10. UΔSale	.0792*	.0173	–.0136	–.0172	–.0486	.0471	.0384	–.0121	–.0114					
	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(710)					
11. UΔEPS	.1048*	.0004	.0179	–.0032	.0066	–.0067	–.0037	–.0514	.0357	.1301*				
	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(710)				
12. UΔROA	–.0148	–.0053	.0179	–.016	.0305	.0542	.014	–.0752	–.0007	.1792*	.1941*			
	(660)	(660)	(660)	(660)	(660)	(660)	(660)	(660)	(660)	(660)	(660)			
13. UΔSGA	–.0408	.0001	–.057	–.0049	–.0217	.032	–.0226	–.072	.0103	.2743*	–.0463	.3685*		
	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(660)		
14. MktValue	–.0487	.0661	–.0167	.0024	.0114	.0146	.0075	.0261	.0291	.0756*	.005	.0694	–.007	
	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(660)	(710)	
15. BkMk	.1042*	–.0782*	.0343	–.0173	.0096	–.0216	–.0008	.0045	–.0268	–.0877*	–.0483	–.1173*	–.0514	–.4042*
	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(710)	(660)	(710)	(710)

*$p < .05$.
Notes: We present correlations as Pearson correlation coefficients. The number of observations for each pairwise correlation appears in parentheses.

TABLE 7
Informational Value of Social Tag Metrics

		1		2		3		4	
	MktValue	–.0406	(.0348)	–.0469	(.031)	–.0498	(.0316)	–.0507	(.0314)
	BkMk	.106***	(.0427)	.104***	(.0384)	.101**	(.0392)	.0993**	(.0393)
	UΔSale			.017**	(.0071)	.0158**	(.007)	.0155**	(.007)
	UΔEPS			.113***	(.0378)	.115***	(.0385)	.12***	(.0386)
	UΔSGA			–.0087	(.0057)	–.0079	(.0053)	–.0085	(.0054)
	UΔMedia			–.0055	(.0053)	–.0048	(.0051)	–.0045	(.005)
	UΔBlogs			–.0035	(.0058)	–.0011	(.0063)	–.0011	(.0063)
H_1	UΔVolBK					–2.4e–04	(.006)	–1.1e–04	(.006)
H_2	UΔSocialPop					.0141**	(.0064)	.0134**	(.0064)
H_{3a}	UΔPositive					.0071*	(.0041)	.0081*	(.0042)
H_{3b}	UΔNegative					–.0143**	(.0058)	–.0137**	(.0058)
H_4	UΔCatDominance					.0045	(.0057)	.0097**	(.0047)
H_5	UΔConnectedness					.0095**	(.0047)	.004	(.0043)
H_6	UΔCatDominance × ST[a]							.0138***	(.0049)
H_7	UΔConnectedness × ST[a]							–.0099**	(.0041)
	Constant	.458	(.32)	.516*	(.285)	.543*	(.29)	.551*	(.289)
	N	710		710		710		710	
	R-square	.07		.0971		.117		.126	
	Number of firms	44		44		44		44	
	F-statistics	10.9		8.98		10.4		11.5	

*p < .10.
**p < .05.
***p < .01.
[a]ST = 1 if the brand is classified as a strong brand and –1 if otherwise.
Notes: Robust standard errors appear in parentheses. Model includes four year dummies and brand fixed effects.

these associations with socially popular keywords rather than just the volume of a brand's tags. We argue that having a stronger association with socially popular keywords can help a brand enhance its popularity and brand familiarity to (potential) consumers. This will further help the brand increase its probability of being included in the consideration set and indirectly build on its popularity. To the best of our knowledge, this is the first empirical finding that a stronger connection of a brand to socially popular topics and trends is positively related to firm valuation.

Next, the favorability of associations is significantly related to stock market evaluation of a brand, as we predicted in H_{3a} and H_{3b}. We find that the unanticipated increase in positive social tags (UΔPositive) is marginally positively related to stock returns (β = .0081, p < .1) and the unanticipated increase in negative social tags (UΔNegative) is significantly negatively related to stock returns (β = –.0137, p < .05). Thus, the stock market views both positive and negative brand evaluations as informative, but it reacts more significantly to negative brand evaluations. This result is consistent with previous findings that the market is more sensitive to negative information disclosures than to positive ones (Kothari, Shu, and Wysocki 2009) and that negative UGC has a stronger impact on the stock market (Tirunillai and Tellis 2012).

We also find that the competitive overlap of brand associations is significantly related to firm value and that its impact on stock returns is moderated by brand strength. First, on average, the financial market evaluates stronger connections to primary associations in the product category (UΔCatDominance) positively (β = .0097, p < .05). As we predicted in H_4, stronger connections to primary associations in the product category indicate a brand's category leadership and dominance. However, the positive relationship is qualified by brand strength. As we expected in H_6 (β = .0138, p < .01), the financial market reacts more favorably to enhanced category leadership of strong brands than that of weak brands. Essentially, the financial market views category dominance to be more critical information for a strong brand's prospects (e.g., whether the brand can maintain category leadership) than for a weak brand's prospects.

We find that, on average, a shock in competitors' connectedness to the focal brand's associations (UΔConnectedness), which indicates that competitors share a focal brand's associations, is not significantly related to stock return, contrary to our prediction in H_5 (β = .004, p > .10). However, the stock market differentially evaluates the information contained in competitors' connectedness: greater connectedness of competitors is less favorably evaluated for strong brands, whereas it could indicate a more positive signal for weak brands, as we predicted in H_7 (β = –.0099, p < .05). Thus, we conclude that the financial market interprets a stronger connection to competitors' associations as more favorable information for weak brands than for strong brands.

Robustness tests. We further tested whether the findings from Table 7 are robust to different operationalizations of the variable and outliers. First, we tested the sensitivity of moderating relationships to different operationalizations of brand strength (for more detailed results, see Web Appendix A5-1). The moderating effect of brand strength on both CatDominance and Connectedness is consistently significant (for all cases, p < .05) across the model with brand strength based on the Interbrand database. In the model with brand strength based on the volume of blog posts, the moderating effect of brand strength on Connectedness is significant ($p < .05$), but CatDominance is not significant although the direction is positive consistent. In the model with brand strength based on the volume of media citations, the moderating effect of brand strength on CatDominance is significant ($p < .05$), but although the direction is consistently negative, Connectedness is not significant. Second, we tested whether the results in Table 7 are sensitive to extreme values or outliers (for more detailed results, see Web Appendix A5-2). We ran the same analysis by deleting ±1%, ±2%, and ±3% outliers from our sample and found that the results are consistent with Table 7. Third, we tested whether the results in Table 7 are consistent with different operationalizations of unanticipated change in social tag metrics (for more detailed results, see Web Appendix A5-3). We employed first-differenced terms of each social tag metric as measures of unanticipated change in social tag metrics, and the results are consistent with Table 7 (the only difference is the significance of the volume of positive tags, which is not significant in this model and is marginally significant in Table 7). Fourth, we tested the sensitivity of the results to different operationalizations of socially popular tags (for more detailed results, see Web Appendix A5-4). The significant positive relationship between socially popular tags and abnormal returns holds for the top 100, 200, and 500 socially popular keywords. (For the top 1,000 keywords, the relationship is marginally significant, and for the top 1,500 keywords, the relationship is not significant.)

Mediation. Because we posit that social tag metrics capture customers' interpretations of a firm's marketing actions (e.g., communication campaigns, new product introductions), we further tested whether social tag metrics mediate the impact of firms' marketing actions on stock returns. As a proxy measure for a firm's action, we employ unanticipated sales growth (UΔSale), unanticipated change in media citations (UΔMedia), and unanticipated change in SG&A expenditures (UΔSGA). We investigate whether (1) each social tag metric can be explained by the measures capturing a firm's marketing actions and (2) each social tag metric mediates the impact of firms' marketing actions on stock returns.

First, we found that an unanticipated sales growth is positively related to an unanticipated increase in category dominance (UΔCatDominance) ($\beta = .048$, $p < .05$), indicating that stronger category dominance may partly stem from a brand's sales performance. To formally investigate whether category dominance mediates the impact of sales growth on stock returns, we conducted a Sobel mediation test (Baron and Kenny 1986). The impact of sales growth on stock returns is partially (marginally significantly) mediated by the moderating effect of category dominance and brand strength (z-statistic = –1.648, $p < .10$). Second, an unanticipated change in media citations (UΔMedia) is marginally significantly and positively related to an unanticipated change in socially popular tags (UΔSocialPop) ($\beta = .0668$, $p < .10$), indicating that stronger associations with socially popular tags may stem from media disclosure. Yet the mediation of socially popular tags for the impact of media disclosure on stock returns is not significant. Third, an unanticipated change in media citations (UΔMedia) is negatively related to unanticipated change in positive tags (UΔPostive) ($\beta = –.0854$, $p < .05$), indicating that most of the positive tags are not from the media disclosure. However, the mediation of positive tags for the impact of media disclosure on stock returns is not significant.

Discussion

Informational Value of Social Tag Metrics

Brand familiarity. We find that an increase in connectedness to socially popular keywords can be viewed as a positive signal of a brand's future cash flows, whereas dynamics in volume of bookmarks does not contain significant information for investors. This result indicates that investors favorably evaluate an effort to position a brand associated with socially popular concepts, issues, and events. We argue that an increase in stronger associations with socially popular keywords is a reflection of enhanced brand awareness or brand familiarity (e.g., Keller 1993) and brand salience (e.g., Ehrenberg, Barnard, and Scriven 1997). It is worthwhile to note that the financial market does not react to mere volume metrics such as media disclosure and UGC volume in the longer time period. We reason that this is mainly due to the use of a quarterly aggregated volume metric. Considering that in a shorter time frame (at the daily level), researchers have found market inefficiency for the UGC volume metric (Tirunillai and Tellis 2012), we conclude that a shock in social attention possibly caused by media exposure or UGC can be viewed as significant information for a short-term stock return. However, in the long run, the market does not evaluate the quarterly change in volume metrics as informative; rather, it is the association of the brand with socially popular trends that is reflective of brand familiarity.

Favorability of brand associations. We find that the valence of social tags can serve as a good proxy measure for favorability of brand associations, to which the financial market reacts. Our findings suggest that investors interpret a shock in negative evaluations captured by social tags as informative about brands' future prospects; however, investors overlook a shock in positive evaluations reflected in social tags. This suggests that the negative brand information explains not only the short-term market inefficiency but also the longer-term period firm valuation.

Competitive overlap of brand associations and differentiation. We find that the competitive overlap of brand associations can be captured by the extent to which social tags are shared with competitors. Furthermore, our finding suggests that the financial market interprets the competitive overlap of brand associations differentially for strong brands versus weak brands: enhanced category dominance captured by stronger connection with primary associations in the category can be viewed as a more positive signal for future cash flow for strong brands than for weak brands. This indicates that by developing and occupying primary associations in the category, strong brands can benefit from the brand image as a category leader. However, for weak brands, it is not significantly important to be strongly connected to primary content in the category.

In evaluating information about weak brands, the more critical information for the financial market is the increase in connections with competitors' associations. The level of connectedness to competitors' associations captured by social tags is more favorably related to the firm value of weak brands than is the case for strong brands. We argue that such an increase in connectedness to competitors' associations can be viewed as a positive sign for weak brands' future prospects by indicating potential future cash flow attributable to associations with the strong brands in the category.

Managerial Implications

The proposed brand performance metrics, based on social tags, provide marketing practitioners with a methodology to mine and track social tags to create proxy measures of CBBE from UGC. Although customers frequently express their thoughts and perceptions about firms/brands in UGC, brand managers often find it difficult to collect and digest all such information. The current social media metrics based on volume and valence may be a good starting point to track CBBE, but they do not fully capture associative information connected to a brand. This research shows how the attributes and associations created by online users for a brand/firm can serve as a basis for proxy measures of CBBE elements. For example, the competitive overlap of brand associations, a critical component of CBBE, is not frequently monitored in practice due to the difficulty of the data collection process. By employing the social tag metrics we propose herein, managers can assess their brand image and brand positioning by tracking and monitoring their points of parity and points of difference (e.g., Keller, Sternthal, and Tybout 2002). Because users of a social tagging system can process a diverse set of content, including other forms of UGC such as blogs, microblogs, and reviews, the metrics we propose based on social tags are more inclusive of all UGC than any single form of UGC. We believe that our new metrics and measures can be good complements to annual brand surveys by enabling marketers to track the dynamics of CBBE.

Our findings provide marketing practitioners with insights into the kinds of marketing activities and communication strategies that are beneficial for developing a strong CBBE. We suggest a differential brand asset management strategy according to brand strength. For strong brands, it is more critical to focus on developing and occupying the primary associations in the category and strengthening dominant position in the category by creating "prototype" or "category exemplar" image. Marketing managers of weak brands should invest in creating more content or comparable attributes that can be linked to their competitors, making the brands more assimilated and comparable to the competitors. For example, when promoting new products, marketers of weak brands may take advantage of an assimilation strategy to hijack the brand associations from the market leaders. Keller, Sternthal, and Tybout (2002, p. 5) also note that "savvy marketers can hold off a competitor's point of difference by creating competitive points of parity." In contrast, managers of strong brands should be more selective in promoting their marketing activities and creating brand-related content. Because strong brands already have a well-refined set of healthy brand associations, the manager's goal should be to create content that will expand and bolster the current assets by developing primary associations in the category that are not easily hijacked by competitors.

We find that, in the long run (quarterly, in our case), the market appreciates and incorporates change in brand associative structure (e.g., associative strength with socially popular keywords, favorability of brand associations, competitive overlap of brand associations) rather than mere shock in the content volume. Thus, maintaining current brand image by managing negative associations appropriately and developing strong, favorable associations is crucial for both strong and weak brands. In addition, marketing managers should note that they can always position their brands more favorably to potential customers by associating with socially popular topics, issues, and concepts.

Contributions, Limitations, and Further Research

Our work contributes to the marketing literature in four respects. First, we conceptualize how associative information in UGC can act as proxy measures in explaining variations in stock returns. Unlike most previous studies on UGC that focus on volume and valence metrics, we further investigate the dynamics in the associative structure of textual information about a brand. To the best of our knowledge, this is the first study to establish the relationship between the information in the brand associative network and firm value.

Second, our results suggest that in the long run (at least for a quarterly window), firm valuation is more strongly related to the types of content and associative relationships than to mere volume of the content. Rather than the mere volume of the content, the volume of socially popular keywords, association valence, and competitive overlap of brand associations is more significantly related to firm value. Our findings suggests that by using customer mindset measures that capture the associative structure of brand image, we can more fully understand how the market appreciates brand equity.

Third, our work contributes to research streams in brand equity by exposing the moderating role of brand strength in the relationship between the competitive overlap of brand associations and firm value. The results suggest that for strong brands, strengthening category leadership by developing primary associations in the category, while maintaining current brand image/positioning that competitors cannot imitate, is deemed a promising signal for future cash flow. In contrast, for weak brands, creating more comparable references and attributes to competing brands in the category can be viewed as a viable strategy. The finding suggests that different brand management strategies are needed for strong versus weak brands.

Finally, the proposed social tag metrics present a new way to track CBBE using networks of user-generated keywords. To the best of our knowledge, our work is the first to quantify the information contained in social tags and investigate their informational value in the context of firm valuation. We show that social tags can be a more appropriate source to infer brand associative networks than other forms of UGC because they provide the semantic network structure of keywords, which enables us to construct metrics capturing the elements of CBBE: brand familiarity, favorability of brand associations, and competitive overlap of brand associations. We hope that the proposed proxy measures based on social tagging data will serve as a new, efficient, and effective method to track CBBE.

This article has several limitations that invite further research. Our analysis is based on a quarterly time frame because (1) accounting measures are available on a quarterly basis and (2) it is not easy to observe systematic brand associative structure change using a more granular window (e.g., weekly level). However, it would be fruitful to construct metrics using a more granular window and investigate the explanatory power of social tag metrics in daily or weekly stock returns. In addition, our model does not exclusively include a brand's specific marketing actions, such as change in a brand's communication message, new product announcements, and innovations. Rather, we capture an integrative picture of customers' perceptions of those marketing activities with social tag metrics under the reasoning that the impact of those activities are reflected in the associative structure of tags. It would be worthwhile to investigate the chain of marketing activities and customer perceptions/reactions captured by social tags, sales, and firm value.

From a methodological viewpoint, the meanings of brand associations in tagging data are sometimes ambiguous. A firm name such as "Blockbuster" can be used as a brand name or as a general descriptor. We excluded those brands from our analysis because the results can be misleading. One way to resolve this problem might be to consider intertag relationships and classify only relevant tags. Future studies relying on computational linguistics techniques to resolve such ambiguity in the data would be highly valuable. Finally, although our sample extensively covers 44 brands in 14 product categories, because we do not have sufficient social bookmarking data for small brands, our sample only represents relatively large and well-established brands. As more data become available from increasing social tagging activity on the Internet, it would be worthwhile to investigate social tags and keywords for small brands as well.

REFERENCES

Aaker, David A. and Robert Jacobson (2001), "The Value Relevance of Brand Attitude in High-Technology Markets," *Journal of Marketing Research*, 38 (November), 485–93.

——— and J. Gary Shansby (1982), "Positioning Your Product," *Business Horizons*, 25 (3), 56–62.

Alba, Joseph W. and J. Wesley Hutchinson (1987), "Dimensions of Consumer Expertise," *Journal of Consumer Research*, 13 (4), 411–53.

Ames, Morgan and Mor Naaman (2007), "Why We Tag: Motivations for Annotation in Mobile and Online Media," in *CHI '07: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: Association for Computing Machinery, 971–98.

Baron, Reuben M. and David A. Kenny (1986), "Moderator-Mediator Variables Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology*, 51 (6), 1173–82.

Bharadwaj, Sundar G., Kapil R. Tuli, and Andre Bonfrer (2011), "The Impact of Brand Quality on Shareholder Wealth," *Journal of Marketing*, 75 (September), 88–104.

Brucks, Merrie (1985), "A Typology of Consumer Knowledge Content," in *Advances in Consumer Research*, Vol. 13, Richard J. Lutz, ed. Provo, UT: Association for Consumer Research, 58–63.

Buchanan, Lauranne, Carolyn J. Simmons, and Barbara A. Bickart (1999), "Brand Equity Dilution: Retailer Display and Context Brand Effects," *Journal of Marketing Research*, 36 (August), 34–55.

Burke, Raymond R. and Thomas K. Srull (1988), "Competitive Interference and Consumer Memory for Advertising," *Journal of Consumer Research*, 15 (1), 55–68.

Carhart, Mark (1997), "On Persistence in Mutual Fund Performance," *Journal of Finance*, 52 (1), 57–82.

Chevalier, Judith and Dina Mayzlin (2006), "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research*, 43 (August), 345–54.

Cohen, Joel B. and Kanul Basu (1987), "Alternative Models of Categorization: Towards a Contingent Processing Framework," *Journal of Consumer Research*, 13 (4), 455–72.

EBizMBA.com (2010), "Top 15 Most Popular Social Bookmarking Websites," (accessed April 3, 2014), [available at http://www.ebizmba.com/articles/social-bookmarking-websites].

Ehrenberg, Andrew S.C., Neil Barnard, and John Scriven (1997), "Differentiation or Salience," *Journal of Advertising Research*, 37 (November/December), 82–91.

———, Gerald J. Goodhardt, and T. Patrick Barwise (1990), "Double Jeopardy Revisited," *Journal of Marketing*, 54 (July), 82–91.

Fama, Eugene (1998), "Market Efficiency, Long-Term Returns, and Behavioral Finance," *Journal of Financial Economics*, 49 (3), 283–306.

——— and Kenneth French (1993), "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33 (1), 3–56.

Faust, Katherine (1997), "Centrality in Affiliation Networks," *Social Networks*, 19 (2), 157–91.

Fleiss, J.L. (1971), "Measuring Nominal Scale Agreement Among Many Raters," *Psychological Bulletin*, 76 (5), 378–82.

Foster, George (1977), "Quarterly Accounting Data: Time-Series Properties and Predictive-Ability Results," *Accounting Review*, 52 (1), 1–21.

Fountain, Trevor and Mirella Lapata (2010), "Meaning Representation in Natural Language Categorization," in *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, Niels Taatgen and Hedderik van Rijn, eds. Amsterdam: Cognitive Science Society, 1916–21.

Fu, Wai-Tat, Thomas Kannampallil, Ruogo Kang, and Jibo He (2010), "Semantic Imitation in Social Tagging," *ACM Transactions on Computer-Human Interaction*, 17 (3), 1–37.

Gregan-Paxton, Jennifer and Deborah Roedder John (1997), “Consumer Learning by Analogy: A Model of Internal Knowledge Transfer,” *Journal of Consumer Research*, 24 (3), 266–84.

Griffiths Thomas L., Mark Steyvers, and Joshua B. Tenenbaum (2007), “Topics in Semantic Representation,” *Psychological Review*, 114 (2), 211–44.

Han, Young Jee, Joseph C. Nunes, and Xavier Drèze (2010), “Signaling Status with Luxury Goods: The Role of Brand Prominence,” *Journal of Marketing*, 74 (July), 15–30.

Hoberg, Gerard and Gordon Philips (2010), “Product Market Synergies and Competition in Mergers and Acquisitions: A Text-Based Analysis,” *Review of Financial Studies*, 23 (10), 3773–3811.

Ho-Dac, Nga N., Stephen J. Carson, and William L. Moore (2013), “The Effects of Positive and Negative Online Customer Reviews: Do Brand Strength and Category Maturity Matter?” *Journal of Marketing*, 77 (November), 37–53.

Hoeffler, Steve and Kevin L. Keller (2003), “The Marketing Advantages of Strong Brands,” *Journal of Brand Management*, 10 (6), 421–45.

Interbrand (2009), “Best Global Brands 2009,” (accessed April 3, 2014), [available at http://www.interbrand.com/en/best-global-brands/best-global-brands-2009.aspx]

Jacoby, Jacob, George J. Syzabillo, and Jacqeline Busato-Schach (1977), “Information Acquisition Behavior in Brand Choice Situations,” *Journal of Consumer Research*, 3 (4), 209–216.

John, Deborah Roedder, Barbara Loken, and Christopher Joiner (1998), “The Negative Impact of Extensions: Can Flagship Products Be Diluted?” *Journal of Marketing*, 62 (January), 19–32.

———, ———, Kyeongheui Kim, and Alokparna Basu Monga (2006), “Brand Concept Maps: A Methodology for Identifying Brand Association Networks,” *Journal of Marketing Research*, 43 (November), 549–63.

Kaplan, Andreas M. and Michael Haenlein (2010), “Users of the World, Unite! The Challenges and Opportunities of Social Media,” *Business Horizons*, 53 (1), 59–68.

Keller, Kevin Lane (1987), “Memory Factors in Advertising: The Effect of Advertising Retrieval Cues on Brand Evaluations,” *Journal of Consumer Research*, 14 (3), 316–33.

——— (1993), “Conceptualizing, Measuring, and Managing Customer-Based Brand Equity,” *Journal of Marketing*, 57 (January), 1–22.

———, Brian Sternthal, and Alice Tybout (2002), “Three Questions You Need to Ask About Your Brand,” *Harvard Business Review*, 80 (9), 80–89.

Körner, Christian, Dominik Benz, Andreas Hotho, Markus Strohmaier, and Gerd Stumme (2010), “Stop Thinking, Start Tagging: Tag Semantics Emerge from Collaborative Verbosity,” in *Proceedings of the 19th International Conference on World Wide Web*. New York: Association for Computing Machinery, 521–30.

Kothari, S.P., Susan Shu, and Peter D. Wysocki (2009), “Do Managers Withhold Bad News?” *Journal of Accounting Research*, 47 (1), 241–76.

LeRoy, Stephen (1989), “Efficient Capital Markets and Martingales,” *Journal of Economic Literature*, 27 (4), 1583–1621.

Liu, Yong (2006), “Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue,” *Journal of Marketing*, 70 (July), 74–89.

Luo, Xueming (2009), “Quantifying the Long-Term Impact of Negative Word of Mouth on Cash Flows and Stock Prices,” *Marketing Science*, 28 (1), 148–65.

MacInnis, Deborah J. and Kent Nakamoto (1992), “Factors That Influence Consumers’ Evaluations of Brand Extensions,” working paper, Karl Eller School of Management, University of Arizona.

Medin, Douglas L. and Marguerite M. Schaffer (1978), Context Theory of Classification Learning,” *Psychological Review*, 85 (3), 207–238.

Mika, Peter (2007), “Ontologies Are on Us: A Unified Model of Social Networks and Semantics,” *Web Semantics: Science, Services and Agents on the World Wide Web*, 5 (1), 5–15.

Mizik, Natalie and Robert Jacobson (2004), “Stock Return Response Modeling,” in *Assessing Marketing Strategy Performance*, Christine Moorman and Donald R. Lehmann, eds. Boston: Marketing Science Institute, 29–46.

——— and ——— (2008), “The Financial Value Impact of Perceptual Brand Attributes,” *Journal of Marketing Research*, 47 (February), 15–32.

——— and ——— (2009), “Financial Markets Research in Marketing,” *Journal of Marketing*, 46 (May), 320–24.

Nedungadi, Prakash (1990), “Recall and Consumer Consideration Sets: Influencing Choice Without Altering Brand Evaluations,” *Journal of Consumer Research*, 17 (3), 263–76.

Netzer, Oded, Ronen Feldman, Jacob Goldenberg, and Fresco Moshe (2012), “Mine Your Own Business: Market Structure Surveillance Through Text Mining,” *Marketing Science*, 31 (3), 521–43.

Pink, Daniel H. (2005), “Folksonomy,” *The New York Times*, (December 11), (accessed April 3, 2014), [available at http://www.nytimes.com/2005/12/11/magazine/11ideas1-21.html].

Robu, Valentin, Harry Halpin, and Hana Shepherd (2009), “Emergence of Consensus and Shared Vocabularies in Collaborative Tagging Systems,” *ACM Transactions on the Web*, 3 (4), 14:1–14:34.

Rosch, Eleanor (1973), “Natural Categories,” *Cognitive Psychology*, 4 (3), 328–35.

Smallwood, Denise E. and John Conlisk (1979), “Product Quality in Markets Where Consumers Are Imperfectly Informed,” *Quarterly Journal of Economics*, 93 (1), 1–23.

Srivastava, Rajendra K. and Allan D. Shocker (1991), “Brand Equity: A Perspective on Its Meaning and Measurement,” Report No. 91-124, Marketing Science Institute.

Strohmaier, Markus, Christian Körner, Roman Kern (2010), “Why Do Users Tag? Detecting Users’ Motivation for Tagging in Social Tagging Systems,” in *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*. Palo Alto, CA: Association for the Advancement of Artificial Intelligence, 339–42.

Sujan, Mita and James R. Bettman (1989), “The Effects of Brand Positioning Strategies on Consumers’ Brand and Category Perceptions: Some Insights from Schema Research,” *Journal of Marketing Research*, 26 (November), 454–67.

Tellis, Gerard J. and Joseph Johnson (2007), “The Value of Quality,” *Marketing Science*, 26 (6), 758–73.

Tirunillai, Seshadri and Gerard J. Tellis (2012), “Does Chatter Really Matter? Dynamics of User-Generated Content and Stock Performance,” *Marketing Science*, 31 (2), 198–215.

Web3 Technology (2012), “Market Share Quarterly Trends for Social Widgets for Websites,” (accessed April 3, 2014), [available at http://w3techs.com/technologies/history_overview/social_widget/ms/q].

White, Halbert (1980), “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity,” *Econometrica*, 48 (4), 817–38.

Wooldridge, Jeffrey (2009), *Introductory Econometrics*. Mason, OH: Thompson Higher Education.