King's Research Portal

DOI:
10.1177/0022242921995173

Document Version
Peer reviewed version

Link to publication record in King's Research Portal

Citation for published version (APA):
Roland Rust, William Rand, Ming-Hui Huang, Andrew T. Stephen, & Timur Chabuk (2021). Real-Time Brand Reputation Tracking Using Social Media. *JOURNAL OF MARKETING*, *85*(4), 21-43. https://doi.org/10.1177/0022242921995173

Citing this paper
Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights
Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

•Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
•You may not further distribute the material or use it for any profit-making activity or commercial gain
•You may freely distribute the URL identifying the publication in the Research Portal

Take down policy
If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 13. Jul. 2025

Real-Time Brand Reputation Tracking Using Social Media

***Roland T. Rust, William Rand, Ming-Hui Huang, Andrew T. Stephen, Gillian Brooks, and Timur Chabuk**

January 19, 2021

***Roland T. Rust** is Distinguished University Professor and David Bruce Smith Chair in Marketing, and Executive Director of the Center for Excellence in Service at the Robert H. Smith School of Business, University of Maryland. Address: 3451 Van Munching Hall, University of Maryland, College Park, MD 20742. Phone: 301-405-4300. Fax: 301-405-0146. Email: rrust@rhsmith.umd.edu.

William Rand is Associate Professor of Marketing at the Poole College of Management, North Carolina State University. Address: 2324 Nelson Hall, North Carolina State University, Raleigh, NC 27607. Phone: 919-515-6947. Email: wmrand@ncsu.edu.

Ming-Hui Huang is Distinguished Professor of Electronic Commerce in the Department of Information Management, College of Management, National Taiwan University. Address: Department of Information Management, College of Management, National Taiwan University, 1, Sec. 4, Roosevelt Rd., Taipei 10617, Taiwan. Phone: +886-2-33661185, Email: huangmh@ntu.edu.tw.

Andrew T. Stephen is Associate Dean of Research and L'Oreal Professor of Marketing at the Säid Business School at the University of Oxford. Address: Park End Street, Oxford OX1 1HP, United Kingdom. Email: andrew.stephen@sbs.ox.ac.uk.

Gillian Brooks is Assistant Professor in Marketing at King's Business School at King's College London. Address: Bush House, 30 Aldwych, London WC2B 4BG, United Kingdom. Email: gillian.brooks@kcl.ac.uk.

Timur Chabuk is Vice President of Machine Learning and Advanced Analytics at Perceptronics Solutions, Inc. Address: 3141 Fairview Park Drive, Suite 415, Falls Church, VA 22042. Phone: 703-342-4660. Fax: 703-342-4661. Email: timc@percsolutions.com.

The first three authors contributed equally to the paper. This research was supported by the University of Oxford's Centre for Corporate Reputation, and the Center for Excellence in Service at the University of Maryland, and grants (104-2410-H-002-142-MY3, 106-2410-H-002-056-MY3, 107-2410-H-002-115-MY3) from the Ministry of Science and Technology, Taiwan. Thanks to OssaLabs for their collaboration in building and maintaining the database, to YouGov for access to data used in this study, and to James Davies and the IT staff at the Oxford's Säid Business School for their help in setting up the website. We would also like to thank Bradley Senft who helped with the development and implementation of the software to collect and process the social media. We also thank Rupert Younger, Rowena Olegario, and participants at the Oxford Corporate Reputation Forum and the American Marketing Association Winter Conference for their helpful comments.

Real-Time Brand Reputation Tracking Using Social Media

Abstract

How can we know what stakeholders think and feel about brands in real-time and over time? Most brand reputation measures are at the aggregate level (e.g., the Interbrand "Best Global Brands" list) or rely on customer brand perception surveys on a periodical basis (e.g., the Y&R Brand Asset Valuator). To answer this question, brand reputation measures must capture the voice of the stakeholders (not just ratings on brand attributes), must reflect important brand events in real-time, and must connect to a brand's financial value to the firm. This paper develops a new social media-based brand reputation tracker by mining Twitter comments for the world's top 100 brands using Rust-Zeithaml-Lemon's value-brand-relationship framework, on a weekly, monthly, and quarterly basis. The paper demonstrates that brand reputation can be *monitored* in real-time and longitudinally, *managed* by leveraging the reciprocal and virtuous relationships between the drivers, and *connected* to firm financial performance. The resulting measures are housed in an online longitudinal database and may be accessed by brand reputation researchers.

Key Words: Brand Reputation Tracker, Value Driver, Brand Driver, Relationship Driver, Social Media Mining, Corporate Reputation, Abnormal Return, Customer Equity, Twitter

Statement of Intended Contribution

How can we make brand reputation managerially actionable and financially accountable in real-time? The brand reputation tracker developed in this paper provides a solution to this question. Based on the customer equity framework, we develop a three-driver tracker (i.e., value, brand, and relationship) that comprises how stakeholders think, feel, and talk about a brand.

We collected longitudinal Twitter data for 100 leading global brands, from July 1, 2016 to the end of 2018, on a weekly basis. Volumes, sentiments, levels, and unanticipated components of the three drivers and their sub-drivers were all measured.

We first establish the causal relationship among the three drivers via a dynamic multivariate panel vector autoregression (VAR) model that allows a firm to manage brand reputation based on its comparative brand advantage. A firm can leverage the Brand-Value reciprocal relationship or the Brand-Relationship-Value virtuous circle, based on which driver for which the firm has better leverage.

We further establish the financial accountability of the drivers. Using a stock return response model for abnormal stock returns, we show that the unanticipated components of the three drivers provide additional information to the firm's abnormal stock returns.

Our study contributes to brand management by making brand reputation managerially actionable and financially accountable in real-time and over time, connects social media mining and marketing, bridges the corporate reputation literature with marketing actions, renews the customer equity literature with new social media tracking methodology, and makes brand reputation measures accessible in an almost real-time manner.

The tracker is validated using social media data from Facebook and Instagram, consumer perceptual data from YouGov, and aggregate brand data from Interbrand, Forbes, and Brandz.

The resulting social media-based brand reputation tracker is highly time-sensitive and context-specific, allowing firms to respond quickly to market stimuli. The tracker connects the voice of stakeholders to brand reputation drivers that drive firm financial performance.

Real-Time Brand Reputation Tracking Using Social Media

Social media are playing an increasingly important role in enabling conversations about brands (Fossen and Schweidel 2019; Hewett et al. 2016; Kubler, Colicev, and Pauwels 2020). In the social media age, listening online to how brands are talked about is critical for brand management because the comments are from people who care about brands, and they are real-time and dynamic.

Many brand measures currently exist, but they are mostly at the aggregate level, are survey-based, and are typically only available annually. Examples of these are the InterBrand "Best Global Brands" list, the Forbes "The World's Most Valuable Brands" list, the Kantar Millward Brown "BrandZ Top 100 Global Brands" list, as well as many others. These brand measures rank leading companies in terms of their overall brand value.

Other brand measures (e.g., the Y&R Brand Asset Valuator, BAV) measure multiple dimensions of customer brand perceptions using surveys, but the actionability of the dimensions is limited, given that the measured dimensions (e.g., differentiation, relevance, esteem, and knowledge) do not map in a straightforward way to strategic business decisions. Nevertheless, researchers who have attempted to research brand perceptions over time have resorted to such measures (and BAV in particular) to conduct their research (e.g., Huang and Dev 2020; Mizik 2014; Mizik and Jacobson 2008; Stahl et al. 2012; Tavassoli, Sorescu, and Chandy 2014).

To provide a brand measure that exploits this new social environment, we develop a real-time longitudinal brand tracker using social media. The tracker provides a new window into what stakeholders think and feel about brands, in their own voice. Such data provide timely, actionable information about how firms should manage brands. The tracker facilitates both longitudinal analysis and exploration of brands in a real-time and granular way.

This new metric is compiled by mining Twitter data for 100 leading global brands for comments on specific drivers and sub-drivers based on the customer equity framework (Rust, Zeithaml, and Lemon 2000; Rust, Lemon, and Zeithaml 2004). The database is available to all academic researchers for performing research on the antecedents and consequences of brand reputation changes over time, at a level of granularity not previously available. Table W1 in the

Web Appendix compares the proposed brand reputation tracker with existing brand measures to illustrate the gap.

Our brand reputation tracker contributes to brand management by providing actionable data and analytics in a straightforward manner that facilitates strategic brand decisions. Results from our dynamic multivariate panel vector autoregression (VAR) model show that the three drivers have a Brand-Value reciprocal relationship as well as a Brand-Relationship-Value virtuous circle. A firm thus can leverage either the reciprocal relationship or the virtuous circle, based on whichever driver(s) for which the firm has better leverage. We further demonstrate that the three drivers have real-time, short-term, and longer-term impact, collectively and individually, on abnormal stock returns. The findings provide important implications for managing brands based on the dynamics and the tempos of the drivers.

The tracker also builds on the literature connecting social media mining and marketing. Data mining of Twitter feeds and other social media, for example, have been used to measure brand sentiment (Hewett et al. 2016) and other marketing-relevant metrics (e.g., Schweidel and Moe 2014; Tirunillai and Tellis 2012, 2014). By applying our methodology to three social media platforms—Twitter, Facebook, and Instagram, that are distinct in data and nature of interaction, we demonstrate that the three sets of trackers converge—evidence that social media mining can be used to track brand reputation in real-time, and its fluctuations reflect important brand events.

Our work also adds to the literature and management of corporate reputation by bridging brand reputation and corporate reputation, and by providing actionable drivers for managing corporate reputation. Brand reputation is similar to corporate reputation when a firm uses a branded house strategy (e.g., Google), and is a component of corporate reputation when a firm uses a house of brands strategy (e.g., P&G). While the corporate reputation literature has established the importance of corporate reputation as a strategic asset (Ferguson, Deephouse, and Ferguson 2000; Fombrun and Shanley 1990; Lange, Lee, and Dai 2011; Rindova et al. 2005) and as a driver of financial performance and firm value (Pfarrer, Pollock, and Rindova 2010; Raithel and Schwaiger 2015), that literature has tended not to focus on marketing actions. We show that the time series of brand reputation, drivers, and sub-drivers captures important brand and firm events; for example, the fluctuations of Facebook's Brand and Value drivers time series coincide with its unauthorized account licensing scandal in March 2018, while Google's innovative sub-driver time series captures its late 2018 announcement of updating many algorithms. This shows that brand reputation, a key element of, and sometimes equal to, corporate reputation, can be monitored using the tracker and managed by the drivers, which helps drive firm financial performance.

The tracker also contributes to the customer equity literature. To date, most customer equity research has either relied on cross-sectional surveys for data collection (e.g., Gao, Melero-Polo, and Sese 2020; Ou et al. 2014; Ou, Verhoef, and Wiesel 2017; Rust, Lemon and Zeithaml 2004; Vogel, Evanschitzky, and Ramaseshan 2008) or sacrificed granularity by using aggregate acquisition and retention statistics (e.g., Gupta, Lehmann, and Stuart 2004; Kumar and Shah 2009; Villanueva, Yoo, and Hanssens 2008; Wiesel, Skiera, and Villanueva 2008). We show that the tracker correlates significantly positively with other well-known aggregate brand rankings, such as Interbrand, Forbes, and Brandz, and correlates significantly positively with YouGov's daily brand measures of brand word-of-mouth and brand buzz. This demonstrates that the customer equity literature can be renewed with our new social media tracking methodology that links modern social media marketing actions to customer value, and facilitates longitudinal analysis and exploration of customer equity drivers in a more granular way. We also broaden the conceptual nature of customer equity to include all stakeholders, rather than just current and future customers of the brand.

In the following sections, first, we conceptualize brand reputation based on multiple conceptual sources. Second, we develop our social media tracking method to track and monitor brand reputation. Third, we present empirical evidence based on 130 weeks' tracking data that show this brand reputation tracker can be used to monitor and manage brand reputation and competition dynamics, and is accountable for a firm's abnormal stock returns. Fourth, we generalize the tracker to multiple social media platforms and validate the tracker using other brand measures. Finally, we discuss implications for brand managers and researchers.

Conceptualizing Brand Reputation

Brand Reputation

We define brand reputation as the overall impression of what stakeholders think, feel, and talk about a brand. This is typically due to brand events that affect firm financial performance. This definition has the following characteristics: 1) it is about all stakeholders (current and potential customers, employees, partners, and investors), not just the current or potential customers, 2) it has thinking, feeling, and talking components (not just knowledge about brands), 3) it can reflect actual brand events (e.g., controllable marketing activities or uncontrollable public events about brands), and 4) it connects to firm financial performance.

We highlight "stakeholders" to indicate a broader view of customers, which is an integration of the corporate reputation and the marketing literature. This view is customer-centric, but is broader, considering all stakeholders as customers, including current (e.g., current and churning current customers), potential (e.g., competitors' customers and non-customers), internal (e.g., employees) and external customers (e.g., investors and partners). For example, Berry (2000) considers both external customers and internal employees as relevant for building service brand equity. Hanssens, Rust, and Srivastava (2009) similarly urge the need to take a broader view of the stakeholders of marketing strategy by including the investor community as a customer. Nguyen, Calantone, and Krishnan (2020) demonstrate that social media emotional word-of-mouth influences investors' decisions on holding a firm's stocks. This broader view of customers reflects that brand reputation can be perceived by non-relationship brand stakeholders who can influence the brand's financial performance.

Brand reputation should be based on whatever stakeholders say about a brand, meaning what is explicitly expressed about their thinking and feeling, not what is implicitly inferred. Stakeholders on social media can talk about anything about a brand. It can be brand experience, opinions about brand events, or simply personal sentiments about a brand. It can be positive, neutral, or negative in varying degrees. The overall impression of a brand may be summarized by what stakeholders say about a brand on social media (Hewett et al. 2016).

Changes in brand reputation typically result from actual brand events. These brand events can be controllable marketing actions and activities, as well as uncontrollable public events about brands. This characteristic emphasizes the actionability of brand reputation, allowing marketers to actively manage a brand's reputation and to track the reputation for risk and crisis management. Such actionability is the focus of models of return on marketing (Rust, Lemon, and Zeithaml 2004).

Brand reputation should be value relevant, that is, connecting to a firm's financial performance. This value relevance reflects investors' expectations about the financial value of current, potential, internal, and external customers to the firm. Value relevant brand reputation thus is a corporate asset and a driver of firm financial performance, as defined in the management literature (e.g., Ferguson, Deephouse, and Ferguson 2000).

Below, we compare brand reputation with other related concepts in the management and the marketing literatures to illustrate the conceptual nature of brand reputation. Figure 1 illustrates the relationship between brand reputation and other related concepts. It shows that our brand reputation concept lies in the intersection between the concepts of corporate reputation, brand equity, and customer equity. It has consequences to all of the firm's stakeholders (not just customers), focuses on brand thinking and feeling, and emphasizes marketing actions to drive firm value.

Corporate Reputation

In the management literature, corporate reputation is generally defined as an overall appraisal of a company by its stakeholders, which is the result of the company's past actions and predictions about the company's future (Ferguson, Deephouse, and Ferguson 2000). Such an overall appraisal can be thinking and/or feeling based, for example, the quality or capability (Boivie, Graffin, and Gentry 2015), the admiration (Dowling 2016), or knowledge and emotions (Hall 1992). With its broad scope, taking all stakeholders into consideration (Argenti and Druckenmiller 2004), and with the appraisal being at the corporate level, corporate reputation has a broader and higher-level focus on components such as leadership, social responsibility, product and service, workplace/employee, and management/financial performance, and their consequences on firm performance (Parker, Krause, and Devers 2019; Raithel and Schwaiger 2015; Roberts and Dowling 2002).

Compared to corporate reputation, which is the overall appraisal of a firm held by stakeholders, and has consequences for firm performance, brand reputation is similar to corporate reputation when the company uses a branded house strategy (e.g., FedEx, Google, and Apple), while it is a component of corporate reputation when the company uses a house of brands strategy, i.e., having multiple brands to represent the company (e.g., P&G).

Brand Equity

The brand equity literature emphasizes customers' overall impression about a brand, even though there is disagreement on the scope of the impression. Some take a broader view. Aaker's (1995) brand equity definition includes mainly brand loyalty, brand awareness, perceived quality, and brand associations, constituting various assets and liabilities associated with a brand and value derived from them. Keller's (1993) customer-based brand equity defines brand equity in terms of the individual consumer's brand knowledge, which influences the consumer's reaction to the brand's marketing mix. Brand knowledge is broad, including all knowledge about a brand that the consumer has, such as higher-level brand awareness and brand image, and all lower-levels associations, such as brand recall and brand associations. Alternatively, some take a narrower view. Rust, Zeithaml, and Lemon (2000) and Rust, Lemon, and Zeithaml (2004) define brand equity as customers' subjective or emotional appraisal of a brand, above and beyond its objectively perceived value. This inconsistency in the scope of brand equity motivates some studies trying to bridge brand equity with customer equity (e.g., Gani and Grobler 2019; Leone et al. 2006).

Brand reputation reflects both the knowledge and emotions held by all stakeholders about a brand. Our broader view aims to capture how a broader set of stakeholders thinks, feels, and talks about a brand, not limited to whether they are current or potential customers.

Customer Equity

Reflecting a customer lifetime value view, the existing customer equity literature focuses on the contribution of a brand's single stakeholder, i.e., customers, to a firm. Rust, Lemon, and Zeithaml (2004) define customer equity as the total of the discounted lifetime values summed over all of the firm's current and potential customers (p. 110). Blattberg and Deighton (1996) define customer equity as the discounted profit stream, and explore customer equity in terms of the optimal trade-off between acquiring and retaining customers. Gupta, Lehmann, and Stuart (2004) similarly define the value of the customer base as the expected sum of discounted future earnings.

To enhance customer value, the existing studies focus on either customer equity drivers (mostly perceptual) or marketing investments. For customer equity drivers, it has been shown that the three drivers in the customer equity framework—value, brand, and relationship (Rust, Zeithaml, and Lemon 2000) can improve customer value, as captured by loyalty intentions (Vogel, Evanschitzky, and Ramaseshan 2008), customer loyalty (Ou, Verhoef, and Wiesel 2017; Ou et al. 2014), and customer experience quality (Go, Melero-Polo, and Sese 2020). For marketing investments, the research stream on marketing-finance interface further links customer equity to shareholder value (Gupta, Lehmann, and Stuart 2004; Kumar and Shah 2009; Schulze, Skiera, and Wiesel 2012; Skiera, Bermes, and Horn 2011; Wiesel, Skiera, and Villanueva 2008).

Brand reputation shares commonness with customer equity in that both can be driven by the customer equity drivers and linked to firm financial performance. The focus is on what firms do, in terms of customer equity drivers or marketing investments, to influence firm financial performance.

The Social Media Tracking Method

We use a multi-stage, theory-data iterative process to develop the tracker, detailed below and shown in Figure 2. We briefly summarize these steps and explain the logic behind them, and describe this process in detail in the next sections.

- *Theory-driven text mining*. Use the customer equity framework to establish a 3-driver tracker. This makes the brand reputation outcome theory-driven and thus explainable.
- *Managerially actionable sub-drivers*. Develop the sub-drivers of brand reputation based on managerial actionability. This makes the drivers managerially relevant and actionable.

- *Customized dictionaries*. Generate dictionaries for the sub-drivers based on real-world stakeholders' own words. This extracts what they think and feel about a brand in conversation and in context.
- *Real-time brand reputation tracking*. Collect data in real-time and from multiple social media platforms. This allows quick responses from firms and makes the methodology cross-platform applicable.
- *Leverageable driver synergies*. Provide evidence for the dynamics and tempos of the three drivers. This provides managerial and theoretical insight as to the internal relationships between the drivers and how a firm can leverage the synergies of the three drivers.
- *Financially accountable brand reputation*. Establish the accountability of the tracker for firm abnormal returns. This makes brand reputation and its drivers financially accountable, rather than just time-series of brand reputation fluctuations.

Establish a Three-Driver Tracker

We employ the driver structure from the customer equity framework (Rust, Zeithaml and Lemon 2000; Rust, Lemon and Zeithaml 2004) to develop our tracker. This framework organizes the factors driving customer lifetime value and contribution to the firm into three main drivers, which themselves may be broken down further into sub-drivers. Value equity is the rational and objective aspects of a brand, such as quality and price. Brand equity is the subjective feeling that a customer has about the brand, such as brand sentiment and brand image. Relationship equity is the ties between the customer and the brand, above and beyond the value equity and brand equity, such as brand community building and personal connection.

The following considerations contribute to the choice of this framework: First, the conceptual attraction of this framework has been well-recognized in the academic community, and recognized with several article and book awards. Second, the three customer equity drivers have been validated conceptually and empirically in many subsequent studies, using data from multiple countries, incorporating both perceptual survey and behavioral data, considering industries, firms, and consumer characteristics, and gauged using various firm performance variables (e.g., loyalty intentions, future sales, customer experience quality) (Gao, Melero-Polo, and Sese 2020; Leone et al. 2006; Ou, Verhoef, and Wiesel 2017; Ou et al. 2014; Vogel, Evanschitzky, and Ramaseshan 2008). Third, it was designed to map to strategic expenditures, and thus has high managerial actionability. The drivers and sub-drivers have been shown to link to return on marketing, an important characteristic that helps link brand reputation to firm financial performance. Its managerial relevance is reflected in this framework being applied at many leading companies worldwide. Fourth, the value and brand drivers together capture the thinking and feeling aspects of brand reputation. It is the consensus that a brand metric should include both aspects (Huang and Dev 2020; Lovett, Peres, and Shachar 2013; Vogel, Evanschitzky, and Ramaseshan 2008).

Choosing the Social Media

There are many different social media platforms that are used to discuss brands, for example Twitter, Facebook, Instagram, and many others. To construct a dynamic tracker of sentiment about brands on social media, we choose Twitter for the following considerations: 1) Most Twitter accounts are public, meaning that conversations on Twitter have a larger impact on public perception of the brand, while many other social media platforms, such as Facebook, default to private communications; 2) most brands maintain an active presence on Twitter, which means that brand conversations are continuously updated and are available for public access; and 3) Twitter provides a publicly available Application Programming Interface (API) that can identify conversations about the brands, for example, using username “@coach,” rather than “coach” to identify conversations about the brand ensures precision (which is the number of relevant tweets retrieved divided by the number of all tweets).

Choosing Brands to Monitor

The choice of brands to be monitored is based on various prominent industry rankings on brands. These rankings included Forbes’ World Best Brands, BrandZ’s Top 100 Most Valuable Brands, Interbrand’s Best Global Brands, CoreBrand’s Top 100 Brand Power Rankings, Credit Suisse Research Institute’s Great Brands of Tomorrow, Ad Age’s Social Media Brand Ranking

Top 10, UTA Brand Studio's Brand Dependence Index, and Reputation Institute's Global Reputation Pulse U.S. Top 15.

Once all the brands are tabulated, any brand that appears twice or more across the lists is added to our database as a brand to be included in the tracker. Table W2 in the Web Appendix lists the brands included in the tracker. The database consists of 100 global brands across a broad range of industries such as manufacturing, wholesale trade, retail trade, transportation and warehousing, information, finance and insurance, professional, scientific, and technical services, and accommodation and food services. Both Internet brands and traditional brands, and both corporate brands and individual brands are included.

Collecting Tweets Using Twitter Username

After the list of brands is established, we identify the top Twitter username (i.e., tweeter handle) associated with each brand. In deciding which handle to use, we apply the rule of the "top returned handle," that is, the brand handle (ignoring sub-handles or regional handles) returned in the top search result by searching the brand name on Twitter.com. If no corporate brand is found, we check @brandname to see if it is a valid handle and double-check if users mention @brandname tweets at least once in the last month. The technical details of the Twitter data collection are shown in the Web Appendix, and Table W2 in the Web Appendix lists the brand names and handles used for collecting the data.

Table 1 summarizes the final 11 sub-drivers for the three brand reputation drivers, including their conceptual descriptions and the final positive and negative dictionaries used in the data collection. Specifically, they are: Value driver (price, service quality, and goods quality), Brand driver (cool, exciting, innovative, and social responsibility), and Relationship driver (community, friendly, personal relationships, and trustworthy).

The 11 sub-drivers and their dictionaries are theoretically derived and empirically validated by multiple rounds of data collection and evaluation. They capture nicely the social media language and technology, while preserving the conceptual nature of the three brand reputation drivers as laid out in Rust, Lemon, and Zeithaml's (2004) framework. For example, for the Value driver, they have quality and price as the sub-drivers, and we further refine the quality sub-driver into service quality and goods quality, a reflection of the service economy and the distinctiveness of service quality from goods quality. The dictionaries of the three sub-drivers contain keywords that are unique in a social media setting and are from stakeholders' daily language, such as "joy" for the price sub-driver, "lazi" for the service quality sub-driver, and "beauty" for the goods quality sub-driver.

For the Brand driver, RLZ have corporate citizenship and ethical standards as part of the sub-drivers, and we have social responsibility as one of the sub-drivers. Our Brand driver is the most social media-centric driver, with three of the four sub-drivers reflecting stakeholders' usage of social media language in expressing their thinking and feeling about brands, such as cool, exciting, and innovative. The dictionaries of the sub-drivers also reveal the language stakeholders use, such as "sexi" for the cool sub-driver, "thrill" for the exciting sub-driver, "intellig" for the innovative sub-driver, and "give" for the social responsibility sub-driver.

For the Relationship driver, RLZ have community as one of the sub-drivers, and we add friendly, personal relationships, and trustworthy to capture that the new information and communication technologies connect stakeholders more closely to companies and their brands. These sub-drivers include both the interaction and communication process of a relationship, as well as the trustworthy outcome of a relationship. The dictionaries of the sub-drivers are unique in suggesting new terms when communicating with stakeholders, such as "famili" for the community sub-driver, "open" for the friendly sub-driver, "intim" for the personal relationships sub-driver, and "transpar" for the trustworthy sub-driver.

Empirical Evidence

The dataset covers the week of July 1, 2016 to the week of December 31, 2018, 130 weeks in total.[1] The brand panel data contain 13,000 brand-week observations of 100 unique brands.[2] We measure volume and sentiment on the three drivers and sub-drivers. Table 2 presents the

---

[1] The data files are available for weekly, monthly and quarterly data. We use weekly data for the subsequent analyses. The database may be accessed from the website of the Centre for Corporate Reputation at Oxford's Saïd School of Business.

[2] There are 7 Internet brands of Alibaba, Amazon, Facebook, Google, Twitter, Yahoo, and eBay (910 brand-week observations) that are scaled separately from the remaining 93 brands in the dataset due to Twitter tweets of Internet brands being much more numerous, which may distort comparisons between Internet and traditional brands.

mean, standardization deviation, minimum, maximum, and correlations among the overall brand reputation, drivers and sub-drivers. All pairs of correlations are significant at the .000 level but are moderate in effect, indicating a balance between representativeness and uniqueness.

Tracking Brand Reputation and Competition

Monitoring brand engagement. Brand tweet volume can be viewed as how engaged stakeholders are with a brand. The more discussion a brand can generate, the more engaged it is with its stakeholders. Volume of social media discussion (e.g., tweets and retweets on Twitter) is considered to capture social media engagement (Colicev et al. 2018). It has also been shown to impact brand financial results (Kumar et al. 2013).

In our dataset, the average number of tweets collected per week per brand was 14,102 (SD = 46,310, min = 0, max = 1,660,963), which is substantial, indicating that the brand engagement on social media is high. This varies substantially by brand, however. For example, the brands with the largest mean weekly tweet volume were Amazon (mean = 1,660,963, the week of February 23, 2018), T-Mobile (mean = 1,180,385, the week of November 18, 2016), and Google (mean = 824,432, the week of February 23, 2018).[3] From the Black Friday of November 18 to 20, 2016, T-Mobile rolled out a Magenta Friday promotion, offering two additional lines free to both existing and new customers. T-Mobile has 1.4M Twitter followers, and this promotion no doubt generated hot discussion. By comparison, examples of brands with the smallest mean weekly tweet volumes were HSBC, Kraft-Heinz, and Canon, the latter of which has zero tweets. This variation is not surprising, given that the decisions to tweet about a given brand will be driven by various factors, including some that are related to the brand, but also factors that are person-related or intrinsic to the individual (e.g., Toubia and Stephen 2013).

Table W3 in the Web Appendix shows that among the three drivers across all brands, in terms of average volume, the Brand driver (N = 387) has a higher volume than the Value driver (N = 303) and the Relationship driver (N = 165). This pattern in engagement volume among the three drivers may imply that the Brand driver captures brand events more quickly and closely.

[3] The first day of the week is Friday, following the stock week practice, e.g., the Center for Research in Security Prices (CRSP) data and Fama and French data.

Monitoring Brand Sentiment. In Table W3 in the Web Appendix, we show the frequencies of net, positive, and negative words, as defined in our dictionaries, and calculate the positive to negative ratio for brand reputation, drivers, and sub-drivers. The higher the ratio, the more positive sentiment of the reputation, drivers, and sub-drivers. We can see that, in terms of sentiment ratio (proportion of positive to negative volume), the general sentiment is positive (all greater than 1). For brand reputation, the brands in the tracker, in general, have positive sentiment (3.91). This is understandable because they are top brands worldwide. For drivers, the Brand driver is the most positive (13.09), followed by the Relationship driver (5.58), with the Value driver having the least positive sentiment (2.34). For sub-drivers, the top three positive sub-drivers are exciting (20.28), cool (15.00), and innovative (14.88), whereas the least positive sub-drivers are price (1.45), friendly (3.65), and community (4.36).

For sub-drivers of the Value driver, price (1.45) is relatively ambivalent, and service quality (7.35) and goods quality (5.23) are both quite positive. This indicates that people talk about price for both positive and negative reasons, but mainly talk about quality when it is positive.

For sub-drivers of the Brand driver, the sentiment of all sub-drivers is generally positive; especially for exciting, which has the highest ratio (20.28) among all sub-drivers, indicating it is a powerful driving force for brand reputation. The second most influential sub-driver is innovative (14.88). It is the most talked about sub-driver, and is highly positive.

For sub-drivers of the Relationship driver, the most positive sub-driver is personal relationships (12.38), suggesting the importance of establishing a personal relationship with stakeholders. When people talk about this, it is overwhelmingly positive.

Monitoring competition. The brand reputation tracker can be used for monitoring brand fluctuations and tracking competitive dynamics at the brand reputation, driver, and sub-driver levels. We illustrate this use using two competitive dyads: two technology service brands, Facebook and Google, and two technology goods brands, Apple and Samsung, for 2018. Results and discussions of the Apple and Samsung dyad are presented in the Web Appendix: Monitoring Competition: Apple vs. Samsung, and Figures W1, W2a-W2c, and W3a-W3i.

Figure 3 shows the time series of their brand reputation (blue line for Facebook and red line for Google), Figures 4a-4c show the time series of the three drivers, and Figures 5a-5d show the time series of selected sub-drivers.

The brand reputation time series show that Facebook appears to have higher brand reputation than Google for the first three quarters of 2018. In the fourth quarter, the gap decreases.

The Brand (Figure 4a, blue line) and Value (Figure 4b, blue line) drivers time series of Facebook show that there was a negative spike from the week of March 19, with the Brand driver scores plummeting from .575 to .045 to -.218 in two weeks. This coincided with the revelation of the scandal that Facebook permitted the unauthorized licensing of 30 million people's accounts to Cambridge Analytica, a data firm used by Donald Trump's 2016 Presidential campaign to target voters. From the week of September 14 to the end of the year, Facebook's Brand driver scores reached a long depression with an average of -.332, compared with an average of .029 the month before it. This corresponded with the event that in September, fifty million Facebook accounts and sensitive personal data were hijacked. The larger scale and more severe data leakage of this negative event are reflected in the more enduring plunge of the Brand driver, indicating that stakeholders are quite concerned about the consequences of this significant personal data hijacking. According to Facebook's official news, they saw that unusual activities began on September 14, and on September 28 the news came out. The brand reputation tracker captures this negative event in real time, as well as its carryover effect.

The Relationship driver time series of Google (Figure 4c, red line) reveals that Google in general underperforms Facebook in this driver. This is understandable since Facebook is a social networking platform for people to establish and maintain their relationships. Nevertheless, in the last week of September and the first week of October, Google had its second highest Relationship driver score (.240) of the year and surpassed Facebook in this driver. September 27th was Google's 20th anniversary and the company also updated many algorithms of the important services it provides.

We further investigate which sub-drivers are most accountable for the ups and downs of the brand reputation, shown in Figures 5a-5d. Figure 5a shows that Facebook outperformed Google in the exciting sub-driver, whereas Figure 5b shows that Google did a better job in the innovative sub-driver, especially in the fourth quarter of 2018 when it updated many algorithms. Both sub-drivers capture stakeholders' differential perceptions about the two brands.

Figure 5c shows that Facebook mainly bested Google in the personal relationships sub-driver, but the brand is not considered more trustworthy than Google, as shown in Figure 5d. This indicates that the trustworthy sub-driver can be an opportunity for Google to capitalize on, but should be a pain point for Facebook to deal with, especially with all the data leakage negative events.

Managing the Dynamics of Brand Reputation Drivers

We have shown that the tracker reflects important brand events and can be used to monitor competition. In this section, we further examine the internal relationship of the three drivers for managing brand reputation.

Dynamic multivariate VAR model. We use a rigorous dynamic multivariate panel vector autoregression (VAR) model, estimated with generalized method of moments (GMM), to simultaneously estimate the three drivers as a system of equations (Abrigo and Love 2015; Love and Ziccino 2006). The estimator is dynamic, as the current realization of the endogenous variables (i.e., value_t, brand_t, and relationship_t) is influenced by their past values (i.e., value_{t-1}, brand_{t-1}, and $\text{relationship}_{t-1}$). The inclusion of the lagged one-period endogenous variables considers the cumulative effect of the drivers over time.

In the model, the predictors include the lagged one-period values of the three endogenous drivers ($\mathbf{Y}_{\text{it-1}}$). The three drivers are Helmert transformed (i.e., forward orthogonal deviation) to remove the brand-specific fixed effects. Equation (1) shows the dynamic VAR model:

(1) $\mathbf{Y}_{\text{it}} = \mathbf{Y}_{\text{it}-1}\boldsymbol{\alpha} + \mathbf{u}_{\text{i}} + \mathbf{e}_{\text{it}}$,

where

i = Brand (there are 100 brands)

t = Week (there are 130 data weeks)

$\mathbf{Y}_{\text{it}}$ = A (1 *x* 3) vector of endogenous variables (i.e., value, brand, and relationship drivers)

$\mathbf{u}_{\text{i}}$ = A (1 *x* 3) vector of endogenous variable-specific brand fixed-effects

$\mathbf{e}_{it}$ = A (1 *x* 3) vector of idiosyncratic errors

$\boldsymbol{\alpha}$ = A (3 *x* 3) matrix of parameters for endogenous variables to be estimated

First, we examine whether each of the three drivers is stationary using a Fisher-type test (Choi 2001). The test has the null hypothesis that all the brand time series contain a unit root. It assumes the data are generated by a first-order autoregressive (AR(1)) process; thus, we specify an augmented Dickey Fuller (ADF) unit-root test on each brand with one lag of the first-differenced driver to remove the higher-order autoregressive components of the series. To mitigate the impact of cross-sectional dependence, we also follow Levin, Lin, and Chu (2002)'s suggestion to demean the data. All test statistics for the three drivers, respectively, are significant at .001 level, which reject the null hypothesis of having a unit root. Hence, the test results support that the three drivers are stationary.

Second, we carried out model selection tests to determine the optimal lag-order for the model. The results suggest that the first-order panel VAR minimizes MBIC (-226.449), MAIC (-26.400), and MQIC (-93.451), compared to the second-order (MBIC = -158.123; MAIC = -24.758; MQIC = -69.459) and the third-order (MBIC = -78.305; MAIC = -11.622; MQIC = -33.972) models. This is expected, given that we have weekly data and discussion about a brand on Twitter changes rapidly and frequently.

Third, we estimated the first-order panel VAR using GMM-style instruments as in Holtz-Eakin, Newey, and Rosen (1988). Table 3 presents the results. We find that the Value driver is influenced by its own lagged value (.214, p = .000), the lagged Brand driver (.111, p = .000), and the lagged Relationship driver (.122, p = .000). The Brand driver is influenced by its own lagged value (.289, p = .000), and is marginally influenced by the lagged Value driver (.030, p =.063). The Relationship driver is influenced by its own lagged value (.286, p =.000) and the Brand driver (.068, p = .013). Hansen's J statistic is near zero, confirming that the model is not over-identified.

The Granger causality tests[4] confirm that the Value driver marginally Granger-causes the Brand driver (χ^2 = 3.461, p = .063), the Brand driver Granger-causes the Value driver (χ^2 =

[4] The Granger causality test is referred to more accurately as a prediction test that is a necessary but not sufficient condition for causality.

11.951, p = .001) and the Relationship driver (χ^2 = 6.144, p = .013), and the Relationship driver Granger-causes the Value driver (χ^2 = 13.560, p = .000).

We calculate the impulse response function (IRF) confidence intervals using 200 Monte Carlo draws based on the estimated model. Figures W4a-W4d show the relevant IRF figures. The shaded area is a 95% confidence band. The IRF figures show that in general the inter-driver effects level off in about 5-6 weeks.

For the impact of the Brand driver on the Value driver, a shock on the Brand driver creates a short-term (lagged one-week) surge on the Value driver, and this surge gradually levels off in five weeks (Figure W4a). For the impact of the Value driver on the Brand driver, a shock on the Value driver has a real-time positive impact on the Brand driver. Although it levels off quickly (Figure W4b), its full effect dissipates gradually over 4-5 weeks.

For the impact of the Relationship driver on the Value driver, a shock on the Relationship driver has a short-term positive impact on the Value driver (Figure W4c). Its impact is smaller than the impact of the Brand driver (Figure W4a) but is about equally persistent.

For the impact of the Brand driver on the Relationship driver, a shock on the Brand driver has a real-time positive impact on the Relationship driver, which levels off slower than the impact of the Value driver on the Brand driver, and persists longer, for about 5-6 weeks (Figure W4d).

Dynamics of the three drivers. The customer equity framework considers that the three drivers together constitute the bonds that hold the customer to the brand, but it does not specify how the three drivers causally relate to each other. Our empirical examination thus provides original empirical evidence regarding the dynamics of the three drivers. The three rectangular boxes and the white arrows in Figure 6 illustrate the mutual impacts of the drivers. The outer gray arrows depict the financial impact of the three drivers, which will be discussed in the next section.

Two relationships emerge. First, we find a reciprocal relationship between the Brand and the Value drivers, with the impact from the Brand driver to the Value driver being stronger than the reverse. Second, we find a virtuous circle among the three drivers, from the Brand driver to the Relationship driver, from the Relationship driver to the Value driver, and finally from the

Value driver back to the Brand driver. The IRF figures further reveal different tempos of the carryover effects among the three drivers.

Together, the two relationships expand our knowledge of how the customer equity drivers relate to each other over time, and provide rich implications for managing brand reputation. We discuss their managerial implications in the discussion section.

Establishing the Financial Accountability of the Tracker

In an earlier section, we show that the time series of the brand reputation tracker can capture important brand events. In this section, we further demonstrate that the unanticipated components of the tracker provide additional information to stakeholders for the firm's abnormal stock returns.

To do so, we match the brand tracker data with the firm's financial data from the Center for Research in Security Prices (CRSP). After the matching, we obtain 8,710 firm-week observations with 67 single-brand firms that trade in the U.S. stock market. In this matching, individual brands from firms that follow a house of brands strategy (i.e., one firm has multiple brands) are excluded (e.g., Pampers) for consistency. Table W4 in the Web Appendix summarizes the industry characteristics of the brands in the tracker, based on their 2-digit North American Industry Classification System (NAICS) codes. It shows that we have 40.30% (N = 27) manufacturing brands and 59.70% (N = 40) service brands. Both the manufacturing and the service brands are dominated by technology and information brands, reflecting the nature of the modern information economy.

In calculating abnormal returns, we first estimate a firm's expected stock returns using the Carhart four risk factors (Carhart 1997; Fama and French 1993) to adjust stock returns for the risk factors and to demonstrate that the drivers provide additional explanations for abnormal returns. Our calculation is similar to that of Nam and Kannan (2014) and Srinivasan et al. (2009), as shown in Equation (2):

$$(2) \quad R_{it} - R_{ft,t} = \alpha_i + \beta_i(R_{mt} - R_{rf,t}) + s_i SMB_t + h_i HML_t + u_i UMD_t + \varepsilon_{it}$$

where R_{it} is firm i's actual stock return in week t, $R_{ft,t}$ is the risk-free rate of return in week t. We obtain daily stock returns data from the CRSP and collapsed them into weekly stock returns to match our weekly brand data. The three Fama-French factors, the risk-free market return rate, $R_{mt} – R_{rf,t}$, the return difference between small and big-firm stocks, SMB_t, and the return difference between high and low book-to-market stocks, HML_t, are accessed from Fama and French's data library. The momentum factor (UMD_t) is the return difference between portfolios of past winners and losers (Aslam et al. 2008).

We estimate the impact of the unanticipated component of the drivers on the firm's abnormal return using Equation (3):

$$(3)\quad AR_{it} = \sum_{d=1}^{3}\sum_{s}^{2}(\beta_{d,t-s} U\Delta X_{id,t-s}) + \sum_{d=1}^{2}\sum_{g=d+1}^{3}\sum_{s=0}^{2}(\gamma_{dg,t-s} U\Delta X_{id,t-s} U\Delta X_{ig,t-s}) + \sum_{s=0}^{2}(\theta_{t-s} U\Delta X_{i1,t-s} U\Delta X_{i2,t-s} U\Delta X_{i3,t-s}) + \mathbf{C}_{i1}\boldsymbol{\delta}_1 + \mathbf{C}_{i2}\boldsymbol{\delta}_2 + \varepsilon_{it},$$

where

i =	Brand (1 to 67, i.e., single-brand firms)
t =	Week (there are 130 data weeks)
AR_{it} =	Abnormal return for firm i in week t ($AR_{it} = (R_{it} – R_{ft,t}) – Eret_{it}$, where expected return, $Eret_{it,}$ is the predicted value of $(R_{it} – R_{ft,t})$ in Equation (2))
$U\Delta X$ =	Unanticipated component of the drivers
d =	Driver (1 to 3, i.e., brand, value, and relationship drivers)
s =	No. of week lag (1 to 2)
g =	Index for the other brand in the two-way interaction
t-s =	Index for time with lag(s), i.e., the current, lag one week, or lag two weeks
$\boldsymbol{\delta}_1$ =	A (4 *x* 1) vector of coefficients for the industry control variable
$\boldsymbol{\delta}_2$ =	A (2 *x* 1) vector of coefficients for the year control variable
$\mathbf{C}_{i1}$ =	A (1 *x* 4) vector of industry dummy variables, with manufacturing industry as the base
$\mathbf{C}_{i2}$ =	A (1 *x* 2) vector of year dummy variables, with year 2016 as the base
ε_{it} =	The error term
$\beta_{d,t-s}$, $\gamma_{dg,t-s}$, θ_{t-s} =	Parameters to be estimated

UΔX, the unanticipated component of driver is the standardized residual estimated by a fixed-effect panel model for each driver, using its lagged one-week value as the predictor. We include lagged one-week and two-week drivers as predictors to capture the immediate, short-delayed, and longer-delayed effects of the drivers.[5] Industry sector and year dummy are included as control variables. Industry sector is the set of industry dummies, ranging from 1 to 5 (the manufacturing sector as the baseline). Year dummy is the set of year dummies, ranging from 2016 to 2018 (year 2016 as the baseline).[6] Equation (3) is estimated using a feasible generalized least squares panel model, specifying a heteroskedastic error structure and panel-specific autocorrelation. This allows flexible autocorrelation across brands and a brand-specific AR(1) process for the error in each brand.

Accountability of drivers. The estimation of Equation (2) shows that the market risk factor ($R_{mt} – R_{ft,t}$) has a significant positive effect on stock returns (.801, p = .000), whereas the SMB factor has a significant negative effect (-.001, p = .000). The other two factors are not significant. The positive coefficient for the market risk factor shows that each firm's stock returns covary with the risk-free market returns, and the negative coefficient for the SMB factor indicates that big firms have higher returns than small firms. The constant is insignificant (-.000, p = .857), consistent with the efficient market hypothesis.

We then estimate Equation (3) to check the accountability of the residuals of the drivers for a firm's abnormal return. Table 4 presents the results for the main effect model and the interaction model, respectively. For the main effect model, the residual of the Brand driver has a real-time positive impact on abnormal returns (.001, p = .034), the residual of the Value driver has a short-term positive impact (.001, p = .085), and the residual of the Relationship driver has real-time (-.002, p = .028) and short-term (-.002, p = .001) negative impacts, but a longer-term positive impact on abnormal returns (.001, p = .079). The information sector has higher abnormal returns than other sectors (.001, p = .013). No significant year effect is found.

[5] In discussing the results, we always refer the effect of current value of the driver (i.e., t) as "real-time" effect, lagged one-week value of the driver (i.e., t-1) as "short-term" effect, and lagged two-week value of the driver (i.e., t-2) as "longer-term" effect.
[6] We do not include a constant term for Equation (3) because the constant is estimated in Equation (2). Nevertheless, since the constant term in Equation (2) is insignificant, whether to include a constant term in Equation (3) does not change the results.

Results from the interaction effect model show that the residual of the Brand*Value interaction has a real-time positive impact on abnormal returns (.001, p = .089), the residual of the Value*Relationship interaction has a short-term positive impact on abnormal returns (.001, p = .062), but the residual of the Brand*Relationship interaction has a real-time negative impact on abnormal returns (-.001, p = .052).

Accountability of driver sentiment. We then take the sentiment of the drivers into consideration by estimating Equation (3), but separating the residual of the positive and negative sentiments of the drivers into two models.

For the negative sentiment model, we find that the residual of the negative Value driver has a short-term negative effect (-.002, p = .041), the residual of the negative Relationship driver has a longer-term negative effect (-.003, p = .006), and the residual of the negative Brand driver has no impact. The residual of the negative Brand*Value interaction has a longer-term negative impact (-.001, p = .041), and the residual of the negative Relationship*Value interaction has a real-time negative impact (-.001, p = .095).

For the positive sentiment model, we find that the residual of the positive Value driver has a short-term positive effect (.002, p = .025), the residual of the positive Relationship has a short-term negative effect (-.004, p = .000), and the residual of the positive Brand driver has no impact. The residual of the positive Brand*Relationship interaction has a negative effect (-.002, p = .051), but the residual of the positive Brand*Value*Relationship interaction has a marginal positive effect (.000, p = .108).

Together, the results from the two sentiment models show that the negative sentiment of drivers matters more, and more consistently. Furthermore, the marginal positive impact of the three-way interaction from the positive sentiment model confirms the Brand-Relationship-Value virtuous relationship between the drivers found in the VAR model, indicating that this virtuous relationship can be accountable for a firm’s abnormal returns.

Summary. The analysis demonstrates that the three drivers provide additional information for a firm’s risk-adjusted abnormal stock returns in real-time, short-term, and longer-term. The results show that the Brand driver has a real-time impact and is the dominant driver for abnormal returns, the Value driver has a short-term impact and synergizes with the other two drivers, and the Relationship driver has a longer-term impact and its positive sentiment synergizes with the other two drivers. This pattern of impact is consistent with the dynamics of the three drivers observed in the previous section.

The impact of the Brand driver is more real-time, reflecting that the driver captures stakeholders' immediate sentiments to brand events or activities, as demonstrated by Hewett et al. (2016) that online word-of-mouth echoes fast and wide in an "echoverse" of the brand's communication.

The impact of the Value driver is more short-term, reflecting that quality and cost do not fluctuate as frequently as brand feelings. This driver reflects the knowledge aspect of a brand, which according to the brand equity literature (e.g., Keller 1993), can be expected to be more stable than emotional reactions to brand events. Its positive and negative sentiments provide separate information for a firm's abnormal returns, indicating the need for monitoring both the positive and negative discussions about a brand. Its synergy with the Brand driver also indicates that the objective aspects of a brand's reputation (e.g., price, quality) need to be associated with positive feelings with the brand (e.g., cool, exciting) to benefit a firm's abnormal returns.

The impact of the Relationship driver is longer-term and hinges more on the positive sentiment, reflecting that relationships take time to play out, but the stock market may be myopic with respect to longer-term marketing impacts (e.g., Huang and Trusov 2020; Mizik and Jacobson 2007). Once a good reputation on this driver is built, it benefits a firm's abnormal returns in the longer-term, as shown in the longer-term Brand-Relationship-Value synergy for the positive sentiment of the drivers.

Together, the results suggest that the residual of the drivers provides information value for a firm's abnormal returns, immediately or in a delayed manner, individually and collectively. The dark curved arrows in Figure 6 summarize the analysis. Thus, by monitoring the fluctuations of the drivers, stakeholders can have a more accurate picture about a firm's financial performance.

Validation

We validate the brand reputation tracker using three approaches. First, we replicate our methodology using two additional social media platforms, Facebook and Instagram, each of which has idiosyncratic features: Facebook focuses on social networking, and Instagram focuses on photo sharing. Second, we establish a nomological relationship with the survey-based YouGov brand data, from which we demonstrate that the tracker is related to YouGov's brand word-of-mouth (WOM) and brand buzz, and leads to YouGov's purchase intention. Third, we demonstrate that the tracker correlates significantly positively with three aggregate annual brand measures, Interbrand, Forbes, and Brandz, showing that the tracker not only converges with the aggregate annual measures, but also provides more granular information (both in terms of time interval and drivers) for brand reputation.

Replicating the Tracker Using Other Social Media Platforms

Data. We collected data from Facebook and Instagram, from January 1 to June 30, 2018 (i.e., 26 weeks), for the seven Internet brands in the tracker. The data were collected using Crimson Hexagon, with a method similar to what was described in the methodology section. However, the data were more difficult and problematic to collect and analyze, because many brand posts on Facebook are not publicly available, and Instagram concentrates on visual data.

We applied the same dictionaries of the sub-drivers to the two social media platforms. For Facebook, we focused on post contents that are available on the firm's own brand pages. For example, for the Amazon brand, its Facebook page is https://www.facebook.com/Amazon, and a sample post content is: "Thank you Amazon for excellent customer service and speedy response! Keep up the great work!" For Instagram, we collected captions and comments on photos that mention the brand handles. For example, the Amazon brand's Instagram handle is @amazon, and a sample post is: "@amazon…I ordered a waffle iron two days ago and they delivered it to this tiny island 30 miles out to sea so quickly."

In calculating the driver and sub-driver scores, an initial screening of the data reveals that data for the Yahoo brand are problematic, because the brand posts a lot of news articles, resulting in 20-40 times more posts than the other Internet brands. We thus drop the Yahoo brand from the calculation and subsequent analysis.

Results. We use multiple methods, including descriptive statistics, correlation analysis, and repeated and mixed-measures analysis of variance, to replicate the tracker with the two social media platforms. Table 5 presents the descriptive statistics and correlation analysis for the three social media platforms, along the overall brand reputation and the three drivers.

The descriptive statistics show that Twitter has much higher volume than the other two platforms. This is because Twitter has more publicly available content (500M daily) and has more brand content than other platforms. Although Facebook also has a high volume of posts (comparable with Twitter), most of the posts are not on public pages, and thus are not readily available. Instagram's posts (95M daily) are mostly not brand related. The descriptive statistics confirm empirically our choice of Twitter as the social media platform of the Tracker, as articulated previously.

The correlation analysis shows that the three social media platforms converge at both the brand reputation and the driver levels. All correlations are significantly positive, except the Relationship driver between Twitter and Facebook.

A repeated and mixed-measures analysis of variance, by taking into consideration the time series nature of the measures and the brand differences, further suggests that our tracker can be replicated using other social media platforms. We treat the brand reputation (and its three drivers), respectively, as within-brand repeated measures that are between platforms. We do not find significant platform differences for brand reputation (F = .96, p = .386, df = 2), but we find significant brand differences (F = 378.63, p = .000, df = 5) and brand*platform differences (F = 87.31, p = .000, df = 10). The results for the three drivers are the same; all platform differences are insignificant, but brand differences and brand*platform differences are significant. The findings confirm that our tracker can be generalized to other social media platforms and suggest that brands have different social media strategies.

Summary. The results of descriptive statistics and correlation analysis provide general support for the robustness of the tracker across social media platforms, which can be text-based or visual-based. Among the three social media platforms, Twitter is more suitable for monitoring and tracking brand reputation, due to its substantially more publicly available brand-related content. Furthermore, the tracker can be replicated using the other two social media platforms at both the brand reputation and the driver levels—evidence for the generalizability of the tracker as a social media-based brand reputation tracker. This replication also supports that our conceptualization and methodology are robust even if the nature of individual social media varies. The results of the repeated and mixed-measures analysis of variance provide stronger evidence by considering the tracker's time series nature, social media characteristics, and brand differences.

These multiple approaches to replication consistently support the robustness of the Twitter-based brand reputation tracker: the three-driver framework and the methodology we develop here can be generalized to other social media platforms.

Establishing a Nomological Relationship with YouGov's BrandIndex

Data. We purchased access to YouGov's BrandIndex data, and matched 71 non-Internet brands[7] that are common between the two data sets for the data period (i.e., 130 weeks, 9,230 brand-week observations). YouGov's BrandIndex interviews a consumer panel about their opinions regarding three broad sets of brand metrics: brand health, media, and purchase funnel metrics. A detailed description of their methodology can be seen at YouGov's company website (https://today.yougov.com/solutions/syndicated/brandindex). Essentially, for the three broad sets of brand metrics, the media metrics are conceptually similar to our tracker. It contains brand WOM (i.e., whether the consumer has recently spoken about the brand) and brand buzz (i.e., whether the consumer has heard anything positive or negative about the brand). The purchase funnel metrics, such as purchase intention, are more appropriate as the non-financial outcome of the tracker. Thus, we establish the nomological relationship of the tracker with YouGov's BrandIndex by conceptualizing its brand WOM and brand buzz as concurrent variables with the tracker, and its purchase intention as the outcome variable, as shown in Figure 7.

Results. A simple correlation analysis shows that the tracker's overall brand reputation correlates significantly with BrandIndex's brand WOM (.355, p = .000), brand buzz (.317, p = .000), and purchase intention (.248, p = .000). All three brand reputation drivers of our tracker also all correlate significantly with the three BrandIndex variables.

[7] Internet brands are scaled separately, as noted earlier.

We then conducted panel regression analysis to establish the causality between the three BrandIndex variables. The results show that brand WOM has a lagged impact (.221, p = .000), while brand buzz has a concurrent impact (.072, p = .027) on purchase intention.

After establishing that the causal chain is likely to be from brand WOM to brand buzz to purchase intention, we ran a panel VAR model to explore the dynamics among BrandIndex's brand WOM and brand buzz, and the tracker's overall brand reputation. Results of the Granger causality test show that it is more likely for brand reputation (χ^2 = 18.542, p = .000) and WOM (χ^2 = 7.690, p = .006) to Granger-cause brand buzz, indicating that the tracker's brand reputation is likely to be a concept that encompasses BrandIndex's brand WOM and brand buzz.

Lastly, we ran a panel regression analysis using the current and lagged one-week values of the tracker's overall brand reputation to predict BrandIndex's purchase intention. Results of the analysis, with standard errors adjusted for brand heterogeneity, show that the lagged brand reputation (not the concurrent one) significantly predicts consumers' intentions to purchase the brand (.019, p = .017).

Summary. We establish the nomological relationship between our Tracker and YouGov's BrandIndex, with YouGov's conceptually similar metrics of brand WOM and brand buzz correlating significantly with the tracker, and with the lagged tracker being a significant predictor for consumers' intentions to purchase the brand.

Converging with Other Aggregate Brand-related Measures

There are many other brand-related measures, as listed in Table W1 in the Web Appendix. We compare the brand tracker with three other aggregate brand-related measures for 2018: the Interbrand, Forbes, and Brandz lists (chosen for their availability) to check whether they converge. The correlational analysis shows that the three lists are highly correlated, with correlation coefficients all greater than .851 (p < .000). This indicates that they are very similar, even if they claim to use different methods of brand evaluation. The correlations of the three lists with the overall brand reputation scores are significant (p < .000), but differentiable, since the correlation coefficients range from .079 (Forbes), .081 (BrandZ), to .127 (Interbrand), indicating our tracker converges as well as differentiates from the other aggregate measures.

If we look at the ranking, rather than the brand value, the correlation between the three other brand measures is more discriminable. The Interbrand rank is still highly correlated with the Forbes' rank (.802), but is more discriminable from Brandz's rank (.404). The correlation between Forbes' and BrandZ's ranking also become more discriminable (.581). Meanwhile, their correlations with our tracker become higher (.335 for Interbrand; .356 for Forbes; .179 for BrandZ), suggesting that these measures converge better by using rankings.

We further explore the relationship between the 2018 Interbrand ranking and the three drivers. We regress the Interbrand ranking on the three drivers and find that the Brand driver predicts Interbrand ranking most closely (.279, p = .000), followed by the Relationship driver (.109, p = .015). The Value driver does not predict Interbrand ranking, when all three drivers are considered.

In sum, the analysis using the Interbrand ranking provides evidence that the two ranking systems correspond for the brands included, but our tracker provides more granular measures both in cross-sectional dimension (multiple brand drivers and sub-drivers) and longitudinal dimensions (weekly fluctuations).The brand reputation tracker not only provides real-time information about a brand's performance, but also is more granular at the driver and sub-driver levels, providing additional actionable information about a brand's performance, beyond the typical annual, aggregate-level brand ranking system.

Discussion

We demonstrate the various uses of brand reputation tracker data, using various methods and approaches. Our approach can be used to monitor and manage a brand's reputation over time, both at the driver and sub-driver levels. The data can also be used to manage brand competition by tracking the ups and downs of drivers for major competitors, and the sub-driver analyses can provide detailed insights about how to enhance or improve brand reputation. The accountability of the residual of the brand reputation drivers for abnormal returns confirms the financial implications of using the tracker to monitor and manage brands. Finally, the validation against other social media platforms and other brand-related measures provides evidence for the generalizability and external validity of the tracker. Below we discuss implications for managers and practitioners of the tracker, and propose a research agenda for researchers to leverage the availability and the methodology of the tracker.

Implications for Managers

Managing the Brand-Value Reciprocity. A firm can manage the reciprocal relationship between the Brand and the Value drivers. Depending on which driver a firm has the comparative advantage, the firm can selectively manage one of the drivers first, and then let the effect carry over to the other driver. For example, Apple has a stronger reputation in the Brand driver as being innovative, but a weaker reputation in the Value driver, due to its premium price; hence, Apple can prioritize leveraging its stronger Brand driver (and the innovative sub-driver), and let stakeholders understand that the innovativeness is worthy of a premium price (e.g., with better service quality). Figure W4a in the Web Appendix shows that this Brand-to-Value carryover effect takes one-week to takeoff but lasts for 5-6 weeks. Alternatively, Samsung has a stronger reputation in the Value driver as being affordable, while still innovative; hence, Samsung can prioritize leveraging its stronger Value driver (i.e., being affordable with high goods quality), and by doing so move stakeholder perceptions of its innovativeness, due to the reciprocal relationship. Figure W4b shows that this Value-to-Brand carryover effect is real-time and lasts for 4-5 weeks.

Managing the Brand-Relationship-Value Virtuous Circle. A firm can manage the virtuous circle among the three drivers, from the Brand driver to the Relationship driver, from the Relationship driver to the Value driver, and finally from the Value driver back to the Brand driver. For example, Apple has a strong foothold on all three drivers (shown in Figures W2a-W2c in the Web Appendix), and thus Apple is in a good position to leverage this virtuous circle. For a firm that is good at one or two drivers, it can manage the Brand-Value reciprocity, as discussed above. For a firm that is good at the Relationship driver (e.g., a mature brand with existing loyal customers, but having nothing new to be talked about on social media), the firm can manage the Relationship driver to get to the Value driver (e.g., generate discussion about its new product, new service, or new price), so that the firm can subsequently leverage the Brand-Value reciprocity. Figure W4c shows that the Relationship-to-Value carryover effect takes one week to takeoff but lasts for 5-6 weeks, and Figure W4d shows that the Brand-to-Relationship carryover effect is real-time and lasts for 4-5 weeks.

Managing Brands based on Drivers' Temporal Impact on Financial Returns. We find that the three drivers tend to impact a firm's financial returns at different tempos: The Brand driver has a real-time impact, the Value driver takes one week, and the Relationship driver takes two weeks to play out.

Given that the Brand driver reflects brand sentiment that can fluctuate easily with brand events, and given its dominant impact among the three drivers, a firm can boost this driver using brand events and activities, such as a new product launch, and expect a real-time impact on financial returns.

Given that the Value driver reflects brand knowledge, which does not change as easily as brand sentiment, a firm can leverage this driver as a relatively more stable foundation of its brand reputation. Its synergy with the Brand driver supports this strategy that brand sentiment can be manipulated by brand events, while brand knowledge can settle the sentiment into more enduring brand knowledge that can stabilize the effects of brand sentiment's ups and downs on financial returns. Brand crisis management is one example such that in the case of an unexpected negative brand event (e.g., Facebook's account data leakage), given stakeholders' knowledge about Facebook, the impact of the temporarily negative brand sentiment would be settled if a stakeholder has more positive knowledge about Facebook than negative knowledge.

Given that the Relationship driver reflects brand relationship, a firm can leverage this driver for long-term returns that are less subject to temporal fluctuations, though the firm needs to be patient with branding efforts for building relationship. Although the observation that relationship takes time to play out is established in the existing customer relationship literature, its long-term synergy with the other two drivers is not yet widely recognized, and can be leveraged for a stable brand reputation and its impact on financial returns.

Implications for Data Trackers and Providers

Our theory-data iterative approach to developing the tracker illustrates the importance of theory, and the need for data providers to pursue academic collaborations. Most data providers that track data or provide raw or summary data do not have a sound theory to guide them about what data to track and what summary data would be valuable. For example, Crimson Hexagon tracks data but does not provide raw data, whereas YouGov provides data, but let clients guide them about what data to collect. With the black-box machine learning approach continuing to advance for data tracking, making sense of data will be a pressing issue. Our methodology illustrates the importance of theory-based data tracking.

Implications for Researchers

The longitudinal brand data are available for free access to the academic research community. The data can be used in a variety of ways. In this section, we provide a list of possible research agendas along with a sampling of specific research questions for future research that can successfully leverage the data. Table 6 lists the research agendas and specific research questions. We discuss these in detail below.

- *Social Media Lens to Brand Reputation.* The tracker is built using social data from Twitter and is demonstrated to be generalizable to other social media platforms, which uniquely reflects how stakeholders talk and think about brands on social media. Given the growing importance of social media mentions for brands, one potential use of the data is to understand brand reputation based on social media activities. In the example analyses, we find that stakeholders talk about Samsung as "Cool" whereas Pepsi as not "Cool." Such wordings are distinct from how brands are portrayed in traditional media but are caught uniquely by the tracker.
- *Cross-Platform Application*. The tracker is applicable to multiple social media platforms that are distinct in data type (e.g., text or photo) and interaction pattern (e.g., one-to-many or one-to-one). Although different platforms have very different purposes (e.g., Facebook is mostly for social interaction between friends, while Twitter is more of a "broadcast" platform), we find our approach produces meaningful results even for very different platforms.
- *Longitudinal Research on Brand Reputation.* Given the time series nature of the data, one area of research that can make use of the tracker is to explore variation in brand drivers and sub-drivers over time. This is important, because drivers of brand reputation are unlikely to

be stable over time for all brands and in fact might vary considerably. The within-brand (over time) and between-brand temporal volatilities and dynamics of these measures are worth exploring as a use of the data. Ultimately, the dynamic richness of the data should open up new empirical possibilities for researchers interested in understanding how brands evolve, in a multitude of ways, between brands, across categories, and over time.

- *Granular Investigation of Brand Reputation*. The granularity of the tracker in both the time (weekly, monthly, and quarterly) and driver (drivers, sub-drivers, and sentiments) dimensions enables researchers to examine brand reputation in a much more granular way. On this front, the numerous drivers and sub-drivers allow for novel theory building and testing opportunities in relation to impacts of exogenous shocks on brands. For example, a theory could predict an impact of a shock or event on one driver but not another, and empirical evidence for this could be sought using our tracker, where there are theoretically meaningful differences between the affected and unaffected drivers that, indirectly at least, shed light on the underlying mechanism for an observed effect.
- *Brand-Related Events for Brand Reputation Variation.* Another potentially fruitful avenue for researchers is to identify brand-related exogenous events for brand reputation variation over time to see how specific events that are relevant to given brands in the tracker affect brand reputation (overall and for each of the drivers and sub-drivers). One obvious application is to examine how brands are impacted by negative and positive exogenous events such as product recalls, crises and scandals, major announcements, changes in C-level executives, product launches, and other potentially significant strategic marketing actions. Although prior research has at times considered such topics, including in the context of social media (e.g., Borah and Tellis 2016), more work is needed to increase our understanding, in the finer granularity provided by the sub-drivers.
- *Brand, Stakeholder, and Firm Characteristics for Brand Reputation Driver Variation.* One theoretically fruitful area is to link subtle and interesting brand, stakeholder, and firm characteristics to the variation in the brand reputation drivers. Using those data to account for the fluctuations of brand reputation drivers allows researchers to develop new models and theories about why a certain brand “attracts” or “offends” stakeholders on social media, as a

function of those characteristics. We demonstrate a nomological relationship of the tracker with YouGov data, which can be one of the many approaches to link the driver variations to those characteristics.

- *Brands in Novel Classifications*. One potential way that researchers can benefit from our tracker is to classify brands based on their scores on the three brand reputation drivers or the larger set of sub-drivers, or on statistically estimated properties such as the extent to which time series within brands are correlated/cointegrated, stationary/evolving, and so on. Additionally, machine learning techniques for classification could be used to achieve a similar outcome. This can lead to the identification of new types of classes or groups for brands that might have interesting theoretical and practical implications. This could also lead to new research questions pertaining to understanding differences between brands that score high versus low on various drivers of interest.
- *Brand Reputation Drivers and Marketing/Financial Outcomes*. We demonstrate some applications of the data by linking brand reputation drivers to firm abnormal stock returns. Our analysis illustrates the potential available for researchers interested in the marketing-finance interface. Given the time series nature of our data, more complex models could be developed that allow for marketing/financial metrics to not only be considered as being influenced by brand reputation but also to have an effect on changes in the various drivers in our data. This would be interesting, as it would allow us to understand the extent to which brand reputation drives, for example, financial outcomes versus how much reputation is instead driven by firm performance.

Contribution

As opposed to most aggregate ratings or rankings of brand value, the brand reputation tracker enables a more granular investigation of the components of brand reputation. Unlike survey-based attitudinal brand measures, the Tracker is designed to map more directly to competing strategic marketing expenditures.

We tie to the expanding literature on mining (“listening in on”) social media to obtain brand reputation insights. By developing a methodology that is applicable across multiple social media platforms and providing a longitudinal database that is granular enough to guide marketing actions, we make it much easier for researchers to tie social media posts to brand reputation.

We provide a new resource for corporate reputation research. To date, most corporate reputation research has been in the management and strategy literatures, and has placed less emphasis on marketing actions. This database enables corporate reputation research to link more naturally to marketing.

Finally, we contribute to the research of brand reputation by making the brand reputation tracker data used in this paper available, to facilitate longitudinal brand research. Further extending over additional time periods or additional brands, can be a valuable resource for future brand research, and increase our understanding of how brands work.

Limitations

Our tracker is based on Twitter tweets. This gives the tracker many advantages over other social media. However, the social media environment is not static; for example, Twitter might go out of business, or their terms of service making it impossible to apply our brand tracker on the platform. In such an eventuality, researchers may need to migrate the tracker to a different social media platform. We have demonstrated the generalizability of the methodology using data from Facebook and Instagram, suggesting that our approach may be usable on other platforms as well.

Historical data are more difficult to collect. We suggest that brands that are not included in the tracker or that want to add more actionable sub-drivers should be forward-looking and follow our methodology and access the underlying data using the public free streaming API to build their own tracker. Purchasing historical data from data providers is still an option, as we did to backfill the missing data.

Although we can mine millions of tweets automatically, there is still the need to update the usernames manually. Our collection is limited to English tweets. People’s use of keywords for talking about brands may also change over time and across contexts, and thus the dictionaries need to be updated periodically. We expect that with more advanced machine learning, usernames and sub-driver dictionaries can be updated automatically.

We start from 14 sub-drivers and refine them into a smaller set of 11 sub-drivers based on the multiple-stage, theory-data iterative process. These sub-drivers are shown to be applicable to brands in the dataset. Sub-drivers are directly actionable and thus a firm can explore more potentially actionable sub-drivers for further enhancing the marketing relevance of the tracker to its brand.

Conclusions

The brand reputation tracker is a longitudinal data base of brand reputation driver and sub-driver data, for 100 top global brands, based on mining Twitter tweets. Our study contributes to the literature by making brand reputation financially accountable and managerially actionable in real-time and over time. The tracker is highly time-sensitive and context-specific, allowing firms to respond fast to market stimuli. The final goal is to provide a database resource that any academic researcher can access and/or extend. We anticipate that this should increase the amount of research done on brand reputation over time, increasing our knowledge of the antecedents and consequences of the components of brand reputation. We also hope and expect that the tracker will give marketing more importance in the broader corporate reputation literature.

References

Aaker, David A. (1995), *Building Strong Brands*, New York: The Free Press.

Abrigo, Michael R. M. and Inessa Love (2015), "Estimation of Panel Vector Autoregression in Stata: A Package of Programs," University of Hawaii working paper.

Aksoy, Lerzan, Bruce Cooil, Christopher John Groening, and Timothy L. Keiningham (2008), "The Long-Term Stock Market Valuation of Customer Satisfaction," *Journal of Marketing*, 72 (4), 105-122.

Argenti, Paul A., and Bob Druckenmiller (2004), "Reputation and the Corporate Brand," *Corporate Reputation Review*, 6 (4), 368-374.

Berry, Leonard L. (2000), "Cultivating Service Brand Equity," *Journal of the Academy of Marketing Science*, 28 (1), 128-137.

Blattberg, Robert C. and John Deighton (1996), "Manage Marketing by the Customer Equity Test," *Harvard Business Review*, 74 (4), 136-144.

Boivie, Steven, Scott D. Graffin, and Rich J. Gentry (2015), "Understanding the Direction, Magnitude, and Joint Effects of Reputation When Multiple Actors' Reputations Collide," *Academy of Management Journal*, 59 (1), 188-206.

Borah, Abhishek and Gerard J. Tellis (2016), "Halo (Spillover) Effects in Social Media: Do Product Recalls of One Brand Hurt or Help Rival Brands?" *Journal of Marketing Research*, 53 (2), 143-160.

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

Choi, In (2001), "Unit Root Tests for Panel Data," *Journal of International Money and Finance*, 20 (April), 249-272.

Colicev, Anatoli, Ashwin Malshe, Koen Pauwels, and Peter O'Connor (2018), "Improving Consumer Mind-Set Metrics and Shareholder Value through Social Media: The Different Roles of Owned and Earned," *Journal of Marketing*, 82 (1), 37-56.

Davis, Ernest and Gary Marcus (2015), "Commonsense Reasoning and Commonsense Knowledge in Artificial Intelligence," *Communications of the ACM*, 58 (9), 92-103.

Dowling, Grahame R. (2016), "Defining and Measuring Corporate Reputations," *European Management Review*, 13 (3), 207-223.

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

Ferguson, Tamela D., David L. Deephouse, and William L. Ferguson (2000), "Do Strategic Groups Differ in Reputation?" *Strategic Management Journal*, 21 (12), 1195-1214.

Fombrun, Charles and Mark Shanley (1990), "What's in a Name? Reputation Building and Corporate Strategy," *Academy of Management Journal*, 33 (2), 233-258.

Fossen, Beth L. and David A. Schweidel (2019), "Measuring the Impact of Product Placement with Brand-Related Social Media Conversations and Website Traffic," *Marketing Science*, 38 (3), 481-499.

Gani, Ariel Nian and Andreas Grobler (2019), "Linking Brand Equity and Customer Equity: A System Dynamics Perspective," working paper.

Gao, Lily, Iguacel Melero-Polo, and F. Javier Sese (2020), "Customer Equity Drivers, Customer Experience Quality, and Customer Profitability in Banking Services: The Moderating Role of Social Influence," *Journal of Service Research*, 23 (2), 174-193.

Gupta, Sunil, Donald R. Lehmann, and Jennifer Ames Stuart (2004), "Valuing Customers," *Journal of Marketing Research*, 41 (1), 7-18.

Hall, Richard (1992), "The Strategic Analysis of Intangible Resources," *Strategic Management Journal*, 13 (2), 135-144.

Hanssens, Dominique M., Roland T. Rust, and Rajendra K. Srivastava (2009), "Marketing Strategy and Wall Street: Nailing Down Marketing's Impact," *Journal of Marketing*, 73 (November), 115-118.

Hewett, Kelly, William Rand, Roland T. Rust, and Harald van Heerde (2016), "Brand Buzz in the Echoverse," *Journal of Marketing*, 80 (3), 1-24.

Holtz-Eakin, Douglas, Whitney Newey, and Harvey S. Rosen (1988), "Estimating Vector Autoregressions with Panel Data," *Econometrica*, 56 (6), 1371-1395.

Huang, Ming-Hui and Chekitan S. Dev (2020), "Growing the Service Brand," *International Journal of Research in Marketing*, 37 (2), 281-300.

Keller, Kevin L. (1993), "Conceptualizing, Measuring and Managing Customer-Based Brand Equity," *Journal of Marketing*, 57 (1), 1-22.

Kubler, Raoul V., Anatoli Colicev, and Koen H. Pauwels (2020), "Social Media's Impact on the Consumer Mindset: When to Use Which Sentiment Extraction Tool?" *Journal of Interactive Marketing*, 50 (May), 136-155.

Kumar, V. and Denish Shah (2009), "Expanding the Role of Marketing: From Customer Equity to Market Capitalization," *Journal of Marketing*, 73 (6), 119-136.

Kumar, V., Vikram Bhaskaran, Rohan Mirchandani, and Milap Shah (2013), "Creating a Measurable Social Media Marketing Strategy: Increasing the Value and ROI of Intangibles and Tangibles for Hokey Pokey," *Marketing Science*, 32 (2), 194-212.

Lange, Donald, Peggy M. Lee, and Ye Dai (2011), "Organizational Reputation: A Review," *Journal of Management*, 37 (1), 153-184.

Leone, Robert P., Vithala R. Rao, Kevin Lane Keller, Anita Man Luo, Leigh McAlister, and Rajendra Srivastava (2006), "Linking Brand Equity to Customer Equity," *Journal of Service Research*, 9 (2), 125-138.

Levin, Andrew, Chien-Fu Lin, and Chia-Shang James Chu (2002), "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties," *Journal of Econometrics*, 108 (1), 1-24.

Love, Inessa and Lea Ziccino (2006), "Financial Development and Dynamic Investment Behaviour: Evidence from Panel VAR," *The Quarterly Review of Economics and Finance*, 46, 190-210.

Lovett, Mitchell J., Renana Peres, and Ron Shachar (2013), "On Brands and Word of Mouth," *Journal of Marketing Research*, 50 (4), 427-444.

Mizik, Natalie (2014), "Assessing the Total Financial Performance Impact of Brand Equity with Limited Time-Series Data," *Journal of Marketing Research*, 51 (6), 691-706.

Mizik, Natalie and Robert Jacobson (2007), "Myopic Marketing Management: Evidence of the Phenomenon and its Long-Term Performance Consequences in the SEO Context," *Marketing Science*, 26 (3), 361-379.

Mizik, Natalie and Robert Jacobson (2008), "The Financial Value Impact of Perceptual Brand Attributes," *Journal of Marketing Research*, 45 (1), 15-32.

Nam, Hyoryung and P. K. Kannan (2014), "The Informational Value of Social Tagging Networks," *Journal of Marketing*, 78 (July), 21-40.

Nguyen, Hang, Roger Calantone, and Ranjani Krishnan (2020), "Influence of Social Media Emotional Word of Mouth on Institutional Investors' Decisions and Firm Value," *Management Science*, 66 (2), 887-910.

Ou, Yi-Chun, Lisette de Vries, Thorsten Wiesel, and Peter C. Verhoef (2014), "The Role of Consumer Confidence in Creating Customer Loyalty," *Journal of Service Research*, 17 (3), 339-354.

Ou, Yi-Chun, Peter C. Verhoef, and Thorsten Wiesel (2017), "The Effects of Customer Equity Drivers on Loyalty across Services Industries and Firms," *Journal of the Academy of Marketing Science*, 45 (3), 336-356.

Parker, Owen, Ryan Krause, and Cynthia E. Devers (2019), "How Firm Reputation Shapes Managerial Discretion," *Academy of Management Review*, 44 (2), 254-278.

Pfarrer, Michael D., Timothy G. Pollock, and Violina P. Rindova (2010), *Academy of Management Journal*, 53 (5), 1131-1152.

Porter, Martin F. (1980), "An Algorithm for Suffix Stripping," *Program*, 14 (3), 130-137.

Raithel, Sascha and Manfred Swhwaiger (2015), "The Effects of Corporate Reputation Perceptions of the General Public on Shareholder Value," *Strategic Management Journal*, 36 (6), 945-956.

Rindova, Violina P., Ian O. Williamson, Antoaneta P. Petkova, and Joy Marie Sever (2005), "Being Good or Being Known: An Empirical Examination of the Dimensions, Antecedents, and Consequences of Organizational Reputation," *Academy of Management Journal*, 48 (6), 1033-1049.

Roberts, Peter W. and Grahame R. Dowling (2002), "Corporate Reputation and Sustained Superior Financial Performance," *Strategic Management Journal*, 23 (12), 1077-1093.

Rust, Roland T., Katherine N. Lemon, and Valarie A. Zeithaml (2004), "Return on Marketing: Using Customer Equity to focus Marketing Strategy," *Journal of Marketing*, 68 (January), 109-127.

Rust, Roland T., Valarie A. Zeithaml, and Katherine N. Lemon (2000), *Driving Customer Equity: How Customer Lifetime Value is Reshaping Corporate Strategy*, New York: The Free Press.

Schulze, Christian, Bernd Skiera, and Thorsten Wiesel (2012), "Linking Customer and Financial Metrics to Shareholder Value: The Leverage Effect in Customer-Based Valuation," *Journal of Marketing*, 76 (March), 17-32.

Schweidel, David A. and Wendy W. Moe (2014), "Listening in on Social Media: A Joint Model of Sentiment and Venue Format Choice," *Journal of Marketing Research*, 51 (4), 387-402.

Skiera, Bernd, Manuel Bermes, and Lutz Horn (2011), "Customer Equity Sustainability Ratio: A New Metric for Assessing a Firm's Future Orientation," *Journal of Marketing*, 75 (May), 118-131.

Srinivasan, Shuba and Dominique M. Hanssens (2009), "Marketing and Firm Value: Metrics, Methods, Findings, and Future Directions," *Journal of Marketing Research*, 46 (3), 293-312.

Srinivasan, Shuba, Koen Pauwels, Jorge Silva-Risso, and Dominique M. Hanssens (2009), "Product Innovations, Advertising, and Stock Returns," *Journal of Marketing*, 73 (1), 24-43.

Stahl, Florian, Mark Heitmann, Donald R. Lehmann, and Scott A. Neslin (2012), "The Impact of Brand Equity on Customer Acquisition, Retention, and Profit Margin," *Journal of Marketing*, 76 (4), 44-63.

Tavassoli, Nader T., Alina Sorescu, and Rajesh Chandy (2014), "Employee-Based Brand Equity: Why Firms with Strong Brands Pay Their Executives Less," *Journal of Marketing Research*, 51 (6), 676-690.

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.

Tirunillai, Seshadri, and Gerard J. Tellis (2014), "Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation," *Journal of Marketing Research*, 51 (4), 463-479.

Toubia, Olivier and Andrew T. Stephen (2013), "Intrinsic vs. Image-Related Utility in Social Media: Why Do People Contribute Content to Twitter?" *Marketing Science*, 32 (3), 368-392.

Villanueva, Julian, Shijin Yoo, and Dominique M. Hanssens (2008), "The Impact of Marketing-Induced Word-of-Mouth Customer Acquisition on Customer Equity Growth," *Journal of Marketing Research*, 45 (1), 48-59.

Vogel, Verena, Heiner Evanschitzky, and B. Ramaseshan (2008), "Customer Equity Drivers and Future Sales," *Journal of Marketing*, 72 (6), 98-108.

Wiesel, Thorsten, Bernd Skiera, and Julian Villanueva (2008), "Customer Equity: An Integral Part of Financial Reporting," *Journal of Marketing*, 72 (March), 1-14.

Figure 1. Conceptual Sources of Brand Reputation.

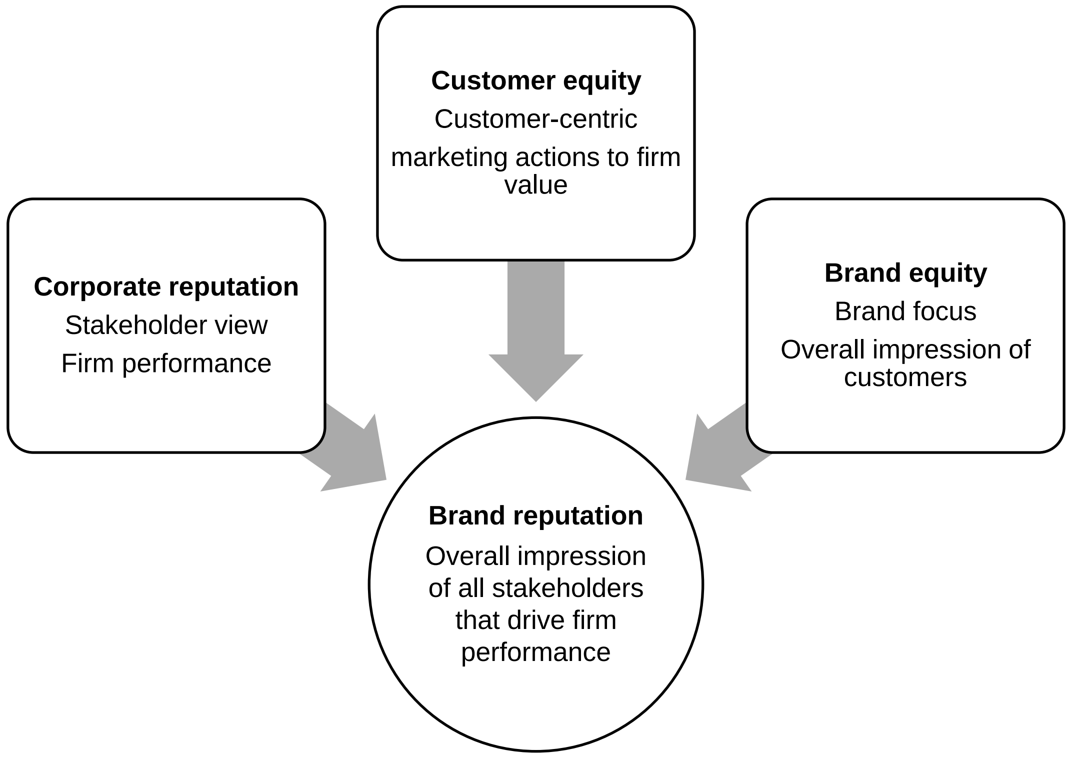


Figure 2. A Multi-Stage, Theory-Data Iterative Process to Develop the Tracker.

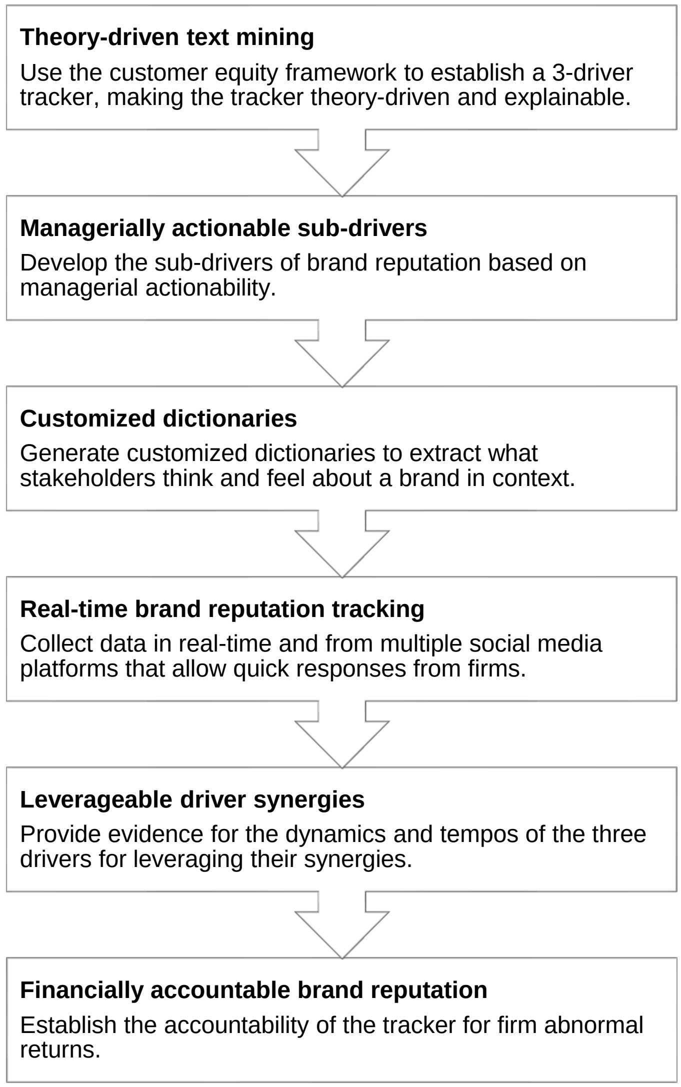


Figure 3. Brand Reputation: Facebook vs. Google 2018.

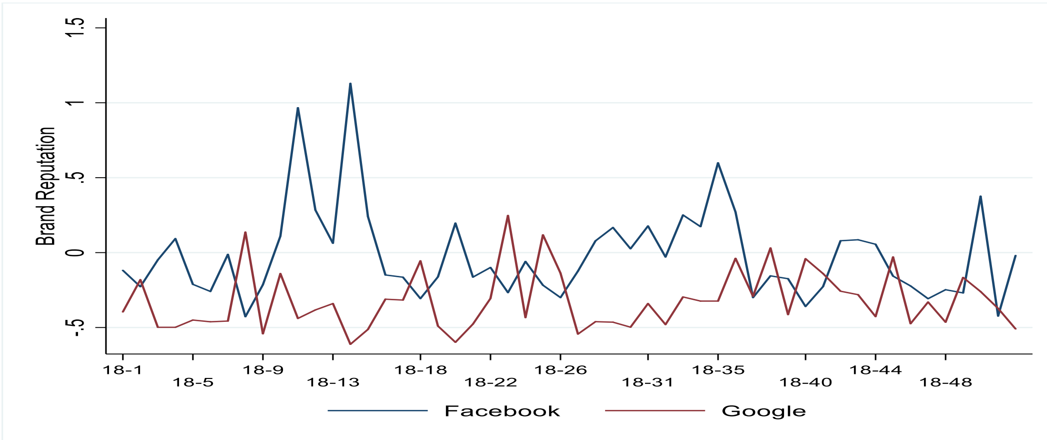


Figure 4a. Brand Driver

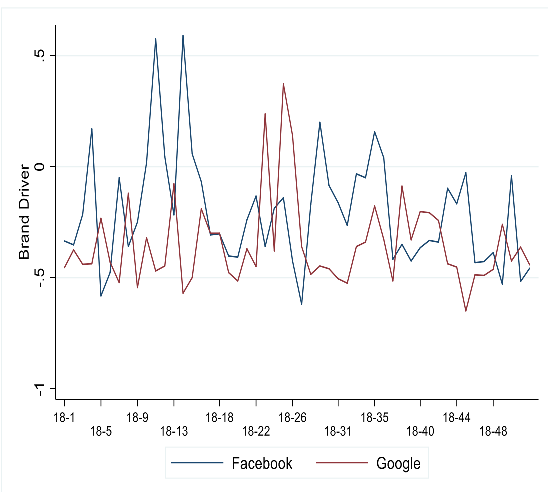


Figure 4b. Value Driver

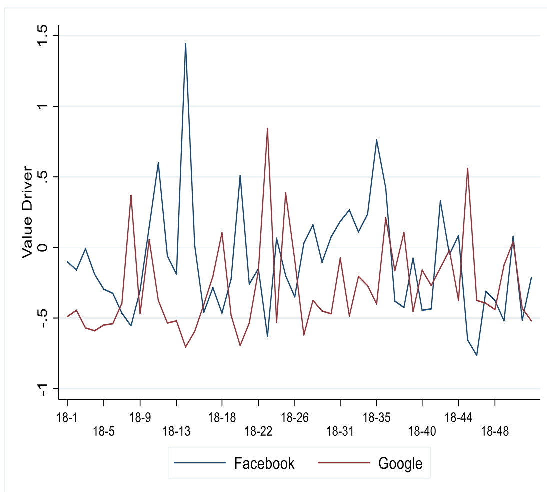


Figure 4c. Relationship Driver

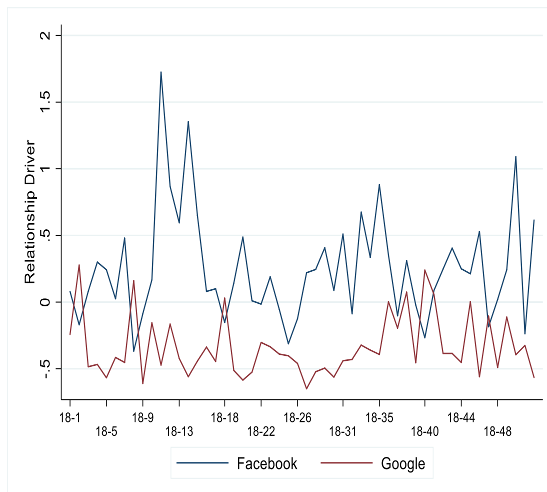


Figure 5a. Exciting Sub-Driver

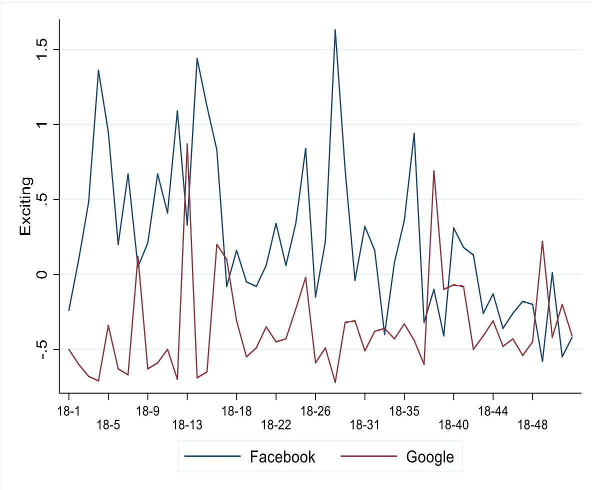


Figure 5b. Innovative Sub-Driver

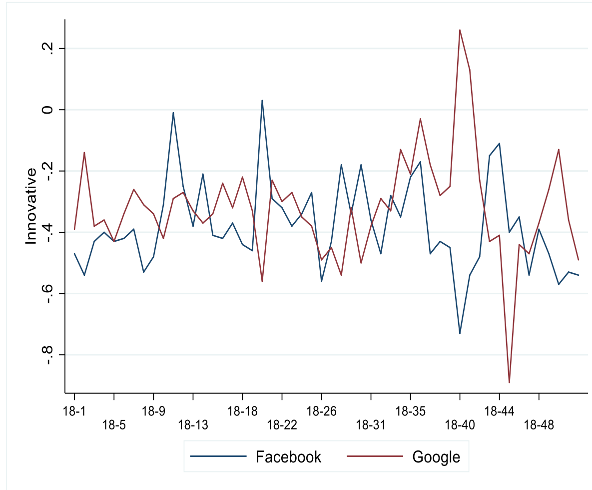


Figure 5c. Personal Relationship Sub-Driver

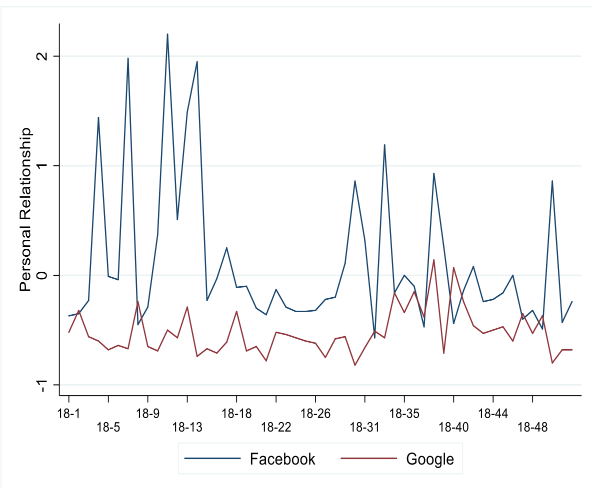


Figure 5d. Trustworthy Sub-Driver

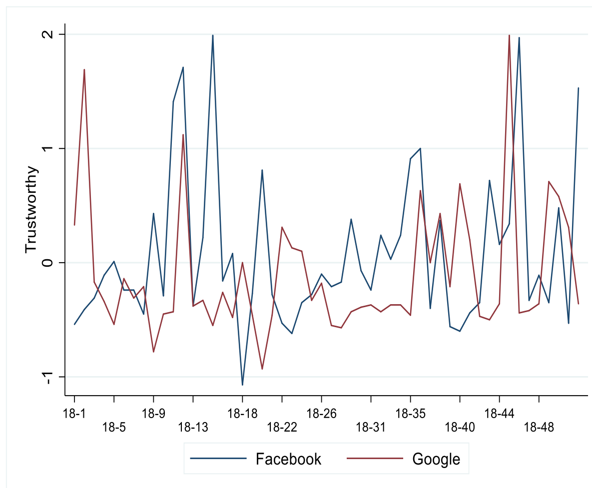
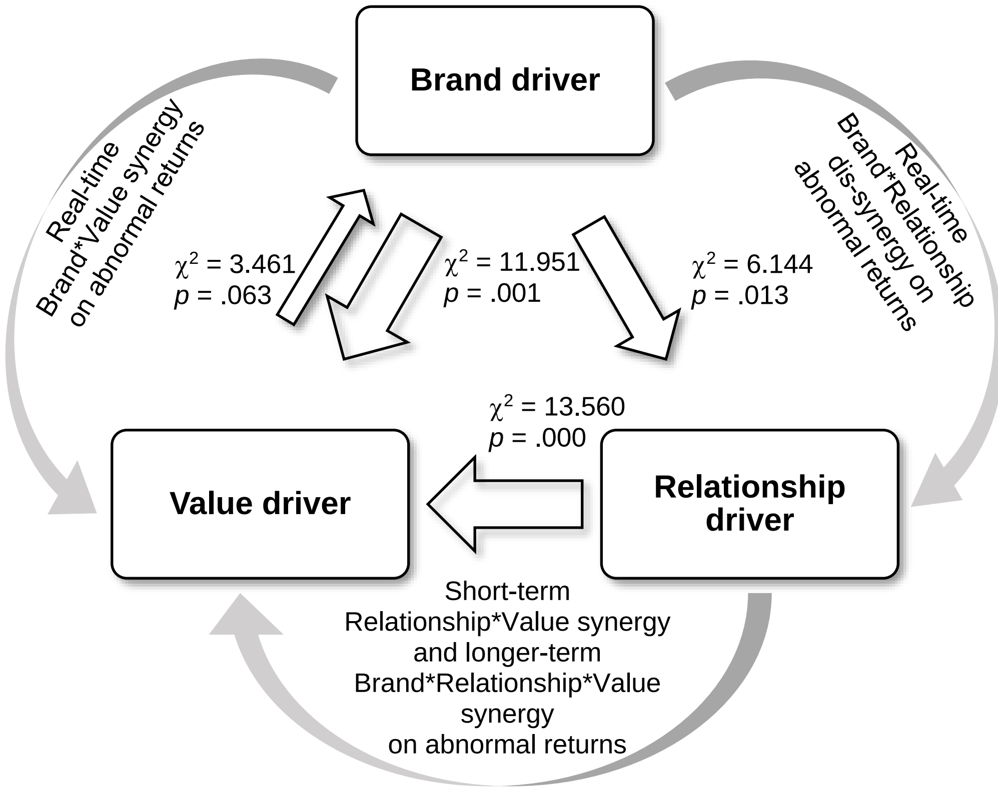
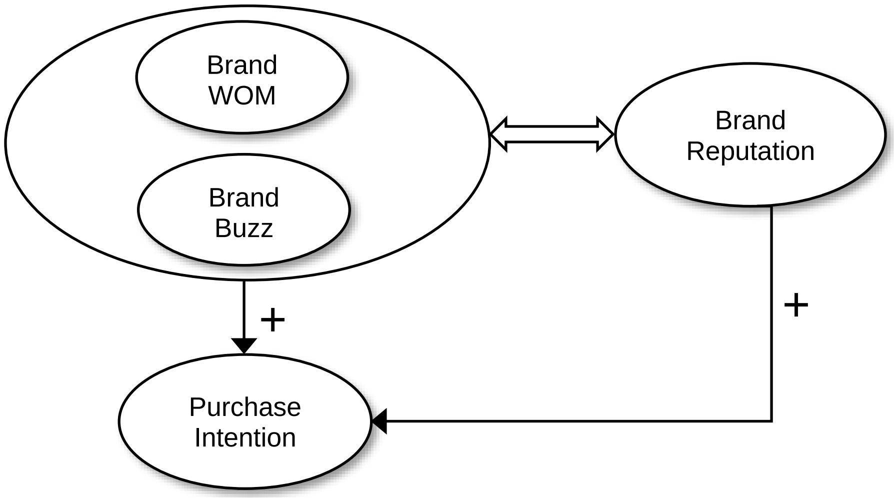


Figure 6. The Mutual Impacts of Brand Reputation Drivers, and their Synergies on Abnormal Stock Returns.



Note. Thicker arrows indicate stronger relationships obtained from the VAR model for all brands. Curved arrows indicate the synergy/dis-synergy obtained from the financial model. The longer-term Brand*Relationship*Value synergy on abnormal returns is obtained from the sentiment model.

Figure 7. Nomological Relationship between the Tracker and YouGov's BrandIndex.



Note. Variables on the left-hand side are from YouGov, and the brand reputation is from the tracker. The BrandIndex measures are based the brand's consumers, whereas our tracker is based all stakeholders.

Table 1. Brand Reputation Drivers, Sub-Drivers, Descriptions, and Dictionaries.

Driver	Sub-Driver	Description	Positive Dictionary	Negative Dictionary
Value	Price	Is the brand known for low prices, such as being cheap, affordable, having deals, bargains, discounts, and sales?	Cheap, afford, inexpens, deal, low, bargain, thrifti, reason, econom, frugal, joy, discount, pleas, sale	Expens, pricey, costly, overpr, unfair, rich, excess, extravag, high, exclus, outrag
	Service quality	Does the brand provide high quality service, such as being competent, helpful, fast, knowledgeable, understanding, with patient and respect?	Help, great, fast, knowledge, attent, understand, easi, polit, patient, respect, prompt, compet	Rude, frustrat, terribl, slow, careless, incompet, disrespect, aw, lazi, irrit, horribl, angri
	Goods quality	Does the brand create high quality products, such as durable, functional, strong, beautiful, and valuable?	Quality, durabl, function, excel, perfect, use, beauty, strong, valu, sturdi, luxuri, worth, long-last, best, satisfi, impress, uniqu, clean	Junk, bad, poor, wast, ugli, breakabl, worthless, flimsi, useless, disappoint, shoddi, mediocr, garbag, short-liv
Brand	Cool	Is the brand known for being trendy, hip, awesome, cool, stylish, and sexy?	Trendi, hip, awesom, cool, modern, stylish, current, sexi	Ordinari, lame, ancient, averag
	Exciting	Does the brand bring a sense of excitement to its products / services, such as being fun, exciting, inspiring, and stimulating?	Fun, excit, inspir, happi, thrill, stimul, live, interest	Bore, dull, uninspir, tire, bland
	Innovative	Does the brand new, smart, technologically advanced, intelligent, innovative, creative, novel, and cutting edged?	New, smart, invent, advanc, cut, futurist, intellig, progress, innov, technolog, creative, novel, cutting-edg	Old, old-fashion, tradit, uninterest, outdate
	Social responsibility	Is the brand caring, benevolent, giving and beneficial?	Benevol, give, benefici	Greedi, uncar, irrespons, evil, profit
Relationship	Community	Does the brand generate a sense of community, such that people are involved, together, and harmonious with the brand, and can communicate and social with the brand?	Famili, involv, commun, social, togeth, harmoni	Cold, sad, selfish
	Friendly	Is the brand nice, pleasant, warm, kind, open, and accommodating?	Nice, friendli, pleasant, kind, warm, welcom, trustworthi, open, accommod	Mean, unpleas, unhelp, unfriendli, aloof, nasti, arrog
	Personal relationships	Does the brand connect personally with its stakeholders by being special, personal, intimate, and close?	Connect, special, person, intim, close, profession, comfort	Cold, distant, imperson, disconnect
	Trustworthy	Is the brand honest, reliable, dependable?	Honest, reliable, good, depend, trust, transpar, safe, honesti, principl, honor	Dishonest, unreli, cheat, shadi, untrustwo, deceit, decept, lie

Note. The goods quality sub-driver applies to goods brands only, while the service quality sub-driver applies to all brands.

Table 2. Descriptive Statistics: Median, Minimum, Maximum, and Correlations among Brand Drivers and Sub-drivers.

Variables	Mean	SD	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Value driver																		
1.Price	171	3,086	-91,415	79,450	1.00													
2.Service quality	413	1,531	-5,368	81,851	.12	1.00												
3.Goods quality	186	1,075	-3,627	62,281	.29	.43	1.00											
Brand driver																		
4.Cool	140	639	-1,976	36,454	.10	.54	.41	1.00										
5.Exciting	347	1,433	-15,620	92,119	.15	.58	.42	.53	1.00									
6.Innovative	944	15,066	-5,009	1,120M	.09	.47	.42	.52	.45	1.00								
7.Social responsibility	117	592	-13,730	28,238	.17	.54	.45	.48	.50	.36	1.00							
Relationship driver																		
8.Community	121	798	-10,371	31,975	.06	.43	.18	.34	.42	.28	.33	1.00						
9.Friendly	143	637	-10.989	24,851	.04	.48	.37	.47	.46	.43	.39	.29	1.00					
10.Personal relationships	182	870	-1,189	36,573	.12	.57	.41	.48	.49	.43	.45	.40	.44	1.00				
11.Trustworthy	214	810	-9,470	33,035	.18	.64	.45	.54	.58	.43	.59	.38	.50	.54	1.00			
12.Value driver	303	1,994	-31,303	80,651	.71	.71	.87	.46	.51	.41	.50	.33	.38	.48	.56	1.00		
13.Brand driver	387	3,882	-3,763	284,245	.16	.69	.57	.81	.80	.75	.75	.44	.57	.60	.69	.60	1.00	
14.Relationship driver	165	584	-1,534	18,910	.13	.70	.52	.60	.65	.52	.59	.69	.74	.79	.80	.58	.76	1.00
15.Brand reputation	285	1,648	-10,175	97,036	.39	.80	.73	.71	.74	.64	.70	.55	.64	.71	.78	.83	.90	.89

Notes: The unit of analysis is brand-week. Variables 1-11 are sub-drivers, variables 12-14 are drivers, and variable 15 is the overall brand reputation. All variables have 13,000 observations, except goods quality sub-driver having 8,580 observations. All correlation coefficients are significant at .000. The “M” in the innovative sub-driver denotes 1,000. Mean, SD, Min, and Max are calculated based on raw scores, and correlations are calculated based on the normalized scores, shown in Equation (W4). Negative minimums indicate the number of negative keyword mentions is greater than the number of positive keyword mentions.

Table 3. The Mutual Impacts of Brand Reputation Drivers: Multivariate Dynamic Panel Model.

Endogenous variables **Predictors**	**Value $_t$** α (z-value)	**Brand $_t$** α (z-value)	**Relationship $_t$** α (z-value)
Endogenous causality			
Value $_{t-1}$	.214(6.48)***	.030(1.86)†	.023(1.31)
Brand $_{t-1}$	.111(3.46)***	.289(8.91)***	.068(2.48)**
Relationship $_{t-1}$	.122(3.68)***	.029(1.19)	.286(9.90)***
Model statistics			
No. observations	12,800		
No. of brands	100		
Avg. no. of weeks	128		
Hansen's J.	≅ 0 (df = 0, p = .)		
Granger causality Wald test	**χ^2(prob.)**	**χ^2(prob.)**	**χ^2(prob.)**
Value $_t$		3.461(.063)	n.s.
Brand $_t$	11.951(.001)		6.144(.013)
Relationship $_t$	13.560(.000)		

Notes: "t" denotes the current value, and "t-1" denotes the lagged one-week value of the respective variables. Instruments include all variables in the equation. *** $p < .001$ ** $p < .01$ * $p < .05$ † $p < .1$

Table 4 Accountability of the Unanticipated Component of Brand Reputation Drivers for Abnormal Stock Returns.

	I		II	
	Main Effect Model		**Interaction Model**	
Predictor	**Estimate**	**z-value**	**Estimate**	**z-value**
Brand Drivers				
U∆Brand $_{t}$	.001**	2.12	.002**	1.96
U∆Brand $_{t-1}$	.001	1.00	.001	1.05
U∆Brand $_{t-2}$	.000	.30	-.000	-.30
U∆Value $_{t}$	.000	.72	.000	.33
U∆Value $_{t-1}$	.001*	1.72	.001	1.13
U∆Value $_{t-2}$	-.001	-.95	-.001	-.97
U∆Relationship $_{t}$	-.002**	-2.19	-.000	-.06
U∆Relationship $_{t-1}$	-.002***	-3.37	-.003***	-3.29
U∆Relationship $_{t-2}$	.001*	1.76	.001	.57
U∆(Brand*Value) $_{t}$			.001*	1.70
U∆(Brand*Value) $_{t-1}$			-.000	-.66
U∆(Brand*Value) $_{t-2}$			.000	.01
U∆(Value*Relationship) $_{t}$			-.001	-1.47
U∆(Value*Relationship) $_{t-1}$			.001*	1.87
U∆(Value*Relationship) $_{t-2}$			.000	.61
U∆(Brand*Relationship) $_{t}$			-.001*	-1.94
U∆(Brand*Relationship) $_{t-1}$			.000	.48
U∆(Brand*Relationship) $_{t-2}$			.001	.84
U∆(Brand*Value*Relationship) $_{t}$			.000	1.17
U∆(Brand*Value*Relationship) $_{t-1}$			-.000	-1.43
U∆(Brand*Value*Relationship) $_{t-2}$			-.000	-1.13
Industry				
Wholesale/retail	.001	1.07	.001	1.04
Transport/warehouse	-.001	-.37	-.001	-.36
Information/finance/ professional/scientific	.001**	2.47	.001**	2.35
Accommodation/food	.000	.20	.000	.19
Year				
2017	.000	.70	.000	.58
2018	-.001	-1.21	-.001	-1.27
Model Details				
Adjusted R-square				
Wald χ^2 (df)		39.56(15)***		52.66(27)***

Notes: “t” denotes the current week value, “t-1” denotes the lagged one-week value, and “t-2” denotes the lagged two-week value of the respective variables. “U∆” denotes the unanticipated component of brand drivers, estimated as the standardized residual using the lagged one-week value of the respective variable (i.e., t-1) as the predictor in a fixed-effect panel model.
* $p < .1$ ** $p < .05$ *** $p < .01$

Table 5. Descriptive Statistics and Correlations across Three Social Media Platforms.

Variables	Mean	SD	Min	Max	Twitter	Facebook	Instagram
Brand reputation							
Twitter	3,172	4,158	-166	41,078	1.00		
Facebook	75	118	0	845	.16	1.00	
Instagram	134	111	0	446	.36	.78	1.00
Brand driver							
Twitter	2,854	3,346	-2,325	23,678	1.00		
Facebook	36	57	-2	381	.23	1.00	
Instagram	174	146	1	623	.24	.77	1.00
Value driver							
Twitter	4,958	8,340	-3,013	80,651	1.00		
Facebook	143	220	-3	1,364	.29	1.00	
Instagram	144	128	0	473	.51	.84	1.00
Relationship driver							
Twitter	1,705	2,354	-1	18,910	1.00		
Facebook	46	93	-1	915	**-.02**	1.00	
Instagram	86	70	0	354	.21	.65	1.00

Note. Bold entries are NOT significant at .05 level. Mean, SD, Min, and Max are calculated based on raw scores, and correlations are calculated based on the normalized scores. Negative value means that the negative sentiment is larger than the positive sentiment.

Table 6. Agenda of Future Research Questions Using the Tracker.

Research Agenda	Specific Research Questions
1. Social media lens to brand reputation	• Why does a certain brand "attract" or "offend" consumers on social media? • How do people talk about brands on social media and what language do they use in representing what brand reputation means? • How to bridge traditional measures to the social media lens of brand reputation? • What contributes to the volume and sentiment variations across brands and over time?
2. Granular investigation of brand reputation	• What are the gains and losses of a brand's reputation along the three drivers over time and why? • What are the underlying mechanisms (i.e., theories and hypotheses) that can explain why an event having differential impacts on a brand's reputation at the driver and sub-driver levels?
3. Longitudinal research on brand reputation	• How do brand reputation drivers and sub-drivers vary and co-vary over time? • What are the within-brand (over time) and between-brand temporal volatilities and dynamics? • How do brands evolve, in a multitude of ways, between brands, across categories, and over time?
4. Brand-related events for brand reputation variation	• What are the brand events and strategic marketing actions that affect brand reputation, given event, brand, and economy characteristics? • How do brand-specific and general events impact brand reputation in both the short- and long-term? • How are brands impacted by negative and positive events such as product recalls, crises and scandals, changes in C-level executives, and product launches?
5. Brand, customer, and firm characteristics for brand driver variation	• How do brand, customer, and firm characteristics, individually and collectively, account for the fluctuations of the brand reputation drivers? • How can the success of a brand be predicted (e.g., direct, moderate, or mediate effects) by the brand reputation drivers and sub-drivers? • Why does a certain brand "attract" or "offend" consumers on social media, as a function of those characteristics? • How does product diffusion vary as a function of brand reputation?
6. Brands in novel classifications	• How can brands be classified based on their scores on the three brand reputation drivers or the larger set of sub-drivers? • What is the best statistical approach for classification to use in subsequent analyses using standard multivariate statistical analysis techniques, such as cluster analysis and multidimensional scaling? • What is the machine learning approach for classification pertaining to understanding differences between brands that score high versus low on various drivers of interest?
7. Brand reputation drivers and marketing/financial outcomes	• What are the returns on brand reputation in terms of marketing/financial outcomes, such as customer (re)purchases, short- and long-term financial performance? • What are the differential impacts of brand reputation drivers and sub-drivers on various stakeholders? • What are new approaches to brand valuation? • What is the best time-series modeling approach to considering the endogeneity of marketing/financial outcomes and brand reputation drivers?

Web Appendix

A COMPARISION WITH EXISTING BRAND MEASURES

Table W1. A Comparison with Existing Brand Measures.

	Granular level	Aggregate level	Conceptual source	Customer's voice/ Stakeholders	Availability
Brand Reputation Tracker	Weekly, monthly, quarterly	• 100 global brands • At the overall brand reputation, driver, and sub-driver levels • Brand sentiments (positive and negative)	Rust et al.'s (2000, 2004) customer equity framework	• Stakeholders' own voice on Twitter • Stakeholders, (current and prospective) customers, employees, and investors	Publicly available
Y&R's BrandAsset Valuator (BAV)	Quarterly, annually	• Brand strength (relevance and differentiation) and brand stature (esteem and knowledge)	Keller's (1993) customer-based brand equity (CBBE) model	• Consumer perception surveys • Customers	For paid clients only
YouGov's BrandIndex	Daily, annually	• An overall ranking of 25 global healthiest brands	Calculate overall brand health index (net of positive and negative feedback) based on 6 metrics: impression, quality, value, satisfaction, recommend, and reputation	• Consumer's brand perceptions based on consumer panel interview • North America only • Consumers	Brand health index released annually, and daily data are only available for paid clients
BrandZ's top 100 most valuable global brands	Annually	• An overall ranking of 100 brands	Calculate the proportion of a brand's contribution to its firm's financial value	• Firm financial value based on expert judgment and calculation • Brand contribution based on consumer interviews • Customers	Publicly available
Interbrand's Best Global Brands	Annually	An overall financial brand value ranking for 100 brands	• A firm's overall financial returns • The brand's contribution to the financial returns • Brand loyalty	• Expert evaluation based on financial data, consumer goods data, and text analytics and social listening • Stakeholders	Publicly available
Forbes' World Best Brands	Annually	An overall financial value ranking for 100 brands	A brand's contribution to the firm's financial value	• Financial revenue and earnings based • Investors	Publicly available

TECHNICAL DETAILS OF THE TWITTER DATA COLLECTION

Tracking brand reputation on social media is dynamic. Brands change their major handles occasionally. We monitor the brand handles periodically to update the brand handles used. For example, in the two stages of data collection (detailed below), Kraft-Heinz changed its handle from @KraftHeinz to @KraftHeinzCo, Michelin added @Michelin as a corporate brand handle in stage 2, and Samsung added @Samsung handle in stage 2.

From there, we established a streaming API collection of all tweets about or from that username. The streaming API works by telling Twitter that we want to filter all tweets being issued for a list of keywords. These tweets were retrieved and stored in a relational database by OssaLabs, a social media analytics company that provides in-depth Twitter analysis.

Stage 1: Generating Dictionaries and Initial Data Collection

We identify 14 sub-drivers for the three brand reputation drivers, through a synthesis of expert opinions, literature support, and real-world observation of what words people use on social media to discuss brands. Rust, Lemon, and Zeithaml (2004) and Rust, Zeithaml, and Lemon (2000) provide a detailed discussion of possible sub-drivers.

For the 14 sub-drivers, we then use a “bag-of-words” approach to creating dictionaries for each of the sub-drivers, which is theory-driven (based on the customer equity framework), making the brand reputation outcome explainable, as opposed to other black-box machine learning models. We rely on crowd sourcing using MTurkers to generate keywords that they would use when writing a tweet about a brand. This approach collects naïve individuals’ mental presentations about the content domain of each sub-driver (Davis and Marcus 2015). We first ask them to imagine a scenario where they had recently interacted with a company and had a positive experience with a brand about one of the sub-drivers. We then ask them to list five keywords that they would use to describe this scenario, and to write a short textual description of the experience. This scenario is repeated with the negative version of the same question. For each of the 14 sub-drivers, we have up to 50 participants complete the procedure. Based on the responses, we create a dictionary taking any keyword that is identified by at least three participants as being related to that sub-driver. This result in 28 dictionaries, one for positive sentiment and one for negative sentiment for each sub-driver.

We then stem the keywords and the actual tweet text using the Porter stemmer (Porter 1980) to make the dictionary comparable to the tweet text. Equation (W1) shows the calculation for positive (or negative) sub-driver based on its positive (or negative) dictionary:

(W1) $\text{Positive subdriver}_{\text{i,j,k}} = \sum_{\text{l}=1}^{|\text{k}|} \text{match}(\text{word}_l, \text{positive subdriver dictionaries}_{i,j,k})$, and
$\text{Negative subdriver}_{\text{i,j,k}} = \sum_{\text{l}=1}^{|\text{k}|} \text{match}(\text{word}_l, \text{negative subdriver dictionaries}_{i,j,k})$,

where *i* is brand, *j* is sub-driver, *k* is tweet, $word_l$ is the *lth* stemmed word in tweet *k*, and *match* is an indicator function that returns 1 if $word_l$ is in the dictionary for sub-driver *j* and 0 otherwise. For instance, the following tweet garners a positive "Cool" sub-driver score due to the words, "fun" and "awesome":

> *Had so much fun at @intel with @fairlightex @mattfromintel (: was awesome for you to show my nephew Virtual Reality thank you!*

Once this is done, we count the numbers of positive and negative tweets for each sub-driver in a given time period (week, month, or quarter) that each brand has. Equation (W2) shows that for brand *i*, for sub-driver *j*:

(W2) $\text{Positive subdriver}_{\text{i,j}} = \sum_{\text{k}=1}^{\text{n}} \mathbf{I}(\text{Positive subdriver}_{\text{i,j,k}})$, and $\text{Negative subdriver}_{\text{i,j}} = \sum_{\text{k}=1}^{\text{n}} \mathbf{I}(\text{Negative subdriver}_{\text{i,j,k}})$,

where $\boldsymbol{I}$ is the indicator function, returning 1 if positive (or negative) $\text{sub-driver}_{i,j,k} > 0$ and 0 otherwise. We then compute a net score for each sub-driver by subtracting the number of negative tweets from the number of positive tweets about the sub-driver, shown in Equation (W3):

$$\text{(W3)}\,\text{Net subdriver}_{i,j} = \text{Positive subdriver}_{i,j} - \text{Negative subdriver}_{i,j}$$

We then normalize the scores for each sub-driver across brands using standard z-scoring to facilitate tracking temporal changes. Equation (W4) shows that for brand *i* and sub-driver *j*:

$$\text{(W4)}\,\text{Scaled positive subdriver}_{i,j} = \frac{\text{Positive subdriver}_{i,j} - \overline{\text{Positive subdriver}_j}}{s(\text{Positive subdriver}_j)},$$

where $\overline{\text{Positive subdriver}_j}$ is the mean of the positive sub-driver$_j$ and s(Positive subdriver$_j$) is the sample standard deviation.

After calculating the positive, negative, and net scores for each sub-driver, we move one level up and calculate the similar three sets of scores for the three brand reputation drivers, which is the average of their respective sub-drivers, and one level further up to calculate the overall brand reputation as the average of the three brand reputation drivers.

Based on the method described above, we collected data for the tracker from the week of June 26, 2016 to the week of June 26, 2017 for one year. To reflect our broader view of stakeholders, we do not limit the brand tweets to be from the brand's customers.

Stage 2: Reevaluating the Theoretical-driven Drivers

After the initial one-year data collection, we were able to evaluate empirically the appropriateness of the keywords for sub-drivers and the representativeness of the sub-drivers for brand reputation drivers.

First, we adjusted the keywords of each sub-driver to better capture the uniqueness of each sub-driver. There are keywords mentioned by participants that appear in more than one sub-driver, such as "good" and "bad." We assigned such duplicated keyword to a sub-driver based on the principles: 1) a keyword that matches a sub-driver name is assigned to that sub-driver, 2) a keyword with a higher mentioned frequency of a sub-driver is assigned to that sub-driver, and 3) a tied keyword is assigned to the sub-driver that has fewer keywords. This procedure intends to balance between keywords representativeness and uniqueness of the sub-driver.

Second, we also updated brand handles, because some brands change their major brand handles during the period. For example, we replace the brand handle @Alibabatalk with @AlibabaGroup because Alibaba added that corporate brand handle after the initial data collection.

Third, we dropped one sub-driver (i.e., Internet/mobile) from the Relationship driver, because it appears to capture whether a brand uses web and mobile applications in general, rather than how a brand connects with stakeholders using web and mobile applications.

After the adjustment, we backfilled the first-year data using Crimson Hexagon and continued collecting new data until the end of 2018.

After the data collection was completed, we reevaluated the sub-drivers and adjusted sub-drivers for the three drivers based on two considerations: First, we want the meaning of the sub-drivers to be connotatively consistent with their respective driver. For example, compared with other sub-drivers, “loyalty program” is a company-centered sub-driver, whereas the other sub-drivers in the Relationship driver are stakeholder-centered perceptions, which explains why it is a poor fit with other sub-drivers. Second, we want the sub-drivers to have desirable statistical properties. For example, sustainability is conceptually too close to social responsibility, tending to generate a collinearity issue, and thus is dropped from the Brand driver.

A final step of the reevaluation is to calculate the precision and recall of the 11 sub-drivers, which are the standard measures for accuracy of classification. Precision (i.e., the positive predictive value) is the number of relevant tweets retrieved divided by the number of all tweets. Recall is the number of relevant retrieved tweets divided by the total number of relevant tweets (i.e., sensitivity). For each of the 11 sub-drivers, we analyze 200 relevant tweets (i.e., 2,200 tweets in total), with 100 tweets matching the positive dictionary and with 100 tweets matching the negative dictionary of each sub-driver. The results show that, on average, the precision and recall across all sub-drivers are .70 and .85, respectively. For each driver, the precision and recall are .69 and .85 for the Value driver, .71 and .85 for the Brand driver, and .67 and .84 for the Relationship driver. This performance is comparable and even better than many approaches that use dictionary-based emotional text-mining (e.g., LIWC), that have a range of precision from .17

to .63 and a range of recall from .38 to .89 (Gill et al. 2008; McDonnell 2015; Roberts et al. 2012).

REFERENCES

Gill, Alastair J., Robert M. French, Darren Gergle, and Jon Oberlander (2008), "The Language of Emotion in Short Blog Texts," CSCW'08: Proceedings of the 2008 ACM conference on Computer supported cooperative work, 299-302.

McDonnell, Michelle (2015), "Comparison of Text Analysis Programs for Identification of Emotional Expression," Loma Linda University Electronic Theses, Dissertations & Projects. 308, [available at http://scholarsrepository.llu.edu/etd/308].

Roberts, Krik, Roach, Michael A., Johnson, Joseph, Guthrie, Josh, and Harabagiu, Sanda M. (2012), "EmpaTweet: Annotating and Detecting Emotions on Twitter," Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12), 3806-3813.

BRANDS IN THE TRACKER

Table W2. Brands Included in the Analysis (in Alphabetical Order).

Brand	Twitter Username	Brand	Twitter Username	Brand	Twitter Username
3M	@3M	Fox	@FOXTV	Nintendo	@NintendoAmerica
Adidas	@Adidas	Gap	@Gap	Nokia	@Nokia
Alibaba	@Alibabatalk	GE	@generalelectric	Nordstrom's	@Nordstrom
Amazon	@amazon	General Mills	@GeneralMills	Oracle	@Oracle
American Airlines	@AmericanAir	Gillette	@Gillette	Pampers	@Pampers
ABC	@ABCNetwork	Google	@google	Pepsi	@Pepsi
American Express	@AmericanExpress	Gucci	@Gucci	Philips	@Philips
Apple	@Tim_cook	H&M	@hm	Ralph Lauren	@RalphLauren
AT&T	@ATT	Harley-Davidson	@harleydavidson	Revlon	@revlon
Audi	@Audi	Hermes	@Hermes_Paris	Rolex	@ROLEX
Bank of America	@BofA_News	Hershey	@Hersheys	Samsung	@Samsung
Barnes & Noble	@BNBuzz	Home Depot	@HomeDepot	SAP	@SAP
Bayer	@Bayer	Honda	@Honda	Shell	@Shell
BMW	@BMW	HP	@HP	Siemens	@Siemens
Budweiser	@Budweiser	HSBC	@HSBC_Press	Sony	@Sony
Canon	@Canon_Camera	IBM	@IBM	Starbucks	@Starbucks
CBS	@CBSTweet	Ikea	@IKEA	Sunoco	@Sunoco
Cisco	@Cisco	Intel	@intel	Target	@Target
Coach	@Coach	Johnson & Johnson	@JNJNews	Texas Instruments	@TXInstruments
Coca-Cola	@CocaCola	John Deere	@JohnDeere	Thomson-Reuters	@Thomsonreuters
Colgate-Palmolive	@Colgate	Kellogg	@KelloggUS	Time Warner	@TWC
Costco	@Costco	Kraft Heinz	@KraftHeinz	T-Mobile	@Tmobile
Dannon	@Dannon	L'Oreal	@Loreal	Toyota	@Toyota
Dell	@Dell	Louis Vuitton	@LouisVuitton	Twitter	@Twitter
Dior	@Dior	Macy's	@Macys	UPS	@UPS
Disney	@Disney	MasterCard	@Mastercard	Vauxhall	@vauxhall
eBay	@eBay	Mattel	@Mattel	Verizon	@Verizon
Ericsson	@Ericsson	McDonald's	@McDonalds	Visa	@Visa
ESPN	@ESPN	Mercedes-Benz	@MercedesBenz	Volkswagen	@VW
Exxon	@ExxonMobil	Michelin	@MichelinUSA	Walgreens	@Walgreens
Facebook	@Facebook	Microsoft	@Microsoft	Wal-mart	@Walmart
FedEx	@FedEx	Nestle	@Nestle	Wells Fargo	@WellsFargo
Ford	@Ford	Nike	@Nike	Whole Foods	@WholeFoods
				Yahoo	@Yahoo

Note: The seven pure online brands (Alibaba, Amazon, Facebook, Google, Twitter, Yahoo, and eBay) are scaled separately due to their disproportional high number of tweets, compared to offline brands.

Table W3. Tweet Volume and Sentiment of Drivers and Sub-Drivers.

Variables	Net Volume (SD)	Positive Volume (SD)	Negative Volume (SD)	Sentiment ratio (+/-)
Value driver	**303(1,994)**	**529(1,825)**	**226(1,229)**	**2.34**
Price	171(3,086)	553(2,150)	382(2,503)	1.45
Service quality	413(1,531)	478(1,650)	65(242)	7.35
Goods quality	186(1,075)	230(1,195)	44(184)	5.23
Brand driver	**387(3,882)**	**419(3,912)**	**32(136)**	**13.09**
Cool	140(639)	150(656)	10(45)	15.00
Exciting	347(1,433)	365(1,453)	18(158)	20.28
Innovative	944(15,066)	1,012(15,117)	68(299)	14.88
Social responsibility	117(592)	147(621)	31(293)	4.74
Relationship driver	**165(584)**	**201(690)**	**36(145)**	**5.58**
Community	121(798)	157(853)	36(225)	4.36
Friendly	143(637)	197(725)	54(261)	3.65
Personal relationships	182(870)	198(896)	16(62)	12.38
Trustworthy	214(810)	251(905)	37(233)	6.78
Brand reputation	**285(1,648)**	**383(1,705)**	**98(433)**	**3.91**

Note: Due to service brands not having the goods quality measure, the averages of the Value driver are not equal to the averages of the sums of its three sub-drivers.

INDUSTRY CHARACTERISTICS OF BRANDS IN THE TRACKER

Table W4. Summary of Industry Characteristics of Brands in the Tracker.

Industry Characteristics	No. of Brands	%	Brands
31.Manufacturing	5	7.46	General Mills, Hershey, Kellogg, Kraft Heinz, Nike
32.Manufacturing	3	4.48	Exxon Mobil, Revlon, Shell
33.Manufacturing	19	28.36	3M, Apple, Canon, Cisco, Coach, Dell, John Deere, Ericsson, General Electric, HP, Harley-Davidson, Honda, Intel, Philips, Mattel, Nokia, Sony, Texas Instruments, Toyota
42/44/45.Wholesale trade/Retail trade	13	19.40	Amazon.com, Barnes & Noble, Costco, eBay, Home Depot, Macy's, Nordstrom's, Ralph Lauren, Sunoco, Target, Walgreens, Wal-Mart, Whole Foods
48/49.Transportation and warehousing	3	4.48	American Airlines, FedEx, UPS
51.Information	13	19.40	AT&T, CBS, Disney, Facebook, Fox, Google, Microsoft, Oracle, SAP, Time Warner, Twitter, Verizon, Visa
52/54.Finance and insurance/ Professional, scientific, and technical services	9	13.43	Alibaba, American Express, Bank of America, HSBC, IBM, Mastercard, Thomson-Reuters, Wells Fargo, Yahoo
72.Accommodation and food services	2	2.99	McDonald's, Starbucks
Total	**67**	**100.00**	

Notes: The entries show the brand NAICS classification at the beginning of the data period. Some firms change their NAICS classification during the data period. For example, Thomson Reuters changed its NAICS classification from 51 to 52 on the week of April 25, 2017. MasterCard changed its classification from 54 to 56 on the week of July 19, 2016, and then changed from 56 to 51 on the week of January 12, 2017. Only single-brand firms (i.e., corporate brands) are included in the financial analysis.

MONITORING COMPETITION: APPLE VS. SAMSUNG

To demonstrate the generalizability of the monitoring and diagnostic value of the brand reputation tracker, we use another renowned dyad in the high-tech goods industry, Apple versus Samsung, to further illustrate their competitive dynamics, shown in Figure W1 (brand reputation), Figure W2a-W2c (drivers), and Figure W3a-W3i (selected sub-drivers).

The brand reputation time series (Figure W1) shows that Samsung appeared to have higher brand reputation than Apple for the first half of the year, especially in March and April, whereas Apple appeared to have higher brand reputation than Samsung in the second half of the year, with two peaks in September and November.

Samsung's high brand reputation in March reflected the brand event that its flagship smartphones Galaxy S9 and Galaxy S9+ were launched officially in March 16. The two new models introduced innovative features of dual aperture camera, slow-motion video, and augment reality (AR) emojis for consumers to capture and share special moments and express themselves. Figures W2a reveals that the Brand driver was the major source of March reputation surge, followed by the Value driver but to a lesser degree. Figure W3a further reveals that stakeholders talked about Samsung as being innovative, likely due to those new features, as indicated by the surges of goods quality (Figure W3e) and price (Figure W3f) sub-drivers.

Apple's twin-peak brand reputation in September and November reflected its two major brand events. On September 21, Apple released three new iPhones (iPhone XS, iPhone XS Max, and iPhone XR), which are equipped with innovative features of A12 Bionic processor, up to 512 GB internal storage, dual rear cameras, and dual SIMs. In the first week of November, Apple released new iPad Pros, MacBook Air, and Mac mini and in November 14 released the 15-inch MacBook; with features such as high-resolution Retina display, faster performance, Touch ID fingerprint recognition, portability improvements and USB-C connectivity, to bring those backseat family members back to the front. The two peaks in reputation were captured by all three drivers, indicating its brand reputation has a more solid foundation across all three drivers. The November release appeared to satisfy those existing computer product customers, reflected in the surge of the Relationship driver (Figure W2c) in November and December. Figure W3b further reveals that the release of new iPhone models is discussed as very exciting, whereas the release of new computer products creates surge in the personal relationship (Figure W3g) and friendly (Figure W3i) sub-drivers.

Additional comparisons of the Value driver (Figure W2b) and its three sub-drivers reveal that the brand reputation tracker nicely reflected that Apple is a premium brand emphasizing service quality, whereas Samsung is an affordable brand emphasizing price. Figures W3e and W3f shows that people talked about Samsung's price and goods quality positively during the March peak, while they mainly talked about Apple's service quality in the two peaks (Figure W3d), reflecting their respective positioning in the market.

Figure W1. Brand Reputation: Apple vs. Samsung 2018.

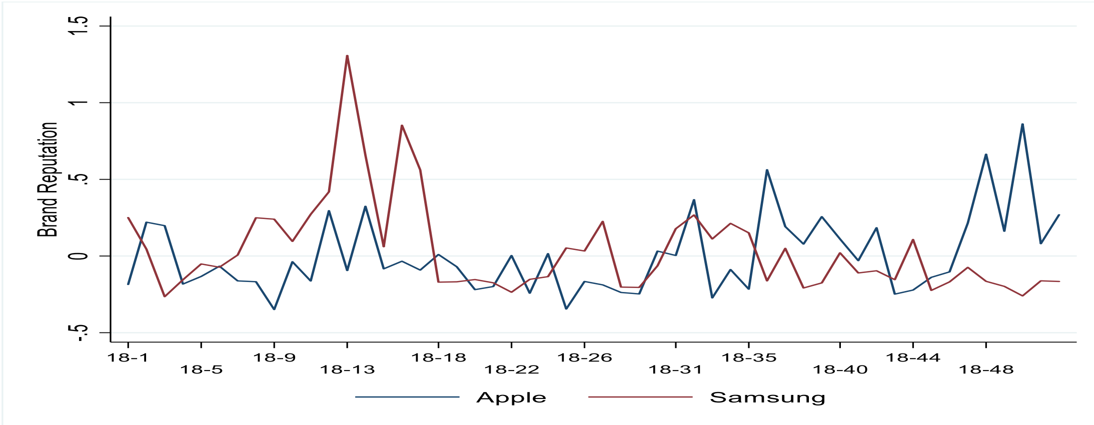


Figure W2a. Brand Driver

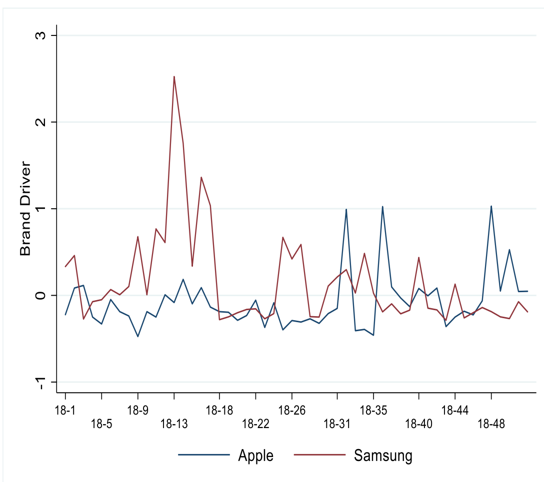


Figure W2b. Value Driver

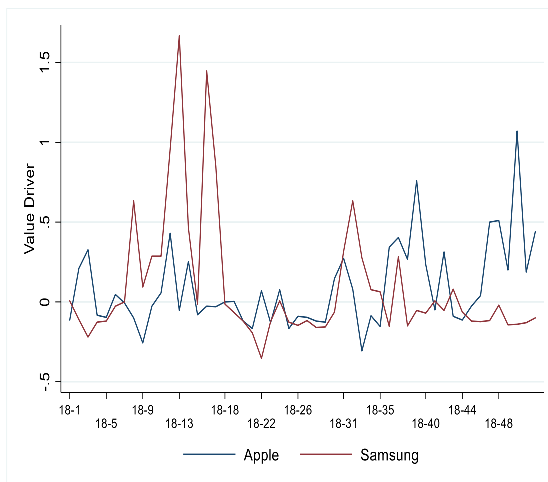


Figure W2c. Relationship Driver

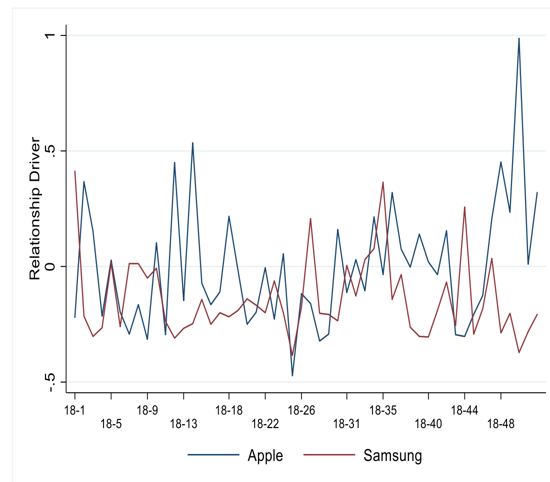


Figure W3a. Innovative Sub-Driver

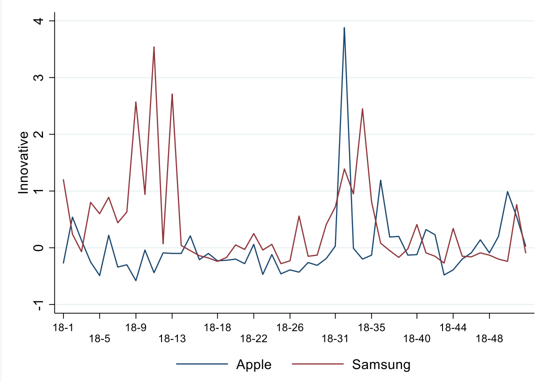


Figure W3b. Exciting Sub-Driver

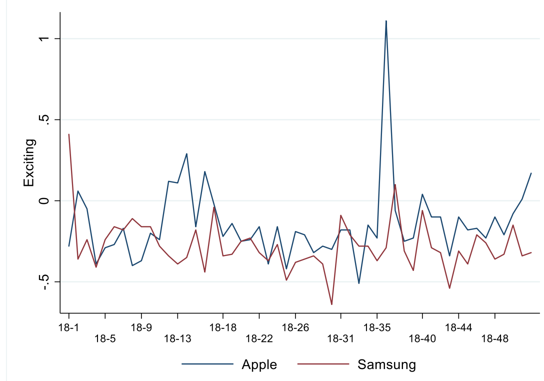


Figure W3c. Cool Sub-Driver

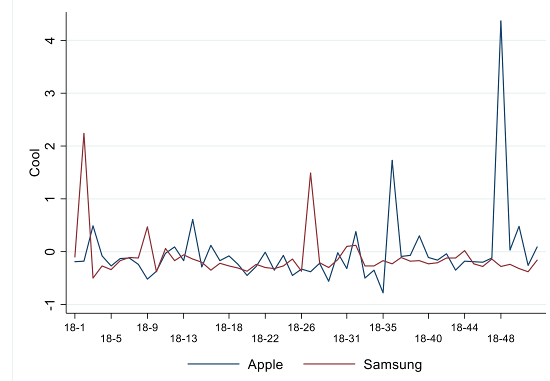


Figure W3d. Service Quality Sub-Driver

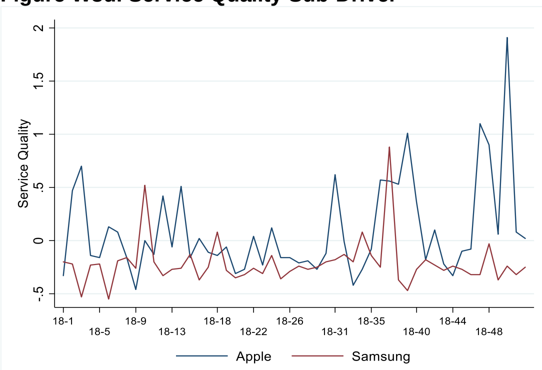


Figure W3e. Goods Quality Sub-Driver

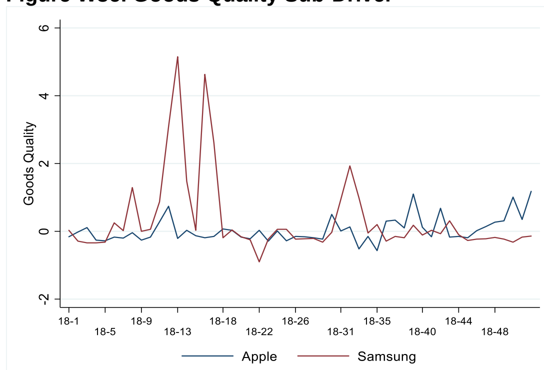


Figure W3f. Price Sub-Driver

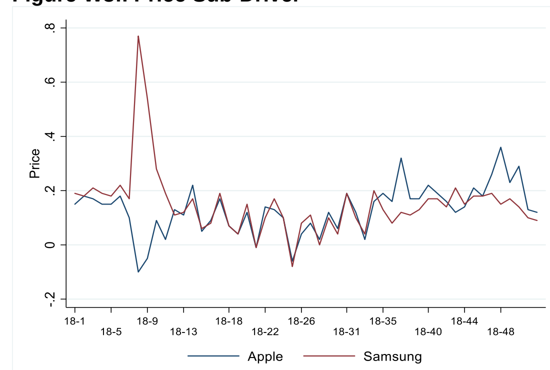


Figure W3g. Personal Relationship Sub-Driver

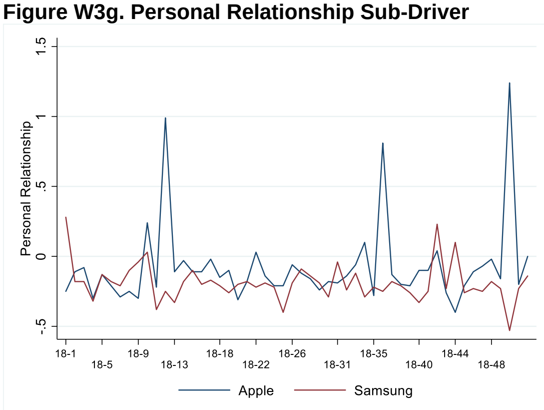


Figure W3h. Trustworthy Sub-Driver

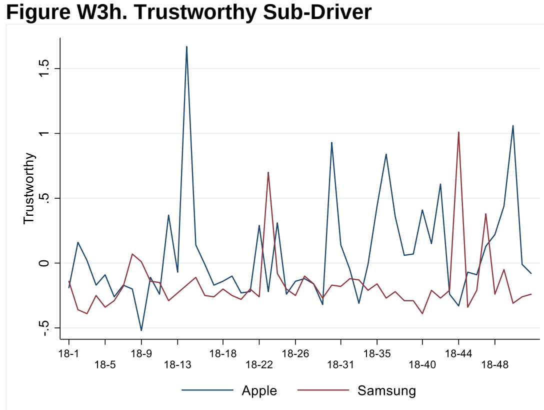
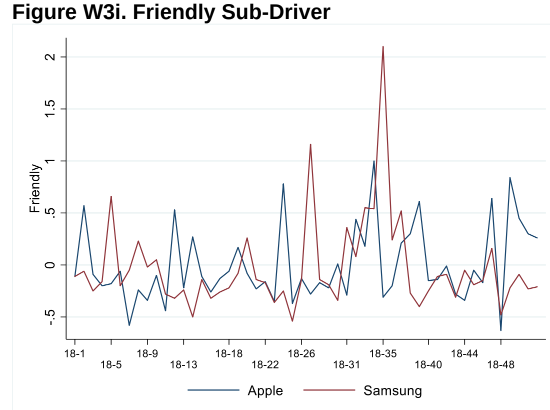


Figure W3i. Friendly Sub-Driver



IMPULSE RESPONSE FUNCTIONS OF THE DRIVER DYNAMICS

Figure W4a. The Impulse Response Function of the Brand Driver on the Value Driver.

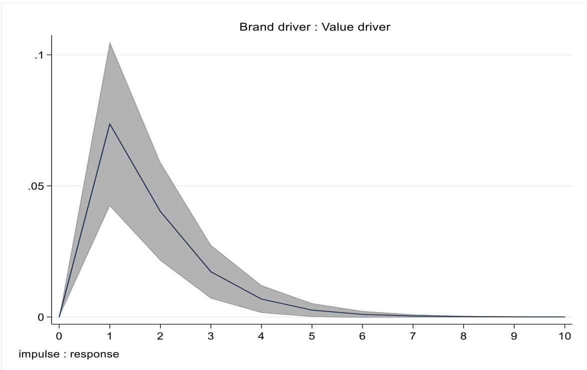


Figure W4b. The Impulse Response Function of the Value Driver on the Brand Driver.

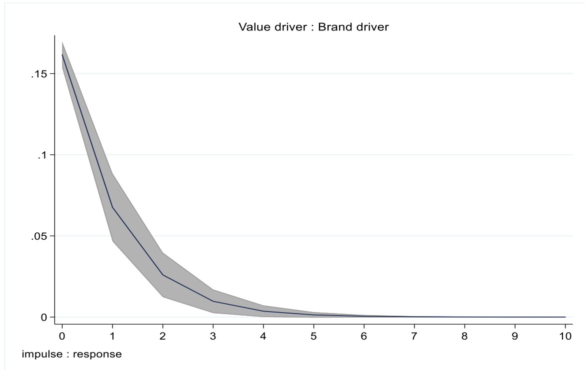


Figure W4c. The Impulse Response Function of the Relationship Driver on the Value Driver.

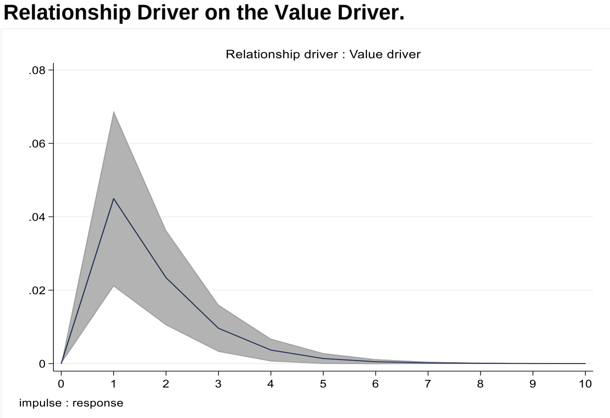


Figure W4d. The Impulse Response Function of the Brand Driver on the Relationship Driver.