CrossMark

Financial Disclosure and Customer Satisfaction: Do Companies Talking the Talk Actually Walk the Walk?

Ronald J. Balvers[1] · John F. Gaski[2] · Bill McDonald[2]

Abstract Using the emerging technology of large-scale textual analysis, this study examines the use of the term 'customer satisfaction' and its variants in the annual reports issued by publicly traded U.S. corporations and filed with the Securities and Exchange Commission as Form 10-K. We document the frequency of the term's occurrence in 10-Ks over the 1995–2013 period and the differences in usage across industries. We then relate the term's usage in 10-Ks to subsequent scores from the American Customer Satisfaction Index (ACSI) to determine whether management's discussion of customer satisfaction in financial disclosures is credible. The commitment of management to shareholders versus, more broadly, stakeholders is a central question in business ethics, and the integrity of management communication is a fundamental construct in the American Marketing Association's Statement of Ethics. We document a complex relation between management's discussion of customer satisfaction and subsequently reported satisfaction. We find that the general use of customer satisfaction (and similar terms) in 10-K documents is negatively correlated with subsequent ACSI scores. However, for retail firms, when the phrase is located near words indicating measurement or monitoring of the phenomenon, the empirical relation is reversed and becomes positive.

✉ John F. Gaski
  jgaski@nd.edu

  Ronald J. Balvers
  balvers@mcmaster.ca

  Bill McDonald
  mcdonald.1@nd.edu

[1] McMaster University, Hamilton, ON L8S 4M4, Canada

[2] University of Notre Dame, Notre Dame, IN 46556-5646, USA

Customer satisfaction is a core construct in marketing management, probably the most central of all as it is intrinsic to the venerated marketing concept itself (e.g. Kotler 1972, pp. 17–23). Although partially aligned, the interests of shareholders and customers usually are not fully aligned, presenting an ethical dilemma for management who must trade-off the fiduciary responsibility of generating value for shareholders versus the ethical obligations towards customers. In this paper, we consider how management, in their most critical periodic financial disclosure, communicates the importance of customer satisfaction for the firm.

The bridge between customer satisfaction, marketing, finance, and ethics is well established in the extant literature. For example, the American Marketing Association (AMA) Statement of Ethics emphasizes the commitment of marketers to "the highest standard of professional ethical norms and values" in the context of a firm's stakeholders, with an emphasis on customers.[1] Our focus on a company's discussion of customer satisfaction in its mandated financial disclosures is underscored in the AMA statement of ethics, which repeatedly emphasizes the importance of trust, honesty, and integrity in communications with stakeholders. Additionally, Ittner et al. (2009) establish that customer satisfaction is value-relevant for shareholders, thereby reinforcing the pertinence of a customer focus in the finance realm.

The more general issue underlying our research is the fundamental debate—spanning ethics and economics—

---

[1] The AMA Statement of Ethics can be found at: https://archive.ama.org/archive/AboutAMA/Pages/Statement%20of%20Ethics.aspx.

concerning the domain of a firm's most fundamental objective: shareholders or, more inclusively, all stakeholders. The imperative of serving the interests of all stakeholders in the firm, as emphasized by the stakeholder theory of business ethics (Freeman 1984; Hasnas 2013), implies an attention to customers that transcends the objective of value maximization advocated by the shareholder theory of business ethics (Friedman 1962; Jensen 2002). As Jensen (2002) argues, long-run value maximization must necessarily take all stakeholders into account. However, Hasnas (2013) interprets stakeholder theory as implying that management's first priority among stakeholders should be its customers.

When firms discuss customer satisfaction in their financial disclosures, we consider three different contexts that might occur: (1) They might be engaging in 'cheap talk,' i.e. misrepresenting the firm's true commitment to customer satisfaction in order to maximize current value; (2) they might be legitimately signaling a commitment to the construct; or (3) they might simply be discussing an unresolved issue concerning customer complaints. To compare the hypothesized scenarios, we use textual analysis to study how the context of customer satisfaction-oriented comments impacts subsequent survey measures of a firm's customer satisfaction. We provide a theoretical model for a signaling equilibrium where we can distinguish between cheap talk in a pooling equilibrium and credible signals in a separating equilibrium, and we include an empirical measure that attempts to capture those firms that are simply discussing an extant problem with customer satisfaction.

Because serving customers well is crucial for any business, whether it adheres to a value maximization view or a stakeholder view, it is natural for corporate executives to discuss customer satisfaction in their mandated financial disclosures. We focus on the annual report (Form 10-K) which is the most value-critical of the disclosures filed with the Securities and Exchange Commission (SEC). The information involving customer satisfaction appearing in a 10-K might be assessed directly by individual investors, as in Lawrence (2013) who finds that the content of financial disclosures affects individual investment decisions, or might filter through stock analysts, as in Luo et al. (2010), who show that customer satisfaction impacts analyst recommendations. More generally, 10-K verbiage should reflect the overall external message and brand imaging of the firm regarding customer satisfaction. Michalisin (2001) and Michalisin and White (2000) substantiate the validity of assertions made in annual reports as accurate representations of the firm and its managers.

Focusing on a large sample of 10-K filings from 1995 to 2013, we use textual analysis to examine whether firms' inclusion of the phrase 'customer satisfaction' (and its variants) is correlated with measured outcomes of customer satisfaction based on the well-known American Customer Satisfaction Index (Fornell et al. 1996). The question is: Is more discussion of customer satisfaction in managements' communications to shareholders associated with increased customer satisfaction as measured by the ACSI Index? Importantly, we also examine a more refined view of this question: Are managers using a cheap signal to lay false claim to the positive attribute of satisfied customers, or are they simply acknowledging existing problems? And, if management declares any kind of monitoring of the attribute, does this make the signal more effective?

To preview results, we find that use of 'customer satisfaction' (or similar phrases) in 10-K filings is substantially different across industries, but it is not clear if a company's rhetorical sensitivity to the issue actually translates into satisfied customers. For the longitudinal perspective, the proportion of companies using the focal 'customer satisfaction' terms in their 10-K filings has increased during the 1995–2013 interval. To find whether this increase in attention is associated with increases in customer satisfaction itself, i.e. fulfillment of the ideal, we then match American Customer Satisfaction Index (ACSI) scores with the 10-K filings. After controlling for other factors that might impact the overall customer satisfaction score, we find that the incidence of 10-K customer satisfaction language is significantly related to the American Customer Satisfaction Index score—in a nuanced way predicted by a signaling environment.

These results suggest that the voluntary use by firms of monitoring and measurement of customer satisfaction is a credible signal that is associated with higher subsequent customer satisfaction. Simply discussing negative issues surrounding customer satisfaction, not surprisingly, is associated with, on average, lower satisfaction scores. And, simply discussing customer satisfaction, with no credible commitment, appears to be 'cheap talk,' and is associated with a lower subsequent measure of customer satisfaction.

Previous Research

Customer Satisfaction and Stock Valuation

Formal measurement of the customer satisfaction construct has a rich history in the managerial and consumer behavior literatures. Pioneering efforts, typically applied to *consumer* satisfaction, include Anderson (1973), Cardozo (1965), Czepiel and Rosenberg (1977), and Olshavsky and Miller (1972). The measurement context transited to the mass or aggregate satisfaction level with Hustad and Pessemier (1973) and Lundstrom and Lamont (1976), and was extended longitudinally by Barksdale et al. (1976) and

Barksdale and Darden (1972). Barksdale and colleagues measured generalized consumer satisfaction with business at two-year intervals from 1971 through 1975 using a national U.S. sample survey. They consistently found somewhat negative attitudes toward business.

Beginning in 1983 and continuing annually, Gaski and Etzel (1986, 2005; Gaski 2008) measured and reported U. S. consumer sentiment toward the establishment of marketing practice using a national sample and 20-item composite scale. Their work reveals aggregate consumer sentiment toward marketing (*qua* satisfaction) that is ratcheting higher on a gradual trajectory. A decade later, Fornell et al. (1996) created the ACSI to assess consumer satisfaction with the products of about 200 individual firms. From telephone survey responses and multiple-indicator measures, satisfaction is also aggregated across industries, sectors, and the entire U.S. economy. Over the project's history, ACSI consumer satisfaction scores show a slight positive trend. It is the Fornell et al. ACSI research that provides a measurement environment and data source for this study.

Fornell et al. (2006) also examine the relation between stock market valuation and customer satisfaction, concluding that trading rules based on ACSI data can produce long-term abnormal returns. Aksoy et al. (2008) confirm these findings. In contrast, using more comprehensive methods from finance and accounting, both Jacobson and Mizik (2009) and Ittner et al. (2009) find that the ACSI data do not provide information that facilitates profitable trading. Although these results suggest controversy over whether customer satisfaction measures can be used to trade profitably in the stock market, Gruca and Rego (2005), Ittner et al. (2009), and Luo et al. (2010) show that customer satisfaction is value-relevant in providing information on future cash flows and operating income, and impacting the assessments of stock analysts. While this line of research is distinct from the present study, it provides the linkage between customer satisfaction and firm value that underlies our focus on 10-K communications. These studies establish that customer satisfaction is related to a firm's stock value, which motivates the research premise that customer satisfaction is an attribute meriting attention in management's communication with shareholders.

The conceptual framework for our empirical analysis presumes that discussion of customer satisfaction in the 10-K document provides a public signal of management's commitment to the concept. Importantly, this does not require that consumers carefully digest the corporate filings, nor that management is even signaling to consumers per se. Many channels potentially provide an avenue for the 10-K communication to serve as a general signaling mechanism. For example, management rhetoric in the 10-K could simply be a consistent representation of the firm's culture as broadly espoused in its public documents, advertising, and brand building efforts. Erdem and Swait (1998), citing the importance of consistency, suggest that a firm's communications with shareholders should parallel its message in other commercial venues. Alternatively, given the linkage to stock valuation, the firm's signal could be interpreted via stock analysts, with the ubiquity of business news media a further information conduit. As noted earlier, Michalisin (2001) and Michalisin and White (2000) validate that management rhetoric can provide credible signals about underlying firm attributes.

The Present Study's Differential Contribution

While the cited research has targeted the relation between customer satisfaction and financial outcomes, no previous research has systematically investigated the role of public statements about customer satisfaction made by management in required disclosures and their link with exogenous measures of customer satisfaction. Recent work has pioneered the analysis of management statements through annual reports or 10-K textual analysis, and some of it has connected text characteristics, such as ethics affirmation, to objective results (Loughran et al. 2009). The present project, however, is the first to employ large-scale textual analysis to methodically identify customer satisfaction-related verbiage in financial reports and relate it to metrics of realized customer satisfaction. When interpreted in a signaling environment, the results have implications for whether management can credibly signal a commitment to customer satisfaction, and whether under some circumstances their signals are 'cheap talk' possibly designed to mislead stakeholders. As highlighted by the AMA, the integrity of management communication is a fundamental ethical issue for all firms.

Textual Analysis

The method we refer to as textual analysis is used under various names across many disciplines—e.g. computational linguistics, natural language processing, or information retrieval.[2] The popularity of the technique has trended upward with the availability of more potent technology and computational access to large text databases. In the early 1900s, the technique was frequently used by journalists looking for patterns in media, while the mid-twentieth century saw applications to political speeches emerge (Dovring 1954–1955). The increasing availability

---

[2] In advertising, the term 'textual analysis' sometimes has been used in a different context to focus on the linguistic interpretation of text appearing in ads (see, for example, Stern 1996). Our usage of the term is confined to the computational analysis of large text sources.

of computers contributed to rising interest in textual analysis beginning in the 1950s. The use of textual analytic techniques continued to spread across the disciplines as computers became more powerful and large bodies of text became more prevalent (Stone et al. 1966), but the arrival of the internet and corresponding importance of parsing documents for content to facilitate search engines caused a rapid increase in the method's popularity and a corresponding increase in the sophistication of techniques. A contemporary and tangential offshoot of the genre in the business research context includes Forster et al. (2009), Holder-Webb and Cohen (2012), Patelli and Pedrini (2013), and Yuthas et al. (2002) who examine documents for tone or specific language.

Although the computational assessment of a document for context and intent is a task challenging the boundaries of artificial intelligence, in our application we simply parse a document for targeted phrases. Using the programming technique known as 'regular expressions,' it is relatively straightforward to isolate a predetermined set of phrases as opposed to attempting to deeply parse a document for context and meaning. Ultimately, in the present study, we are interested in relating the frequency and context of customer satisfaction-oriented words with subsequent survey measures of the construct.

Theoretical Motivation and Hypotheses

Theoretical Perspective

In important public communications, such as the 10-K report, management must be aware of the impact its statements have on the stakeholders of the firm. In view of its insider perspective, the 10-K message content will be watched closely. If management wants to be believed it should avoid 'cheap talk,' i.e. messages that are not credible because the incentives to follow through are either absent or unclear to stakeholders. Applying Spence (1973), a statement provides a credible signal of good intent if offering it would have repercussions too costly for anyone with an intention to mislead.

Our theory, presented formally in the Appendix, considers the incentives of management from a value-seeking perspective to boost short-term profitability by overselling the quality of a product and the satisfaction it should provide customers. The cost of such opportunistic behavior is that, the larger and clearer the overselling proves to be in hindsight relative to realized customer satisfaction, the more long-term sales and profitability will be hurt. Thus, realizing management's basic incentives, stakeholders will believe a stated emphasis on customer satisfaction if deviating from this statement sufficiently reduces future profitability.

We position the firm's commitment to customer satisfaction, together with the customers' expected level of satisfaction, in the context of a signaling framework. Our approach resonates with the theory of brand equity in Erdem and Swait (1998). They use an asymmetric information environment, in which firms know more about the quality of their product than do customers, to assess brand signaling by firms. As in their model, our theory also emphasizes that the impact of the signal depends critically on its credibility.

A firm's actual commitment to satisfying customers—as reflected in its corporate culture, personnel performance measures, training, resource allocation, production processes, etc.—is only indirectly observable by the ultimate customer. The resulting informational asymmetry between management's inside knowledge of their degree of commitment to customer satisfaction and customers' outside observations produces the incentive to signal information as first formulated by Spence (1973): To overcome the adverse selection problem presented by Akerlof (1970)—that only the worst quality types will be available to the uninformed—the agent with superior information must provide credible signals to convince uninformed agents about quality (or other attributes unobservable to outsiders). In their paper on motion pictures, Basuroy et al. (2006) provide a market-oriented example of a similar asymmetric information environment that leads to signaling.

Declaring the importance of customer satisfaction in a 10-K filing is an essentially costless signal. With costless signals it is impossible to create a 'separating' equilibrium (in which 'better' firms can be distinguished from 'worse'). In Spence's scenario, 'better' job candidates cannot differentiate themselves from less productive ones unless they send the costly signal of overeducating themselves. A parallel example is provided by Leland and Pyle (1977) in which entrepreneurs who have a project they know to be good cannot convince investors to finance it by cheap talk, i.e. simply stating they have a good project. They must reveal their knowledge of good value by sending a costly signal, namely by committing to finance a substantial part of the project themselves.

In our model, if a credible 10-K signal of commitment to customer satisfaction is absent, the relation between management's discussion of customer satisfaction and realized customer satisfaction is not clear. The mere fact that management discusses customer satisfaction in public disclosures might simply indicate that it is an important component of value or an ongoing issue, but does not credibly signal that management is committed to improving or maintaining customer satisfaction. In the empirical analysis to follow, we incorporate textual analysis to distinguish the simple signal of discussing customer

satisfaction from the case where it is addressed in the context of measurement or monitoring. In this signaling environment, management subjecting itself to objective benchmarks for customer satisfaction provides a credible signal of commitment to the ideal. This signal becomes credible because publicly available measurement of the realized level of customer satisfaction makes deviations between promised and realized customer satisfaction costly. These deviations make the signal more costly because they are plainly observable after the fact and would lead to either expensive corrective action or disaffected customers and loss of future sales.

In the absence of a purely ethical inhibition against providing misleading information to customers, management thus has a value-seeking incentive to overstate its commitment to customer satisfaction if this action has a strong positive impact on current demand and profitability, and a weak impact on future profitability when found out. However, in this case management rhetoric will not be believed unless it makes its commitment to customer satisfaction more credible by also transparently committing to monitoring and measuring realized customer satisfaction. This has the effect of increasing the negative impact on future profitability when realized satisfaction falls short of the promised level: Lower future profit reflects the costs of providing refunds, additional service, follow-up calls, and employee dissatisfaction, or risking loss of future sales, when customers discover that promised satisfaction is clearly not delivered.

In this view, the extent to which a firm biases its marketing message varies positively with the impact promises have on current profitability and negatively with the impact unmet promises have on future sales. When accountability for the firm's promises is increased, as is the case when the firm commits to monitoring and measuring customer satisfaction, the promised and realized customer satisfaction become better aligned.

Hypotheses

The following hypotheses can be inferred from the signaling equilibrium:

1. The relationship between ACSI scores and customer satisfaction statements without credible commitment is indeterminate.
   The lack of an attempt to monitor or measure customer satisfaction makes it clear to consumers that any prior promises cannot easily be verified after the fact, at least in terms of mechanisms discussed in the firm's 10-K. Talk is cheap in this setting, and the firm is in practice unaccountable for earlier promises. Firms may pay lip service to customer satisfaction but this is not likely to be tightly linked to performance, and consumers should rationally ignore such talk.

2. Firms with a credible promise of higher customer satisfaction generate higher future ACSI scores.
   The reason is that management of these firms has committed to not exaggerating promises because monitoring makes a mismatch between prior claims and realized satisfaction too noticeable, with repercussions of lost future profitability.

3. In industries with reduced incentives for management to overpromise (low profit margins and heavy reliance on repeat sales), the impact of credibly promised customer satisfaction on future ACSI scores is stronger.
   In this type of market environment, firms have less to gain and more to lose from overpromising. Thus, promised and realized customer satisfaction are more closely aligned, in which case firms also have the incentive to make themselves accountable by making the realized level of customer satisfaction transparent. In the opposite case—firms with high profit margins and infrequent repeat sales—even fully committing to monitoring and measuring consumer satisfaction may not be sufficient to make the firm's promises credible. In this case the firms' statements cannot support a separating equilibrium and, instead, a pooling equilibrium emerges in which a firm can offer any 'cheap talk' but will not be believed by rational consumers and, all else equal, such firms would be expected to provide a level of consumer satisfaction that is average for the group. We then can neither predict the promised level of consumer satisfaction nor the degree of monitoring.

We test these three hypotheses in the following by using as a proxy for promised customer satisfaction the incidence of phrases indicating customer satisfaction in 10-K reports. However, in some cases firms could be discussing customer satisfaction simply as an artifact of existing issues that have been publicly revealed. Thus, the 10-K customer rhetoric variable to be used excludes cases where negative words occur nearby (and the excluded cases are used to create another variable to represent that alternative condition). Credibility of the promise is proxied by identifying the cases in which occurrences of the phrase 'customer satisfaction' are found together with words indicating measurement or monitoring. Realized customer satisfaction is measured by using the subsequent ACSI score of a firm. Industries are divided into retail and non-retail to proxy for, respectively, low incentives for management to overpromise (low profit margins and reliance on repeat sales) and high incentives for management to overpromise.

Data and Methods

The 10-K Sample

Publicly traded companies are required by the SEC to "disclose information on an ongoing basis" (see www.sec.gov). For companies based in the United States, most of these ongoing periodic disclosures come in the form of a 10-K, 10-Q, or 8-K filing. The 10-K is management's annual report on a company's business and financial condition, which is distinct from the annual report to shareholders. (The shareholder annual report is typically a subset of the 10-K.) The 10-Q form is filed in the first three quarters of a company's fiscal year and is an unaudited update of the information contained in a previous 10-K. The 8-K provides current reports on twenty-eight events itemized by the SEC, such as "Election of Directors," "Regulation FD Disclosure," or "Notice of Delisting." This study focuses on the 10-K filings because, relative to a 10-Q, they are more likely to contain new information and because the 10-Ks are similar in structure across most firms, as dictated by Regulation S-K. We do not include amended filings in the sample.[3]

Electronic filings were phased in on the SEC's Electronic Data Gathering, Analysis, and Retrieval (EDGAR) website (www.sec.gov) beginning in 1994 and were required of all firms (except for rare hardship exemptions) as of May 1996. We download all 10-K filings for the years 1995 through 2013. This produces an initial sample of 198,991 unique 10-Ks. We then match the 10-K filings with the ACSI survey data and with financial data from CRSP (University of Chicago's Center for Research in Security Prices) and Compustat. Because the 10-K universe includes filings for trusts, ETFs, and other privately held firms, we refine the broader dataset by including only those companies that can be matched to the CRSP dataset. This assures that the dataset is focused on firms with publicly traded equity. We refer to the broader sample of 10-Ks as the CRSP sample ($N = 95,745$). The sample of firms in the ACSI survey is relatively small. We will refer to the subset of the CRSP data with complete ACSI scores and financial data used for control variables as the ACSI sample ($N = 1,998$). We have excluded from the 10-K sample four firms (Sento Corp., Sandy Corp., NFO Worldwide, and Pivotal Corp.) whose primary business is consulting others on customer satisfaction, which would obviously create a false measure for our target phrase counts.

Parsing the 10-Ks

We use the EDGAR master index to identify and retrieve all 10-K filings from 1995–2013. All non-textual characters, HTML, and tables are removed. The parsing process is described in detail in the Appendix and is based on Loughran and McDonald (2011). We then use the programmatic parsing technique 'regular expressions' to identify and count target phrases in the document.

Terminology and Target Phrases

Because we deal with groups of words and phrases, we use the following notational convention. For a targeted phrase such as 'customer satisfaction' we are actually interested in the stated phrase plus all substantially similar variants, for example, plural or possessive forms, or cases where the word 'consumer' appears instead of 'customer.' In all cases we use an italicized word or phrase to represent the full set of analogous words or phrases. Thus our italicized terms *Customer Satisfaction*, *Negative*, and *Monitoring* are used as shorthand for the targeted terms representing a set of closely related words. The full domain for each targeted search term follows.

Customer Satisfaction—This variable is derived from the word list of Michalisin and White (2000):

(1) 'customer' or 'consumer' followed by any of the following tokens—'driven,' 'focus,' 'involvement,' 'relation,' 'satisfaction,' 'service,' 'support,' or 'value.' Note that basic extensions of these words will also be captured in this count, e.g., 'relation' also captures 'relations'.

(2) 'Help', 'respon*', 'satisfy', or 'serving' followed by either 'customer' or 'consumer'. Note that the asterisk within 'respon*' represents variants such as 'responds' or 'response'.

(3) For firms whose SIC code indicates they are in the healthcare industry, we extend the prior two searches to include the token 'patient,' along with 'customer' or 'consumer.' (Note that while the 10-K sample includes firms from the healthcare industry, the ACSI sample does not.)

Negative—To refine the context of our *Customer Satisfaction* variable we tabulate whether any negative words occur within plus or minus 300 characters from any of the identified phrases from items (1)–(3) above.[4] Specifically, the negative count is incremented if any of the following

---

[3] Specifically we consider 10-K, 10-K405, 10KSB, and 10KSB40 filings. The 10-K405 variant is based on a check box on the first page of the filing indicating that the current filing does not include a "disclosure of delinquent filers pursuant to Item 405." This option was eliminated in 2003 due to confusion and inconsistencies in selecting the 405 option. Until 2003, small businesses filed the 10KSB and 10KSB40 forms. In this paper, we label all of these basic variants simply as '10-K.'.

[4] The mean number of words per sentence in a 10-K is about 25 with the 95th percentile occurring at approximately 40 words. The average word length is about 5.4 characters. Thus, using a buffer of plus and minus 300 characters around the target phrase should easily capture something slightly more than a typical sentence.

terms appear within the targeted range: 'complain', 'declin*', 'decreas*', 'issue', 'low', 'reduc*', 'problem', and 'drop'. Note that the terms are designed as word stems that allow for various simple inflections of the targeted word (e. g., the token 'declin*' captures 'decline', 'declining', and 'declined'). It is likely that firms discussing customer satisfaction in a negative context are discussing past service failures, and that this is not a reflection of management policy that relates to subsequent satisfaction measures.

When we initially report descriptive results for the CRSP and ACSI samples, we will report counts for the variables *Customer Satisfaction* and *Negative*. For purposes of the correlations and regressions we will normalize the counts based on document length and take the natural logarithm of one plus the normalized value to ameliorate the skewness typical of counts. In addition, because we want to be able to separate out those cases where *Customer Satisfaction* is referenced in a negative context, the *Customer Satisfaction* used in the transformation is reduced by the *Negative* count.

Monitoring—Based on the notion of credible signals posited by our theory, we separately tabulate all occurrences where a word suggesting actual measurement or monitoring of customer satisfaction occurs within plus or minus 300 characters of the *Customer Satisfaction* target. The *Monitoring* targets are 'assess,', 'demonstrat*', 'measur*', 'monitor', 'questionnaire', 'rating', 'score', and 'survey'. Again, all of the targets are word stems that allow for extended variants. For purposes of the regressions we use the variables *Monitoring-Intercept*, which is a dummy variable set equal to one if any *Customer Satisfaction* target occurs along with at least one of the measurement tokens, and *Monitoring-Slope*, which is set equal to the *Customer Satisfaction* variable in cases where *Monitoring-Intercept* equals one.

Control Variables

In addition to the target phrase counts, we incorporate control variables in our regression tests. We include three variables common in the finance literature and an attribute of the 10-K document:

- *Size*—the market capitalization, i.e., stock price times the number of shares outstanding, on the day before the filing date (in billions of dollars).
- *Market-to-book*—the market equity of the firm divided by the book value of the firm (as defined in Fama and French 2001) winsorized at the 0.01 level.
- *Beta*—the slope coefficient from the regression of the CRSP value-weighted stock index on the returns of the firm's stock for the 252 trading days (approximately one year) prior to the 10-K filing date. At least 60 observations of daily returns must be available to be included in the sample.

- *Document Length*—the number of characters in the document after tables and extraneous characters (see parsing discussion) have been removed. In the correlations and regressions, the log of document length is used.

The American Customer Satisfaction Index Sample

The Fornell et al. (1996) research involves a three-item, 10-point reflective measure of consumer satisfaction with a given firm's products, administered to several hundred questionnaire respondents (initially 250 per firm in 1994–1995) via national probability sample. Target firms (the 200 referenced previously) come from over 40 industries in seven comprehensive economic sectors: manufacturing/non-durables, manufacturing/durables, transportation-communications-utilities, retail, finance and insurance, other services, and public administration or government. Scores are expressed within the confines of a 0–100 scale and validation evidence is presented in the original reference (pp. 8–9). Although the ACSI orientation has been the aggregation across industry, sector, and national economy, the firm-level metrics are of interest as raw material here. For firms with multiple measured products we use a simple average of the primary product groups (e.g. Ford and Lincoln). Single products that are relatively minor volume contributors for their firms are excluded (e.g. Iams pet food for P&G).

After examining the temporal and cross-sectional patterns of usage for *Customer Satisfaction* in the full 10-K sample, we then match the 10-K sample to the firms included in the ACSI data downloaded from http://www.theacsi.org, which began in 1996. We are able to match a total of 1,998 firm/year 10-K filings with ACSI scores and refer to this as the ACSI sample.

Note that in our regression analysis to follow, we match the 10-K phrases for a given year with the firm's ACSI score in the subsequent year. This separation assures that the observed 10-K document and financial control variables are determined before the measured outcome. Thus our ACSI variable is denoted as $ACSI_{t+1}$. This one year lag is also why we begin the 10-K sample in 1995 to align with the first year of ACSI data in 1996.

Results and Discussion

Descriptive Statistics

We focus on the two samples previously outlined. First, we descriptively examine the occurrence of the category of phrases captured by the *Customer Satisfaction* variable

Fig. 1 Average ACSI$_{t+1}$ score and percentage of firms' 10-Ks containing at least one occurrence of *Customer Satisfaction* (including variants) by year
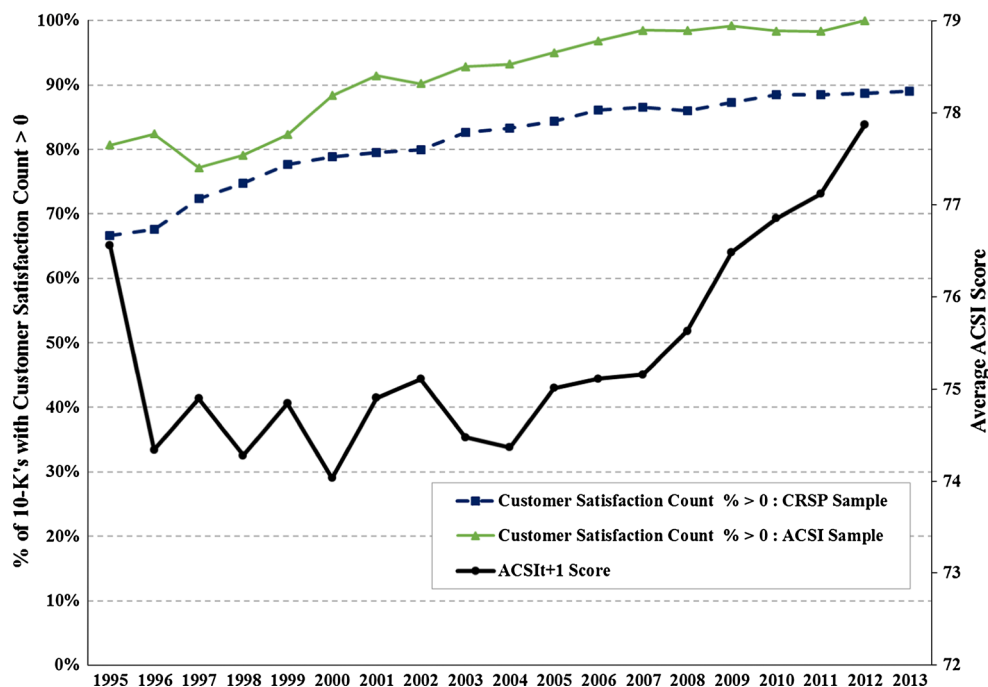
over the years 1995–2013 in the CRSP sample ($N = 95,745$). Although not the primary focus of the paper, this assessment provides an interesting documentation of a critical marketing-managerial term in financial filings over the past 19 years. We then consider the ACSI sample of firm/year observations containing both the 10-K data and the ACSI scores ($N = 1,998$). Because we link 10-K filings to subsequent ACSI scores, the ACSI sample concludes in 2012. The CRSP sample contains approximately four to seven thousand firms each year, although this number is lower for 1995–1996 because at that time the electronic filing requirements were just being implemented by the SEC.

Figure 1 plots the average customer satisfaction score for the ACSI sample subset and, for both of the samples, the proportion of 10-Ks with *Customer Satisfaction* occurring at least once in the document. Thus, for example, of the 6,360 10-K filings in the year 2000, approximately 79 % contained at least one instance of the terms captured by the *Customer Satisfaction* variable, whereas for the 103 ACSI firms in that year, about 88 % used the term.

The average ACSI score drops substantially from the first year (1996) to the second year in the series, with a slight upward movement in the last eight years. The large decrease across these first two years is not attributable to a change in the included firms; the number reflects a shift in scores within a relatively homogeneous sample. Other than the initial drop and a more encouraging positive trend over the past decade, the time series of ACSI scores does not reflect any clear temporal patterns, with the average score ranging between about 74 and 78 on the 100-point scale.

For the ACSI sample, the presence of any reference to *Customer Satisfaction* appears to trend upward over the sample period, ranging from an initial proportion of about 80 % to 100 % in the final year. The trend in the proportional occurrence for the CRSP sample is also positive and moves from an average proportion (of one or more occurrences) of about 67 % to about 80 %. Clearly, in recent years the vast majority of firms provide at least some discussion of customer satisfaction in their annual filings.[5]

Table 1 provides descriptive statistics for the targeted phrases captured in the *Customer Satisfaction* variable, disaggregated by sample. In this table, the reported *Customer Satisfaction, Negative,* and *Monitoring* variables are based on the raw counts and not transformed, as previously discussed, for the regressions. Specifically, *Customer Satisfaction* is the full count, but later when we transform the variable for the correlations and regressions we net out the *Negative* count. For the count data, we report statistics that focus on the right tail of the distribution, since the left tail is truncated at zero.

Over the entire CRSP sample *Customer Satisfaction* appears at least once in 81.2 % of the filings, while for the ACSI sample approximately 92 % of the 10-K documents include the term. This is expected because of the ACSI sample composition, which consists of large firms with national brands. For the *Negative* measure, where firms use negative terms near their discussion of customer

---

[5] The t-statistic for a simple trend regression of Customer Satisfaction in the CRSP sample is 4.00. A similar regression for the ACSI sample also produces a significant trend with a t statistic of 5.33.

Table 1 Descriptive statistics

Panel A	>0 (%)	Mean	Median	90th%	Max	N
Customer satisfaction count						
CRSP sample	81.24	6.63	4	17	234	95,745
ACSI sample	92.14	11.19	7	26	93	1,998
Negative count						
CRSP sample	61.90	2.50	1	7	65	95,745
ACSI sample	78.68	4.59	3	12	40	1,998
Monitoring count						
CRSP sample	45.72	1.28	0	4	53	95,745
ACSI sample	64.86	2.47	1	7	41	1,998

Panel B	Mean	Median	SD	N
ACSI sample				
$ACSI_{t+1}$	75.86	76	6.23	1,998
Control variables				
Size ($ billions)	29.34	11.51	50.96	1,998
Market-to-book	4.22	2.08	9.92	1,998
Beta	0.90	0.83	0.49	1,998
Document Length	463,312	342,030	402,448	1,998

something about attention to the concept, their impact is not likely a simple linear one. That is, Penske is not likely 234 times more concerned with customer satisfaction than firms with only one mention of the term. As a result of the distributional characteristics for these counts, we use the natural logarithm of one plus the count divided by document length for Customer Satisfaction and other count variables when examined in the correlations and regressions. This type of transformation, used to attenuate the positive skewness in phrase counts and account for document length, is very common in the natural processing literature (see, for example, the discussion of term weighting in Manning and Schütze 2003). We also use log transformations for the Size variable included in the regressions, which is consistent with its application in the finance literature, and for Document Length, due to its skewness.

Panel B of Table 1 reports descriptive statistics for the ACSI variable and the control variables in the ASCI sample of 1,998 elements. The mean value for the ACSI customer satisfaction variable is 75.86, with reported scores ranging from 49 for PG&E Corporation (an electric utility) in the year 2000 to 91 for H. J. Heinz Company (food products) in the year 2004. Not surprisingly the control variable measures reflect the relatively large size of firms included in the ACSI sample, with a mean market capitalization (price × number of shares outstanding) of $11.51 billion.

In Table 2, we report the correlations among the ACSI scores, the transformed count variables, and the control variables used in the regressions. Our theoretical framework suggests that the relation between ACSI scores and our customer satisfaction counts is nuanced and depends on context. Thus, the simple correlations between ACSI scores and the counts must be conditioned on this limitation. At this level, the simple correlation between customer satisfaction counts and the ACSI scores is negative, and, given that the transformed Customer Satisfaction variable is net of negative counts, suggests that there is some evidence of 'cheap talk' in the collective sample.

Presumably when customer satisfaction is being discussed in a negative context, the firm is more likely discussing existing issues versus attempting to signal, thus the negative correlation between the Negative variable and the ACSI variable is not surprising. As explained in the theoretical discussion, the sign for Monitoring depends on the sample context, but we can see from Table 2 that the simple correlation across the entire sample is negative. In terms of the control variables, larger firms' and growth firms' measured values (i.e. high market-to-book) are positively correlated with the ACSI score, while risk, as measured by beta, and document length have relatively large and negative correlations. Loughran and McDonald (2014) argue that 10-K document length is a proxy for

satisfaction, more than three-fourths of the sample contains at least one occurrence. One of the firms with the highest proportion of Customer Satisfaction and the highest proportion of Negative is Verizon Communications. (From the year 2006 forward, Verizon appeared in the top twenty for combined Customer Satisfaction and Negative score more frequently than any other firm.)

Approximately 65 % of the ACSI firms also include at least one case where Customer Satisfaction is near one of our Monitoring terms (e.g. Pepsico, in one of its 10-Ks, states "We have implemented customer satisfaction measures to evaluate the success of these initiatives.") Thus, of those firms in the ACSI sample that mention Customer Satisfaction at least once, more than half indicate some form of monitoring or measurement of the construct. Firms with high proportions of monitoring terms include Amazon, Kohl's, and JetBlue Airways.

Notice that the tabulated counts are heavily skewed. More than 90 % of the counts for the 10-K sample are twenty or less. Given the nature of the distribution, we separately examine all 10-Ks having a phrase count of greater than ten. The case with a count of 234 for Customer Satisfaction is Penske Automotive's 2002 10-K filing, for which customer satisfaction is a key measure in the services they provide to automotive dealers throughout the United States. In Penske's 2002 10-K, management chose to include all license agreements with each of their major automobile franchisees, thus repeatedly restating customer satisfaction objectives. Clearly, although large counts say

Table 2 Correlations ($N = 1{,}998$)

	$ACSI_{t+1}$	Cust Sat	Negative	Monitoring	Size	Market-to-book	Beta
Customer Satisfaction	−0.035						
Negative	−0.057	0.661					
Monitoring	−0.056	0.572	0.582				
Size	0.091	0.036	−0.027	−0.031			
Market-to-book	0.075	−0.012	−0.009	−0.019	0.057		
Beta	−0.189	0.233	−0.200	0.186	−0.109	0.071	
Document Length	−0.197	−0.313	−0.313	−0.041	0.035	−0.130	−0.034
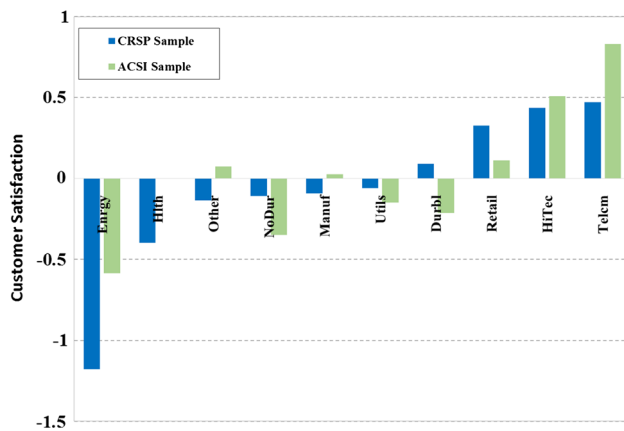
Fig. 2 *Customer Satisfaction* (transformed and standardized) by industry and sample

readability, thus this result would support the notion that firms trying to obscure their financial data also do not score well in terms of customer satisfaction.

In sum, the measured levels of *Customer Satisfaction* in both the full 10-K sample and the ACSI sample show a positive trend over almost two decades of data. The ACSI firms tend to have higher levels of *Customer Satisfaction* counts than the broader CRSP sample. Of the 90 % of firms in the ACSI sample with a *Customer Satisfaction* count of at least one, about two-thirds of those also discuss some type of measurement or monitoring of the construct.

Industry Effects

Figure 2 presents the standardized value of *Customer Satisfaction* within ten industry groupings for both the CRSP and ACSI samples.[6] That Energy is relatively low in its focus on the phrase and Retail high is consistent with respective industry reputations. The popularity of the

---

[6] We use the 10-industry classification scheme provided by Ken French at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

locution for the Telecommunications and Technology groups foreshadows our later regression results, i.e. use of the phrase does not necessarily mean a firm is successful at achieving higher customer satisfaction scores. If we rank based on the Fama and French 49 industries, the lowest measure of the *Customer Satisfaction* variable is in the tobacco industry, while the highest is in the entertainment industry.

Regression Results

Consistent with our general hypothesis, we examine whether the presence of the *Customer Satisfaction* measure in a given year's 10-K report is related to the level of realized customer satisfaction measured in the subsequent year. The first-order impact of mentioning customer satisfaction is ambiguous, likely a determinant of the importance of customer relations in a specific industry and the history of the industry in delivering on the implicit or explicit promise of customer satisfaction. We argue that although the unqualified signal is ambiguous, credible signals of commitment to customer satisfaction should be positively related to the measured outcomes in industries where customer satisfaction is more important.

That industry matters is apparent from the differences shown in Fig. 2 (e.g. the *t* statistic for the mean differences between Energy and Retail is 61.31). To identify industries where customer behavior is most likely to resonate with the crux of our theory, we consider a simple division. Specifically, for retail firms, satisfaction of customers is paramount, but these businesses are also likely to have less reason to overpromise due to the large cost of deception relative to the gains from deception: They rely heavily on repeat business which would be negatively affected by deception, and they have relatively small profit margins, thus reducing the benefit from overpromising. Also, apart from self-evident closeness to the customer, their market structure is less concentrated than most other industries (see, for example, Pryor 2001). Pragmatically, a simple division of retail and non-retail also avoids arbitrary specificity and produces reasonable sample sizes in both

groups. Additionally, Fornell et al. find that retail is distinguished from other commercial sectors on some key satisfaction relations in the ACSI data (1996, p. 14).

In the first column of Table 3 we report results for the simple regression of the $ACSI_{t+1}$ on the control variables. All of the regressions include a constant and time dummies, with robust standard errors clustered by year. Consistent with the simple correlations, both *Size* and *Market-to-book* have a significantly positive impact on subsequent ACSI scores. The ACSI sample is already limited to relatively large firms, and apparently even within this group, larger firms are more successful at generating higher customer satisfaction as reported by consumers. The *Market-to-book* coefficient suggests that growth companies tend to perform better in the realm of customer satisfaction. More volatile firms, as measured by *Beta*, are less likely to score well. Finally, firms producing more lengthy filings are more likely to have lower customer satisfaction scores. All of the control variables are statistically significant at any traditional level.

In column (2) we first include the *Customer Satisfaction* and *Negative* variables. References to customer satisfaction not appearing in a negative context have a negative impact on ACSI scores, with the estimated coefficient being significant at the .10 level. To the extent those discussions of customer satisfaction which do not take place in a negative context are more likely to be signaling, then the results suggest that, across the entire sample, the discussion is seen as 'cheap talk.' A one standard deviation increase in the count of *Customer Satisfaction* for an average size document is associated with approximately a 2.5 point decrease in the subsequent ACSI score, given an average ACSI score across all firm/years of 75.9. The coefficient on *Negative* indicates that, as expected, firms simply discussing existing problems in the area of customer satisfaction are more likely to have lower ACSI scores.

From our theoretical model we are interested in separating out discussions of customer satisfaction to identify those that take place in the context of monitoring, i.e. signals that are more likely to have credibility. As previously posited, we are more likely to see a separating equilibrium produced by credible monitoring signals in industries that correspond more closely to the assumptions underlying the model. Thus, in columns (3) and (4) we add variables for the intercept and slope of *Monitoring*, as previously defined, and separate the sample into retail and non-retail firms. For retail firms, the coefficient on *Customer Satisfaction* continues to be negative and is now significant at the 0.05 level. In the retail sample, the impact of negative occurrences is negative as expected but not significant. However, consistent with the notion of credible signals, both the intercept and slope coefficients on the *Monitoring* variable are significant at the .05 level. The

intercept coefficient indicates that firms with any discussion of actually measuring customer satisfaction score are significantly higher on subsequent ACSI surveys, with the slope suggesting that this effect increases with the frequency of customer satisfaction comments.

For the retail sample, the coefficients on two of the control variables, *Size* and *Beta*, are reversed from the column (2) results. Many non-retail firms tend to have lower betas, and perhaps the differential effect reflects these mean differences. (The average beta for non-retail firms is 0.87; for retail it is 1.02.)

For the non-retail firms, the coefficient on *Customer Satisfaction* remains negative but is no longer significant, while the coefficient on *Negative* is significant and mirrors the result in column (2). For this set of industries, neither of the monitoring coefficients is significant.

Our more refined examination of the customer satisfaction signal suggests that simply discussing the concept is not effective and might reflect past issues in the area, but efforts featuring actual attempts to objectively measure progress on this front, beyond mere lip service, do produce higher subsequent ACSI scores for retail firms. The more costly signal of creating objective, measurement-based commitments to customer satisfaction also creates a credible indicator of higher satisfaction scores. It is a matter of management transcending its own rhetoric and imposing a real standard or test.

Conclusion

Given the importance of customer satisfaction as a normative principle in marketing/management theory and practice, the simple time series aspect of our aggregate results for annual reports is encouraging. That ethical firms should focus on a customer's well-being is a central premise of the AMA's Statement of Ethics. Across a broad sample of firms, the use of *Customer Satisfaction* in management's mandated reporting to shareholders has increased since 1995. Across the CRSP sample of essentially all firms with equity trading on a U.S. exchange, more than 80% of the firms discuss the topic. For firms having customer satisfaction actively monitored in the ACSI research product, almost all address the subject.

Our primary results, concerning the relation between the *Customer Satisfaction* locution in 10-K filings and subsequent survey-based measures of customer satisfaction, are more subtle. We find that "talk is cheap." Ignoring the context provided by our *Monitoring* variables, the simple relation between *Customer Satisfaction* and subsequent ACSI scores is negative. This result is not necessarily surprising. Loughran et al. (2009) observe a similar phenomenon when examining use of the term *ethics* in 10-K filings prior to the requirement

Table 3 Estimated coefficients (t statistics) for the $ACSI_{t+1}$ regressions

Variables	All firms		Retail	Non-retail
	(1)	(2)	(3)	(4)
Customer Satisfaction		−0.416	−0.941	−0.412
		(−1.91)	(−2.21)	(−1.49)
Negative		−0.751	−0.253	−0.741
		(−3.93)	(−0.62)	(−2.97)
Monitoring-Intercept			17.167	−6.254
			(2.86)	(−1.46)
Monitoring-Slope			1.584	−0.495
			(2.82)	(−1.23)
Size	0.234	0.244	−0.482	0.563
	(3.79)	(3.75)	(−2.90)	(6.60)
Market-to-book	0.061	0.055	0.092	0.032
	(4.36)	(3.96)	(2.37)	(2.51)
Beta	−3.081	−2.666	2.388	−3.246
	(−7.27)	(−6.00)	(3.77)	(−8.21)
Document Length	−2.392	−2.935	−1.712	−3.416
	(−15.97)	(−20.53)	(−4.56)	(−24.38)
Intercept	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Observations	1,998	1,998	377	1,621
R-squared (%)	16.25	18.01	36.50	24.27

The dependent variable is the subsequent year's ACSI survey score. Variables are defined in the text. The t statistics are based on robust standard errors clustered by year

of ethics codes by the Sarbanes–Oxley Act of 2002. They find that a firm most likely to use that particular term in its 10-K filings is also one most likely to be classified as a 'sin stock,' is more likely to have subsequent shareholder lawsuits filed against it, and is more likely to have dictatorial corporate governance. Essentially they conclude that firms discussing ethics are typically ones having underlying issues in this area. Similarly, it may be that many firms discussing customer satisfaction are ones for which the topic has become an unavoidable headline issue—for example, problems with customer relations at cable companies over the past decade. However, simply declaring that customer satisfaction is important is a costless signal. Not surprisingly, we also find that firms using language relating to customer satisfaction in a negative context tend to have lower subsequent ACSI scores.

More notably, we document that retail firms using *Customer Satisfaction* in the context of measuring or monitoring the construct have higher subsequent ACSI scores. Thus, retail firms that go beyond expressing a popular ideal and actually declare attempts to measure their progress are more likely to receive higher ACSI scores. For example, McDonald's Corp., a firm whose average ACSI score is less than 63 over the sample period (while the

sample has a cross-sectional average score of 75), seems to make broad sweeping statements in support of customer satisfaction ("…essential to driving the Company's strategies of delivering customer satisfaction and increasing market share"), but never indicates specific mechanisms for measuring the construct. At the other end of the spectrum, Barnes and Noble, with an average ACSI score of more than 84, discusses at length independent measures of customer satisfaction for their retail stores and online operations. For discussions of customer satisfaction to produce positive outcomes, they must be credible. This result underscores the statement by Erdem and Swait (1998, p. 137) that "…credibility of a brand signal is perhaps its most important characteristic."

Collectively our results show that management's rhetoric on customer satisfaction is indeed associated with measured levels of customer satisfaction. The theoretical link requires that a firm is sufficiently concerned with reputation relative to current profitability, as in the retail industry which relies on repeat sales and has low profit margins. However, even then talk *alone* is cheap. For retail firms, corporate 'talk' (10-K usage of customer satisfaction-related terms) only maps positively into 'walk' (ACSI scores) when firms not merely pay lip service to what stakeholders would like to hear, but when they also back the rhetoric by committing to monitoring the outcomes. In contrast, in non-retail industries—insofar as they are marked by higher profit margins and less emphasis on repeat business—credible signaling is apparently more challenging, and there is no significant link between measured levels of customer satisfaction and rhetoric, irrespective of the monitoring context.

The contrast in our results for non-retail firms and retail firms is consistent with the traditional shareholder theory of business ethics of Friedman (1962) and Jensen (2002). While management's rhetoric may be reliable when it is in the firm's self-interest to maintain a reputation (retail firms), there is no indication that the non-retail firm is concerned with the customer segment of its stakeholders beyond the extent to which its satisfaction affects the bottom line for the shareholders.

Appendix

Parsing the 10-K Text Files

All 10-K SEC complete text document filings (as specified in footnote 4) are downloaded for each year/quarter. The

text version of the filings provided on the SEC server is an aggregation of all information provided in the browser-friendly files also listed on Edgar for a specific filing. For example, IBM's 10-K filing on 20120228 lists the core 10-K document in html format, ten exhibits, four jpg (graphics) files, and six XBRL files.[7] All of these files are also contained in a single text file with the embedded HTML, XBRL, exhibits, and the ASCII-encoded graphic.[8] In the IBM example, of the 48,253,491 characters contained in the file, only about 7.6 % account for the 10-K text including the exhibits and tables. The HTML coding accounts for about 55 % of the file. The XBRL tables have a very high ratio of tags to data and account for about 33 % of the text file. The remaining 27 % of the file is attributable to the ASCII-encoded graphics. In many cases, ASCII-encoded pdfs, graphics, xls, or other binary files that have been encoded can account for more than 90 % of the document.

In parsing the 10-K documents we exclude markup tags, ASCII-encoded graphics, and tables. We exclude tables, because they are usually not the focus of textual analysis.

Parsing Details

Each raw text file downloaded from EDGAR is parsed using the following sequence:

1. Remove ASCII-Encoded segments—All document segment <TYPE> tags of GRAPHIC, ZIP, EXCEL, and PDF are deleted from the file. ASCII-encoding is a means of converting binary-type files into standard ASCII characters to facilitate transfer across various hardware platforms. A relative small graphic can create a substantial ASCII segment. Filings containing multiple graphics can be orders of magnitude larger than those containing only textual information.
2. Remove <DIV>,<TR>,<TD>, and <FONT> tags—Although we require some HTML information for subsequent parsing, the files are so large (and processed as a single string) that we initially simply strip out some of the formatting HTML.
3. Remove all XBRL—all characters between <XBRL …> … </XBRL> are deleted.
4. Remove SEC Header/Footer—All characters from the beginning of the original file thru </SEC-HEADER> (or </IMS-HEADER> in some older documents) are deleted from the file. In addition the footer "———— END PRIVACY-ENHANCED MESSAGE————" appearing at the end of each document is deleted.
5. Remove tables—all characters appearing between <TABLE> and </TABLE> tags are removed.

    Note that some filers use table tags to demark paragraphs of text, so each potential table string is first stripped of all HTML and then the number of numeric versus alphabetic characters is compared. For this parsing, only table encapsulated strings where *numeric chars/(alphabetic + numeric chars)* >15 % are excluded.
    In some instances, Item 7 and/or Item 8 of the filings begins with a table of data where the Item 7 or 8 demarcation appears as a line within the table string. Thus, any table string containing "Item 7" or "Item 8" (case insensitive) is *not* deleted.

6. Remove all carriage returns—EDGAR files use a linefeed (/n) as the "newline" character. Because of the way we process very large files the actual string we parse can contain the Windows end-of-line specification (i.e., carriage return followed by a linefeed – \r\n). In order to simplify subsequent parsing all carriage returns are deleted from the document, i.e., all documents will have only a linefeed as the "newline" character.
7. Remove Markup Tags—remove all remaining markup tags (i.e. <…>).
8. Re-encode reserved HTML characters (character entity references)—In order to encode a broad set of universal characters within the limitations of ASCII coding many characters are encoded. For example, the "&" symbol can be encoded as "&" or "&". For items listed below we replace the encode items with a character(s). The remaining encoded items are deleted.

    a. "<" or "&#60" -> "LT" - note we use LT instead of "<" to avoid any confusion with markup tags.
    b. ">" or "&#62" -> " GT "
    c. "&NBSP;" or " " -> " "
    d. """ or "&#34" -> """"
    e. "&APOS;" or "&#39" -> """"
    f. "&" or &#38" -> "&"

---

g. All Regular Expression \t and \v items are deleted.

h. All remaining ISO 8859-1 symbols and characters are deleted.

A Theoretical Model of Signaling and Delivering Customer Satisfaction

We present a stylized model focusing on how firms decide to signal their commitment to customer satisfaction (CS) and how these signals might be interpreted. For virtually any product the quality dimension is both essential and not always observable in advance. When a firm markets its products, conveying the quality is critical. Naturally, the firm can receive a higher price for a product that is perceived to be of higher quality, and for that reason may aggressively promote product performance and reliability, service commitment, special features, and other product characteristics that foretell a high level of satisfaction. However, there are real costs to the firm of promising more than it expects to deliver. If satisfaction falls short of promises the firm has essentially two choices. It may accept poor word-of-mouth reporting and a decrease in future sales or it may decide to provide additional service in the form of refunds, repairs, service calls, etc. In either case, it will have to deal in various costly ways with dissatisfied customers.

We present first the perspective of a representative consumer who derives utility from a basket of goods/services as well as the quantity and quality aspects of a featured product. The formulation is a special case of the Lancaster (1966) approach in which satisfaction is an aggregate result of different attributes of the products consumed. The utility function of the representative consumer is

$$U = g(q)s + Q. \tag{1}$$

Here Q represents an aggregate utility index of the quantity of all goods and their characteristics, excluding only the featured product, its quantity represented by q. The utility from this product is an increasing, concave function $g(q)$ of the quantity q, times the level of CS derived per unit of the item, s.[9] The consumer makes decisions subject to a budget constraint

$$Y = pq + Q, \tag{2}$$

with Y representing total income and p the price of the featured product. The aggregate index Q is the numeraire and thus has a price of 1, or equivalently, both Y and p are expressed in real terms, i.e. units of the aggregate index. Maximizing expected utility subject to the budget constraint, using the constraint to eliminate Q, gives

$$\frac{\text{Max}}{q}[g(q)E(s) + Y - pq], \tag{3}$$

where $E(s)$ reflects the consumer's expectation of the satisfaction with the featured product which is not known a priori. The first-order condition for q from differentiation of Eq. (3) then provides the (inverse) demand function for the featured good

$$p = g'(q)E(s). \tag{4}$$

The concavity of $g(q)$ guarantees that $g'(q)$ decreases in q. Thus, the price the firm can charge depends negatively on the quantity it projects to sell and positively on the level of CS the representative consumer expects to receive.

The firm producing the featured product will maximize profits subject to the demand function from Eq. (4). Profits are given by

$$\prod = q[p - c(s, \ s^*)]. \tag{5}$$

Define s as the level of CS the firm actually provides and s^* as the level of CS the firm promises to provide. The unit cost of production, c, depends positively on the level of CS the firm actually delivers, s, reflecting the cost of providing quality and service, and positively on the level of CS promised by the firm relative to what is delivered, $s^* - s$, reflecting the implicit guarantee provided, i.e., the firm's cost of dealing with disappointed consumers when CS falls short of expectations created by the firm. These costs include providing refunds, additional service, follow-up calls, and employee dissatisfaction, as well as the less tangible loss of reputation. For analytical convenience, we posit a specific functional form of the unit cost function[10]:

---

[9] The utility of the product may also depend on the level of satisfaction promised by the firm, s^*. Promised satisfaction may affect the consumer in two ways: First, it presents an implicit guarantee that suggests some compensation if satisfaction falls short of the promise. Second, it provides a prior indication of the value of the "match"—the extent to which satisfaction follows from the product fitting in properly with the consumer's overall consumption plan, given that the plan is based in part on promised/promoted quality. Because this feature does not affect the key implications of our theoretical model, we avoid formally introducing this complication. Note that the

Footnote 9 continued
traditional satisfaction paradigm is based on the relation of actual to expected performance (e.g. Anderson and Sullivan 1990), but this research extends theory an additional step to the actual-expected *satisfaction* relation (per Eq. 9).

[10] Note that we assume that expected sales are a positive linear function of expected CS. The "½" preceding the second term is inserted to simplify the subsequent derivation and has no impact on the subsequent results. The quadratic term in Eq. (6) is the simplest function that provides a convex cost of promised CS not being met. The quadratic functional form also makes it costly for delivered CS to *exceed* promised CS but this is merely for mathematical convenience and is not relevant in equilibrium. One may set the cost equal to zero for all cases where delivered exceeds promised, without affecting the model solution.

$$c(s, s^*) = \delta s + (1/2)\gamma(s^* - s)^2, \tag{6}$$

where δ is a positive parameter reflecting the expense needed to increase CS, and γ is a positive parameter representing the level of cost incurred by overpromising.

To focus on the issues of promising CS, we assume for convenience that both the firm's production level and the firm's provision of CS are determined exogenously, with the latter also stochastic from an outsider's perspective:

$$q = \bar{q}, \quad s \sim N(\bar{s}, \sigma_s^2). \tag{7}$$

The level of CS provided depends on the firm's commitment to service. It is stochastic with a mean level of $\bar{s}$ known by all, including the consumers, and a realization s only known in advance by the firm.

The firm thus chooses its promised level of CS to maximize profits

$$\underset{s^*}{\text{Max}} \, \bar{q} \Big[g'(\bar{q}) E(s) - \delta s - (1/2)\gamma(s^* - s)^2 \Big]. \tag{8}$$

The level of CS expected by the consumer could be written more precisely as $E(s) \equiv E(s|s^*)$. The expected level depends on the consumer's rational inference from relevant information produced by the firm. For now we postulate that the expected level of CS depends linearly on s^*, which we verify later:

$$E(s) = \alpha + \beta s^* \tag{9}$$

Accordingly, the first-order condition for s^* based on the expression in (8) and subject to (9) becomes

$$s^* = s + [\beta g'(\bar{q})/\gamma]. \tag{10}$$

Each of the terms in Eq. (10) is known to the consumer once the firm promises a level of CS. In the separating signaling equilibrium the level of CS the firm decides to provide is revealed fully to the consumer so that $E(s) \equiv E(s|s^*) = s$. Hence, we can use Eq. (10) to verify the linear relation postulated in Eq. (9). Equations (9) and (10) are identical, confirming the postulated linear dependence, if and only if $\alpha = -\beta g'(\bar{q})/\gamma$ and $\beta = 1$. Thus,

$$s^* = s + b, \quad b = g'(\bar{q})/\gamma. \tag{11}$$

The promised level of support equals the CS the firm knows it can provide, *plus* a bias b that is related to the firm's current profitability from selling its output $g'(\bar{q})$ relative to its cost level arising from unmet expectations γ, representing the incentive to exaggerate the product's desirability, especially for firms that experience higher profit margins. (In customary marketing language, this ratio captures the way that reluctance to dissatisfy by over-inflating customer expectations restrains any tendency to overpromise or deceive.) Note that the consumer market understands that the firm's promises are biased and corrects this bias in assessing future support.

It is useful to point out that the separating equilibrium may not exist in cases in which b (current profitability from selling its output relative to the cost of unmet expectations) is large; a large bias may impose too large a signaling cost. The profit level in signaling equilibrium becomes negative due to the term $-1/2\gamma(b)^2$ (given Eqs. 8 and 11) so that the signaling equilibrium cannot be realized.

The fact that the firm has superior information about the level of CS it intends to provide implies that it has a need to send a credible signal about this intent in order to generate the appropriate level of demand and to be able, from Eq. (4), to charge a price that is proportional to the level of CS that consumers rationally expect. To send this signal the firm incurs a real cost. The efficient promised level of CS is equal to the level of CS the firm knows it can deliver and that the firm would choose to provide if it could otherwise credibly convey its level of CS. However, consumers know the firm's incentive to overpromise, according to Eq. (10), in order to enhance demand. If the firm promises exactly what it can deliver, consumers will expect a lower CS than that and the product price will be lower. By overpromising the firm incurs a real cost from its implicit commitment to service—disappointed consumers.

Testable Hypotheses

Recall Eq. (11) for the promised level of CS, and obtain the actual level of CS directly as

$$s = s^* - b, \quad b = g'(\bar{q})/\gamma. \tag{11}$$

The formal theoretical hypotheses are derived directly from Eq. (11′):

$$\left.\frac{\mathrm{d}s}{\mathrm{d}s^*}\right|_b > 0, \quad \left.\frac{\mathrm{d}s}{\mathrm{d}b}\right|_{s^*} > 0. \tag{12}$$

Conditional on a given level of bias, a credible promise of higher CS is associated with accordingly higher observed CS; conditional on a given promised level of CS, a higher incentive to bias is associated with lower observed CS. The conditioning is important. In particular, an increase in promised CS by itself need not imply higher delivered CS if the cause is a higher incentive to bias: $\mathrm{d}s/\mathrm{d}s^* = 0$ if higher b causes the higher promise.

From the statistician's perspective we observe a survey-based measure of actual $CS \equiv s$ (the ACSI score referenced earlier). Given the firm-years in our sample, we account for the variation across firms without explicitly utilizing information in the time series of individual firms. Hence, statistically we explain each firm's level of average CS conditional on the latest observation of the firm's promised CS.

In moving from theory to empirical investigation, the challenge is to find proper empirical proxies for the level of

promised CS and the bias: $s*$ and b. Higher *promised* CS, $s*$, increases future CS for given profit incentives (that is, for given bias b). For the theoretical result to apply, the proxy for $s*$ must reflect a credible commitment by the firm to provide more CS. A firm's frequent mention of CS in conjunction with a stated commitment to ensure its realization provides a credible promise of CS, backed by higher inputs that will truly generate higher CS. Such a promise is credible because, for given profit motives and higher support capital, the positive signal provides higher sales without leading subsequently to customer disappointment and associated costs. Firms with higher support capital and lower immediate profit incentives provide credible signals, observed as discussion of CS together with clear statements on how it can be measured.

The key issue is that mention of CS must be in the context of genuine commitment, in our case indicated by measurement, so that it reflects an implicit guarantee to provide more CS: If in public statements the firm reports on a committed level of CS, it should have quantified this in some way. On the other hand, if the firm merely wants to communicate that CS is important for the bottom line or discuss ongoing issues with CS, it is not essential to quantify the statement. This leads to hypothesis 1 which states that, in the absence of a monitoring commitment, there is no clear link between promised CS and realized CS. Thus, from the model, promised CS is represented by the discussion of CS in the context of measurement. In this case, higher promised CS correlates with higher realized CS (Hypothesis 2). However, to explain actual CS in Eq. (11´), the bias—$b = g'(\bar{q})/\gamma$—matters as well. Although there are no direct measures of the bias, we propose a simple proxy. The bias is larger when the benefit of consumers' expected support in promoting sales, $g'(\bar{q})$, is larger relative to the cost of overpromising, captured by γ. Omitting an accurate measure of bias in a regression explaining actual CS by promised CS imparts a downward econometric bias on the promised CS coefficient. Since promised CS ($s*$) is higher, not just when actual CS (s) is higher but also when the bias (b) is higher, its coefficient captures an average of both effects: Actual CS is naturally higher when promises are based on actual resources for providing CS, but actual CS is not affected when the promises stem from a bias to overpromise. The econometric bias can be addressed if we partition the sample based on the incentive to overpromise. We would expect the posited effects to be greatest in industries in which the incentive to overpromise is small (hypothesis 3). These are industries where firms have little to gain from overpromising and much to lose. Thus, we will consider separate regressions for firms classified as retail, where profit margins are small but results depend crucially on repeat business, versus non-retail, in expectation that the hypothesized relation should be more apparent for retail firms.

In summary, we propose three empirical hypotheses from the theory. First, the link between the discussion of CS in a context without commitment and observed CS is ambiguous (Hypothesis 1). Likely, CS is discussed without a credible signal of measurement in cases where it has been an historical issue, so we might expect a negative relation. However, it is also possible that mention of CS without measurement simply reflects the value relevance of CS, in which case the sign could be positive. Second, mention of CS in a measurement context predicts higher subsequent realizations of CS (Hypothesis 2). Third, the link between frequency of mention of CS in a measurement context and observed CS should be stronger for industries in which the econometric bias due to overpromising is less (Hypothesis 3), which we will attempt to identify by partitioning firms into retail and non-retail.

References

Akerlof, G. (1970). The market for 'Lemons': Qualitative uncertainty and the market mechanism. *Quarterly Journal of Economics, 89*, 488–500.

Aksoy, L., Cooil, B., Groening, C., Keiningham, T. L., & Yalçin, A. (2008). The long-term stock market valuation of customer satisfaction. *Journal of Marketing, 72*, 105–122.

Anderson, R. E. (1973). Consumer dissatisfaction: The effect of disconfirmed expectancy on perceived product performance. *Journal of Marketing Research, 10*, 38–44.

Anderson, E. W., & Sullivan, M. W. (1990). The antecedents and consequences of customer satisfaction for firms. *Marketing Science, 12*(Spring), 125–143.

Barksdale, H. C., & Darden, W. R. (1972). Consumer attitudes toward marketing and consumerism. *Journal of Marketing, 36*, 28–35.

Barksdale, H. C., Darden, W. R., & Perreault Jr, W. D. (1976). Changes in consumer attitudes toward marketing, consumerism, and government regulation: 1971–1975. *Journal of Consumer Affairs, 10*(Winter), 117–139.

Basuroy, S., Desai, K. K., & Talukdar, D. (2006). An empirical investigation of signaling in the motion picture industry. *Journal of Marketing Research, 43*, 287–295.

Cardozo, R. N. (1965). An experimental study of consumer effort, expectation, and satisfaction. *Journal of Marketing Research, 2*, 244–249.

Czepiel, J. A., & Rosenberg, L. J. (1977). Consumer satisfaction: Concept and measurement. *Journal of the Academy of Marketing Science, 5*(Fall), 403–411.

Dovring, K. (1954–1955). Quantitative Semantics in 18th century Sweden. *Public Opinion Quarterly*, *18*(Winter), 389–394.

Erdem, T., & Swait, J. (1998). Brand equity as a signaling phenomenon. *Journal of Consumer Psychology, 7*, 131–157.

Fama, E., & French, K. (2001). Disappearing dividends: Changing firm characteristics or lower propensity to pay. *Journal of Financial Economics, 60*, 3–43.

Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996). The American Customer Satisfaction Index: Nature, purpose, and findings. *Journal of Marketing, 60*, 7–18.

Fornell, C., Mathas, S., Morgenson, F. V, I. I. I., & Krishnan, M. S. (2006). Customer satisfaction and stock prices: High returns, low risk. *Journal of Marketing, 70*, 1–14.

Forster, M., Loughran, T., & McDonald, B. (2009). Commonality in codes of ethics. *Journal of Business Ethics, 90*(2), 129–139.

Freeman, R. E. (1984). *Strategic management: A stakeholder approach*. Boston: Pitman.

Friedman, M. (1962). *Capitalism and freedom*. Chicago: University of Chicago Press.

Gaski, J. F. (2008). The index of consumer sentiment toward marketing: Validation, updated results, and demographic analysis. *Journal of Consumer Policy, 31*(2), 195–216.

Gaski, J. F., & Etzel, M. J. (1986). The index of consumer sentiment toward marketing. *Journal of Marketing, 50*, 71–81.

Gaski, J. F., & Etzel, M. J. (2005). National aggregate consumer sentiment toward marketing: A thirty-year retrospective and analysis. *Journal of Consumer Research, 31*, 859–867.

Gruca, T. S., & Rego, L. L. (2005). Customer satisfaction, cash flow, and shareholder value. *Journal of Marketing, 69*, 115–130.

Hasnas, J. (2013). Whither stakeholder theory? A guide for the perplexed revisited. *Journal of Business Ethics, 112*, 47–57.

Holder-Webb, L., & Cohen, J. (2012). The cut and paste society: Isomorphism in codes of ethics. *Journal of Business Ethics, 107*(4), 485–509.

Hustad, T. P., & Pessemier, E. A. (1973). Will the real consumer activist please stand up: an examination of consumers' opinions about marketing practices. *Journal of Marketing Research, 10*, 319–324.

Ittner, C. D., Larcker, D. F., & Taylor, D. J. (2009). The stock market's pricing of customer satisfaction. Rock Center for Corporate Governance Working Paper No. 62.

Jacobson, R., & Mizik, N. (2009). The financial markets and customer satisfaction: Reexamining possible financial market mispricing of customer satisfaction. *Marketing Science, 28*, 810–819.

Jensen, M. C. (2002). Value maximization, stakeholder theory, and the corporate objective function. *Business Ethics Quarterly, 12*(2), 235–256.

Kotler, P. (1972). *Marketing management: Analysis, planning, and control* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall Inc.

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy, 74*, 132–157.

Lawrence, A. (2013). Individual investors and financial disclosure. *Journal of Accounting and Economics, 56*, 130–147.

Leland, H. E., & Pyle, D. H. (1977). Informational asymmetries, financial structure, and financial intermediation. *Journal of Finance, 32*, 371–387.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance, 66*, 35–65.

Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *Journal of Finance, 69*, 1643–1671.

Loughran, T., McDonald, B., & Yun, H. (2009). A wolf in sheep's clothing: the use of ethics-related terms in 10-K reports. *Journal of Business Ethics, 89*, 39–49.

Lundstrom, W. J., & Lamont, L. M. (1976). The development of a scale to measure consumer discontent. *Journal of Marketing Research, 13*, 373–381.

Luo, X., Homburg, C., & Wieseke, J. (2010). Customer satisfaction, analyst stock recommendations, and firm value. *Journal of Marketing Research, 67*, 1041–1058.

Manning, C. D., & Schütze, H. (2003). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.

Michalisin, M. D. (2001). Validity of annual report assertions about innovativeness: An empirical investigation. *Journal of Business Research, 53*, 151–161.

Michalisin, M. D., & White, G. P. (2000). Validity of annual report assertions about quality: An empirical study. *The Mid-Atlantic Journal of Business, 36*, 103–122.

Olshavsky, R. W., & Miller, J. A. (1972). Consumer expectations, product performance, and perceived product quality. *Journal of Marketing Research, 9*, 19–21.

Patelli, L., & Pedrini, M. (2013). Is tone at the top associated with financial reporting aggressiveness? *Journal of Business Ethics, 3*, 1–17.

Pryor, F. L. (2001). New trends in U.S. industrial concentration. *Review of Industrial Organization, 18*, 301–326.

Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics, 87*(3), 355–374.

Stern, B. B. (1996). Textual analysis in advertising research: Construction and deconstruction of meanings. *Journal of Advertising, 25*(Fall), 61–73.

Stone, P. J., Dunphy, D. C., Smith, M. S., & Ogilvie, D. M. (1966). *The general inquirer: A computer approach to content analysis*. Cambridge, MA: MIT Press.

Yuthas, K., Rogers, R., & Dillard, J. F. (2002). Communicative action and corporate annual reports. *Journal of Business Ethics, 41*(1–2), 141–157.