Not All Investors Know:

The Investment Value of Spamming Consumer Opinions[*]

Dawoon Kim[†]

Preliminary and Incomplete – please do not cite or distribute

April 2019

Most recent version available here

Abstract

This paper investigates whether there is significant investment value embedded in consumer opinion spamming. Using nearly 45 million reviews from Amazon.com, I measure the likelihood that a review is spam (its "spamicity") via machine learning and find that spamming is prevalent across time and products. I show evidence that spamming is typically conducted to conceal unfavorable opinions and thus contains hidden information about product quality. The positive association between stock returns and abnormal review scores disappears when such reviews are subject to high spamicity. Furthermore, I examine the investment value of spam detection by demonstrating that a portfolio that goes long on stocks with high abnormal review scores and low spamicity and short on stocks with low abnormal review scores and high spamicity earns abnormal returns of 1.17% to 1.23% per month. Finally, I show evidence that there are sophisticated investors who exploit the information embedded in spamicity. I interpret these results pertaining to the costs of processing consumer opinions in the context of information-signaling theories.

Key Words: Consumer Opinions; Opinion Spamming; Noise in Information; Stock Returns

JEL: G12, G14, G23, L15

---

[*] I am very grateful to the members of my dissertation committee for their guidance, support, and helpful comments: Andrew Karolyi (chair), Murillo Campello, Hyunseob Kim, and Scott Yonker. I am also grateful to Warren Bailey, David Ng, Justin Murfin, Matthew Baron, Adrian Aycan Corum, Jawad Addoum, Fang Liu, Byoung-Hyoun Hwang, Pamela Moulton, Gaurav Kankanhalli, Alan Kwan, and Rajiv Banker (discussant) as well as conference and seminar participants at the 2019 Conference on Emerging Technologies in Accounting and Financial Economics, Cornell University, Babson College, Auburn University, Villanova University, State University of New York at Buffalo, University of Missouri, Nanyang Technological University, University of Melbourne, University of New South Wales for valuable feedback. All errors are my own.

[†] Cornell S.C. Johnson College of Business, Cornell University; Email: dk722@cornell.edu; Personal Website: https://sites.google.com/view/dawoonkim

"The greatest deception men suffer is from their own opinions" - Leonardo da Vinci

1. Introduction

The growth of internet shopping has given rise to new tools that firms and consumers use to interact with each other. One such tool is the consumer product review. Not only do consumer opinions contained in product reviews convey information about products to potential consumers, they can also help firms and their investors make investment decisions based on the market's perception of product quality. While there may be considerable informational value in consumer opinions, however, there may also be an incentive to influence those opinions by means of spam reviews,[1] also known as opinion-spamming ("spamming" hereafter). Detecting spam reviews is a non-trivial exercise. In this paper, I measure the probability that a review is spam using machine-learning techniques adapted from computer science and find evidence that a significant amount of valuable information is hidden in consumer spamming.

As internet usage has grown (seemingly exponentially) so has e-commerce.[2] E-commerce helps firms streamline their distribution channels and find consumers. Consumers also benefit from e-commerce due to the ease of transactions and enhanced search capabilities. What may be equally, if not more, important is communication between firms and consumers online. The presence of information asymmetry between consumers and firms causes firms to rely heavily on advertising to signal their quality of their products (e.g., Milgrom and Roberts, 1986). This information asymmetry is even more severe for e-commerce as physical quality inspection before making a purchase

[1] In this paper, I define spam reviews broadly as unsolicited product reviews with commercial intent designed to influence consumer behavior. This includes fake reviews and other forms of biased reviews sent with commercial intent.

[2] For instance, U.S. manufacturing firms' aggregate shipment value from e-commerce grew from $1 trillion in 2004 to $3.6 trillion in 2014 (a 260% increase) and it is about 61.4% of their total shipment value in 2014 (an 89.1% increase from 2004). See multi-sector e-commerce data from the United States Census Bureau available here.

is practically impossible.[3] In this context, consumer opinions provide critical means for e-commerce vendors to reduce information asymmetry by providing product quality information to potential consumers.[4]

Consumer opinions may be as important to investors as they are to firms insofar as they are associated with firms' future cash flows. For example, recent studies by Fornell, Morgeson, and Hult (2016) and Huang (2018a) find evidence that consumer opinions contain information that is not fully captured by the stock market. The underlying reason that publicly available consumer opinions are not fully reflected in financial markets is yet to be investigated. In this paper, I hypothesize that the noise diluting consumer opinions that arises from spam reviews is an important factor. If spamming is conducted to dilute unfavorable opinions, their presence will convey information about product quality that is not apparent from observed review scores. Given that spamming tends to distort consumer opinions, spam detection will provide more accurate signals of product quality that may benefit investors who can exploit such information.

Traditional signaling models have focused more intently on the cost of transmitting information from firms to investors than on the cost of gathering and processing data as the primary source of information about a firm. Because managers gather the relevant information they need to operate a firm, the marginal cost of gathering such information by investors has been deemed

[3] Not surprisingly, internet advertising has become the fastest growing sector in the advertising industry. For example, the aggregate revenue from internet advertising in the U.S. grew from $2.7 billion in 2004 to $49.5 billion in 2014 (a 1,733% increase), an amount that is about 52.8% of the total advertising revenue in 2014 (a 1,190% increase from 2004). See PWC's internet advertising revenue report, available here.

[4] Consequently, the success of e-commerce platforms has grown in tandem with the volume of consumer opinion supplied. For instance, Amazon.com's share of the online sales market in the U.S. has risen to 25.8% and, while it attracted 3.74 reviews per product in 2004, this figure increased to 5.35 reviews per product in 2014 (a 43% increase). Also, Yelp.com has now secured more than 70 million reviews of various products and services; its figure of 1.52 reviews per business in 2004 and had risen to 11 reviews per business in 2014 (a 622.9% increase). See Huang (2018b).

low (Merton, 1987).[5] In light of ever-advancing information technology and tools for big data analysis, however, the cost of gathering and processing information from a myriad of databases may be at least as important (Dugast and Foucault, 2018).

The cost of gathering and processing information will be amplified if it is accompanied by a significant amount of noise, rendering it difficult to extract relevant information cleanly even from publicly available data.[6] Therefore, investment value might exist in information derived from the product market (such as consumer opinions) when spamming is non-negligible. Consequently, perfect informational efficiency can be attained only if sophisticated investors with superior information-processing capability extract the costly information and reveal it to the market. This process will be incomplete, though, as long as such activities are costly (Grossman and Stiglitz, 1980; Diamond and Verrecchia, 1981; Verrecchia, 1982; Kyle, 1985).

The cost of detecting spam reviews comes from several sources. First, there is no clear way to credibly identify a set of spam reviews to train a machine-learning algorithm or to verify the external validity of any estimates (e.g., Ott et al., 2011; Mukherjee, Liu, and Glance, 2012; Mukherjee et al., 2013). Second, spam detection is multi-faceted insofar as it requires a vast amount of data-processing to understand reviewers' behaviors and manufacturers' spamming incentives (e.g., Jindahl and Liu, 2008; Lim et al, 2011; Xie et al., 2012). Third, it is unclear how much value there is in detecting possible spam reviews. This last point is the focus of this paper.

Using almost 45 million product reviews ("reviews" hereafter) from Amazon.com ("Amazon" hereafter) that were posted online from 2004 through 2014, I first measure the probability

[5] Thus, those models focused on conveying the relevant information to investors because attention is a scarce resource. Hence, firms engage in activities designed to attract increased attention from investors and to reduce information asymmetry between firms and investors. See, for example, Karolyi, Kim and Liao (2018).

[6] There is considerable evidence in economics and computer science pertaining to information noise such as spam reviews (see, among others, Mayzlin, Dover and Chevalier, 2014; and Luca and Zervas, 2016) as well as misleading advertisements, fake news, and phishing emails (for example, Böhme and Holz, 2006; Clarke et al., 2018; Kogan, Moskowitz, and Stein, 2018).

that a review is fake/spam ("spam" hereafter) via supervised machine learning. To construct the training sample where so-called positive examples (generally accepted spam reviews) are labeled, I identify 435,456 nearly identical reviews following the approach of Jindhal and Liu (2007, 2008). These reviews are suspected of being spam because spam reviewers write the same or similar reviews on multiple products sometimes using separate accounts.

The estimated average probability that a review in my sample is a spam review (or what I call "spamicity") during the sample period is 24.6%, and ranges between 17% and 30% across time and product categories. The estimated spamicity of a review is negatively correlated with its abnormal review score ("abnormal score" hereafter) and is positively correlated with abnormal scores of peer products. These findings imply that: (1) spamming decisions are correlated with levels of consumer opinions; (2) large portions of spam reviews are produced to counteract unfavorable opinions; and (3) counteracting unfavorable opinions about a firm's own products is more likely to succeed if competing products receive positive opinions. I also find that spamicity is positively correlated with advertising expenses and a host of firm-level governance quality indices; spamming appears to be a deliberate tactic aimed at improving product-market performance rather than a manifestation of poor firm governance.

In the second part of the paper I show evidence that estimated spamicity provides significant investment value among Amazon product reviews that are associated with public companies. Using 469 unique Compustat firms whose products are sold on Amazon in the U.S., I form tercile portfolios in each month based on spamicity, following Huang (2018a), and show that a spread portfolio that goes long on stocks with low spamicity and that goes short on those with high spamicity yields a Fama-French-Carhart four-factor alpha (Fama and French, 1993; Carhart, 1997)

of around 60 basis points per month.[7] This evidence suggests that spamicity contains information about product quality that may be associated with firms' future cash flows.

To further identify the informational value in spamicity, I form the abovementioned spread portfolios conditional on each abnormal score tercile. I find that spamicity is most informative in the mid-tercile of abnormal scores where signals from monthly abnormal scores are mixed to the greatest extent, yielding a Fama-French-Carhart four-factor alpha of around 98.9–112.2 basis points per month. In addition, spamicity predicts future returns. Fama-MacBeth regression results show that spamicity is negatively correlated with one-month-ahead excess returns. A one-standard-deviation increase in spamicity is associated with a decrease of 26.4 basis points in one-month-ahead excess returns. This effect is amplified by abnormal scores.

I then form a trading strategy based both on scores and spamicity to verify the potential investment value of spam detection. A spread portfolio that goes long on stocks with high abnormal scores *and* low spamicity and that goes short on those with low abnormal scores *and* high spamicity yields a monthly Fama-French-Carhart four-factor alpha of between 1.17% and 1.23% per month.[8] The observed abnormal returns are economically and statistically significant among firms with: (1) high e-commerce sales proxied by average product-level e-commerce sales, (2) large numbers of new products sold on Amazon, and (3) high average product prices. The monthly Fama-French-Carhart model-based alphas among such firms are 1.41%, 1.45%, and 1.34%, respectively. These results imply that the investment value of spam detection is concentrated among firms that are affected to a greater extent by consumer opinions and thus are more likely to suffer

[7] I also replicate Huang (2018a) by forming a spread portfolio that goes long on stocks with high abnormal scores and that goes short on those with low abnormal scores and find a Fama-French-Carhart four-factor alpha of around 70 basis points per month. Huang (2018a) finds a Fama-French-Carhart four-factor alpha of 55.7–73.0 basis points per month in his sample for a period running from 2004 through 2015.

[8] I find no evidence of return reversal beyond a one-month holding period. See Table A2 for results pertaining to abnormal returns over longer holding periods.

from, or have incentives to commit, spamming.

Next, I show evidence that sophisticated investors can discern the credibility of consumer opinions. For instance, a one-standard-deviation increase in spamicity lowers institutional ownership by 46 basis points, which accounts for 1.77% of the standard deviation of institutional ownership in the sample. This effect is concentrated among independent institutional investors such as mutual fund managers and investment advisors, who are more likely to collect various sources of information as opposed to "grey" institutions (e.g., pension funds and endowments), per the definitions given in Ferreira and Matos (2008). Given the large abnormal returns from the spread portfolios on abnormal scores *and* spamicity, the small economic magnitude suggests that not enough institutions take advantage of the potential information harbored in spamicity.

I then offer new evidence that the positive correlation between earnings surprises and abnormal scores found by Huang (2018a) is substantially weaker if a firm exhibits high levels of spamicity in its product reviews. A one-standard-deviation increase in spamicity decreases the positive correlation between abnormal scores and earnings surprises by 50%. This particular effect is concentrated among firms with below-median analyst coverage, suggesting that some analysts may be deriving at least some information from the product market. Consistent with these findings, abnormal returns from spread portfolios on abnormal scores *and* spamicity are concentrated among firms with lower percentages of independent institutional ownership and analyst coverage.

In one final experiment, I investigate the consequences of a new policy adopted by Amazon that altered firms' incentive to engage in spamming. In 2012, Amazon launched its "Global Selling Program," an initiative that allows many Chinese manufacturers to list and sell their products directly on Amazon in the U.S.[9] As a result, the number of Chinese manufacturers/sellers on Amazon

[9] Previously, most Chinese manufacturers had to sell their products abroad through importers. The program significantly lowered the bar for Chinese manufacturers to list their products on Amazon websites around the world. This,

increased substantially (by a factor of 13 from 2012 to 2015).[10] The sudden influx of products from China is found in my sample as well. I use universal barcodes such as UPCs (Universal Product Codes) and an EANs (European Article Numbers) where the first few digits indicate the country where the product was barcoded. The ratio of the products barcoded in China to those in the U.S. increased by more than 158.2% between 2011 and 2013.[11]

The influx of such products appears to have increased competition and thus in turn increased incentives to engage in spamming. There is a disproportionate increase in spamicity among firms whose products included above-median Chinese products in their product categories. I find evidence that only independent institutional investors that held shares of firms with higher exposure to the policy change adjust their holdings against spamicity after 2012. This finding is consistent with the hypothesis that only a few sophisticated investors are willing and able to invest in analyzing the information embedded in consumer opinions reflected in spamicity. Less sophisticated investors are unlikely to invest resources to detect spamming and therefore cannot fully infer such information from consumer opinions. Thus, there remains investment value in spam detection.

To the best of my knowledge, this paper is the first to investigate the implications of spamming on investment returns. The remainder of the paper is organized as follows. In Section 2, I summarize the related literature on consumer opinions and spamming. In Section 3, I describe the data and empirical methods. In Section 4, I explain the spam-detection model used to estimate the spamicity of reviews. In Section 5, I present the empirical results and in Section 6, I conclude the paper.

---

however, differentiates them from products manufactured in China by U.S. manufacturers via original equipment manufacturing (OEM).

[10] See a news article on the increase in Chinese manufacturers (sellers as well) after using the Amazon program here.

[11] Many products in the sample lack product identifiers such as UPCs or EANs. As the program allowed many generic products to be sold on Amazon, this estimate is expected to be the lower bound.

2. Literature Review

2.1 Consumer Opinions

This paper contributes to the literature on the importance of consumer opinions to firms and investors. The role that positive consumer opinions play in product sales and marketing strategy is well-documented in the marketing and strategy literatures (see, among other studies, Bickart and Schindler, 2001; Godes and Mayzlin, 2004; Chevalier and Mayzlin, 2006; Liu, 2006; Chen and Xie, 2008; Luca, 2011; and Archak, Ghose, and Ipeirotis, 2011). For instance, Chevalier and Mayzlin (2006) find that improved book reviews on Amazon and Barnes & Noble are positively correlated with increases in book sales. Also, using review data from CNET and Amazon, Chen and Xie (2008) find evidence that consumer reviews provide a new tool that sellers use to adjust their marketing strategies in response to consumer reviews. Finally, Archak, Ghose, and Ipeirotis (2011) find that textual contents in consumer reviews can be used to understand consumers' preferences and to forecast changes in product sales.

Just as consumer opinions are related to firm performance, so they also matter to investors. The role of consumer opinions in investment and corporate decision-making is also well-documented (see, among other studies, Ittner and Larker, 1998; Larkin, 2013; Fornell, Morgeson, and Hult, 2016; Huang, 2018a; Huang 2018b). For example, a recent study by Huang (2018a) finds evidence that consumer opinions contain information that is not fully recognized by the market. Using review data from Amazon, he finds that a spread portfolio that goes long on stocks with high abnormal scores and short on those with low abnormal scores earns monthly abnormal returns of between 55.7 and 73.0 basis points. Meanwhile, Huang (2018b) finds evidence that positive consumer opinions ease firms' financial constraints. Using review data from Yelp, he shows that a one-half-star rating increase is associated with a 40% higher probability of securing Small

Business Administration loans.

The investment value of consumer opinions is also related to the literature on the limited attention and information-processing capabilities of investors (see, among other studies, Hong and Stein, 1999; Hong, Lim, and Stein, 2000; Hirshleifer and Teoh, 2003; Hou and Moskowitz, 2005; Peng and Xiong, 2006; Cohen and Frazzini, 2008; DellaVigna and Pollet, 2009; and Dugast and Foucault, 2018). For instance, Cohen and Frazzini (2008) find evidence that the stock market does not fully incorporate news about economically linked firms, and DellaVigna and Pollet (2009) find that earnings announcements issued on Fridays, when investors are more likely to be inattentive, receive a 70% higher delayed response from the market. This paper contributes to this literature by providing a possible channel through which information-processing costs prevent the market from fully incorporating the relevant information. Spamming, along with other sources of fake or misleading information, induces another layer of information asymmetry between producers and consumers of information. The findings I report in this paper suggest that, as a result of such asymmetry, information from the product market may be incorporated in the capital market on a delayed basis.

2.2 Opinion Spamming

I adapt the machine-learning technique used in this paper from the literature on spam detection in computer science (see, for example, Jindahl and Liu, 2007, 2008; Lim et al, 2011; Ott et al. 2011; Xie et al. 2012; Mukherjee, Liu, and Glance, 2012; and Mukherjee et al., 2013). The spam detection problem was first brought to light by Jindahl and Liu (2008). Using review data from Amazon, they find that near-duplicate reviews are likely to be spam, and thus use this finding as a positive training example for spam detection. They show that spamming is a widespread

problem and incorporating reviewer- and product-level characteristics improves the performance of their model. Subsequent work focuses on behavioral patterns of reviewers and shows that spam-detection model performance improves greatly if such patterns are considered (Lim et al., 2011; Xie et al. ,2012).

The prevalence of spam reviews is also recognized by both economists and practitioners.[12] For example, by comparing a given hotel's reviews on Expedia, where only customers can post reviews, with reviews of that hotel on TripAdvisor, where anyone can post reviews, Mayzlin, Dover, and Chevalier (2014) find evidence that hotels with strong incentives to post fake reviews attract more positive reviews while their neighbors attract more negative reviews. Luca and Zervas (2016) also find that restaurants attracting few reviews or recent bad reviews are more likely to commit review fraud as measured by filtered reviews (approximately 16% in their sample) on Yelp. This paper contributes to this literature by providing additional large-scale analyses of determinants of spamming and its economic consequences. Finally, this paper provides a new insight into the economic value of big-data processing (Lee, Ma, and Wang, 2015; Huang, 2018b; Froot et al., 2017; and Zhu, 2018).

3. Data and Hypothesis Development

3.1 Data

I obtain Amazon review data from the Stanford Network Analysis Project (SNAP) for January 2004 through July 2014.[13] SNAP is widely used in computer science, information technology, and marketing for network and sentiment analyses. The raw data contain more than 140 million

[12] Find examples here.
[13] See McAuley et al. (2015), McAuley, Pandey, and Leskovec (2015) and McAuley and Alex (2016) for details on the data. I thank Julian McAuley for sharing the data.

product reviews covering almost all the reviews posted within the sample period. As Amazon merges reviews of nearly identical products, duplicate reviews of similar products by the same reviewer are removed from the sample.[14] Also, I exclude reviews in product categories such as "Books," "Movies and TV," "Kindle Store," "Digital Music," and "Amazon Instant Video," as reviews of these products are expected to provide less information about the manufacturers. The resulting review sample contains 44,699,245 reviews. Each review contains information on 1) the product identifier (ASIN: Amazon's product identification number), 2) reviewer ID, 3) the score given in the review, 4) the title of the review, 5) the contents of the review, 6) the review helpfulness indicator, and 7) the date of the review.

Additionally, I obtain metadata on Amazon products from SNAP. These metadata provide a detailed product description and a snapshot of each product's price.[15] Also, the data include information on each product's related products (e.g., "Also Bought," "Also Viewed," "Bought Together," and "Buy After Viewing").[16] As a complement to each product's review score, I track down the review scores of related products as well.[17]

Also, as the review data do not provide information on manufacturers, I obtain additional metadata on Amazon products via ASINtool.com, where one can acquire detailed product information including manufacturers and brand names, using ASINs. I then manually match those names to firms in Compustat. Following Huang (2018a), I exclude firms that receive fewer than ten total reviews in each month. The resulting sample comprises 469 unique U.S. firms used in the

[14] For example, reviews of product A in the color red can be merged into reviews of the same product B in the color blue and vice versa.
[15] This is the price of a product when the review data were collected.
[16] According to Amazon, this relatedness is determined by Amazon's recommendation algorithms, which measure cosine similarities between customers. See Amazon's industry report here.
[17] Although the metadata do not indicate historical changes in product relatedness, such relatedness is not expected to change much in the sample as the cost of measuring such relatedness among all products is exorbitant. I assume that this relatedness is constant throughout the sample period and track related products' review scores as they appear in the sample.

main analyses.[18] One interesting feature of these metadata is that they contain detailed information on each product's registration numbers, such as UPCs, EANs, MPNs (Manufacturer Part Numbers), and SKUs (Stock-Keeping Units). Although the last two identifiers are not universal among manufacturers, UPCs and EANs are standardized (barcode numbers) where the first few digits of each indicate the country where the product was barcoded.

Next, I obtain data on firm-level characteristics, stock returns, analyst forecasts, and institutional holdings from Compustat, CRSP, IBES, and Factset Lionshares. For monthly analyses, I impute quarterly or annual firm-level observations based on data availability. Additionally, I obtain quarterly firm-level governance and ESG indices from MSCI ESG GovernanceMetrics. Lastly, I obtain sentiment lexicons (a list of words and their sentiment scores) from Hamilton et al. (2016) to measure the sentiment expressed by each review's content from textual analysis.[19]

[Insert Table 1 Here]

In Table 1, I report review, reviewer, and product information. The review sample (Full Review Sample) contains 44,699,245 reviews of 4,977,718 products posted by 26,482,938 reviewers, and the Compustat matched sample (Main Sample)[20] contains 3,827,524 reviews of 250,554 products posted by 3,100,134 reviewers.[21] Panel A of Table 1 displays a year-by-year summary of average scores and the number of reviews, reviewers, and products for each year from 2004 through 2014. All the numbers, except for average scores, grew rapidly in each year, especially in 2013. The per-product number of reviews also peaks in both samples (6.32 and 9.64 reviews per

[18] This excludes financial firms (SIC one-digit "6") and utility firms (SIC two-digit "49").
[19] I thank William L. Hamilton, Kevin Clark, Jure Leskovec, and Dan Jurafsky for sharing the lexicon data, which are available here.
[20] Manufacturer and brand information is not available for all products offered on Amazon. Products for which no manufacturer or brand names are given on the website are not included in the main sample.
[21] These numbers do not reflect reviews written by the same reviewers on the same product, which were excluded from the sample. This reflects either the merging of reviews of nearly similar products by Amazon or technical errors that occur when reviewers (apparently inadvertently) hit the submit button multiple times.

product, respectively).

In Panel B of Table 1 I report the average review scores (or "scores" interchangeably) and the number of products and reviewers by Amazon product category. Average scores are clustered around 4.00 in both samples (with averages of 4.07 and 4.08, respectively). Among the 18 product categories, "Apps for Android" and "Video Games" attracted the highest number of per-product reviews (43.0 and 24.3 reviews per product, respectively) followed by "Electronics" and "Baby" (16.3 and 14.0 reviews per product, respectively). This pattern may occur because these are the types of experience goods that need the highest level of indirect quality inspection by consumers. Meanwhile, the numbers of reviews and reviewers are highly correlated in both samples (ρ = 0.993 and 0.999, respectively) while the correlation is much weaker between the number of reviews and the number of products (ρ = 0.734 and 0.642, respectively). This suggests that the supply of consumer opinions is not a mere function of the number of products consumers can review.

This also reflects a possible determinant of the supply of consumer opinions. For a given score shown to consumers before they purchase, those consumers who perceive that the quality of a product they might purchase is close to the score have little reason to write a review (Li and Hitt, 2008). If there is a significant discrepancy between the score and the perceived quality,[22] however, consumers are more likely to express their opinions until the score is sufficiently close to the perceived quality. Given the skewed high average scores in the sample, low-quality products are expected to accumulate more reviews than high-quality products, as such a discrepancy will be significant for low-quality products. Consistent with this hypothesis, the number of reviews per product and the average scores across product categories shown in Panel A are negatively correlated in both samples (ρ = -0.450 and -0.221, respectively).

[22] Here, quality is rather broadly defined to indicate the capacity of a product to satisfy consumers' expectations given its price and other product features.

The process by which the wisdom of the crowd ensures that indicators of quality such as review scores converge on the true quality in the product market bears some resemblance to how the capital market attains informational efficiency via arbitrageurs, but there exist several differences. First, unlike arbitrageurs who make riskless profits by eliminating inefficiency, reviewers receive little or no reward for their efforts.[23] Second, unlike the market price, which holds across venues, quality indicators can vary in kind. As a result, the process by which efficiency in quality indicators in the product market is ensured may not play out as quickly or as efficaciously as the arbitrage mechanism in the capital market. This may enable spamming to have a more lasting impact on consumer behaviors.

3.2 Hypothesis Development

When information such as consumer opinions that derives from the product market is free of noise, it can be efficiently incorporated into the market price. Thus, there will be no investment value in deriving such information from the product market. This process may not be as efficient as possible, however, if 1) there is a significant amount of noise in the information from the product market and 2) detecting such noise is costly. Subsequently, such information from the product market will continue to provide investment value that is not fully recognized by the capital market.

The premise of this paper stems from evidence that there is value in information derived from the product market (Fornell, Morgeson, and Hult, 2016; Huang, 2018a; Hoberg and Philips, 2018). By narrowing the focus on information conveyed in publicly observable product reviews, I hypothesize that the value of such information comes from the noise that lies within it. To verify

[23] Some reviewers may receive benefits from firms and platforms when they write reviews for professional purposes (e.g., Paid reviews on Amazon). However, this type of review is more closely related to review spamming as it functions more like advertising than impartial expressions of opinion.

this hypothesis, I first replicate Huang (2018a) to show that the information in the sample indeed has significant investment value.

Next, in the presence of noise that occurs in reviews via spamming, such noise will increase the cost of processing information from the product market. There is strong evidence that spamming is a widespread problem on multiple e-commerce platforms (e.g. Jindahl and Liu, 2007, 2008; Luca and Zervas, 2016).[24] Spam detection is likely to be costly, as it requires substantial computation to conduct review-, reviewer-, and other product-related analysis. Additionally, it is difficult to ensure the external validity of any detection algorithm as it is nearly impossible to confirm that certain reviews are spam. I analyze Amazon reviews and measure their spamicity to show evidence of the prevalence of spamming and the potential cost of spam detection.[25]

Despite the noise that dilutes information derived from the product market and the cost of spam detection, the capital market may attain informational efficiency if there are sophisticated investors with superior information-processing capabilities. With the abundance of big data and increased computing power, there is evidence that big data generates efficiency gains in the capital market (Zhu, 2018). I therefore hypothesize that the presence of such sophisticated investors helps reduce the noise that dilutes information derived from the product market. To test this hypothesis, I subdivide the firms in my sample based on the prevalence of such investors.

The presence of sophisticated investors with superior information-processing capabilities may not in itself, however, be sufficient to fully incorporate information derived from the product market into the market price. If the cost of processing information, such as that involved in spam

[24] See, for example, here.

[25] The existence of spam reviews, such as near-duplicate reviews, may also imply that the cost of spam detection is consequential. As it is critical for e-commerce platforms to establish the credibility of their review data, they may have the greatest incentive to remove spam reviews. However, this incentive may not be sufficient if such activities are expected to harm online sales or review activities. This also will add to the cost to investors of processing consumer opinions.

detection, is non-negligible, those investors are unlikely to reveal all their information (Grossman and Stiglitz, 1980). Thus, the information reflected in spamming is expected to be incorporated into the capital market over time, albeit with delays.

4. Review Spam Detection

4.1 Near-Duplicate Reviews

I measure spamicity in consumer opinions to test the above hypotheses. Following Jindhal and Liu (2007, 2008), I define spam reviews as those that are intentionally crafted to induce consumers to act against their own best interests. One empirical challenge in spam detection is that it is almost impossible for humans to detect spam reviews (Ott, Choi, Cardie, and Hancock 2011). Thus, measuring spamicity usually relies either on 1) finding or generating a small subsample (training sample) of reviews that are likely to be spam and developing a classification mechanism that assigns probability estimates on reviews based on that training sample, or 2) examining the distributional irregularities in the entire sample. The former method is related to supervised machine learning while the latter is related to unsupervised machine learning.[26] Although each method has its merits, the latter is difficult to apply to big data because of the substantial computational burden. As such, in this study I rely on supervised machine learning to measure spamicity.

I create a training sample using near-duplicate reviews identified in the review sample by closely following Jindahl and Liu (2007, 2008).[27] Near-duplicate reviews have identical or nearly identical content in the review texts written 1) by different reviewers and/or 2) of different

[26] When conducting supervised machine learning, one should create a training sample by labeling positive examples (with the value of 1) and negative examples (with the value of 0) for regression analysis.
[27] They find that approximately 1% of the reviews in their sample from Amazon have nearly identical review pairs.

products.[28] Using these near-duplicate reviews as positive examples of review spamming, I develop a classification model and estimate the model using logistic regression to assign spamicity estimates.

Near-duplicate reviews are highly suspected of being spam, for three reasons. First, because spam reviewers are likely to deal with multiple spamming tasks, in some cases repeatedly in relation to multiple products, it is likely to be difficult for them to create completely nonidentical reviews every time. Second, because it may be difficult for manufacturers and vendors to come up with spam reviews frequently, it is widely suspected that they often hire anonymous spam reviewers.[29] To ensure that spam review content is consistent and favorable, those writers are expected to receive similar guidelines that result in their producing similar reviews. Third, unless consumers or e-commerce platforms compare all reviews written at different times possibly by different reviewers, it will be hard to detect such spam reviews, especially when spam reviewers change the content slightly.[30]

[Insert Table 2 Here]

To identify near-duplicate reviews, I estimate the 2-gram Jaccard distance between review pairs and choose any pairs with similarity scores above 0.9 where the 1 indicates the distance

[28] Near-duplicate reviews written by the same reviewers and on the same products are discarded from the analysis as they can occur if a reviewer accidently posts the same review multiple times.
[29] Also see, for example, here.
[30] It is possible for duplicates or near-duplicates of reviews to occur independently. For example, short phrases such as "This is fantastic" can appear multiple times without being posted through spamming. Also, consumers may copy and paste from their own previous reviews for multiple purchases. To control for these instances, I exclude from the training sample very short reviews (of fewer than 10 words and/or fewer than 50 characters) and exclude reviews written by the same reviewers on the same product. Although there remain some honest reviews among them, most of the near-duplicates in the training sample are likely to be spam. Jindahl and Liu (2008) find that 97% of their near-duplicate reviews are of products with slightly different characteristics, implying that they are highly likely to be spam.

between identical reviews (Jindahl and Liu 2008).[31, 32] There are 455,467 near-duplicate reviews out of 44,699,245 total reviews in the sample (1.02% to the total).[33] Table 2 displays examples of near-duplicate reviews, indicating that some reviews are written very quickly and can be written several months apart. Moreover, as they are written for separate and often dissimilar products, it is difficult for consumers to detect such reviews unless they systematically compare reviews of separate products. A review helpfulness indicator (*Helpful* in Table 2) seems not to be particularly informative in this regard.

Since near-duplicate reviews are likely be the easiest to detect, the existence of such reviews indicates that the cost of spam detection may be extremely high. To detect and remove such reviews, each review should be compared with all the other reviews in the database. This means that for the raw number of approximately 140 million reviews in Amazon posted from 2004 through 2014, including any duplicate reviews, there are 140 million x (140 million – 1)/2 review pairs to compare. Assuming there can be other types of hidden spam reviews in the sample that are even harder to detect, spam detection will be costly, and we expect to see significant spamming activity.

Figure 1 displays the average scores and the proportion of near-duplicate reviews in the review sample from 2004 through 2014. When a series of spam review scandals on Amazon occurred in 2004,[34] the company promised to resolve the problem. Consistent with this effort, the proportion of near-duplicate reviews in the sample decreased over time. Nevertheless,

[31] To measure the 2-gram Jaccard distance between two texts, one should first dissect each text into a set of adjacent two-word combinations. For example, for the text "I like this product," the set will be {"I-like," "like-this," and "this-product"}. Jaccard distance, then, is the number of elements in the intersection of the two sets over the number of elements in the union of the two sets while treating identical two-word combinations within a set independently.
[32] Since the dataset contains a huge number of reviews, estimating all n * (n – 1)/2 review pairs would be extremely difficult computationally. To ease the computational burden, I first sort reviews lexicographically and measure Jaccard distance among blocks of 1,000 neighboring reviews.
[33] See Figure A1 for the distribution of the Jaccard distances between review pairs in the review sample.
[34] See for example, here.

approximately 1% of the total number of reviews in the sample are near-duplicate reviews. Interestingly, the correlation between average scores and the proportion of near-duplicate reviews is -0.904, suggesting that spamming is more prevalent for products with high review scores.

[Insert Figure 1 Here]

4.2 Spam Classification Model

Using these 435,456 near-duplicate reviews as positive examples of spamming, I construct a training sample for supervised machine learning. Insofar as I use spam detection in this paper to measure the strength of the intention to affect quality information and thus to alter consumer behavior, negative examples of spamming are chosen from among the reviews that are less likely to be spam reviews. As such, I identify reviews that do not deviate from the average scores by +/−0.1 when those reviews were written.[35] I choose the range of +/−0.1 because the average review scores displayed for consumers reflect the use of a 0.1-point scale. The overall empirical results are robust across various range choices.[36] To further eliminate possible selection bias, I randomly choose 10% of such reviews and use them as negative examples of spamming, resulting in 446,842 such reviews.

Based on the training sample of 882,298 reviews (comprising 435,456 positive and 446,842 negative examples), I estimate the following classification model using logistic regression, following Jindhal and Liu (2008) in using only the information available in the review sample without reference to firm-level characteristics:

$$\textit{Near-Duplicate Review Dummy}_{i,j,k,t} = \alpha + \beta\, \textit{Review}_{i,t} + \gamma\, \textit{Reviewer}_{j,t} + \delta\, \textit{Product}_{k,t} \qquad [1]$$

where $\textit{Near-Duplicate Review Dummy}_{i,j,k,t}$ is a binary variable that takes the value of 1 for near-

[35] The average review score calculated in the sample may differ from the actual review score displayed when the reviewer posts the review.
[36] See Table A4.

duplicate reviews and 0 otherwise for each review i posted by reviewer j of product k on date t; $Review_{i,t}$, $Reviewer_{i,t}$, $Product_{k,t}$ indicate review-, reviewer- and product-centric features on date t when review i was written, respectively. Detailed definitions of all variables are presented in the appendix.

[Insert Table 3 Here]

Model [1] is measured by the following ten-fold cross-validation method (Jindahl and Liu, 2008). First, the training sample is randomly divided into ten groups. Second, I select one of the ten groups and then use the nine remaining groups to estimate model [1] using logistic regression (model training). Third, the estimated coefficients from the logistic regression are used to extrapolate the probability estimate (spamicity) for the initially selected group. Fourth, I carry out the same process nine more times by selecting another one of the ten groups each time. This method is widely used to avoid overfitting of a model in a training sample and to check the external validity of supervised machine learning. As a result, each group is used nine times for model training and one time for spam estimation. I use the average coefficients from those ten logistic regressions to estimate out-of-the-training sample spamicity. In Table 3, I present one of the ten logistic regressions.[37]

Spam reviewers may write spam reviews repeatedly, and therefore the reviewer- and product-level features reported in Table 3 might be the most informative data that can be derived from the reviews. Table 3 displays that reviews written by reviewers who tend to write more first or only reviews of multiple products (Percent of First Reviews and Percent of Only Reviews, respectively), more high-score and fewer dispersed reviews (Avg. Score [Reviewer] and Stddev. of

[37] The coefficients in each regression are similar to each other in both magnitude and statistical significance, and thus only one of the ten results is reported.

Scores [Reviewer], respectively), and more reviews that deviate from the average scores (% of Negative Dev. and % of Positive Dev.), exhibit higher spamicity. Also, in Table 3, I list reviews of products that receive higher and more volatile scores (Avg. Score [Product] and Stddev. of Scores [Product], respectively) as well as products that attract fewer reviews and exhibit wider fluctuations in the number of reviews (Logged # of Product Reviews and Logged Stddev. of # of Reviews, respectively), which tend to result in higher spamicity.

The role of classification model [1] is to identify features that distinguish positive examples from negative examples. The more important question is whether applying model [1] to the training sample makes it possible to predict positive examples that are posted out of the sample. One benefit of the ten-fold cross-validation method is that it enables me to evaluate the out-of-sample predictability of a model. In each logistic regression, where only nine of the ten groups are used, one can verify whether the extrapolation from the estimates can be used to distinguish the positive and negative examples in the remaining group.

The most widely accepted metric for evaluating out-of-sample predictability in supervised machine learning is the Area Under Curve (AUC) of a Receiver Operator Characteristic (ROC) curve. An ROC curve measures the proportion of true-positive rates and false-positive rates from a classification under a given probability threshold. The AUC measures the out-of-sample predictive power of a classification model. Statistically, it indicates the probability that a randomly sampled positive example will be ranked higher in probability than a randomly sampled negative example. An AUC of 1 indicates that a classification model can perfectly predict out-of-sample positive examples without type 1 errors and an AUC of 0.5 indicates that a classification model's performance is no better than that of a random assignment. The average AUC of the out-of-sample prediction is 0.855. This shows that when I apply model [1] to the training sample it demonstrates

out-of-sample predictive capacity.[38]

To further establish the external validity of the spamicity estimates derived from model [1], I conduct a lift-curve analysis. The AUC of 0.855 shows that the model can predict near-duplicate reviews in the training sample. However, as there are no such reviews outside of the training sample, it is unclear whether the spamicity estimates for the reviews outside of the training sample are reasonable. The lift curve, in this regard, shows whether the out-of-the-training-sample spamicity estimates are associated with spam reviews by showing how many reviews with high spamicity estimates are in a category that is expected to be associated with spamming. The results are shown in Figure 2.

[Insert Figure 2 Here]

To obtain the results I report in Figure 2, I randomly divide the out-of-the-training-sample reviews into ten groups and count the number of reviews that are within the top #% of spamicity in each group and fall into either one of four categories: [+] *Deviation (Good Product)*, [–] *Deviation (Good Product)*, [+] *Deviation (Bad Product)*, or [–] *Deviation (Bad Product)*. [+]/[–] *Deviation* indicates reviews that have scores higher/lower than an average score when the review was posted. *Good/(Bad) Product* indicates products that have average scores of 4 or higher/(2 or lower) when each review was posted. Therefore, the lift curves displayed in Figure 2 show the proportion of reviews with the top #% of spamicity estimates in each category. The reference curve indicates random assignments.

[38] See Figures A2 and A3 for detailed depictions of ROC, AUC, and the model's performance. In medical or cybersecurity studies, an AUC of 0.99 or higher is desired as the accuracy of a model is critical for predictive diagnosis. However, because the goal of classification model [1] is not about detecting near-duplicate reviews but about detecting reviews that share characteristics of near-duplicate reviews in the training sample, the AUC of 0.855 is acceptable. The AUC of the similar classification model reported in Jindahl and Liu (2008) is 0.78. The AUCs of model [1] when using only either review-centric, reviewer-centric, or product-centric features are 0.7862, 0.7255, and 0.7212, respectively, indicating that all three features improve model performance (see Figure A3).

For example, in Figure 2 we can see that about 38.2% of reviews that received scores higher than the average scores that are also lower than 2 constitute the top 10% of the reviews with the highest rates of spamicity. Insofar as spamming is intended to alter consumer opinions when quality information about a product is unfavorable, reviews that assign high scores to bad products ([+] *Deviation (Bad Product)*) and reviews that assign low scores to good products ([–] *Deviation (Good Product)*) are more likely to be related to spamming. Consistent with this hypothesis, in Figure 2 we see that such reviews exhibit high spamicity rates derived from the out-of-the-training-sample estimates of model [1].

[Insert Figure 3 Here]

5. Empirical Results

5.1 Prevalence of Opinion Spamming

Having established that the spamicity estimates derived from model [1] are reasonable and do reflect the review spamming, I analyze the economics behind spamming. Figure 3 displays the time trend for average spamicity in the sample. Overall average spamicity in the sample is 24.6%. We can see that, before 2007, average spamicity in the sample decreases along with the proportion of near-duplicate reviews. Since 2008, however, average spamicity has been increasing despite the downward trend in the proportion of near-duplicate reviews. Also, spamicity is higher among reviews assigning high scores (4 or higher) than those assigning low scores (2 or below). One possible explanation of this trend is that spamming is conducted primarily to conceal the poor quality of one's own products rather than attacking others' (competitors') products.

The time trend for spamicity shown in Figure 3 also indicates that spam reviewers may evolve over time. Even though the prevalence of apparent spam reviews (i.e. near-duplicate reviews) decreases throughout the sample period, possibly as a result of efforts to eliminate spam

reviews, average spamicity increases slowly. Because it is more difficult to detect evolving spam reviews, the cost of detecting such reviews is expected to have risen as well.

[Insert Figure 4 Here]

Figure 4 shows average spamicity by product category. Together with Figure 3, it shows that spamming is prevalent across time and product categories.[39] Notably, the four product categories exhibiting the highest spamicity in Figure 4 ("Electronics," "Video Games," "Apps for Android," and "Baby") are the categories receiving the most positive per-product reviews (Table 1). Consumers are more likely to post reviews when there is a significant discrepancy between the score and perceived quality. Such discrepancies are expected to be the highest among the four abovementioned product categories. This also creates an incentive to spam as a means of counteracting unfavorable consumer opinions. This evidence suggests that spamming may undermine the role of the wisdom of the crowd, which ensures that accurate quality information is reflected in the product market.

[Insert Table 4 Here]

5.2 Determinants of Opinion Spamming

Next, I analyze firm-level determinants of spamming. When there is information asymmetry between firms, consumers, and investors, firms that produce high-quality products will try to signal that quality to their consumers and investors (Milgrom and Roberts, 1986). It is unclear, however, how firms react when they know they are selling products of poor quality. It is reasonable to consider spamming as a tactic such firms might undertake, and if so it will increase the level of information asymmetry. To investigate this channel, I estimate the correlation between firm-level

[39] Some sources argue that less than 1% of reviews posted on Amazon are inauthentic. Given the number of near-duplicate reviews alone, however, this number seems to be significantly underestimated. See also an example here.

average spamicity and other characteristics, including scores.

In Table 4, I report average firm-level spamicity along with other firm-level characteristics used in the analyses and compare between firms exhibiting above-median spamicity with those exhibiting below-median spamicity. These results indicate that review characteristics in the high-spamicity and low-spamicity groups differ across average scores and abnormal scores, measured, following Huang (2018a), by each month's average scores subtracted by the previous-12-month average scores. Consistent with the hypothesis that firms engage in spamming to conceal unfavorable opinions, the high-spamicity group receives lower (negative) average abnormal scores than the other group. In addition, consistent with the evidence reported in Table 1 and Figure 4 that spamming should be more prevalent when consumers post more reviews as a reflection of quality discrepancy, the high-spamicity group has a higher per product review average (24.8= 6,903/278.3) than the other group (8.1= 667 / 82.9). Other than the review characteristics, the two groups have similar average overall firm characteristics.

[Insert Table 5 Here]

To investigate the determinants of spamming, I estimate the following panel regression model:

$$Spamicity_{k,t} = \alpha + \beta\, Review_{k,t} + \gamma\, Firm_{k,t} + \eta_k + \theta_t \qquad [2]$$

where $Spamicity_{k,t}$ indicates the firm-level spamicity of new reviews for each firm k in month t; $Review_{k,t}$ indicates each firm k's monthly review characteristics, such as abnormal scores, standard deviations of scores, and abnormal peer scores, which are monthly average scores for products that are included in the "Also Viewed"[40] category on Amazon subtracted by their previous-12-month

[40] Multiple products marketed by a single manufacturer are excluded when measuring the potential competitors' review scores. In this regard, "Also bought" and "Bought Together" products are discarded as they can be

average scores; $Firm_{k,t}$ indicates, closely following Huang (2018a), each firm k's characteristics in month t; *Advertising* is defined by advertising expenses divided by total sales, *R&D*, logged market capitalization (*Logged Market Cap.*), *M/B*, *Gross Profitability* (Novy-Marx, 2013), *F-score* (Piotroski, 2000), logged dollar trading volume (*Logged Dollar Volume*: Brennan et al., 1998), the coefficient of the variation in dollar trading volume over the previous 12 months (*CV of Dollar Volume*: Chordia et al., 2001), monthly stock returns (*Stock Return*), the governance index (*GOV*) and the ESG index (*ESG*); η_k and θ_t indicate firm and time fixed effects, respectively.

In Table 5, I report the estimation results for model [2]. Throughout the remainder of the paper, I use abnormal scores as a quality indicator to capture new information embedded in review scores in each month (as in Huang, 2018a). In columns (1) through (4), I report the results of investigating the correlation between firm-level spamicity and review characteristics. As can be seen in column (1), abnormal reviews are negatively correlated with spamicity, indicating that new consumer opinions that negatively impact average scores are likely to increase the incentive to spam. A one-standard-deviation decrease in abnormal scores is associated with a 1.1% (= 0.014 x 0.78) increase in spamicity, a sizable increase given that it is 4.8% of the average spamicity (22.8%). Also, spamicity is positively correlated with score volatility (as seen in column (2)), abnormal peer scores (as seen in column (3)), and past spamicity (as seen in column (4)).

To obtain the results reported in column (5), I include firm characteristics as I investigate whether a firm's operational and financial incentives affect spamming behavior. These results imply that advertising expenses and size are positively and significantly correlated with spamicity while M/B and Gross profitability are negatively and significantly correlated with spamicity. The results pertaining to Advertising expenses and Gross profitability reported in columns (6) through

complementary goods. "Buy After Viewing" is not used either, as the number of reviews with this information in the review sample is significantly low.

(9), where review characteristics are included as well, are significant. A one-standard-deviation increase in advertising expenses is associated with a 6-basis-point (= 0.012 x 0.05) increase in spamicity, and a one-standard-deviation decrease in Gross Profitability is associated with a 60-basis-point (= 0.028 x 0.20) increase in spamicity. This suggests that firms experiencing poor performance in the product market may have a stronger incentive to engage in spamming.

To show further evidence that spamming decisions are determined largely by unfavorable opinions, I use the abnormal peer score as a possible instrument for the abnormal score of a product. Abnormally positive opinions of competing products will negatively impact positive opinions of the product of interest (inclusion restriction) and they are expected to be orthogonal to any observable firm characteristics (exclusion restriction). The results reported in Column (7) indicate that, using abnormal peer scores as an instrument for abnormal scores, a one-standard-deviation decrease in abnormal scores increases firm spamicity by 47 basis points (= 0.006 x 0.78), which is approximately 2.1% of the average spamicity (22.8%).[41]

Additionally, I investigate whether spamicity decisions are correlated with firm governance. Firms that suffer from poor governance may not allocate their quality control resources optimally and thus may not be able to produce high-quality products. The results reported in columns (8) and (9) show, however, that the correlations between firm-level governance indices and spamicity are positive and significant. These results suggest that spamicity has few implications for poor firm governance, and that spamming decisions are more likely to be made with the product market in mind.

[Insert Table 6 Here]

[41] See Table A1 for first-stage results. The f-statistic from the first stage is 84.51. Consistent with the inclusion restriction, the coefficient of abnormal peer scores on abnormal scores is 0.74. This indicates that when a given product's peers receive abnormally positive scores, the product receives relatively low scores (sensitivity below 1).

5.3 Investment Value of Opinion Spam Detection

The results reported in Table 5 suggest that spamming is correlated with consumer opinions. As spamming is likely to be conducted to compensate for negative opinions, the rate of spamming may covey information on poor product quality and add noise to review scores. This suggests the presence of significant investment value in spam detection. To show such evidence, I follow Huang (2018a) and form tercile portfolios for each month using equal weights or the weight of the number of total reviews (review weight) and investigate their performance using the Fama-French-Carhart four-factor model. In Table 6, I report the performance of the tercile portfolios based on either firm-level abnormal scores or spamicity.

To ensure the credibility of the main sample and to verify the results reported in Huang (2018a), I first replicate the main results of Huang (2018a) in Panel A. Consistent with those results, I find that a spread portfolio that goes long on stocks with high abnormal scores and short on stocks with low abnormal scores earns abnormal returns of 69.8—72.6 basis points per month. This shows that, although the scores are publicly available and relatively easy to process using big data analysis, there is significant investment value in consumer opinions.

I then repeat this analysis using spamicity. Instead of labeling each review as spam or non-spam using subjective thresholds, I use the relative magnitude of spamicity for portfolio construction to show the investment value of spam detection. The results I report in Panel B of Table 6 show that there are comparable abnormal returns when portfolios are constructed based on spamicity. As in Panel A, here a spread portfolio that goes long on stocks with low spamicity and short on those with high spamicity earns abnormal returns of 59.1—63.9 basis points per month. These returns derive from the high-spamicity portfolio where the abnormal returns are significantly negative. The overall evidence suggests that 1) high spamicity is negatively correlated with

stock returns and 2) the information content of abnormal scores and spamicity may overlap, which is consistent with the negative correlation between the two shown in Table 5.

[Insert Table 7 Here]

The equivalent performance of spread portfolios based on abnormal scores and those based on spamicity suggests that the information in each may overlap significantly. To identify the incremental information conveyed by spamicity, I form tercile spamicity portfolios conditional on each abnormal score tercile (Novy-Marx, 2013). In Table 7, I report Fama-French-Carhart four-factor-model abnormal returns of the tercile spamicity portfolios sorted within each abnormal score tercile. The monotone decrease in abnormal returns across spamicity terciles in each abnormal score tercile seen in Table 7 implies that spamicity contains additional information that goes beyond the information embedded in abnormal scores. The difference in abnormal returns between the low- and high-spamicity terciles is the highest in the mid-tercile of abnormal scores, where the signal from the abnormal scores is mixed to the greatest extent. Given the magnitude of the abnormal returns on long/short portfolios based on spamicity, the results reported in Table 7 suggest that there is significant incremental information embedded in spamicity.

[Insert Table 8 Here]

To further verify the incremental information embedded in spamicity and its impact on the information embedded in review scores, I run the following Fama-MacBeth regressions of one-month-ahead excess returns on abnormal scores and spamicity:

$$Excess\ Return_{k,t+1} = \alpha + \beta_1\ Abnormal\ Score_{k,t} + \beta_2\ Spamicity_{k,t} + \beta_3\ Abnormal\ Score_{k,t} \times Spamicity_{k,t} + \gamma\ Firm_{k,t} + \eta_k + \theta_t \quad [3]$$

where $Excess\ Return_{k,t+1}$ indicates one-month-ahead excess stock returns on firm k between month t *and* $t+1$; $Abnormal\ Score_{k,t}$ and $Spamicity_{k,t}$ indicate firm-level monthly average abnormal scores

and spamicity of the new reviews in each month, respectively. $Firm_{k,t}$ indicates firm-level characteristics as in model [2]. η_k and θ_t indicate firm and time fixed effects, respectively. If spamicity has a predictive power for stock returns, I expect to find significant and negative β_2 and β_3 in [3].

The results reported in Table 8 imply that spamicity does have predictive power for stock returns. The results reported in columns (2) and (4) imply that spamicity is negatively correlated with excess stock returns. For instance, the one-standard-deviation increase in spamicity seen in column (5) is associated with a 26.4-basis-point (= -2.640 x 0.1) decrease in one-month-ahead excess stock returns. This effect is amplified if abnormal scores are correlated with spamicity. In untabulated results, I find no significant reversals of the results reported in column (5) for excess returns at *t+2*, *t+3*, and *t+4*.

As can be seen in columns (3) and (6) of Table 8, β_3 in model [3] is statistically and economically significant, suggesting that spamicity adds noise to signals from review scores. A one-standard-deviation increase in spamicity for a given abnormal score is associated with a -1.628% change in one-month-ahead excess stock returns. Additionally, I find suggestive evidence that the return predictability of spamicity indicated in the results reported in column (6) is associated with the amount of information that firms release. Comparing the results reported in columns (7) and (8), we see that the return predictability of spamicity is much lower in magnitude and significance in a month when a firm announces its earnings. This finding is consistent with the hypothesis that information from the product market signals firms' performance in a way that is not fully captured by the capital market. The overall findings suggest that spamicity contains incremental information that can benefit investors.

[Insert Table 9 Here]

To demonstrate the investment value of spam detection in greater detail, I construct double-

sorted portfolios on both abnormal scores and spamicity and form a plausible trading strategy using both abnormal scores and spamicity. In Table 9, I report the performance of the double-sorted portfolios. The results reported in Panel A of Table 9 imply that the performance of the spread portfolio reported in Panel A of Table 6 is worst with respect to stocks with low spamicity. Meanwhile, the performance of the spread portfolio reported in Panel B of Table 6 is worst with respect to stocks with low abnormal scores.

To reflect the fact that high investment value comes from high abnormal scores and low-spamicity portfolios and low investment value comes from low abnormal scores and high-spamicity portfolios, I form a corner spread portfolio that goes long on stocks with high abnormal scores and low spamicity and short on those with low abnormal scores and high spamicity as a possible trading strategy based on both scores and spamicity. I report the results in Panel B. The abnormal returns on the abovementioned portfolios range from 1.173–1.234% per month, which is significant both statistically and economically.

In addition, I form a spread portfolio that goes long on stocks with high abnormal scores and spamicity and short on those with low abnormal scores and spamicity to show the value of spam detection. This is a possible worst-case scenario in which an investor cannot filter out spam reviews. This spread portfolio earns abnormal returns of between 18.3 and 22.6 basis points that are insignificant. The trading strategies associated with Tables 6, 7, and 9 do not incorporate trading costs, so the actual abnormal returns in this worst-case scenario are expected to be non-existent. The performance of this spread portfolio, as seen in the results reported in Panel B, shows that a trading strategy based on scores alone may not earn the significant abnormal returns reported in Panel A of Table 6 unless the noise in those scores reflected in spamicity is detected.

To shed additional light on the importance of information noise reflected in spamicity, I

subdivide the firms associated with Table 9 based on the importance of consumer opinions, and thus on spamming incentives. As consumer opinions are valuable for consumers who purchase products online, firms that rely to a greater extent on e-commerce will value consumer opinions to a greater extent and thus will have stronger incentives to engage in spamming. Similarly, firms that introduce new products online more frequently will have stronger incentives to engage in spamming. Moreover, firms whose products bring higher prices will pay more attention to consumer opinions because consumers will be more cautious about making purchases; those firms will also have stronger incentives to engage in spamming.

If the source of the high abnormal returns generated by the corner spread portfolios whose performance is represented in Table 9 is associated with noise such as spamming, high abnormal returns will be concentrated among firms with stronger incentives to produce such noise. Firms that are willing to produce spam reviews may also become involved in other types of spamming such as phishing emails and misleading advertising. Spamicity, in this case, is assumed to reflect the level of such noise even if it may not be the only source. The value of spam detection will come from partially, if not entirely, eliminating noise in consumer opinions.

[Insert Table 10 Here]

In Table 10, I report the abnormal returns on the corner spread portfolios whose performance is represented in Table 9, which I subdivide into the following groups: *E-commerce Sales*, *Number of New Products*, and *Average Product Price*. To calculate the extent of reliance on e-commerce, I first merge aggregate product-level e-commerce sales data from the United States Census Bureau ("Estimated Annual U.S. Retail Trade Sales–Total and E-commerce") into each firm's product based on their product categories. I then calculate the average of those amounts at the firm level (*E-commerce Sales*). *Average Product Price* is based on price information in the

metadata. Consistent with the hypothesis that the source of high abnormal returns is noise in consumer opinions, the results reported in Table 10 show that abnormal returns are statistically significant only among firms that are expected to rely to a greater extent on e-commerce, introduce new products more frequently on Amazon, and whose average product price is high.

[Insert Table 11 Here]

5.4 Institutional Ownership, Earnings Surprises, and Opinion Spamming

In the final set of tests for this study, I investigate whether there are sophisticated investors with superior information-processing capabilities who can benefit from the investment value embedded in consumer opinions after removing noise. To do so, I estimate the following panel regression model:

$$Institutional\ Ownership_{k,t} = \alpha + \beta_1\ Abnormal\ Score_{k,t} + \beta_2\ Spamicity_{k,t} + \beta_3\ Abnormal\ Score_{k,t} \times Spamicity_{k,t} + \gamma\ Firm_{k,t} + \eta_k + \theta_t \quad [4]$$

where $Institutional\ Ownership_{k,t}$ indicates institutional ownership of firm k in quarter t; $Abnormal\ Score_{k,t}$ and $Spamicity_{k,t}$ indicate the firm-level quarterly average abnormal scores and spamicity of new reviews in each quarter, respectively. $Firm_{k,t}$ indicates firm-level characteristics, as in model [2]. η_k and θ_t indicate firm and time fixed effects, respectively. If institutional investors, as opposed to retail investors, have superior information processing capabilities and expertise in the field, I expect to find significant and negative β_3 in [4]. I further distinguish institutional investors by type into independent institutions and grey institutions, following Ferreira and Matos (2008).

In Table 11, I report the estimation results derived from model [4]. In columns (1) through (3) I show evidence that only β_3 has a negative and significant coefficient, suggesting that institutional investors benefit from the information embedded in consumer opinions in light of the prevalence of spamicity. Considering $Abnormal\ Score_{k,t} \times Spamicity_{k,t}$ in model [4] as a new metric for

product quality, in column (3) I show that certain institutional investors can distinguish the noise from product market information. A one-standard-deviation increase in spamicity given an abnormal score will lower institutional ownership by 46 basis points (= -0.023 x 0.1–0.023 x 0.1) as seen in column (3) and lower the number of institutional investors by 3.15 (= +3.269 x 0.1–34.745 x 0.1) as seen in column (4). Given the economically sizable abnormal returns reported in Tables 6 and 9, however, this effect does not seem to be economically significant. This suggests that the cost of processing consumer opinions is non-negligible and that the benefit may not outweigh the cost for many investors.

In columns (5) and (6) I present further evidence of the existence of sophisticated investors with information-processing capabilities. As independent institutions are more likely to collect information pertinent to firm performance and product quality, β_3 is significant only among independent investors. In untabulated results, I find that foreign institutional ownership and domestic institutional ownership do not generate significant β_3 in combination with grey institutional ownership. Therefore, the results presented in column (3) suggest that not all investors adjust their portfolios based on noise they detect in consumer opinions.[42]

[Insert Table 12 Here]

I also examine whether analysts utilize this information noise when they forecast firm performance. To estimate the forecasting performance of analysts, I calculate the level of earnings surprises (*SUE*) by measuring the difference between the reported earnings per share (EPS) and the average of the most recent EPS forecasts issued by analysts before each earnings announcement.

[42] I also test additional implications of spamicity for firm governance using activist ownership based on institutions that file Form 13D. I find no evidence that lagged institutional ownership, including activist ownership, changes spamicity. Rather, I find evidence that activist institutions that hold more than 1% of a firm's shares also have negative and significant β_3.

I estimate model [4] using firms' earnings surprises instead of institutional ownership as the dependent variable. I find evidence that analysts also partially distinguish the information noise embedded in abnormal scores and spamicity, which I report in Table 12. As can be seen in column (3) of Table 12, a one-standard-deviation increase in spamicity reduces the effects on earnings surprises of abnormal scores by more than 50% (= (0.032–0.162 x 0.1)/0.32). Also, the results reported in columns (4) and (5) show that those reported in column (3) are concentrated among firms with below-median analyst coverage.

[Insert Table 13 Here]

Taking the results reported in Tables 11 and 12 together, the evidence is consistent with the hypothesis that there are sophisticated investors who use the information embedded in consumer opinions after filtering out noise. I show further evidence that the abnormal returns of the corner spread portfolios reported in Table 9 are concentrated only among firms held by sophisticated investors. As such, I subdivide these firms into two groups based on institutional holdings or analyst coverage and repeat the tests that inform Table 10 to produce the results reported in Table 13. These results suggest that, without the presence of sophisticated investors, the cost of processing information to remove noise such as spamming is non-trivial, and thus there remains significant investment value in consumer opinions.

5.5 The Chinese Product Shock of 2012

The evidence reported in the previous sections implies that there are sophisticated investors with superior information-processing capabilities. This does not necessarily mean, however, that these investors acquire information from consumer opinions, let alone detect spamming, as the information embedded in scores and spamicity may also be attained elsewhere. To shed light on

whether observed abnormal returns and institutional holdings are associated with consumer opinions and spamming, I utilize a shock that disproportionately influenced firms' incentive to engage in spamming.

[Insert Figure 5 Here]

In the first quarter of 2012, Amazon launched what it called the "Global Selling Program" in China, allowing Chinese manufacturers to list their products on Amazon and sell their products directly to consumers without the use of importers.[43] This does not mean that these manufacturers could not sell their products in the U.S. Rather, the program allowed local Chinese manufacturers to easily list and sell their products globally on Amazon using their own brand names. As a result, the number of Chinese sellers (Chinese manufacturers) increased dramatically in the U.S.[44]

This effect is also captured in the review sample. The sample includes each product's universal barcodes, such as UPC or EAN information, where (as noted above) the first few digits of the codes indicate where the product was initially barcoded. I find that the ratio of Amazon products sold in the U.S. but barcoded in China to those barcoded in the U.S. increased by more than 111.4% (158.2%) between 2011 and 2012 (2013). Although there is no guarantee that the barcode location is the same as the manufacturing location, it is likely that products barcoded in China are also manufactured in China given the strong dominance of Chinese manufacturers in the global economy. It is also more likely that products barcoded in the U.S. are manufactured in China than vice versa. Thus, the above-reported estimates arguably form the lower bound of the true proportion of products manufactured in China. In Figure 5, I compare products barcoded in China and other countries with products barcoded in the U.S. around 2012, before and after the program was

[43] See Amazon's IR website here.
[44] See also here.

launched.

[Insert Figure 6 Here]

The influx of such products increased competition on Amazon as Chinese manufacturers could produce cheaper substitute products using generic brand names regardless of their quality.[45] After dividing products into two categories based on the number of new products barcoded in China over those barcoded in the U.S. after 2012 ("China exposure"), I find that the average scores for products in categories that were subject to above-median exposure after the program launched experienced significant drops (see Figure 6).

[Insert Figure 7 Here]

The increased level of competition also increased the incentive to engage in spamming. There is a disproportionate increase in spamicity among firms whose products experienced above-median China exposure (Figure 7). I hypothesize that only sophisticated investors who invest in firms with higher exposure to the shock actively adjust their holdings against spamicity after 2012 as the noise in consumer opinions became more prominent. This is consistent with the hypothesis that only sophisticated investors are capable of analyzing the information embedded in consumer opinions reflected in spamicity in light of the cost involved. Because in this case the shock is Amazon-specific, the differential impact on institutional holdings across distinct types of investors indicates that spam detection has unique investment value for certain sophisticated investors.

To test this hypothesis, I estimate the following triple-difference model:

$$\begin{aligned} \textit{Institutional Ownership}_{k,t} = {} & \alpha + \beta_1 \textit{ Abnormal Score}_{k,t} + \beta_2 \textit{ Spamicity}_{k,t} + \beta_3 \textit{ High Exposure}_{k,t} \\ & + \beta_4 \textit{ Abnormal Score}_{k,t} \times \textit{Spamicity}_{k,t} \\ & + \beta_5 \textit{ Abnormal Score}_{k,t} \times \textit{High Exposure}_{k,t} \\ & + \beta_6 \textit{ Spamicity}_{k,t} \times \textit{High Exposure}_{k,t} \end{aligned}$$

[45] Another potential problem is counterfeit products.

$$+ \beta_7 \, Abnormal\ Score_{k,t} \times Spamicity_{k,t} \times High\ Exposure_{k,t}$$
$$+ \gamma \, Firm_{k,t} + \eta_k + \theta_t \qquad [5]$$

where $Institutional\ Ownership_{k,t}$ indicates firm k's institutional ownership in quarter t; $High\ Exposure_{k,t}$ indicates a binary variable for firm-level China exposure that takes the value of 1 if a firm's products generally fall into above-median product categories in terms of the number of products barcoded in China and 0 otherwise. $Firm_{k,t}$ indicates firm-level characteristics similar to [3]. η_k and θ_t indicate firm and time fixed effects, respectively. If sophisticated investors value consumer opinions and spam filtering, I expect to find significant and negative β_7, especially for sophisticated investors.

[Insert Table 14 Here]

In Table 14, I report the results derived from estimating model [5]. Regarding the results reported in columns (1) through (4), only those listed in (2) show a negative and significant β_7. This is consistent with the hypothesis that sophisticated investors are more responsive to China exposure. As can be seen in column (2), a one-standard-deviation additional increase in spamicity decreases the holdings of independent institutions by 57 basis points (= -0.057 x 0.01) after 2012. I further estimate the impact of the program on institutional ownership by estimating the coefficients included in [5] by year. Figure 8 displays the year-by-year estimates of β_7 in [5] using various types of institutional ownership. The results show that only independent institutions after the event year show significantly negative β_7.

[Insert Figure 8 Here]

5.6 Robustness

The overall empirical analyses are robust under varying specifications. To verify whether the observed abnormal returns from spread portfolios based both on abnormal scores and spamicity

result from relevant information about a given product, I analyze the performance of the trading strategies associated with Table 9 beyond a month to check for possible return reversal.[46] The evidence shows that there is no clear sign of reversal although most of the abnormal returns disappear one month after portfolio formation. This result provides suggestive evidence that the results reported in Tables 6 and 9 tell us more about information than about temporary investor attention (Lou, 2014).

Next, I use multiple factor models to check the robustness of the abnormal returns reported in Tables 6 and 9. I find comparable abnormal returns using Fama-French 3- and 5-factor models and find that the overall results do not change significantly. For instance, abnormal returns from a spread portfolio that goes long on stocks with low spamicity and short on those with high spamicity yields a Fama-French-Carhart six-factor alpha (Carhart, 1997; Fama and French, 2015) of between 1.056% and 1.167% per month.[47]

Meanwhile, I check the overall robustness of applying classification model [1] to multiple training samples. Instead of randomly choosing negative examples of spam reviews that do not deviate from the average scores by +/−0.1, I choose several such ranges: reviews that do not deviate from the average 1) by +/−0.01, 2) by +/−0.05, or 3) by +/−0.5; I then include all the reviews. Overall, separate training samples based on randomly chosen negative examples as described above show consistent predictability except for the sample, where all the reviews that are not near-duplicate reviews are used as negative examples.[48] I find abnormal returns comparable to those reported in Tables 6 and 9 in the various training samples.[49]

Finally, I investigate whether there are differences between spam reviews with high scores

[46] See Table A2.
[47] See Table A3 for 6-factor model estimates.
[48] See Figure A4 for AUCs from each classification model estimation
[49] See Table A4.

(Positive spam) and those with low scores (Negative spam). Positive spam is designed to conceal unfavorable opinions or to proactively improve average scores. Meanwhile, negative spam that harms a given product's score presumably is posted by competitors. In light of the negative correlation between abnormal scores and spamicity that I report in Table 5, I expect a large portion of spam reviews to be positive. This potentially reflects the fact that producing spam reviews is costly. We see evidence of this also in Figure 3, which indicates that reviews with higher scores have higher average spamicity throughout the entire sample period. A firm having limited resources with which to engage in spamming may find it preferable to write positive reviews for one's own goods than to write negative reviews about multiple competitors' goods.

Consistent with this hypothesis, I find evidence that the significant abnormal returns reported in Table 9 are concentrated among firms with high positive spamicity (spamicity in reviews with scores of 4 or above) *and* low negative spamicity (spamicity in reviews with scores of 2 or below) rather than among those with low positive spamicity *and* high negative spamicity.[50] This suggests that the return predictability of spamicity indicated by the results reported in Tables 7 and 8 is associated with firms' concealing unfavorable opinions, suggesting that spamming conveys information mostly about the negative quality perceptions of consumers.

Interestingly, I also find evidence that negative spamicity contains information about abnormal returns. When a firm's competing products receive high abnormal scores, a spread portfolio that goes long on stocks with high abnormal scores and spamicity and short on those with low abnormal scores and spamicity yields a Fama-French-Carhart six-factor alpha of -1.457% per month, indicating that negative spamicity also contains some information that is relevant to investors.[51] My findings show that spam review type conveys information that goes beyond that

[50] See Panel A of Table A5.
[51] See Panel B of Table A5.

embedded in review scores.

6. Conclusion

Consumer opinions play a crucial role in the ever-growing world of e-commerce. They reduce information asymmetry between firms and consumers by providing indirect purchase experience. To the extent that consumer opinions influence purchase decisions, they are valuable for firms and their investors. When the wisdom of the crowd ensures that quality information is correctly represented in the product market, there is little cost in processing information from that market. Thus, information transmitted from the product market will be fully recognized by the capital market, assuming the usual conditions for capital market efficiency.

There is also, however, a motivation for firms to intervene in this process when consumer opinions are unfavorable. Such motives incentivize firms to introduce noise through various methods, including misleading advertising, fake news, and review spamming. Because it is difficult to detect these activities and eliminate them from the product market, capital market efficiency will be attained only if sophisticated investors with superior information-processing capabilities reveal this information via their activities. The cost of processing big data that reflects multiple dimensions is, however, non-trivial. Thus, there remains significant investment value in consumer opinions that are accompanied by noise.

In this paper, I first show the prevalence of spamming using supervised machine learning. The non-negligible existence of spam reviews alone strongly indicates that the cost of spam detection is very high. Then, I show that there is significant value in consumer opinions, possibly resulting from noise such as spamming. Even where there are sophisticated investors in the product market, as long as spam detection is costly the transmission of information from that market into

the capital market may not be fully efficient. A potentially important direction that future research might take involves seeking ways to internalize this negative externality from spamming in both the product market and the capital market, which may yield valuable policy implications.

References

Brennan, Michael J., Tarun Chordia, and Avanidhar Subrahmanyam, 1998, "Alternative factor specifications, security characteristics, and the cross-section of expected stock returns", *Journal of Financial Economics* 49 (3), 345-373.

Bickart, Barbara and Robert M. Schindler, 2001, "Internet forums as influential sources of consumer information", *Journal of Interactive Marketing* 15 (3), 31-40.

Böhme, Rainer and Thorston Holz, "The Effect of Stock Spam on Financial Markets", *working paper*, Available at: https://ssrn.com/abstract=897431

Archak, Nikolay, Anindya Ghose, Panagiotis G. Ipeirotis, 2011, "Deriving the Pricing Power of Product Features by Mining Consumer Reviews", *Management Science* 57 (8), 1485-1509.

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

Chen, Yubo and Jinhong Xie, 2008, "Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix", *Management Science* 54 (3), 477-491.

Chevalier, Judith A. and Dina Mayzlin, 2006, "The Effect of Word of Mouth on Sales: Online Book Reviews", *Journal of Marketing Research* 43 (3), 345-354.

Clarke, Jonathan, Hailiang Chen, Ding Du, and Yu Jeffrey Hu, 2018, "Fake News, Investor Attention, and Market Reaction", *working paper*, Available at: https://ssrn.com/abstract=3213024

Cohen, Lauren and Andrea Frazzini, 2008, "Economic Links and Predictable Returns", *Journal of Finance* 63 (4), 1977-2011.

Chordia, Tarun, Avanidhar Subrahmanyam, V. Ravi Anshuman, 2001, "Trading activity and expected stock returns", *Journal of Financial Economics* 59 (1), 3-32.

DellaVigna, Stefano and Joshua M. Pollet, 2009, "Investor Inattention and Friday Earnings Announcements", *Journal of Finance* 64 (2), 709-749.

Diamond, Douglas W. and Robert E. Verrecchia, 1981, "Information aggregation in a noisy rational expectations economy", *Journal of Financial Economics* 9 (3), 221-235.

Dugast, Jérôme and Thierry Foucault, 2018, "Data abundance and asset price informativeness", *Journal of Financial Economics* 130 (2), 367-391.

Fama, Eugene F. and Kenneth R. French, 1993, "Common risk factors in the returns on stocks and bonds", *Journal of Financial Economics* 33 (1), 3-56.

Fama, Eugene F. and Kenneth R. French, 2015, "A five-factor asset pricing model", *Journal of Financial Economics* 116 (1), 1-22.

Ferreira, Miguel A. and Pedro Matos, 2008, "The colors of investors' money: The role of institutional investors around the world", *Journal of Financial Economics* 88 (3), 499-533.

Fornell, Claes, Forrest V. Morgeson III, and G. Tomas M. Hult, "Stock Returns on Customer Satisfaction Do Beat the Market: Gauging the Effect of a Marketing Intangible", *Journal of Marketing* 80 (5), 92-107.

Froot, Kenneth, Namho Kang, Gideon Ozik, and Ronnie Sadka, 2017, "What do measures of real-time corporate sales say about earnings surprises and post-announcement returns?", *Journal of Financial Economics* 125 (1), 143-162.

Godes, David and Dina Mayzlin, 2004, "Using Online Conversations to Study Word-of-Mouth Communication", Marketing Science 23 (4), 545-560.

Grossman, Sanford J. and Joseph Stiglitz, 1980, "On the Impossibility of Informationally Efficient Markets", *American Economic Review* 70 (3), 393-408.

Grullon, Gustavo, George Kanatas, and James P. Weston, 2004, "Advertising, Breadth of

Ownership, and Liquidity", *Review of Financial Studies* 17 (2), 439-461.
Hamilton, William L., Kevin Clark, Jure Leskovec, and Dan Jurafsky, 2016, "Inducing Domain-Specific Sentiment Lexicons from Unlabeled Corpora", *Conference on Empirical Methods in Natural Language Processing*, 595–605.
Hirshleifer, David and Siew Hong Teoh, 2003, "Limited attention, information disclosure, and financial reporting", *Journal of Accounting and Economics* 36 (1-3), 337-386.
Hoberg, Gerard and Gordon M. Philips, 2018, "Text-Based Industry Momentum", *Journal of Financial and Quantitative Analysis, forthcoming*.
Hong, Harrison, Terence Lim, and Jeremy C. Stein, 2000, "Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies", *Journal of Finance* 55 (1), 265-295.
Hong, Harrison and Jeremy C. Stein, 2002, "A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets", *Journal of Finance* 54 (6), 2143-2184.
Hou, Kewei and Tobias J. Moskowitz, 2005, "Market Frictions, Price Delay, and the Cross-Section of Expected Returns", *Review of Financial Studies* 18 (3), 981-1020.
Huang, Jiekun, 2018a, "The customer knows best: The investment value of consumer opinions", *Journal of Financial Economics* 128 (1), 164-182.
Huang, Ruidi, 2018b, "The Financial Consequences of Customer Satisfaction: Evidence from Yelp Ratings and SBA Loans", *working paper*.
Ittner, Christopher D. and David F. Larcker, 1998, "Are Nonfinancial Measures Leading Indicators of Financial Performance? An Analysis of Customer Satisfaction", *Journal of Accounting Research* 36 (1), 1-35.
Jindahl, Nitin and Bing Liu, 2007, "Review Spam Detection", *Proceedings of the 16th International Conference on World Wide Web*, 1189-1190.
Jindahl, Nitin and Bing Liu, 2008, "Opinion Spam and Analysis", *Proceedings of the 2008 International Conference on Web Search and Data Mining*, 219-230.
Karolyi, G. Andrew, Dawoon Kim, and Rose Liao, 2018, "The Theory and Practice of Investor Relations: A Global Perspective", *working paper*.
Kogan, Shimon, Tobias J. Moskowitz, and Marina Niessner, "Fake News: Evidence from Financial Markets", *working paper*, Available at: https://ssrn.com/abstract=3237763
Kyle, Albert S., "Continuous Auctions and Insider Trading", *Econometrica* 53 (6), 1315-1335.
Larkin, Yelena, 2013, "Brand perception, cash flow stability, and financial policy", *Journal of Financial Economics* 110 (1), 232-253.
Lee, Charles M. C., Paul Ma, Charles C. Y. Wang, 2015, "Search-based peer firms: Aggregating investor perceptions through internet co-searches", *Journal of Financial Economics* 116 (2), 410-431.
Li, Xinxin and Lorin M. Hitt, 2008, "Self-Selection and Information Role of Online Product Reviews", *Information Systems Research* 19 (4), 456-474.
Lim, Ee-Peng, Viet-An Nguyen, Nitin Jindal, Bing Liu, and Hady W. Lauw, "Detecting Product Review Spammers using Rating Behaviors", *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, 939-948.
Liu, Yong, 2006, "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue", *Journal of Marketing* 70 (3), 74-89.
Lou, Dong, 2014, "Attracting Investor Attention through Advertising", Review of Financial Studies 27 (6), 1797-1829.
Luca, Michael, 2011, "Reviews, reputation, and revenue: The case of Yelp.Com", *Harvard*

Business School NOM Unit Working Paper 12-016, 1-40.
Luca, Michael and Georgios Zervas, 2016, "Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud", *Management Science* 62 (12), 3412-3427.
Mayzlin, Dina, Yaniv Dover, and Judith Chevalier, 2014, "Promotional Reviews: An Empirical Investigation of Online Review Manipulation", *American Economic Review* 104 (8), 2421-2455.
McAuley, Julian, Christopher Targett, Qinfeng Shi, and Anton van den Hengel, 2015, "Image-Based Recommendations on Styles and Substitutes", *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 43-52.
McAuley, Julian, Rahul Pandey, and Jure Leskovec, 2015, "Inferring Networks of Substitutable and Complementary Products", *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794.
McAuley, Julian and Alex Yang, 2016, "Addressing Complex and Subjective Product-Related Queries with Customer Reviews", *Proceedings of the 25th International Conference on World Wide Web*, 625-635.
Merton, Robert C., 1987, "A Simple Model of Capital Market Equilibrium with Incomplete Information", *Journal of Finance* 42 (3), 483-510.
Milgrom, Paul and John Roberts, 1986, "Price and Advertising Signals of Product Quality", *Journal of Political Economy* 94 (4), 796-821.
Mukherjee, Arjun, Bing Liu, and Natalie Glance, 2012, "Spotting Fake Reviewer Groups in Consumer Reviews", *Proceedings of the 21st international conference on World Wide Web*, 191-200.
Mukherjee, Arjun, Abhinav Kumar, Bing Liu, Junhui Wang, Meichun Hsu, Malu Castellanos, Riddhiman Ghosh, 2013, *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 632-640.
Novy-Marx, Robert, 2013, "The other side of value: The gross profitability premium", *Journal of Financial Economics* 108 (1), 1-28.
Ott, Myle, Yejin Choi, Claire Cardie, and Jeffrey T. Hancock, 2011, "Finding Deceptive Opinion Spam by Any Stretch of the Imagination", *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, 309-319.
Peng, Lin and Wei Xiong, 2006, "Investor attention, overconfidence and category learning", *Journal of Financial Economics* 80 (3), 563-602.
Piotroski, Joseph D., 2000, "Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers", *Journal of Accounting Research* 38. 1-41.
Verrecchia, Robert E., 1982, "Information Acquisition in a Noisy Rational Expectations Economy", *Econometrica* 50 (6), 1415-1430.
Xie, Sihong, Guan Wang, Shuyang, and Philip S. Yu, 2012, "Review spam detection via time series pattern discovery", *Proceedings of the 21st International Conference on World Wide Web*, 635-636.
Zhu, Christina, 2018, "Big Data as a Governance Mechanism", *working paper*.

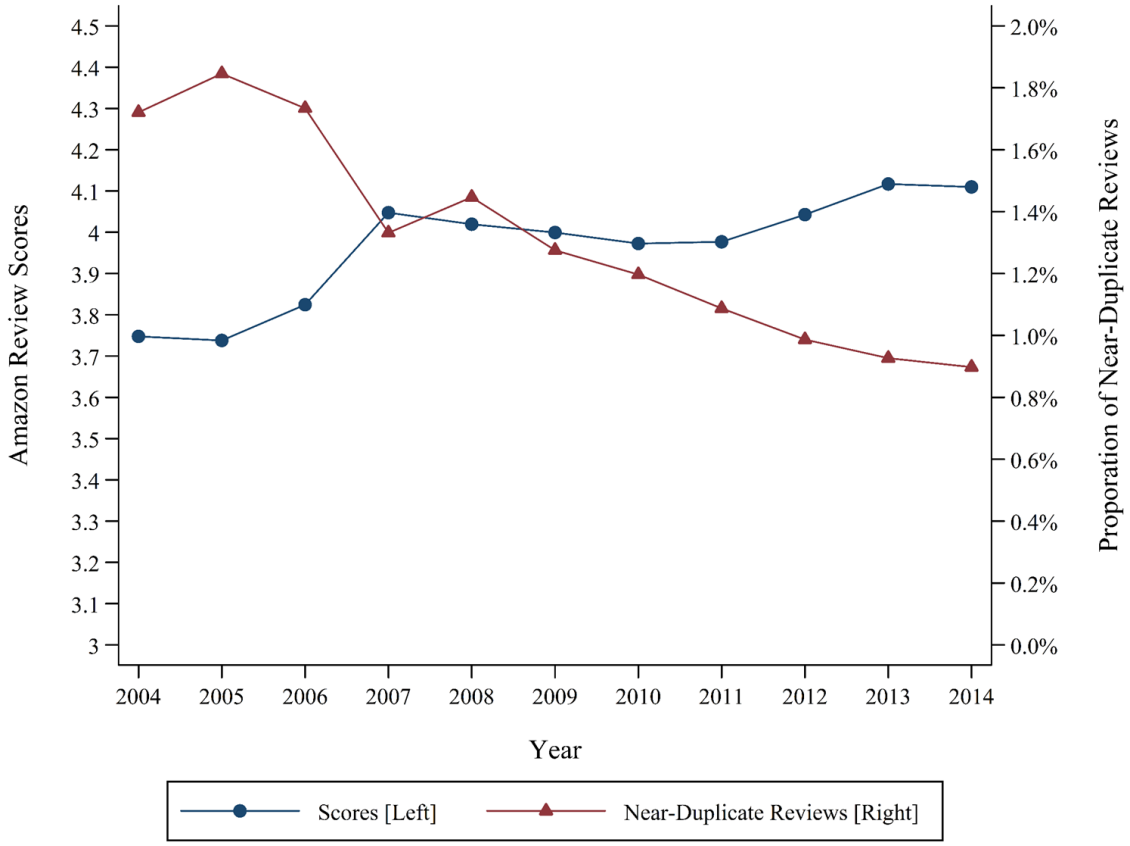


Figure 1. Average Review Scores and Near-Duplicate Reviews. This figure depicts the average Amazon review scores and the proportion of near-duplicate reviews from 2004 through 2014. *Near-Duplicate Review* indicates reviews with nearly identical review pairs in the sample (with similarity scores above 0.9 using 2-gram Jaccard distance) posted by distinct reviewers and/or of distinct products. *Proportion of Near-Duplicate Reviews* indicates the number of near-duplicate reviews divided by the total number of reviews in each year.

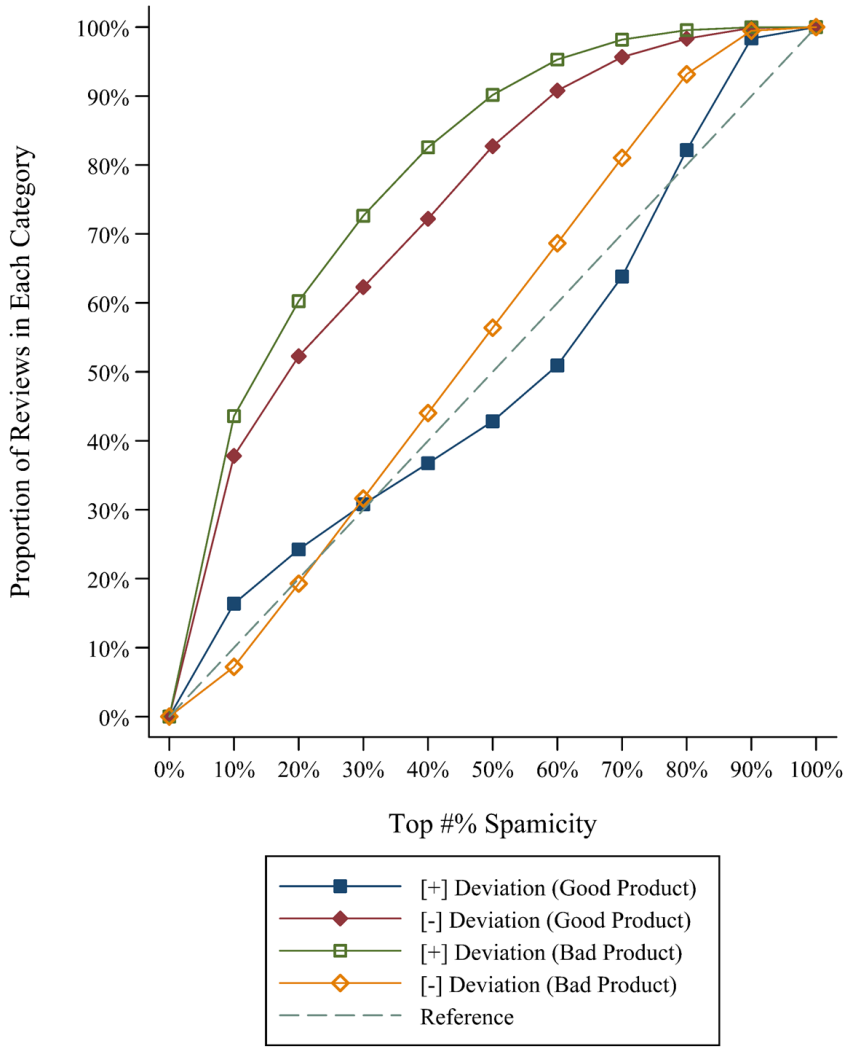


Figure 2. Lift Curves. This figure depicts the lift curves of the out-of-the-training-sample reviews that are not used to estimate spamicity in model [1]. The-out-of-the-training-sample reviews are first randomly divided into ten groups. *Top #% Spamicity* indicates reviews that are among the top #% in spamicity in each group. *Proportion of Reviews in Each Category* indicates the proportion of reviews in each group that fall into one of the following categories: [+] *Deviation (Good Product)*, [–] *Deviation (Good Product)*, [+] *Deviation (Bad Product)*, and [–] *Deviation (Bad Product)*. [+]/[–] *Deviation* indicates reviews that receive scores higher/lower than the average score when each review was posted. *Good/(Bad) Product* indicates products that receive an average score of 4 or higher/(2 or lower) when each review was posted.

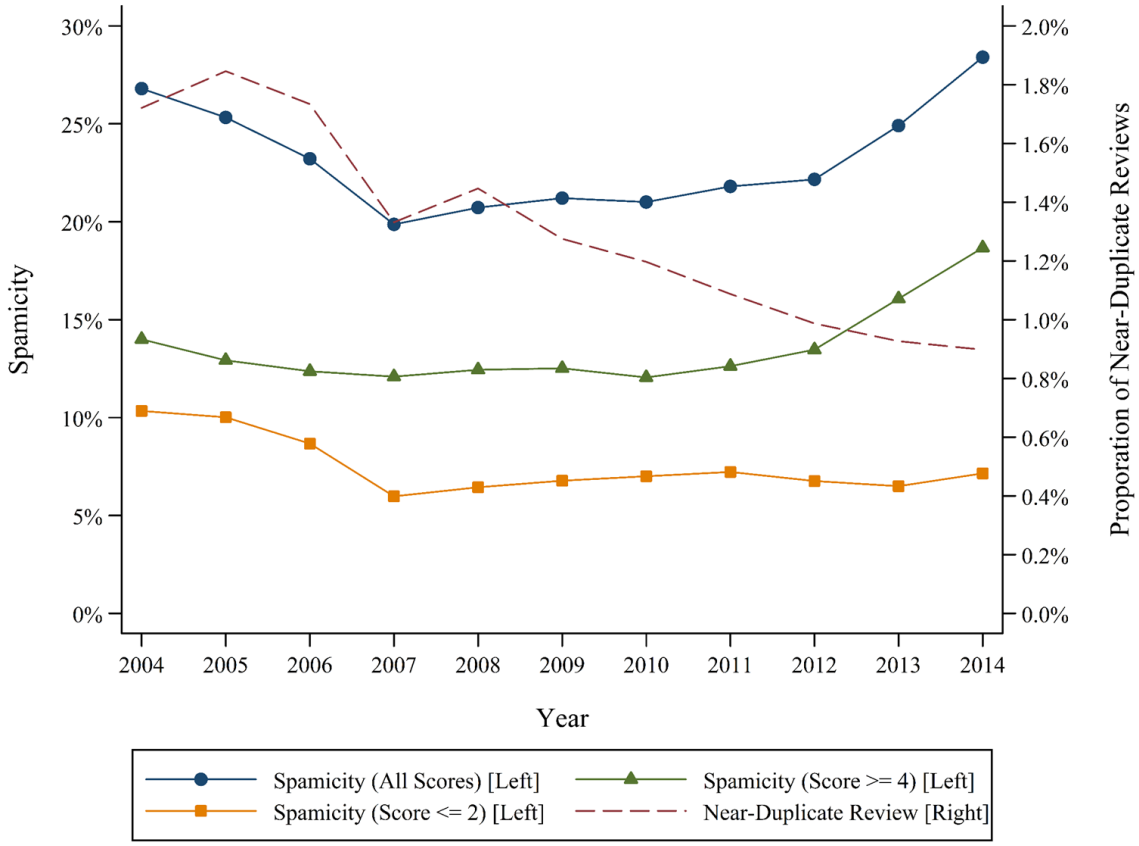


Figure 3. **Average Spamicity across Time.** This figure displays the estimated average spamicity (Spam Probability) of reviews from 2004 through 2014. Spamicity is estimated using *Near-Duplicate Reviews* as positive examples in the training sample. *All Scores* indicates average spamicity among all reviews. *Score >= 4* indicates average spamicity among reviews receiving review scores of 4 or higher. *Score <= 2* indicates average spamicity among reviews receiving review scores of 2 or lower.

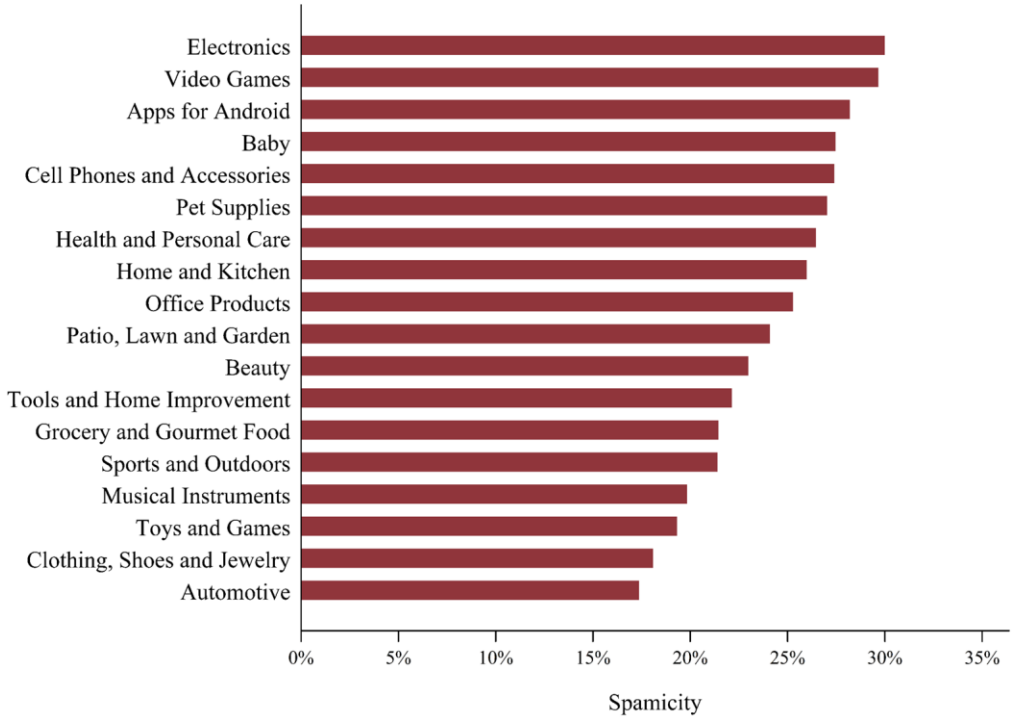


Figure 4. **Average Spamicity across Product Categories.** This figure shows the estimated average spamicity (Spam Probability) of sample reviews from 2004 through 2014 across Amazon product categories. Spamicity is estimated using *Near-Duplicate Reviews* as positive examples in the training sample.

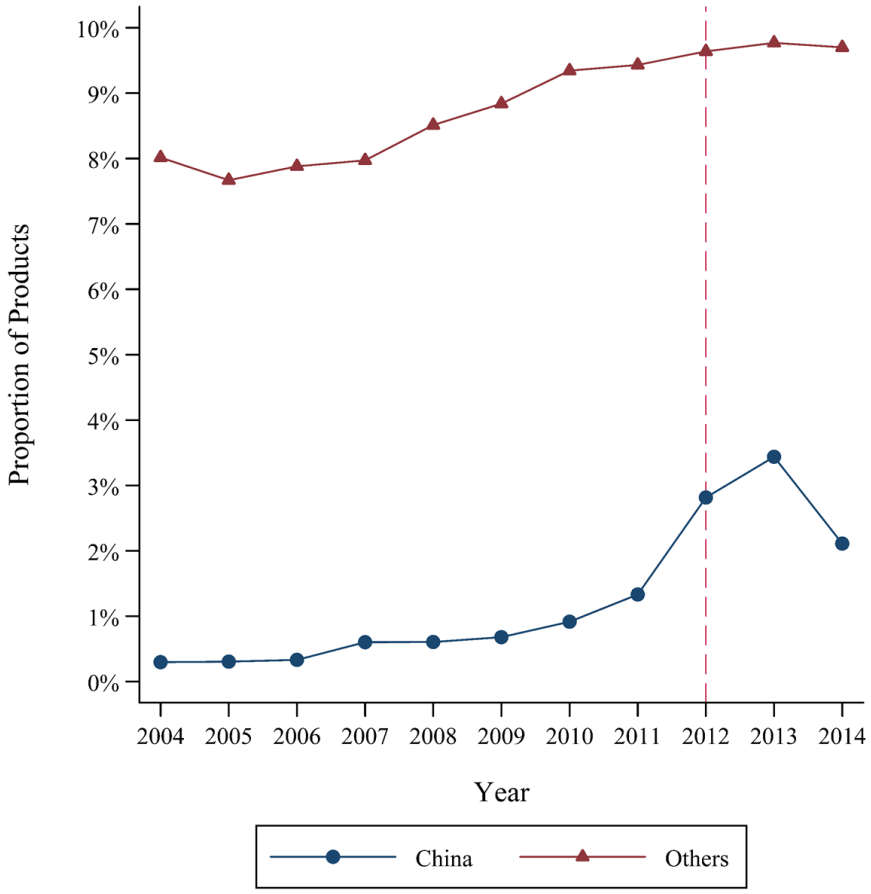


Figure 5. Proportion of Foreign Products on Amazon. This figure shows the proportion of new products available on Amazon that are barcoded outside of the U.S. *China* indicates the ratio of new products on Amazon that are barcoded in China to new products on Amazon that are barcoded in the U.S. as percentages. *Others* indicates new products on Amazon that are barcoded in countries other than China and the U.S. over new products on Amazon that are barcoded in the U.S. as percentages. The vertical dotted line indicates when the "Amazon Global Selling Program" was launched.

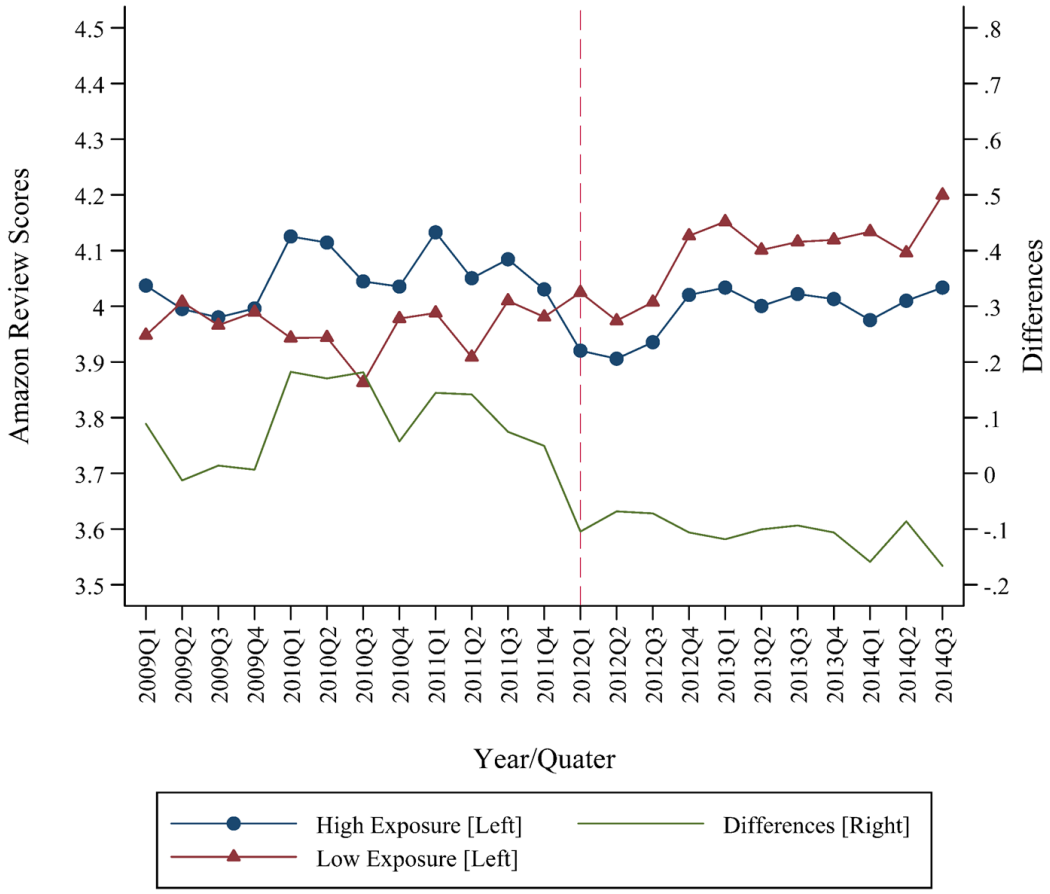


Figure 6. Average Review Scores around the Influx of Foreign Products. This figure shows average review scores across product categories with high and low exposure from the influx of products that are barcoded in China. *High (Low) exposure* indicates product categories with above- (below-)median numbers of new products barcoded in China after 2012. *Differences* indicates differences in scores between the *High* and *Low Exposure* groups. The vertical dotted line indicates when the "Global Selling Program" was launched.

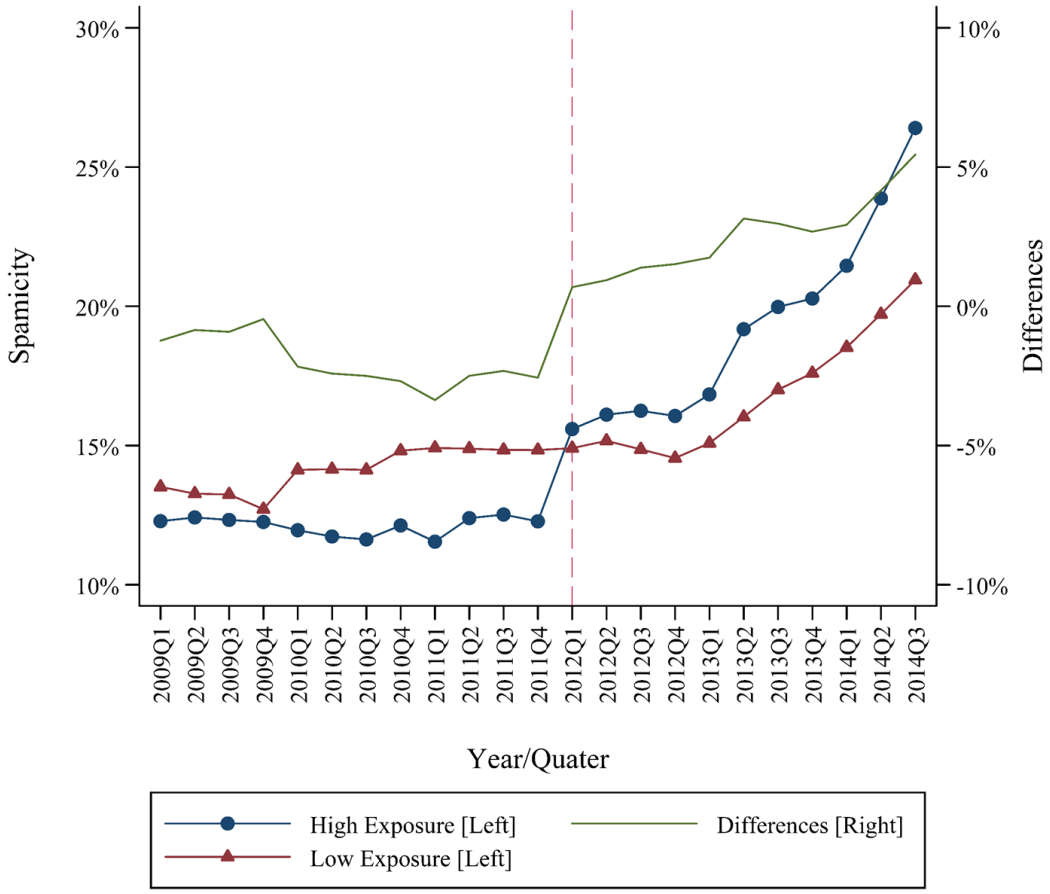


Figure 7. Average Spamicity around the Influx of Foreign Products. This figure shows average spamicity for firms whose products are sold in product categories with high and low exposure from the influx of products that are barcoded in China. *High (Low) exposure* indicates product categories with above- (below-)median numbers of new products barcoded in China after 2012. *Differences* indicates differences in spamicity between products in the *High* and *Low Exposure* groups. The vertical dotted line indicates when the "Global Selling Program" was launched.

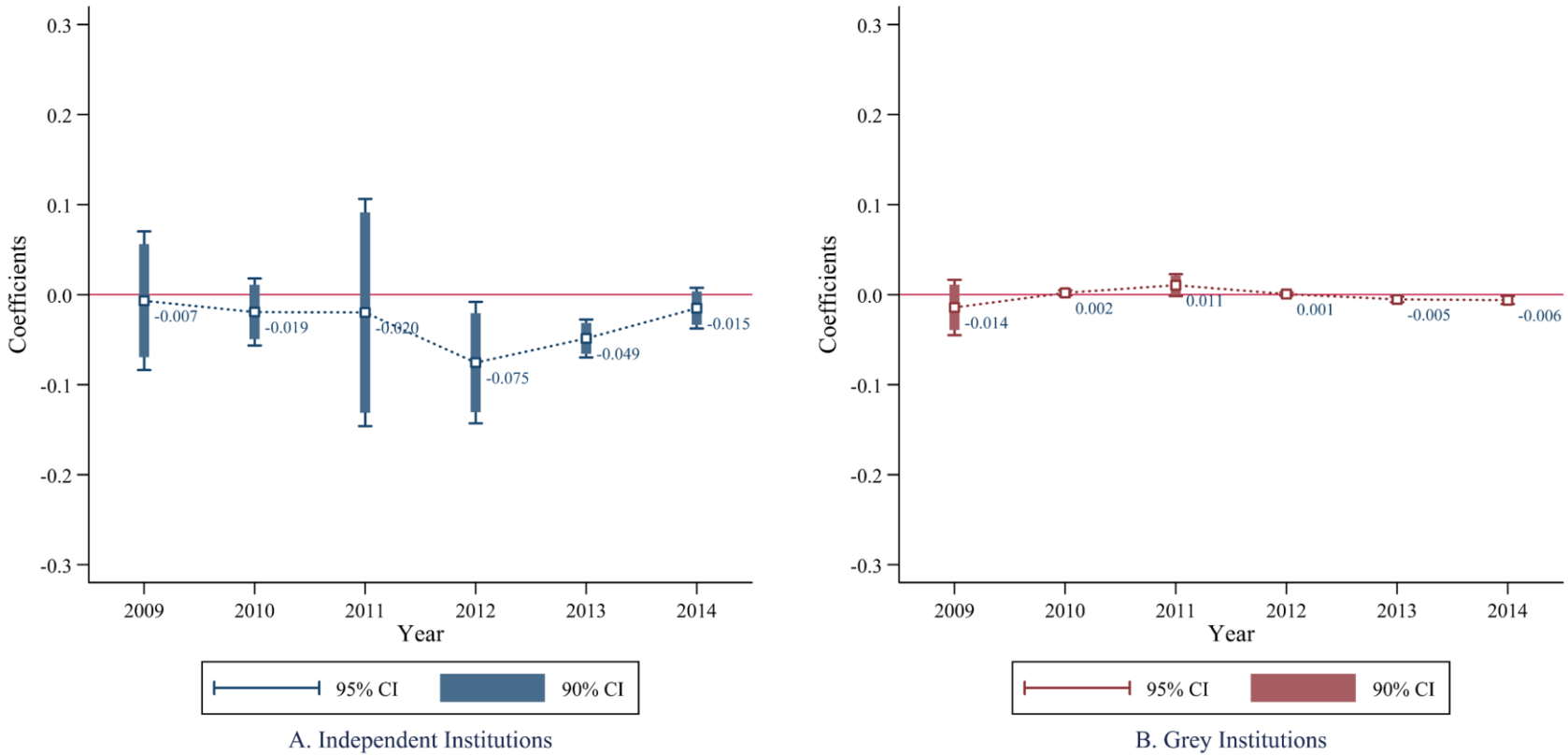


Figure 8. Proportion of Foreign Products available on Amazon. This figure shows year-by-year coefficient estimates of model [5]. *Coefficients* indicates coefficient estimates of β_7 in model [5]. The dependent variable in model [5] in A (B) is independent (grey) institutional ownership as defined by Ferreira and Matos (2008). "95% CI" and "90% CI" indicate 95% and 90% confidence intervals, respectively. The "Amazon Global Selling Program" was launched in 2012.

Table 1. Sample Products and Reviews

In this table I report review characteristics by product category and year from 2004 through 2014. *Full Review sample* indicates the review sample used to estimate model [1] (44,699,245 reviews). *Main Sample* indicates products produced by Compustat firms included in the main analyses (3,827,524 reviews). *Category* indicates product category as defined by Amazon. # of Reviews equals the number of observations. Identical reviews by the same reviewer of the same products are excluded from the sample.

Panel A. By Year

Year	*Full Review sample*				*Main sample* (Compustat)				
	# of Reviews	Avg. Score	# of Products	# of Reviewers	# of Reviews	Avg. Score	# of Products	# of Reviewers	# of Firms
2004	113,515	3.75	30,356	82,974	30,964	3.70	6,158	25,323	158
2005	199,665	3.74	51,464	145,495	46,270	3.70	8,386	38,244	183
2006	281,739	3.82	81,232	208,063	58,356	3.81	10,614	49,831	209
2007	661,623	4.05	187,900	441,847	124,255	4.08	18,360	103,436	246
2008	875,679	4.02	254,261	561,313	148,292	4.01	23,091	122,467	254
2009	1,232,291	4.00	365,313	766,724	183,910	4.03	30,594	151,841	267
2010	1,921,431	3.97	558,099	1,142,758	238,491	3.98	40,385	195,472	276
2011	3,518,431	3.98	896,057	1,907,186	356,767	3.95	56,646	290,843	285
2012	6,795,520	4.04	1,463,547	3,249,284	552,374	4.04	78,927	438,104	294
2013	16,820,953	4.12	2,662,523	6,272,291	1,250,592	4.16	129,714	908,157	301
2014	12,278,398	4.11	2,295,506	4,975,458	837,253	4.16	109,029	631,109	301

Panel B. By Item Category

Category	*Full Review sample*				*Main sample* (Compustat)				
	# of Reviews	Avg. Score	# of Products	# of Reviewers	# of Reviews	Avg. Score	# of Products	# of Reviewers	# of Firms
Apps for Android	2,636,018	4.00	61,249	1,323,483	147,669	4.00	283	132,626	55
Automotive	1,369,462	4.19	319,670	849,811	54,934	4.27	13,051	45,170	115
Baby	902,334	4.12	64,308	524,599	86,906	4.09	4,140	75,318	56
Beauty	2,017,209	4.15	249,067	1,208,262	93,452	4.13	10,447	74,264	108
Cell Phones and Accessories	3,429,238	3.81	319,147	2,254,397	192,856	3.73	11,152	164,575	114
Clothing, Shoes and Jewelry	5,740,430	4.14	1,135,315	3,114,168	398,248	4.29	61,838	342,086	153
Electronics	7,653,271	4.01	470,933	4,146,052	1,320,925	4.10	42,428	1,032,383	188
Grocery and Gourmet Food	1,293,591	4.26	165,815	767,510	39,021	4.24	4,817	33,095	85
Health and Personal Care	2,966,935	4.11	251,883	1,845,694	149,240	4.17	8,521	120,523	164
Home and Kitchen	4,208,457	4.10	408,679	2,493,623	178,924	4.03	11,564	157,939	172
Musical Instruments	487,951	4.24	81,142	333,503	9,172	4.28	1,063	8,857	53
Office Products	1,222,650	3.98	129,238	897,129	279,192	3.90	12,384	214,625	141
Patio, Lawn and Garden	984,337	4.01	105,599	709,950	53,898	4.10	3,661	49,276	81
Pet Supplies	1,229,525	4.11	103,178	738,883	14,265	4.09	1,092	12,795	64
Sports and Outdoors	3,249,098	4.18	478,105	1,983,325	84,020	4.15	11,215	78,849	174
Tools and Home Improvement	1,901,547	4.13	259,417	1,202,443	116,605	4.17	10,647	101,206	154
Toys and Games	2,220,341	4.15	326,235	1,325,509	241,049	4.12	33,251	179,368	135
Video Games	1,186,851	3.98	48,738	764,597	367,148	3.98	9,000	277,179	53

Table 2. Near-Duplicate Review Examples

In this table I report examples of near-duplicate reviews used as positive examples in model [1]. Near-duplicate reviews are defined as reviews that receive another review where the textual contents of the two reviews receive above 0.90 similarity scores calculated using 2-gram Jaccard distance. The training sample includes 435,456 near-duplicate reviews. *ReviewerID* indicates each review's ID and *ASIN* indicates the product identifier. *Helpful* [*A*, *B*] indicates how many, out of *B* consumer feedback responses to a given review, *A* consumers found the review helpful.

ReviewerID†	ASIN†	Helpful	Review	Date	Score
The Same Reviewer on Different Products					
A328AQ********	B004G*****	[1, 1]	I swear I never liked any of my shoes like this one - it feels so comfortable and delivers an excellent fashion statement. Great great buy. I used to be a loyal Nike fan, but this one made me change my party. My next shoe purchase, **very high chances will be another Puma**	3/20/2012	4
A328AQ********	B002F*****	[0, 0]	I swear I never liked any of my shoes like this one - it feels so comfortable and delivers an excellent fashion statement. Great great buy. I used to be a loyal Nike fan, but this one made me change my party. My next shoe purchase, **very high chances will be another adidas**	3/20/2012	4
Different Reviewers on the Same Product					
A1MS73********	B0047*****	[0, 0]	As other buyer says this is a great case, the materials are good quality and fits perfectly on the phone. The pen is very useful, works very well on the phone, maybe other buyers hasn't used it properly so they say that is useless. **Very recommended!**	5/31/2011	5
A3M6MI********	B0047*****	[0, 0]	As other buyer says this is a great case, the materials are good quality and fits perfectly on the phone. The pen is very useful, works very well on the phone, maybe other buyers hasn't used it properly so they say that is useless.	9/3/2011	5
Different Reviewers on Different products					
A1O7CE********	B00F2*****	[19, 25]	After using different weight-loss diets/plans with poor results in the past. I'm glad that I gave this a try. It is truly a 5 star review. I only need to take it twice in a day 30 minutes before a meal. I not only have lost 8 pounds already, but I have more energy than I had before.	3/26/2014	5
A3TFFM********	B00HD*****	[3, 3]	After using different weight-loss diets/plans with poor results in the past. I'm glad that I gave this a try. It is truly a 5 star review. I only need to take it twice in a day 30 minutes before a meal. I not only have lost 8 pounds already, but I have more energy than I had before.	6/5/2014	5
AL9A8W********	B00FS*****	[0, 0]	After using different weight-loss diets/plans with poor results in the past. I'm glad that i gave this a try. It is truly **worth** a 5 star review. I only need to take it twice in a day 30 minutes before a meal. I not only have lost 8 pounds already, but I have more energy than I had before.	4/16/2014	5
A2GT59********	B00G1*****	[0, 1]	After using different weight-loss diets/plans with poor results in the past. I'm glad that I gave this a try. It is truly a 5 star review. I only need to take it twice in a day 30 minutes before a meal. I not only have lost 8 pounds already, but I have more energy than I had before.	7/2/2014	5

† Not fully disclosed (concealed by asterisk)

Table 3. Spam Classification

In this table I report the results of one of the ten logistic regressions carried out to estimate model [1] using ten-fold cross-validation. The dependent variable is *Near-Duplicate Review Dummy*, which takes the value of 1 for near-duplicate reviews and 0 otherwise. *Feedback* indicates feedback given in response to a review. *Helpful Feedback* indicates feedback that labels a review as helpful. *Title* indicates the title of a review. *Position* indicates the relative order of a review for a given product. *First (Only) Review Dummy* takes the value of 1 if a review is the first (only) review for a product and 0 otherwise. *Sent.* indicates aggregate individual review-level sentiment measured by sentiment lexicons as introduced by Hamilton et al. (2016). *Positive (Negative) Sent. Score* indicates the sum of the sentiment scores for words with positive (negative) sentiment. *% of Positive (Negative) Sent.* indicates the number of words with positive (negative) sentiment over the total number of words for a given review. *Score Dummy* takes the value of 1 if the score for a review is 4 or higher, -1 if the score is 2 or lower, and 0 otherwise. *Good and Bad Scores* takes the value of 1 if a reviewer has given both good scores (>= 4) and bad scores (<= 2) and 0 otherwise. *Good and Avg. Scores* takes the value of 1 if a reviewer has given both good scores (>= 4) and the middle score (= 3) and 0 otherwise. *Bad and Avg. Scores* takes the value of 1 if a reviewer has given both bad scores (<= 2) and the middle score (= 3) and 0 otherwise. *All Scores* takes the value of 1 if the reviewer has given scores of every kind and 0 otherwise. *% of Negative (Positive) Dev.* indicates the number of reviews that receive scores lower (higher) than the average scores when the review was posted divided by the total number of reviews posted by a reviewer. Standard errors are clustered at each product level. Numbers in the parentheses indicate t-statistics.

Dependent Variable: *Near-Duplicate Review Dummy*			
Review-centric Features			
Logged # of Feedback	-0.113***	Review Score	1.503***
	(-3.63)		(7.34)
Logged # of Helpful Feedback	0.287***	Score Dummy	-3.153***
	(11.13)		(-7.99)
% of Helpful Feedback	-0.236***	*Reviewer-centric Features*	
	(-7.72)	Percent of First Reviews	3.743***
Logged Length of the Review [Title]	0.148***		(59.14)
	(3.71)	Percent of Only Reviews	0.346***
Logged Length of the Review	-0.215***		(7.24)
	(-2.64)	Avg. Score [Reviewer]	0.102***
Logged Position of the Review [Ascending]	0.501***		(4.55)
	(15.60)	Stddev. of Scores [Reviewer]	-0.236***
Logged Position of the Review [Descending]	-0.075***		(-10.68)
	(-8.57)	Score Dummy [Reviewer]	-0.162***
First Review Dummy	-2.213***		(-7.34)
	(-19.94)	Good and Bad Scores	0.122***
Only Review Dummy	1.075***		(3.12)
	(5.65)	Good and Avg. Scores	-0.161***
% of Positive Sent.	0.258		(-6.63)
	(1.48)	Bad and Avg. Scores	-0.266***
Logged Positive Sent. Score	0.341***		(-2.75)
	(3.80)	All Scores	0.018
% of Negative Sent.	0.585*		(0.50)
	(1.93)	% of Negative Dev.	4.282***
Logged Negative Sent. Score	0.107***		(42.90)
	(7.75)	% of Positive Dev.	4.413***
Logged Positive Sent. Score [Title]	-0.103**		(44.80)
	(-2.57)	*Product-centric Reviews*	
Logged Negative Sent. Score [Title]	0.028***	Avg. Score [Product]	0.319***
	(5.05)		(4.84)
% of Numerals	4.855***	Stddev. of Scores [Product]	0.923***
	(8.17)		(6.15)
% of Capital Letters	0.358***	Logged # of Product Reviews	-0.119***
	(4.20)		(-8.27)
% of Comparatives	-5.955**	Logged Stddev. of # of Reviews	-0.177***
	(-2.57)		(-5.11)
Pseudo R^2		0.237	
N		794,068	

*** p<0.01, ** p<0.05, * p<0.1

Table 4. Summary Statistics

In this table I report firm-level summary statistics for firms in the main sample. *Avg. Score* indicates average monthly review scores. *Abnormal Score* indicates average monthly review scores minus the previous-12-month average review scores, following Huang (2018a). *Avg. Peer Score* indicates average monthly review scores for products that are marked "Also Viewed" by Amazon, excluding products manufactured by the same firm. *Abnormal Peer Score* indicates average peer scores minus the previous 12-month average peer score. *Spamicity* indicates average firm-level spamicity of new reviews in a given month estimated with model [1]. *High (Low) Spamicity* indicates firms with above (below) median spamicity. *Stock Return$_{m\,(q)}$* indicates monthly (quarterly) stock returns. Advertising indicates advertising expenses divided by total sales. *Gross Profitability* indicates the ratio of income before extraordinary items to the book value of assets. *F-score* indicates financial performance as signaled by Piotroski (2000). *Dollar Volume* is the dollar trading volume during the second-to-last month of the previous 12 months, following Huang (2018a). *CV of Dollar Volume* is the coefficient of variations in dollar trading volume calculated over the previous 12 months, beginning in the second-to-last month, following Huang (2018a). *GOV* indicates quarterly averages on the MSCI governance index. *ESG* indicates quarterly averages on the MSCI ESG index. *Independent (Grey)* indicates independent (grey) institutions as defined by Ferreira and Matos (2008). *SUE* indicates the difference between reported quarterly EPS and the median of the most recent EPS forecasts by all analysts prior to earnings announcements divided by stock prices. *Forecast Dispersion* is the standard deviation of the most recent EPS forecasts of all analysts prior to earnings announcements. Review and firm-level characteristics (imputed from quarterly observations) are monthly observations and governance indices; institutional ownership and analyst coverage are quarterly observations. All variables are winsorized at the 1% and 99% levels.

	All Sample				High Spamicity				Low Spamicity			
	N	Mean	Median	Stddev.	N	Mean	Median	Stddev.	N	Mean	Median	Stddev.
Review characteristics												
Avg. Score	18,081	4.02	4.10	0.52	9,041	3.88	3.96	0.53	9,040	4.16	4.22	0.46
Abnormal Score	18,081	0.11	0.03	0.78	9,041	-0.02	-0.01	0.57	9,040	0.24	0.08	0.93
Avg. Peer Score	18,081	4.16	4.21	0.40	9,041	4.10	4.16	0.39	9,040	4.22	4.27	0.41
Abnormal Peer Score	18,081	0.22	0.14	0.75	9,041	0.19	0.14	0.56	9,040	0.26	0.13	0.90
# of Products	18,081	180.6	29.0	584.8	9,041	278.3	48.0	766.6	9,040	82.9	21.0	277.8
# of Reviews	18,081	3,785	86	17,164	9,041	6,903	275	23,304	9,040	667	44	5,163
Spamicity	18,081	0.23	0.22	0.10	9,041	0.31	0.30	0.06	9,040	0.15	0.16	0.05
Firm-level characteristics												
Stock Return$_m$ (%)	18,081	1.14	1.18	9.42	9,041	1.02	1.10	9.51	9,040	1.25	1.24	9.33
Advertising	18,081	0.03	0.01	0.05	9,041	0.03	0.01	0.04	9,040	0.04	0.01	0.05
R&D/TA	18,081	0.04	0.02	0.07	9,041	0.06	0.03	0.08	9,040	0.02	0.01	0.05
Logged Market Cap.	18,081	7.97	7.94	2.31	9,041	8.06	8.02	2.43	9,040	7.88	7.90	2.18
M/B	18,081	2.06	1.69	1.36	9,041	2.04	1.66	1.28	9,040	2.00	1.69	1.30
Gross Profitability	18,081	0.04	0.06	0.20	9,041	0.03	0.06	0.22	9,040	0.05	0.07	0.18
F-Score	18,081	5.23	5.00	1.52	9,041	5.10	5.00	1.54	9,040	5.37	5.00	1.49
Logged Dollar Volume	18,081	14.83	15.10	2.73	9,041	14.93	15.20	2.86	9,040	14.73	14.98	2.58
CV of Dollar Volume	18,081	0.40	0.32	0.25	9,041	0.41	0.32	0.26	9,040	0.38	0.31	0.24
Governance Indices												
GOV	9,207	5.95	5.68	2.51	4,604	5.93	5.60	2.52	4,603	5.97	5.75	2.49
ESG	9,207	4.98	4.92	1.52	4,604	4.95	4.90	1.47	4,603	5.01	4.98	1.58
Institutional Ownership												
# of Institutions	6,475	457	230	535.5	3,238	481.2	237	570	3,237	429.9	220.5	492.5
Institutional Ownership	6,475	0.71	0.77	0.26	3,238	0.70	0.77	0.27	3,237	0.72	0.77	0.24

	All Sample				High Spamicity				Low Spamicity			
	N	Mean	Median	Stddev.	N	Mean	Median	Stddev.	N	Mean	Median	Stddev.
Independent Ownership	6,475	0.68	0.73	0.25	3,238	0.67	0.73	0.25	3,237	0.69	0.73	0.23
Grey Ownership	6,475	0.03	0.03	0.02	3,238	0.03	0.03	0.02	3,237	0.03	0.03	0.02
Analyst Coverage												
# of Analysts	6,465	18.1	14.5	15.7	3,233	19.9	16.0	17.8	3,232	16.3	13.0	12.9
SUE (%)	6,465	0.083	0.046	0.563	3,233	0.106	0.048	0.603	3,232	0.067	0.045	0.559
Forecast Dispersion	6,465	0.019	0.006	0.123	3,233	0.021	0.007	0.137	3,232	0.018	0.004	0.105

Table 5. Determinants of Spam

In this table I report panel and 2SLS regression results derived from model [2]. Below, results derived from Models (1)–(6), (8), and (9) are panel regression estimates while those derived from Model (8) are 2SLS estimates where the instrumental variable is *Average Peer Score* and the instrumented variable is *Abnormal Score.* First-stage estimates are presented in the appendix. Standard errors are double-clustered by firm and year-month. Numbers in parentheses indicate t-statistics.

Dependent Variable: Spamicity	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Abnormal Score	-0.014***	-0.009***	-0.024***	-0.012***		-0.013***	-0.006***	-0.019***	-0.018***
	(-14.15)	(-9.74)	(-15.59)	(-14.74)		(-8.52)	(-6.64)	(-6.62)	(-6.63)
Stddev. of Score		0.031***				0.027***	0.033***	0.025***	0.025***
		(13.24)				(12.93)	(16.44)	(7.08)	(7.22)
Abnormal Peer Score			0.012***			0.005***		0.012***	0.011***
			(6.83)			(3.83)		(4.44)	(4.33)
Lagged Dependent Variable				0.373***		0.355***	0.356***	0.310***	0.316***
				(18.46)		(17.57)	(17.76)	(10.05)	(10.13)
Advertising					0.012*	0.010**	0.010**	0.019**	0.020**
					(1.90)	(2.21)	(2.26)	(2.39)	(2.47)
R&D/TA					-0.063	-0.036	-0.036	0.017	0.009
					(-1.25)	(-1.13)	(-1.13)	(0.22)	(0.11)
Logged Market Cap.					0.009*	0.007**	0.007**	0.010	0.010
					(1.80)	(2.00)	(2.26)	(1.14)	(1.13)
M/B					-0.005**	-0.004**	-0.004***	-0.004	-0.004
					(-2.23)	(-2.59)	(-2.87)	(-1.14)	(-1.14)
Gross Profitability					-0.028**	-0.019**	-0.018**	-0.025*	-0.027*
					(-2.41)	(-2.44)	(-2.47)	(-1.71)	(-1.83)
F-Score					-0.000	0.000	0.000	0.000	0.000
					(-0.39)	(0.35)	(0.45)	(0.35)	(0.47)
Logged Dollar Volume					-0.001	-0.000	-0.000	-0.001	-0.001
					(-0.74)	(-0.56)	(-0.67)	(-0.65)	(-0.54)
CV of Dollar Volume					0.004	0.003	0.003	0.009	0.009
					(0.84)	(0.85)	(0.99)	(1.41)	(1.46)
Stock Return$_m$					-0.000	-0.000	-0.000	-0.000	0.000
					(-1.07)	(-0.73)	(-1.00)	(-0.00)	(0.08)
GOV								0.001*	
								(1.69)	
ESG									0.003**
									(2.01)
Adjusted R^2	0.740	0.749	0.753	0.779	0.733	0.798	0.434	0.804	0.803
Firm fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
Time fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
N	18,081	18,081	18,081	18,081	18,081	18,081	18,081	9,207	9,207

*** p<0.01, ** p<0.05, * p<0.1

Table 6. Calendar-time Score and Spamicity Portfolio Returns

In this table I report calendar-time tercile portfolio regression estimates. For each month, sample stocks are sorted into tercile portfolios based on abnormal scores (*Score*) or spamicity (*Spamicity*). The dependent variable in each regression is excess portfolio returns in the following month. Following Huang (2018a), stocks in each portfolio are weighted equally or weighted by the number of reviews posted in each month. *Alpha* indicates abnormal portfolio returns after considering a Fama-French-Carhart four-factor model. *Long/Short* indicates a spread portfolio that buys the top tercile portfolio and sells the bottom tercile portfolio for Panel A and a spread portfolio that buys the bottom tercile portfolio and sells the top tercile portfolio for Panel B. Numbers in parentheses indicate t-statistics.

Panel A. Score

	Alpha	Market	SMB	HML	MOM	R^2	N
Review Weighting							
T1 (Low Score)	-0.335%	0.965	0.563	0.143	-0.126	0.728	127
	(-1.21)	(12.38)***	(4.07)***	(1.16)	(-1.94)*		
T2	-0.045%	1.014	0.326	0.076	0.001	0.808	127
	(-0.22)	(17.33)***	(3.14)***	(0.82)	(0.02)		
T3 (High Score)	0.363%	1.084	0.068	0.166	-0.097	0.634	127
	(1.10)	(11.70)***	(0.42)	(1.13)	(-1.25)		
Long/Short (High - Low)	0.698%	0.118	-0.495	0.023	0.030	0.071	127
	(2.57)**	(0.23)	(-1.75)*	(0.89)	(2.16)**		
Equal Weighting							
T1 (Low Score)	-0.319%	0.982	0.543	0.172	-0.129	0.701	127
	(-1.07)	(11.72)***	(3.66)***	(1.29)	(-1.84)*		
T2	-0.136%	0.966	0.387	0.033	0.012	0.787	127
	(-0.62)	(15.77)***	(3.56)***	(0.34)	(0.24)		
T3 (High Score)	0.407%	1.082	0.086	0.254	-0.116	0.605	127
	(1.39)	(11.00)***	(0.49)	(1.63)	(-1.42)		
Long/Short (High - Low)	0.726%	0.101	-0.457	0.082	0.013	0.059	127
	(2.55)**	(0.14)	(-1.71)*	(1.15)	(1.76)*		

Panel B. Spamicity

	Alpha	Market	SMB	HML	MOM	R^2	N
Review Weighting							
T1 (Low Spamicity)	0.223%	1.088	0.419	0.152	-0.053	0.842	127
	(1.10)	(19.11)***	(4.13)***	(1.68)*	(-1.11)		
T2	0.176%	1.103	0.422	0.077	0.055	0.750	127
	(0.94)	(14.64)***	(3.14)***	(0.64)	(0.87)		
T3 (High Spamicity)	-0.416%	1.021	0.513	0.154	-0.049	0.675	127
	(-1.75)*	(11.56)***	(3.26)***	(1.09)	(-0.66)		
Long/Short (High - Low)	0.639%	0.067	-0.094	-0.002	-0.004	0.004	127
	(2.55)**	(0.60)	(-0.54)	(-0.06)	(-0.00)		
Equal Weighting							
T1 (Low Spamicity)	0.197%	1.069	0.415	0.116	-0.062	0.853	127
	(1.02)	(19.79)***	(4.31)***	(1.35)	(-1.36)		
T2	0.148%	1.069	0.426	0.035	0.041	0.764	127
	(0.12)	(14.98)***	(3.35)***	(0.31)	(0.69)		
T3 (High Spamicity)	-0.394%	0.995	0.580	0.141	-0.015	0.631	127
	(-1.76)*	(10.33)***	(3.38)***	(0.92)	(-0.18)		
Long/Short (High - Low)	0.591%	0.073	-0.166	-0.025	-0.047	0.010	127
	(2.54)**	(0.47)	(-0.79)	(-0.16)	(-0.56)		

*** p<0.01, ** p<0.05, * p<0.1

Table 7. Calendar-time Spamicity Portfolio Returns Conditional on Score

In this table I report calendar-time 3 x 3 conditional tercile portfolio regression estimates. For each month, sample stocks are first sorted into tercile portfolios based on abnormal scores (*Score*) and then conditionally sorted into tercile portfolios based on spamicity (*Spamicity*) within each score-tercile portfolio. The dependent variable in each regression is excess portfolio returns in the following month. Following Huang (2018a), stocks in each portfolio are weighted equally or weighted by the number of reviews posted in each month. *Alpha* indicates abnormal portfolio returns after considering a Fama-French-Carhart four-factor model. In Panel A I report alpha estimates only. *Long/Short* indicates a spread portfolio that buys the bottom tercile spamicity portfolio and sells the top tercile spamicity portfolio within each score tercile. Numbers in parentheses indicate t-statistics. (H. = High, L. = Low)

	T1 (Low Spamicity)	T2	T3 (High Spamicity)	Long/Short (L. Spamicity - H. Spamicity)
Review Weighting (Score/Spamicity)				
T1 (Low Score)	-0.065%	-0.107%	-0.834%	0.769%
	(-0.47)	(-0.41)	(-4.35)***	(2.22)*
T2	0.571%	-0.154%	-0.551%	1.122%
	(1.29)	(-1.46)	(-1.87)*	(3.30)***
T3 (High Score)	0.682%	0.487%	-0.081%	0.763%
	(1.62)	(0.27)	(-0.17)	(1.79)
Equal Weighting (Score/Spamicity)				
T1 (Low Score)	-0.056%	-0.089%	-0.813%	0.757%
	(-0.50)	(-0.53)	(-3.87)***	(2.46)**
T2	0.376%	-0.172%	-0.613%	0.989%
	(1.44)	(-1.64)	(-1.97)*	(2.80)**
T3 (High Score)	0.694%	0.423%	0.103%	0.591%
	(1.61)	(1.41)	(-0.23)	(1.70)

Table 8. Score, Spamicity and Return Predictability

In this table I report Fama-MacBeth regression estimates of one-month-ahead excess stock returns on abnormal scores and spamicity using model [3]. Coefficients from monthly cross-sectional regression estimates are averaged across time and reported below. *Month of Earnings Announcement* indicates the month of a year when a firm reports its quarterly earnings and other financial outcomes. Numbers in parentheses indicate Fama-MacBeth t-statistics using Newey-West standard errors corrected for twelve lags.

Dependent Variable: Excess Return$_{t+1}$ (%)	(1)	(2)	(3)	(4)	(5)	(6)	*Month of Earnings Announcement*	
							NO	YES
							(7)	(8)
Score	1.114**		4.649**	1.142*		3.852*	4.909	0.345
	(2.44)		(2.49)	(1.67)		(1.69)	(1.57)	(0.28)
Spamicity		-4.413*	-4.078**		-2.640*	-2.538**	-10.956**	-5.997*
		(-1.67)	(-2.01)		(-1.67)	(-2.01)	(-2.00)	(-1.89)
Score x Spamicity			-16.738**			-13.751*	-34.115**	-7.124
			(-2.20)			(-1.85)	(-2.38)	(-1.05)
Advertising				-0.491	-0.492	-0.670*	-0.455	-1.141
				(-1.39)	(-1.52)	(-1.78)	(-0.52)	(-1.33)
R&D/TA				-5.988	-4.274	-5.241	12.204	-4.379
				(-1.64)	(-1.17)	(-1.31)	-1.01	(-0.59)
Logged Market Cap				0.010	-0.005	0.049	-0.033	0.098
				(0.06)	(-0.03)	(0.34)	(-0.11)	-0.14
M/B				0.747***	0.709***	0.727***	0.291	0.684
				(4.02)	(4.04)	(3.95)	-0.65	-1.25
Gross Profitability				6.416*	6.233**	6.698**	7.586	1.679
				(1.97)	(2.05)	(2.10)	(1.61)	(0.43)
F-Score				0.509***	0.495***	0.446***	0.504	-0.208
				(3.29)	(3.61)	(3.33)	-1.47	(-0.46)
Logged Dollar Volume				-0.262*	-0.276*	-0.272*	-0.115	-0.266
				(-1.81)	(-1.76)	(-1.72)	(-0.45)	(-0.42)
CV of Dollar Volume				4.656***	4.460**	4.573**	7.511	7.469**
				(2.83)	(2.45)	(2.55)	-1.46	-2.07
Stock Return$_m$				-0.036	-0.042**	-0.042**	-0.167	-0.273
				(-1.61)	(-2.05)	(-2.25)	(-1.58)	(-1.62)
Average R^2	0.02	0.014	0.052	0.201	0.184	0.218	0.418	0.677
N	18,081	18,081	18,081	18,081	18,081	18,081	11,967	6,114

*** p<0.01, ** p<0.05, * p<0.1

Table 9. Calendar-time Portfolio Returns Double Sorted on Score and Spamicity

In this table I report calendar-time 3 x 3 unconditional tercile portfolio regression estimates. For each month, sample stocks are double-sorted into 3 x 3 tercile portfolios based on abnormal scores (*Score*) and spamicity (*Spamicity*). The dependent variable in each regression is excess portfolio returns in the following month. Following Huang (2018a), stocks in each portfolio are weighted equally or weighted by the number of reviews in each month. *Alpha* indicates abnormal portfolio returns after considering a Fama-French-Carhart four-factor model. In Panel A I report alpha estimates only. *Long/Short* indicates a spread portfolio that buys the top tercile score portfolio and sells the bottom tercile score portfolio for (A), a spread portfolio that buys the bottom tercile spamicity portfolio and sells the top spamicity tercile portfolio for (B), a spread portfolio that buys the top tercile score and bottom tercile spamicity portfolio and sells the bottom tercile score and top tercile spamicity portfolio for (C), and a spread portfolio that buys the top tercile score and spamicity portfolio and sells the bottom tercile score and spamicity portfolio for (D). Numbers in parentheses indicate t-statistics. (H. = High, L. = Low)

	T1 (Low Spamicity)	T2	T3 (High Spamicity)	(B) Long/Short (L. Spamicity - H. Spamicity)
Review Weighting (Score/Spamicity)				
T1 (Low Score)	-0.185%	-0.234%	-0.586%	0.401%
	(-0.37)	(-1.05)	(-2.11)**	(1.33)
T2	0.267%	0.107%	-0.509%	0.776%
	(0.09)	(1.24)	(1.98)*	(-2.13)**
T3 (High Score)	0.587%	0.656%	-0.154%	0.741%
	(1.74)*	(1.88)*	(-0.11)	(1.75)*
(A) Long/Short (H. Score - L. Score)	0.772%	0.890%	0.432%	
	(1.94)*	(2.21)**	(0.81)	
Equal Weighting (Score/Spamicity)				
T1 (Low Score)	-0.075%	-0.227%	-0.655%	0.580%
	(-0.43)	(-1.17)	(-2.21)**	(1.42)
T2	0.087%	0.135%	-0.630%	0.717%
	(0.28)	(1.18)	(-2.17)**	(2.18)**
T3 (High Score)	0.579%	0.537%	0.104%	0.475%
	(2.12)**	(1.64)	(0.91)	(1.56)
(A) Long/Short (H. Score - L. Score)	0.654%	0.764%	0.759%	
	(1.81)*	(2.17)**	(2.16)**	

Panel B. Score x Spamicity Corner Portfolios

	Alpha	Market	SMB	HML	MOM	R^2	N
Review Weighting							
(C) Long/Short (H. Score & L. Spamicity	1.173%	0.066	-0.316	-0.030	-0.279	0.071	127
- L. Score & H. Spamicity)	(2.50)**	(0.16)	(-1.14)	(-0.11)	(-2.57)**		
(D) Long/Short (H. Score & H. Spamicity	0.226%	-0.054	-0.289	-0.119	0.192	0.019	127
- L. Score & L. Spamicity)	(0.06)	(-0.22)	(-0.72)	(-0.39)	(1.07)		
Equal Weighting							
(C) Long/Short (H. Score & L. Spamicity	1.234%	0.081	-0.320	0.020	-0.405	0.090	127
- L. Score & H. Spamicity)	(2.53)**	(0.41)	(-1.24)	(0.02)	(-2.91)***		
(D) Long/Short (H. Score & H. Spamicity	0.183%	-0.037	-0.230	0.025	0.148	0.016	127
- L. Score & L. Spamicity)	(0.06)	(-0.18)	(-0.57)	(0.09)	(0.89)		

*** p<0.01, ** p<0.05, * p<0.1

Table 10. Calendar-time Portfolio based on E-commerce Dependence

In this table I report calendar-time 3 x 3 x 2 tercile portfolio regression estimates. For each month, sample stocks are first divided into high- and low-characteristic groups (*E-commerce Sales, Number of New Products, and Average Product Price*) and then are double-sorted into tercile portfolios based on abnormal scores (*Score*) and spamicity (*Spamicity*). *E-commerce Sales* indicates the firm-level average of e-commerce sales based on each product line's e-commerce sales from the U.S. Census Bureau. *Number of New Products* indicates firm-level monthly average number of new products that first appear in the sample. *Average Product Price* indicates firm-level average prices of a firm's products based on the July 2014 price. The dependent variable in each regression is excess portfolio returns in the following month. Following Huang (2018a), stocks in each portfolio are weighted equally or weighted by the number of reviews posted in each month (only review-weighted Fama-French-Carhart four-factor model alpha estimates are reported below). *Long/Short* indicates a spread portfolio that buys the top tercile score and bottom tercile spamicity portfolio and sells the bottom tercile score and top tercile spamicity portfolio for (A), and a spread portfolio that buys the top tercile score and spamicity portfolio and sells the bottom tercile score and spamicity portfolio for (B). Numbers in parentheses indicate t-statistics. (H. = High, L. = Low)

Panel A. High and Low E-commerce Sales

	Alpha	
(A) Long/Short (H. Score & L. Spamicity - L. Score & H. Spamicity)	0.245%	1.414%
	(0.77)	(2.15)**
(B) Long/Short (H. Score & H. Spamicity - L. Score & L. Spamicity)	0.224%	-0.148%
	(0.30)	(-0.18)
E-commerce Sales	Low	High

Panel B. High and Low Number of New Products

	Alpha	
(A) Long/Short (H. Score & L. Spamicity - L. Score & H. Spamicity)	0.595%	1.454%
	(0.86)	(1.98)*
(B) Long/Short (H. Score & H. Spamicity - L. Score & L. Spamicity)	-0.271%	-0.900%
	(-0.35)	(-1.09)
Number of New Products	Low	High

Panel C. High and Low Average Product Price

	Alpha	
(A) Long/Short (H. Score & L. Spamicity - L. Score & H. Spamicity)	0.430%	1.337%
	(0.89)	(1.86)*
(B) Long/Short (H. Score & H. Spamicity - L. Score & L. Spamicity)	-0.400%	0.631%
	(-0.43)	(0.70)
Average Product Price	Low	High

*** p<0.01, ** p<0.05, * p<0.1

Table 11. Score, Spamicity and Institutional Ownership

In this table I report panel regression estimates of institutional ownership on abnormal scores (*Score*) and spamicity (*Spamicity*) using model [4]. All variables are calculated at the quarterly level. *Indep. (Grey)* indicates independent (grey) institutions as defined by Ferreira and Matos (2008). Standard errors are double-clustered by firm and by year-quarter. Numbers in parentheses indicate t-statistics.

Dependent Variables:	*Institutional Ownership*			*# of Institutions*	*Indep.*	*Grey*
	(1)	(2)	(3)	(4)	(5)	(6)
Score	0.000		0.003	5.283**	0.004	-0.000
	(0.24)		(1.31)	(2.57)	(1.56)	(-1.18)
Spamicity		-0.022	-0.023	3.269	-0.022	-0.004
		(-0.52)	(-0.53)	(0.07)	(-0.51)	(-0.86)
Score x *Spamicity*			-0.023**	-34.745***	-0.024**	0.001
			(-2.03)	(-2.97)	(-2.05)	(0.34)
Advertising	0.016	0.017	0.017	31.434	0.015	0.003
	(0.47)	(0.50)	(0.52)	(1.44)	(0.45)	(1.11)
R&D/TA	0.089	0.090	0.090	25.220	0.077	0.005
	(1.32)	(1.33)	(1.33)	(0.38)	(1.13)	(0.74)
Logged Market Cap.	0.035**	0.035**	0.035**	141.438***	0.033**	0.002**
	(2.57)	(2.59)	(2.61)	(3.78)	(2.39)	(2.17)
M/B	-0.012**	-0.012**	-0.012**	-10.302	-0.011*	-0.002***
	(-2.13)	(-2.14)	(-2.16)	(-1.25)	(-1.90)	(-4.34)
Gross Profitability	0.008	0.008	0.010	125.072	0.014	0.000
	(0.08)	(0.07)	(0.09)	(1.16)	(0.13)	(0.04)
F-Score	-0.000	-0.000	-0.000	-1.298	-0.000	-0.000
	(-0.35)	(-0.36)	(-0.39)	(-1.33)	(-0.35)	(-0.45)
Logged Dollar Volume	0.030***	0.030***	0.030***	-9.577	0.029***	0.001*
	(3.52)	(3.53)	(3.66)	(-1.33)	(3.45)	(1.69)
CV of Dollar Volume	-0.051***	-0.051***	-0.051***	11.334	-0.050***	-0.003**
	(-4.96)	(-4.97)	(-5.00)	(0.93)	(-4.93)	(-2.05)
Stock Return$_q$	-0.001	-0.001	0.001	-45.973	0.002	-0.001
	(-0.05)	(-0.03)	(0.05)	(-1.43)	(0.11)	(-0.33)
Adjusted R^2	0.905	0.905	0.905	0.980	0.890	0.796
Firm fixed effect	YES	YES	YES	YES	YES	YES
Time fixed effect	YES	YES	YES	YES	YES	YES
N	6,475	6,475	6,475	6,475	6,475	6,475

*** p<0.01, ** p<0.05, * p<0.1

Table 12. Score, Spamicity, and Earning Surprises

In this table I report panel regression estimates of the effects of abnormal scores (*Score*) and spamicity (*Spamicity*) on earnings surprises using model [4]. All variables are calculated at the quarterly level. The dependent variable is unexpected earnings (*SUE*). *Analyst Coverage* indicates the previous quarter's number of analysts. *High (Low)* indicates above- (below-)median scores in each category. Standard errors are double-clustered by firm and year-quarter. Numbers in parentheses indicate t-statistics.

				Analyst Coverage	
Dependent Variable: SUE	(1)	(2)	(3)	Low	High
				(4)	(5)
Score	0.016*		0.032**	0.056**	0.006
	(1.88)		(2.34)	(2.33)	(1.63)
Spamicity		-0.094	0.000	-0.024	0.107
		(-0.54)	(0.00)	(-0.08)	(1.60)
Score x *Spamicity*			-0.162**	-0.325**	-0.032*
			(-2.09)	(-2.14)	(-1.74)
Forecast Dispersion	-2.834***	-2.795***	-2.795***	-2.650***	4.636
	(-4.26)	(-3.99)	(-3.99)	(-3.76)	(0.68)
Advertising	0.112	0.219	0.215	0.338	0.006
	(0.65)	(1.09)	(1.08)	(0.69)	(0.20)
R&D/TA	3.573***	3.802***	3.796***	6.794**	0.312
	(3.90)	(3.91)	(3.90)	(2.67)	(1.22)
Logged Market Cap	-0.097**	-0.086**	-0.086**	-0.106	-0.026
	(-2.57)	(-2.17)	(-2.17)	(-1.16)	(-1.37)
M/B	0.045***	0.039***	0.040***	0.072***	-0.001
	(3.25)	(2.95)	(3.02)	(2.72)	(-0.18)
Gross Profitability	0.471	0.779	0.780	2.614*	0.003
	(0.95)	(1.46)	(1.45)	(2.02)	(0.02)
F-Score	0.018*	0.018*	0.018*	0.026	0.002
	(1.98)	(1.84)	(1.86)	(1.53)	(0.66)
Logged Dollar Volume	0.018	0.012	0.012	-0.000	0.021*
	(0.80)	(0.55)	(0.52)	(-0.00)	(1.78)
CV of Dollar Volume	-0.052	-0.058	-0.056	-0.083	-0.005
	(-0.89)	(-0.93)	(-0.91)	(-0.73)	(-0.32)
Stock Return$_{q-1}$	0.009	0.011	0.011	0.000	0.004
	(0.82)	(0.93)	(0.93)	(0.04)	(0.67)
Adjusted R^2	0.174	0.174	0.174	0.195	0.594
Firm fixed effect	YES	YES	YES	YES	YES
Time fixed effect	YES	YES	YES	YES	YES
N	6,465	6,465	6,465	3,232	3,233

*** p<0.01, ** p<0.05, * p<0.1

Table 13. Calendar-time Portfolio based on Institutional Ownership

In this table I report calendar-time 3 x 3 x 2 tercile portfolio regression estimates. For each month, sample stocks are first divided into high- and low-characteristic groups (*Independent Institutional Ownership/Grey Institutional Ownership/Analyst Coverage*) and then are double-sorted into tercile portfolios based on abnormal scores (*Score*) and spamicity (*Spamicity*). *Independent (Grey) Institutional Ownership* indicates the previous quarter's independent (grey) institutional ownership. *Analyst Coverage* indicates the previous quarter's number of analysts. The dependent variable in each regression is excess portfolio returns in the following month. Following Huang (2018a), stocks in each portfolio are weighted equally or weighted by the number of reviews posted in each month (only review-weighted Fama-French-Carhart four-factor model alpha estimates are reported below). *Long/Short* indicates a spread portfolio that buys the top tercile score and bottom tercile spamicity portfolio and sells the bottom tercile score and top tercile spamicity portfolio for (A), and a spread portfolio that buys the top tercile score and spamicity portfolio and sells the bottom tercile score and spamicity portfolio for (B). Numbers in parentheses indicate t-statistics. (H. = High, L. = Low)

Panel A. By Independent Ownership

	Alpha	
(A) Long/Short (H. Score & L. Spamicity - L. Score & H. Spamicity)	1.201%	-0.228%
	(2.24)**	(-0.34)
(B) Long/Short (H. Score & H. Spamicity - L. Score & L. Spamicity)	-0.609%	-0.125%
	(-0.81)	(-0.15)
Independent Institutional Ownership	Low	High

Panel B. By Grey Institutional Ownership

	Alpha	
(A) Long/Short (H. Score & L. Spamicity - L. Score & H. Spamicity)	0.672%	0.280%
	(1.30)	(0.48)
(B) Long/Short (H. Score & H. Spamicity - L. Score & L. Spamicity)	-0.031%	-0.864%
	(-0.04)	(-1.14)
Grey Institutional Ownership	Low	High

Panel C. By Analyst Coverage

	Alpha	
(A) Long/Short (H. Score & L. Spamicity - L. Score & H. Spamicity)	1.298%	0.238%
	(1.80)*	(0.36)
(B) Long/Short (H. Score & H. Spamicity - L. Score & L. Spamicity)	0.862%	-1.197%
	(1.02)	(-1.34)
Analyst Coverage	Low	High

*** p<0.01, ** p<0.05, * p<0.1

Table 14. Shock in Spamming and Institutional Ownership

In this table I report panel regression estimates of the effects of institutional ownership on abnormal scores (*Score*) and spamicity (*Spamicity*) derived from model [5]. All variables are calculated at the quarterly level. *Indep. (Grey)* indicates independent (grey) institutions as defined by Ferreira and Matos (2008). *Post 2012* takes the value of 1 if the year of the observation is 2012 or later and 0 otherwise. *High (Low) Exposure* indicates that a firm's products are sold in item categories with above- (below-)median proportions of products produced by Chinese manufacturers after 2012. Control variables are the same as those for which results are reported in Table 10. Standard errors are double-clustered by firm and year-quarter. Numbers in parentheses indicate t-statistics.

Dependent Variable: *Institutional Ownership*	*Indep.*		*Grey*	
	(1)	(2)	(3)	(4)
Score	0.004	0.001	-0.000	-0.000
	(1.06)	(0.15)	(-0.05)	(-0.59)
Spamicity	0.031	0.069	-0.008	-0.006*
	(1.03)	(0.69)	(-1.45)	(-1.91)
Score x *Spamicity*	-0.018	-0.001	-0.001	-0.002
	(-0.93)	(-0.04)	(-0.35)	(-0.60)
Spamicity x Post 2012	0.047	0.001	-0.004	0.018***
	(0.62)	(0.02)	(-0.37)	(6.21)
Score x Post 2012	-0.003	0.010	0.000	0.000
	(-0.82)	(1.44)	(1.20)	(0.81)
Score x *Spamicity* x Post 2012	-0.006	-0.057**	-0.000	-0.002
	(-0.17)	(-2.98)	(-0.03)	(-0.56)
Adjusted R^2	0.942	0.906	0.866	0.905
Firm fixed effect	YES	YES	YES	YES
Time fixed effect	YES	YES	YES	YES
Controls	YES	YES	YES	YES
Exposure	Low	High	Low	High
N	2,160	2,772	2,160	2,772

*** p<0.01, ** p<0.05, * p<0.1

Appendix. Variable Definitions

In this table I present detailed definitions of the *Review-*, *Reviewer-*, and *Product-centric features* used in model [1]. The review data are taken from SNAP and sentiments are estimated using the sentiment lexicons (lists of words' sentiment scores) published in Hamilton et al. (2016). *Review-centric features* represent individual-review-level information (full-review -sample level). *Reviewer-centric features* are also defined at the full-review-sample level to capture the hidden overall behavioral patterns of reviewers more comprehensively. *Product-centric features* are monthly product-level scores and review characteristics.

Variable	Definition
Review-centric Features	
Logged # representing Feedback	Logged number representing instances of feedback given to a review.
Logged # representing Helpful Feedback	Logged number representing instances of feedback given to a review that indicates that the review is helpful.
% represent ting Helpful Feedback	The proportion of Helpful feedback of the total number of instances of feedback [0 if there are no instances of Helpful feedback].
Logged Length of a Review [Title]	Logged number of characters used in a review title.
Logged Length of a Review	Logged number of characters used in the main text of a review.
Logged Position of a Review [Ascending]	Logged ascending chronological order of a review.
Logged Position of a Review [Descending]	Logged descending chronological order of a review.
First Review Dummy	Dummy variable that takes the value of 1 if a review is the first review for a given product in the sample and 0 otherwise.
Only Review Dummy	Dummy variable that takes the value of 1 if a review is the only review for a given product in the sample and 0 otherwise.
% of Positive Sent.	The number of words with positive sentiment scores over the total number of words in the main text of a review.
Logged Positive Sent. Score	Logged sum of the scores of the words in the main text of a review with positive sentiment.
% of Negative Sent.	The number of words with negative sentiment scores over the total number of words in the main text of a review.
Logged Negative Sent. Score	Logged sum of the scores of the words in the main text of a review with negative sentiment.
Logged Positive Sent. Score [Title]	Logged sum of the scores of the words in the title of a review with positive sentiment.
Logged Negative Sent. Score [Title]	Logged sum of the scores of the words in the title of a review with negative sentiment.
% of Numerals	The number of numeral characters in the main text of a review over the total number of characters in the main text of a review.
% of Capital Letters	The number of capital characters in the main text of a review over the total number of characters in the main text of a review.
% of Comparatives	The number of comparatives in the main text of a review over the total number of words in the main text of a review.
Review Score	The score of a review [1, 2, 3, 4, or 5]
Score Dummy	Indicator variable that takes the value of 1 for scores of 4 or above, -1 for scores 2 or below, and 0 otherwise.
Reviewer-centric Features	
Percent of First Reviews	The number of reviews that are the first review for a given product over the total number of reviews posted by a reviewer.
Percent of Only Reviews	The number of reviews that are the only review for a given product over the total number of reviews posted by a reviewer.
Avg. Score [Reviewer]	The average score of the reviews posted by a reviewer.
Stddev. of Scores [Reviewer]	The standard deviation of the scores of the reviews posted by a reviewer.
Score Dummy [Reviewer]	Indicator variable that takes the value of 1 if the average score of a

Variable	Definition
	reviewer is 4 or above, -1 if it is 2 or below, and 0 otherwise.
Good and Bad Scores	Dummy variable that takes the value of 1 if a reviewer has given both scores of 4 or above and scores of 2 or below and 0 otherwise.
Good and Avg. Scores	Dummy variable that takes the value of 1 if a reviewer has given scores of 3 or above and 0 otherwise.
Bad and Avg. Scores	Dummy variable that takes the value of 1 if a reviewer has given both scores of 3 or below and 0 otherwise.
All Scores	Dummy variable that takes the value of 1 if a reviewer has given scores of all levels and 0 otherwise.
% of Negative Dev.	The number of reviews that receive scores higher than the average score of a product when the review was posted over the total number of reviews posted by a reviewer.
% of Positive Dev.	The number of reviews that receive scores lower than the average score for a product when the review was posted over the total number of reviews posted by a reviewer.
Product-centric Reviews	
Avg. Score [Product]	The average review score for a product in each month.
Stddev. of Scores [Product]	The standard deviation of the review scores of a product in each month.
Logged # of Product Reviews	Logged total number of reviews of a product in each month.
Logged Stddev. of # of Reviews	Logged standard deviation of the total number of reviews of a product in each month.

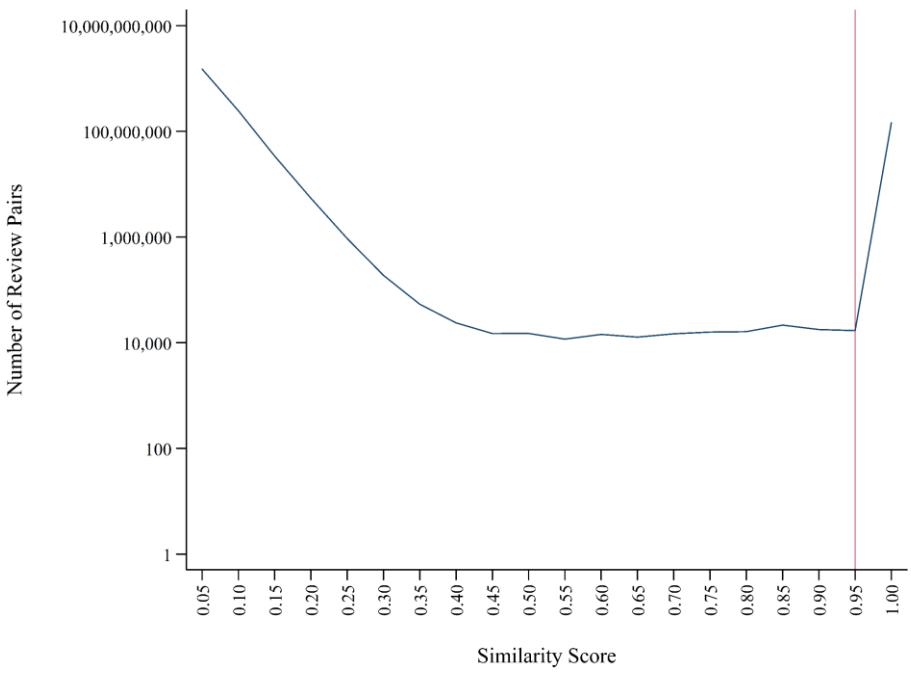


Figure A1. Similarity Scores of Review Pairs. This figure shows how many review pairs in the sample receive the similarity scores represented on the horizontal axis. *Similarity Score* indicates 2-gram Jaccard distances between reviews. There are potentially 44.6 million x (44.6 million - 1)/2 review pairs in the sample. To ease the computational burden, I lexicographically arrange reviews and form 1,000 review-groups and then measure 2-gram Jaccard distances within each group. As a result, there are approximately 44.7 billion review pairs between which to measure the Jaccard distance. The number of review pairs and the number of unique reviews are not equal as there are multiple reviews with the same or similar content in the sample.

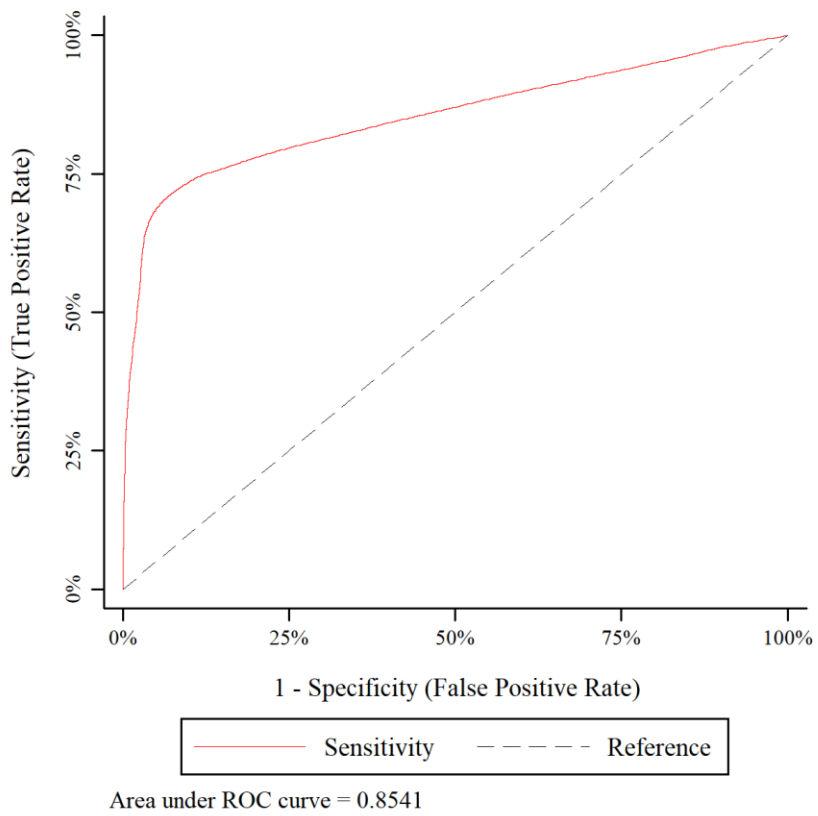


Figure A2. Receiver Operating Characteristic (ROC) Curve. This figure shows one of the ten ROC curves attained after estimating model [1]. *Sensitivity* indicates the proportion of positive examples (near-duplicate reviews) in the training sample that are correctly identified as spam by model [1] under a given probability threshold (1 - x-axis). *1- Specificity* indicates the proportion of negative examples (non-near-duplicate reviews) in the training sample that are falsely identified as spam by model [1] under a given probability threshold (1 - y-axis). Using a ten-fold cross-validation method, the training sample is randomly divided into ten groups and thus yields ten out-of-sample ROC curves. The average Area Under Curve (AUC) of the ten ROC curves is 0.8551.

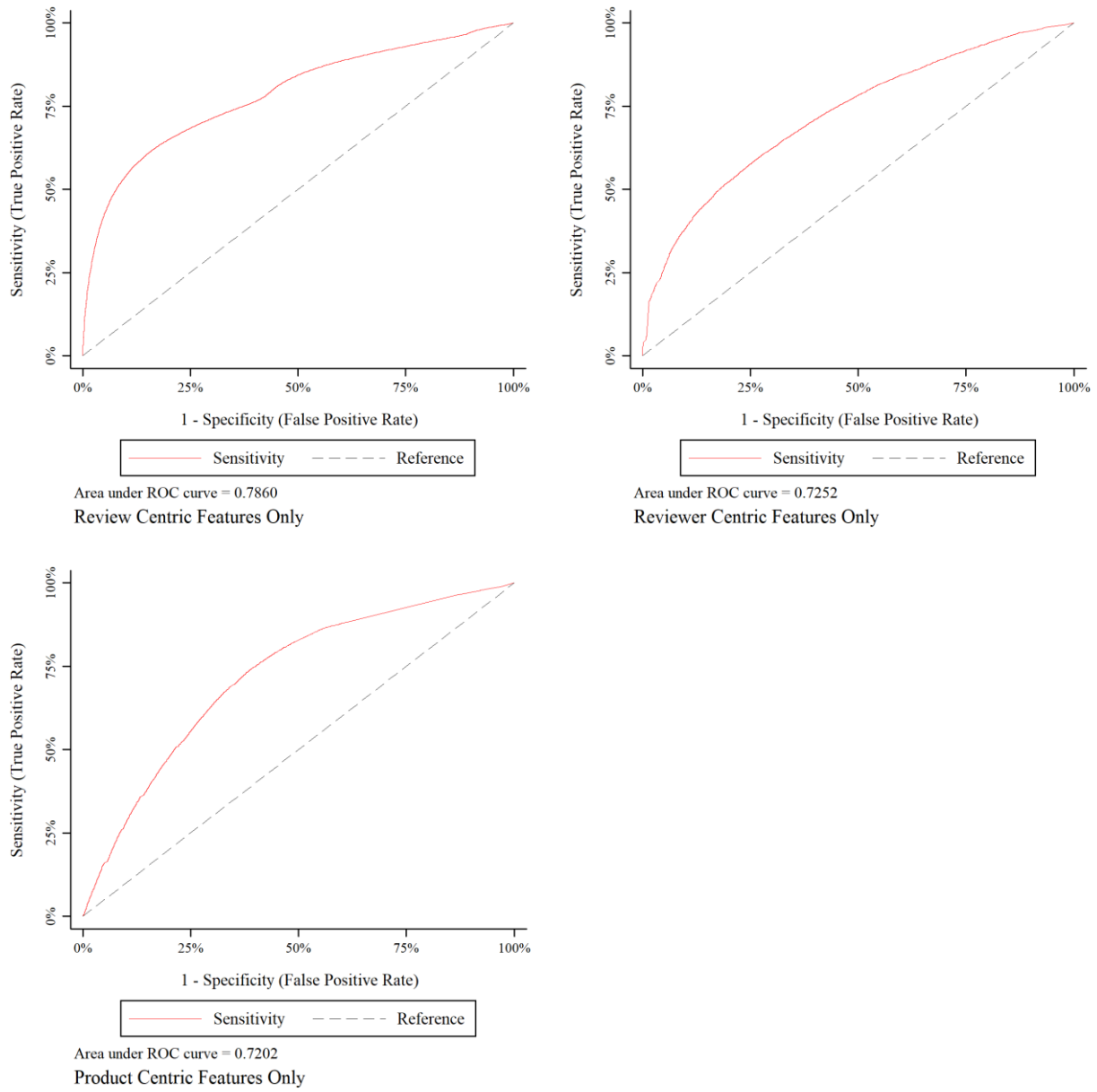


Figure A3. Comparison of Spam Predictability between Review-, Reviewer-, and Product-centric Features. Each graph in this figure shows one of the ten ROC curves in each category after estimating model [1] using one of the review-, reviewer- or product-centric features exclusively. The average Areas Under Curve (AUCs) of the ten groups for review-centric, reviewer-centric, and product-centric features are 0.7862, 0.7255, and 0.7212, respectively.

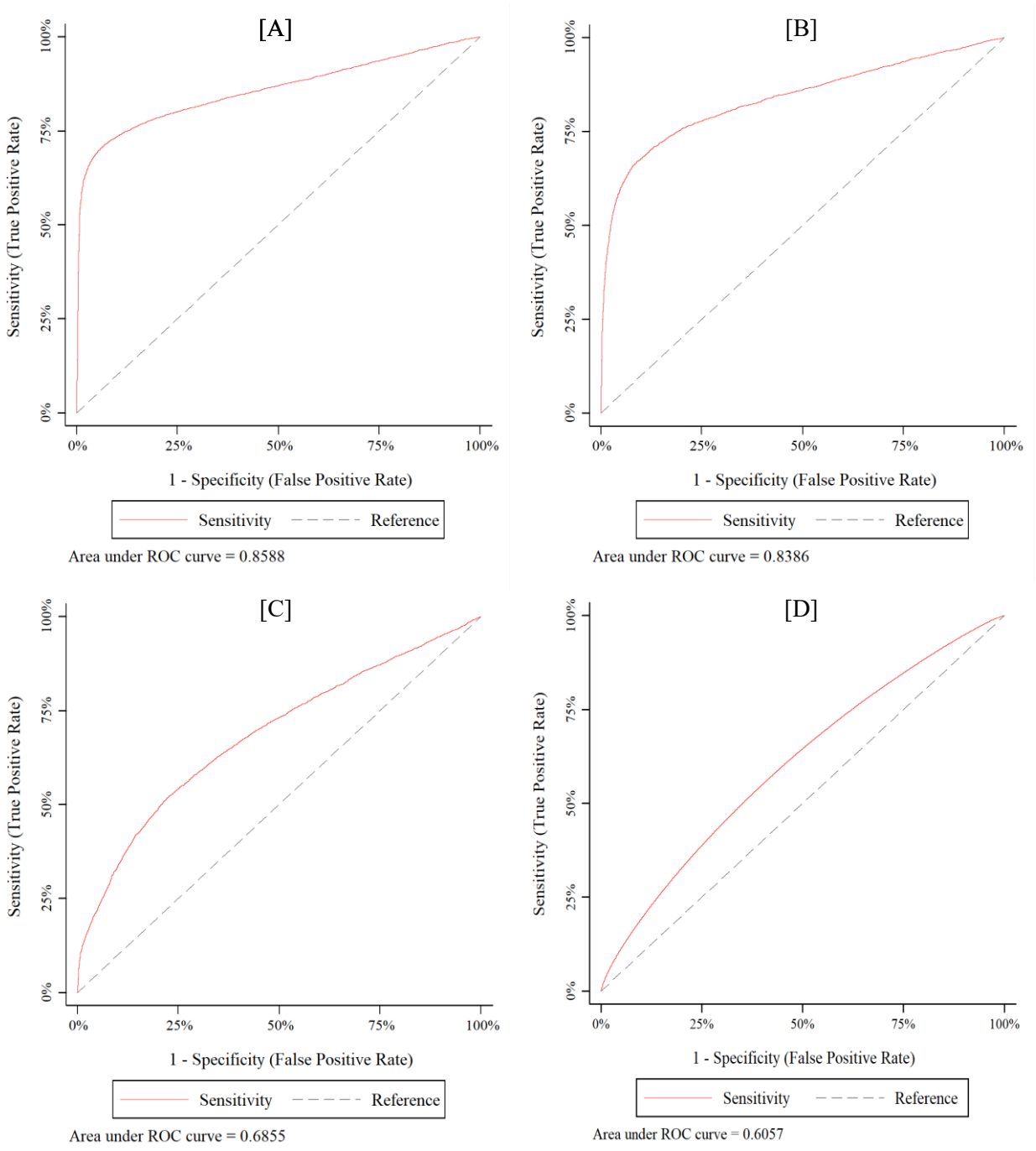


Figure A4. Spam Predictability for separate training samples. Each of the above graphs shows one of the ten ROC curves after estimating model [1] using different negative examples: [A] review scores that do not deviate more than +/- 0.01 from the average, [B] review scores that do not deviate more than +/– 0.05 from the average, [C] review scores that do not deviate more than +/– 0.5 from the average, and [D] all of the non-near-duplicate reviews. The average Area Under Curves (AUCs) of the ten groups are 0.8581, 0.8392, 0.6854, and 0.6055 for [A], [B], [C], and [D], respectively. The replication of results reported in table 9 using separate training samples above are presented in Table A5.

Table A1. First-Stage Estimates

In this table I report the first-stage estimations of the results reported in column (7) of Table 5. The instrumental variable is *Average Peer Score* and the instrumented variable is *Abnormal Score.* Standard errors are double clustered by firm and year-month. Numbers in parentheses indicate t-statistics.

Dependent Variable: Abnormal score	
Abnormal Peer Score	0.740***
	(27.55)
Standard Deviation of Score	-0.794***
	(-42.49)
Lagged Dependent Variable	-0.097
	(-0.90)
Advertising	0.031
	(0.98)
R&D/TA	-0.073
	(-0.38)
Logged Market Cap	0.007
	(0.42)
M/B	0.000
	(0.02)
Gross Profitability	-0.112*
	(-1.88)
F-Score	0.001
	(0.32)
Logged Dollar Volume	-0.002
	(-0.30)
CV of Dollar Volume	-0.034
	(-1.05)
Return	0.000
	(0.11)
Adjusted R^2	0.716
Firm FE	Yes
Year FE	Yes
F-stat	84.51
N	18,081

*** p<0.01, ** p<0.05, * p<0.1

Table A2. Longer-period Score and Spamicity Portfolio Returns

In this table I report longer-period calendar-time 3 x 3 tercile portfolio regression estimates. For each month, sample stocks are double-sorted into tercile portfolios based on abnormal scores (*Score*) and spamicity (*Spamicity*). The dependent variable in each regression is excess portfolio returns in the following month. Following Huang (2018a), stocks in each portfolio are weighted equally or weighted by the number of reviews posted in each month (only Fama-French-Carhart four-factor model alpha estimates are reported below). *Long/Short* indicates a spread portfolio that buys the top tercile score and bottom tercile spamicity portfolio and sells the bottom tercile score and top tercile spamicity portfolio for (C), and a spread portfolio that buys the top tercile score and spamicity portfolio and sells the bottom tercile score and spamicity portfolio for (D). Months [*a* ,*b*] indicates a spread portfolio alpha from month *a* to month *b* after portfolio formation, excluding the first-month returns after the formation month. Each spread portfolio is rebalanced after each month. Numbers in parentheses indicate t-statistics. (H. = High, L. = Low)

Panel A. (C) Long/Short (H. Score & L. Spamicity - L. Score & H. Spamicity)

	Review Weighting		Equal Weighting	
Months [2, 4]	0.30%	(0.23)	0.31%	(0.23)
Months [2, 7]	0.04%	(0.02)	0.06%	(0.03)
Months [2, 10]	0.12%	(0.53)	0.01%	(0.60)
Months [2, 13]	0.03%	(1.17)	0.03%	(1.08)

Panel B. (D) Long/Short (H. Score & H. Spamicity - L. Score & L. Spamicity)

	Review Weighting		Equal Weighting	
Months [2, 4]	-0.31%	(-0.18)	0.21%	(0.13)
Months [2, 7]	-0.03%	(-1.06)	-0.03%	(-0.12)
Months [2, 10]	-0.06%	(-0.89)	-0.02%	(-0.77)
Months [2, 13]	-0.09%	(-1.06)	-0.05%	(-1.08)

*** p<0.01, ** p<0.05, * p<0.1

Table A3. Calendar-time Score and Spamicity Spread Portfolio Returns - Alternative Factors

In this table I report calendar-time 3 x 3 tercile portfolio regression estimates. For each month, sample stocks are double-sorted into tercile portfolios based on abnormal scores (*Score*) and spamicity (*Spamicity*). The dependent variable in each regression is excess portfolio returns in the following month. Following Huang (2018a), stocks in each portfolio are weighted equally or weighted by the number of reviews posted in each month. *Alpha* indicates abnormal portfolio returns after considering a Fama-French-Carhart four-factor model adding two additional factors from Fama-French (2015). *Long/Short* indicates a spread portfolio that buys the top tercile score and bottom tercile spamicity portfolio and sells the bottom tercile score and top tercile spamicity portfolio for (C), and a spread portfolio that buys the top tercile score and spamicity portfolio and sells the bottom tercile score and spamicity portfolio for (D). Numbers in parentheses indicate t-statistics. (H. = High, L. = Low)

	Alpha	Market	SMB	HML	MOM	RMW	CMA	R^2	N
Review Weighting									
(C) Long/Short (H. Score & L. Spamicity	1.056%	0.041	-0.237	0.302	-0.390	0.543	0.315	0.101	127
- L. Score & H. Spamicity)	(2.60)**	(0.25)	(-0.88)	(1.15)	(-3.11)***	(1.42)	(0.71)		
(D) Long/Short (H. Score & H. Spamicity	-0.051%	-0.102	-0.026	-0.507	0.025	-0.115	-0.028	0.039	127
- L. Score & L. Spamicity)	(-0.09)	(-0.63)	(-0.10)	(-1.91)*	(0.19)	(-0.30)	(-0.06)		
Equal Weighting									
(C) Long/Short (H. Score & L. Spamicity	1.167%	0.058	-0.204	0.332	-0.368	0.584	0.168	0.093	127
- L. Score & H. Spamicity)	(2.21)**	(0.36)	(-0.75)	(1.25)	(-2.90)***	(1.51)	(0.38)		
(D) Long/Short (H. Score & H. Spamicity	-0.086%	-0.106	0.043	-0.463	-0.091	-0.243	-0.040	0.051	127
- L. Score & L. Spamicity)	(-0.15)	(-0.63)	(0.15)	(-1.70)*	(-0.70)	(-0.61)	(-0.09)		

*** p<0.01, ** p<0.05, * p<0.1

Table A4. Alternative Training Samples

In this table I report calendar-time 3 x 3 tercile portfolio regression estimates using varying spamicity estimates. In the training sample, negative examples are reviews that do not deviate from the average review score by +/−0.01 (*Range = +/−0.01*), +/−0.05 (*Range = +/−0.05*), and +/−0.5 (*Range = +/−0.5*). Average Spamicity in each Range are 0.51 (*Range = +/−0.01*), 0.34 (*Range = +/−0.05*), and 0.05 (*Range = +/−0.5*). For each month, sample stocks are double-sorted into tercile portfolios based on abnormal scores (*Score*) and spamicity (*Spamicity*). The dependent variable in each regression is excess portfolio returns in the following month. Following Huang (2018a), stocks in each portfolio are weighted equally or weighted by the number of reviews posted in each month. *Alpha* indicates abnormal portfolio returns after considering a Fama-French-Carhart four-factor model. *Long/Short* indicates a spread portfolio that buys the top tercile score and bottom tercile spamicity portfolio and sells the bottom tercile score and top tercile spamicity portfolio for (C), and a spread portfolio that buys the top tercile score and spamicity portfolio and sells the bottom tercile score and spamicity portfolio for (D). Numbers in parentheses indicate t-statistics. (H. = High, L. = Low)

Panel A. Abnormal Score x Spamicity (Range = +/−0.01)

	Alpha	Market	SMB	HML	MOM	R^2	N
Review Weighting							
(C) Long/Short (H. Score & L. Spamicity	1.322%	0.028	-0.469	0.348	-0.210	0.054	127
- L. Score & H. Spamicity)	(2.31)**	(0.07)	(-1.56)	(1.48)	(-1.16)		
(D) Long/Short (H. Score & H. Spamicity	-0.632%	-0.015	-0.008	0.302	-0.029	0.009	127
- L. Score & L. Spamicity)	(-1.58)	(-0.09)	(-0.94)	(0.35)	(-0.19)		
Equal Weighting							
(C) Long/Short (H. Score & L. Spamicity	0.801%	-0.143	0.009	0.323	-0.095	0.033	127
- L. Score & H. Spamicity)	(2.14)**	(-0.88)	(0.48)	(1.13)	(-0.98)		
(D) Long/Short (H. Score & H. Spamicity	-0.181%	-0.063	-0.048	0.150	-0.178	0.021	127
- L. Score & L. Spamicity)	(-1.19)	(-0.03)	(-0.47)	(0.37)	(-1.45)		

Panel B. Abnormal Score x Spamicity (Range = +/−0.05)

	Alpha	Market	SMB	HML	MOM	R^2	N
Review Weighting							
(C) Long/Short (H. Score & L. Spamicity	1.395%	0.027	0.851	0.059	0.056	0.068	127
- L. Score & H. Spamicity)	(2.10)**	(0.14)	(2.58)**	(0.20)	(0.36)		
(D) Long/Short (H. Score & H. Spamicity	0.516%	0.076	-0.688	0.599	-0.138	0.076	127
- L. Score & L. Spamicity)	(0.74)	(0.39)	(-1.97)*	(1.92)*	(-0.84)		
Equal Weighting							
(C) Long/Short (H. Score & L. Spamicity	1.147%	-0.133	0.614	0.303	-0.157	0.038	127
- L. Score & H. Spamicity)	(1.77)*	(-0.73)	(1.90)*	(1.05)	(-1.03)		
(D) Long/Short (H. Score & H. Spamicity	-0.086%	0.007	-0.372	0.711	-0.272	0.083	127
- L. Score & L. Spamicity)	(-0.14)	(0.04)	(-1.18)	(2.52)**	(-1.83)*		

Panel C. Abnormal Score x Spamicity (Range = +/−0.5)

	Alpha	Market	SMB	HML	MOM	R^2	N
Review Weighting							
(C) Long/Short (H. Score & L. Spamicity	1.228%	-0.178	0.695	0.048	-0.270	0.062	127
- L. Score & H. Spamicity)	(1.90)*	(-0.98)	(2.16)**	(0.17)	(-1.78)*		
(D) Long/Short (H. Score & H. Spamicity	0.467%	0.166	-0.239	0.654	-0.198	0.039	127
- L. Score & L. Spamicity)	(0.61)	(0.77)	(-0.63)	(1.93)*	(-1.11)		
Equal Weighting							
(C) Long/Short (H. Score & L. Spamicity	0.904%	-0.203	0.645	0.395	-0.246	0.073	127
- L. Score & H. Spamicity)	(1.67)*	(-1.33)	(2.39)**	(1.64)	(-1.94)*		
(D) Long/Short (H. Score & H. Spamicity	0.237%	-0.032	-0.393	0.574	-0.287	0.071	127
- L. Score & L. Spamicity)	(0.37)	(-0.18)	(-1.23)	(2.01)**	(-1.90)*		

*** p<0.01, ** p<0.05, * p<0.1

Table A5. Score and Spamicity Portfolio Returns - Subsample Analysis

In this table I report calendar-time 3 x 3 x 2 tercile portfolio regression estimates. *Positive Spamicity* indicates the spamicity estimate of a review if it receives scores of 4 or above and 0 otherwise. *Negative Spamicity* indicates the spamicity estimate of a review that takes the value of 1 if it receives scores of 2 or below and) otherwise. For each month, sample stocks are first divided into high- and low-characteristic groups (*Negative Spamicity/Positive Spamicity and Abnormal Peer Score*) and then are double-sorted into tercile portfolios based on abnormal scores (*Score*) and positive or negative spamicity (*Spamicity*). The dependent variable in each regression is excess portfolio returns in the following month. Following Huang (2018a), stocks in each portfolio are weighted equally or weighted by the number of reviews posted in each month (only review-weighted Fama-French-Carhart four-factor model alpha estimates are reported below). *Long/Short* indicates a spread portfolio that buys the top tercile score and bottom tercile spamicity portfolio and sells the bottom tercile score and top tercile spamicity portfolio for (C), and a spread portfolio that buys the top tercile score and spamicity portfolio and sells the bottom tercile score and spamicity portfolio for (D). Numbers in parentheses indicate t-statistics. (H. = High, L. = Low)

Panel A. High and Low Positive/Negative Spamicity

	Positive Spamicity		Negative Spamicity	
(C) Long/Short (H. Score & L. Spamicity - L. Score & H. Spamicity)	1.301%	0.658%	0.051%	0.413%
	(1.94)*	(0.77)	(0.74)	(0.69)
(D) Long/Short (H. Score & H. Spamicity - L. Score & L. Spamicity)	0.067%	0.368%	-0.140%	0.495%
	(0.09)	(0.44)	(-0.34)	(0.63)
Negative Spam† / Positive Spam‡	Low†	High†	Low‡	High‡

Panel B. High and Low Peer Scores

	Positive Spamicity		Negative Spamicity	
(C) Long/Short (H. Score & L. Spamicity - L. Score & H. Spamicity)	0.213%	0.787%	0.243%	0.640%
	(1.00)	(1.05)	(0.73)	(1.41)
(D) Long/Short (H. Score & H. Spamicity - L. Score & L. Spamicity)	-0.127%	0.026%	-0.018%	-1.457%
	(-0.20)	(0.03)	(-0.46)	(-2.31)**
Abnormal Peer Score	Low	High	Low	High