The Impact of Amazon Reviews on Stock Performance

University of Mannheim

Chair of International Finance

Supervisor: Sven Vahlpahl

Frederik Ciupek

Langstraße 72, 68169 Mannheim

Student ID: 1750101

2023

Contents

A. List of Tables

Later Tables in the Appendix are labeled with the prefix A, but not shown here.

B. List of Figures

Later Figures in the Appendix are labeled with the prefix A, but not shown here.

C. List of Abbreviations

ACSI American Customer Satisfaction Index. 3, 4

AMEX American Stoch Exchange. 6

ASIN Amazon Standard Identification Number. 6

CRSP Center for Research on Stock Prices. 10

CV Coeffiicent of Variation of dollar volume. 12

EPS Earnings Per Share). 12

I/B/E/S Instituional Broker's Estimation System. 10, 16

ID Amazon Account Identification. 6

JPX Japanese Exchange Group. 24

NASDAQ National Association of Securities Dealers Automated Quotations. 6

NYSE New York Stock Exchange. 6

R&D Research & Development. 10–12

ROA Return on Assets. 3

S&P 500 Standard & Poor's 500. 19, 27, 34

SIC Standard Industry Classifaction. 7

SUE Standardized Unexpected Earnings). 11

SURGE Standard Unexpected Revenue Growth Estimator). 11

TYO Tokyo Stock Exchange. 20

WRDS Wharton Research Data Service. 10, 16, 28

1. Introduction

In recent years, the evaluation of corporate performance has shifted towards considering intangible assets as key drivers of value creation and economic growth (Lazonick and O'Sullivan, 2000). This change reflects the increasing recognition of intangible assets, such as intellectual property and brand reputation, as more valuable than tangible assets on a company's balance sheet (Fornell et al., 2016). However, accurately assessing the value of intangibles poses challenges as they are not capitalized like other investments. Nevertheless, studies have demonstrated the significant impact of intangible assets on market value and stock returns (Srivastava et al., 1997; Lev and Sougiannis, 1996; Chan et al., 2001; Katsikeas et al., 2016)). For instance, Srivastava et al. (1997) find that intangible assets account for 70% of a firm's market value, as indicated by an average market-to-book ratio of 3.6. Among various intangible assets, customer satisfaction has gained considerable attention in research (Anderson et al., 2004; Fornell, 2001; Fornell et al., 2006; Gruca and Rego, 2005). Consumers, as vital stakeholders of a firm, are the recipients of a company's products and services, but not only provide valuable product-related information that is accessible to other consumers but also consider such information when making purchasing decisions. Accordingly, this has led the investigation of on the impact of customer satisfaction on firm fundamentals, profitability, and stock returns, with predominantly positive correlations (Fornell, 1992; Fornell et al., 2006; Fornell et al., 2016; Fornell et al., 2010; Jacobson and Mizik, 2009; Gruca and Rego, 2005; Aksoy et al., 2008). Notably, consumers possess "serendipitous information": information gained by investors during everyday acitivities about product quality and demand, which, when aggregated across a large number of people, can provide valuable insights into a firm's value (Subrahmanyam and Titman, 1999). Despite the evidence supporting the value of customer satisfaction since the early 1990s (Fornell, 1992, analysts have been slow to incorporate intangible customer-based metrics into performance forecasting (Stuart et al., 2004). Early skepticism and the belief that such information provides no additional value may have contributed to this oversight (Aksoy et al., 2008). However, research has revealed customer satisfaction mispricing in the stock market, demonstrating higher returns and lower risk associated with investments in customer satisfaction (Fornell et al., 2006). Additionally, portfolios based on online customer satisfaction ratings have shown high abnormal returns (Huang, 2018). The market's slow reaction to such information could be attributed to the time required for equity and consumer markets to fully leverage their collective power (Huberman and Regev, 2001; Dellavigna and Pollet, 2009; Hirshleifer et al., 2009), as shown in the ultimate disappearance of the Monday effect, post-earnings announcement drift, and January effect (Marquering et al., 2006; Szakmary and Kiefer, 2004; Heitz et al., 2018. Technological advancements and the rise of online retailers like Amazon, Alibaba, and JB have provided vast datasets of publicly available consumer opinions, significantly reducing the cost of information, meaning markets no longer require sophisticated measurement technologies to acquire customer opinion data. Hedge funds specialized in information processing, for instance, actively trade on customer satisfaction information, albeit to a limited extent (Huang, 2018) This thesis aims to investigate two main aspects: (1) the presence of customer satisfaction mispricing and abnormal returns, building heavily upoon previous research of Huang (2018), and (2) the adaptation of the market to customer satisfaction information, potentially leading to a reduction or disappearance of customer satisfaction mispricing. To achieve this, the methodology outlined by Huang (2018) is closely followed, utilizing a novel dataset. The analysis includes over 27 million customer reviews from Amazon.com, the largest source of Internet consumer reviews, and 200,000 reviews from Japanese Amazon.co.jp, spanning from April 2008 to June 2023. The findings provide compelling evidence of the investment value and return predictability of consumer opinions and suggest a diminishing effect of customer satisfaction mispricing in the modern age.

The structure of the thesis is as follows: Chapter 2 presents a comprehensive literature review, Chapter 3 provides an overview of the data and summary statistics, Chapter 4 presents the empirical results, and finally, Chapter 5 concludes the thesis.

2. Literature Review

This thesis is related to several areas of research. Firstly, it contributes to the expanding literature on marketing that examines the impact of intangible assets, specifically customer satisfaction, on firm value (Anderson et al., 2004; Aksoy et al., 2008; Fornell et al., 1996; Fornell, 2001; Fornell et al., 2006; Fornell et al., 2010; Fornell et al., 2016; Gruca and Rego, 2005, Ivanov et al., 2013; Jacobson and Mizik, 2009. Numerous studies focus on the influence of customer satisfaction on a firm's sales, profitability, and marketing metrics. For example, customer satisfaction has been found to have a negative impact on customer complaints and a positive impact on customer loyalty and usage level (Fornell, 1992; Bolton et al., 2000). It also positively affects word-of-mouth recommendations (Anderson, 1998;Luo, 2009) willingness to pay (Homburg et al., 2005), repeat purchases, cross-selling, and lower transaction costs (Anderson et al., 1994; Reichheld and Sasser, 1990) all of which contribute to profitability and firm value (Fornell et al., 2010). Customer satisfaction also provides valuable information about a firm's future cash flows and revenue growth (Anderson et al., 2004; Gruca and Rego, 2005; Aksoy et al., 2008). Greater cash flows, with lower volatility, enhanced by customer satisfaction, increase firm value and influence stock prices by shaping investor expectations of future cash flows (Srinivasan and Hanssens, 2009. Additionally, Wieseke et al. (2010) find that positive customer reviews have a positive impact on stock analyst recommendations, and Jacobson and Mizik (2009) find a positive impact on Return on Assets (ROA). Recently, the focus has expanded to examine the positive impact of online consumer reviews on a firm's sales and profitability in various product markets such as movies, books, and computers (Chintagunta et al., 2010; Dellarocas et al., 2007; Y. Chen et al., 2011; Mazylin and Chevalier, 2006; Forman et al., 2008; Moe and Trusov, 2011; Zhu and Zhang, 2010; Godes and Mayzlin, 2002)

Newly, there has been increased research on the relationship between customer satisfaction and stock performance (Fornell et al., 2006; Fornell et al., 2010; Fornell et al., 2016; Huang, 2018; Wu et al., 2020; Jacobson and Mizik, 2009; H. Chen et al., 2014; Edmans, 2011; Tirunillai and Tellis, 2012). Using the American Customer Satisfaction Index (ACSI), which reports customer satisfaction scores on a scale of 0 to 100 annually based on phone surveys, researchers have built numerous investing strategies. Early on, Fornell et al. (2006) find that customer satisfaction is an economic asset that brings high returns but surprisingly does not have greater risk. Building on this work, Aksoy et al. (2008) find portfolios based on positive ACSI scores and changes, outperform the market benchmark, and generate abnormal returns. Fornell et al. (2016) discover that using ACSI scores over a 10-year period, customer satisfaction-based investment strategies deliver an abnormal return of 90 monthly basis points and outperform the S&P 500 by 518% between 2004 and 2014. However, other studies have found mixed evidence regarding the predictive power of ACSI for stock performance. Jacobson and Mizik (2009), in particular, argue that ACSI metrics only add value to firms in the computer and internet sectors and, even then, that there is no association between stock returns and ACSI when accounting metrics are considered, hence only acting as a proxy and not holding novel information. Scarcely, reseracher have focused on different customer satisfaction metrics as well. Raithel et al. (2012) use J.D. Power's customer satisfaction measure, and Malshe et al. (2020) use YouGov's measure of customer satisfaction, which is based on daily customer surveys, but only available for paying customers.

Due to the annual nature of the ACSI metric, which provides a limited snapshot of customer satisfaction, and the rapid growth of online platforms and e-commerce, researchers have recognized the need to shift their focus towards online customer reviews for a more comprehensive understanding of customer satisfaction. Online reviews offer real-time, user-generated feedback that reflects the evolving opinions and experiences of consumers over an extended period. For instance, Wu et al. (2020) find that a large dataset of customer product reviews on JD.com [1] positively predicts long-term stock returns at an individual level. In contrast, Tirunillai and Tellis (2012) analyze a dataset of 350,000 online reviews from various websites and find that abnormal returns were influenced by the volume of chatter and reviews but not by numerical customer satisfaction ratings. Building upon the findings of Huang (2018), who studied a dataset of 14.5 million customer product reviews on Amazon.com and discovered that customer satisfaction yields economically and statistically significant abnormal returns, my thesis further explores the idea of consumer opinions value in the market. I incorporate unique Japanese Amazon (Amazon.co.jp) review data and extend the time horizon from 1999 to 2018, providing the most extensive timeframe to date, to study the investment value of such reviews.

Additionally, my research includes a secondary focus on market efficiency and mispricing. In efficient markets, we would expect any information contained in customer satisfaction scores, as they are publicly available, to be quickly reflected in prices. However, attention is a scarce cognitive resource, and limited attention and information processing costs may delay the incorporation of unexpected information from customer reviews into prices (Hong et al., 2000; Hou and Moskowitz, 2005; Peng and Xiong, 2006), resulting in a delayed response (Huberman and Regev, 2001; Dellavigna and Pollet, 2009; Hirshleifer et al., 2009).Fornell et al. (2006) propose that customer satisfaction is and will continue to be systematically undervalued, due to the cost of information attributed to the information gathering outside of ACSI scores. My thesis builds upon that idea, exploring whether the abnormal returns explained above have, have started to diminish due to the reduction of information costs with the proliferation of online reviews, leading to mixed results.

Thirdly, my thesis relates to research on the wisdom of crowds, which suggests that aggregating customer reviews over a long period of time leads to more accurate estimations of product attributes. Differently put, the average

[1] JD.com, also known as Jingdong, is one of China's largest e-commerce platforms, generating $150 billion USD in revenues. It offers a wide range of products, including electronics, clothing, household goods, and more. See https://ir.jd.com/news-releases

estimation of large crowds is likely to be closer to the true value than any individual customer's estimate. Several studies, such as Da et al. (2011), Kelley and Tetlock (2013), H. Chen et al. (2014), and Lee et al. (2015) find evidence that the collective action of large groups of financial market participants conveys information about stock returns and cash flows. However, other researchers find that crowds provide little information about firm fundamentals (Tumarkin and Whitelaw, 2001; Antweiler and Frank, 2001; Das et al., 2005). More recently, crowdsourcing, which involves generating new product ideas by involving users, has shown significant superiority in terms of novelty and customer benefit (Poetz and Schreier, 2012; Bayus, 2013). Additionally, crowd science has demonstrated its potential to make significant contributions to science and deserves more attention (Franzoni and Sauermann, 2014). My thesis, in conjunction with Huang (2018) contributes to this strand of research by further highlighting that aggregated product-related opinions of consumers are an important source of information due to the inherent noise reduction achieved by aggregating opinions.

3. Data and Summary Statistics

This chapter provides information on the data sources and provides figures through tables and graphs of my main variables

3.1 Amazon.com Review data

Amazon.com, the world's largest online retailer, achieved remarkable revenues of $514 billion USD in 2022[2], marking a 500% increase since 2015. Established by Jeff Bezos in 1994, the company has expanded its operations beyond e-commerce to include cloud computing, digital streaming, and artificial intelligence. With over 250 million online reviews on its US website alone, Amazon.com is one of the internet's largest collections of customer product reviews. According to Amazon's guidelines, an Amazon Review is defined as providing "genuine product feedback from fellow shoppers." These guidelines stipulate that Amazon reserves the right to remove reviews under several circumstances, including but not limited to cases where individuals have a financial interest in the product, individuals have close personal relationships with the product's creator, manufacturers misrepresent themselves as impartial consumers, repeated negative reviews are posted by a single customer for the same product, reviews are exchanged for monetary rewards or in-game bonuses, competitors

[2]Annual Report 2022. Retrieved from https://ir.aboutamazon.com/annual-reports-proxies-and-shareholder-letters/default.aspx

post negative reviews, or when artists mutually exchange positive reviews.[3] Customer reviews on Amazon utilize a one-to-five star rating system, with five stars indicating the highest satisfaction level. While customers can include a written review, they may also leave it blank. By recording the date and time of each review, we can track changes in consumer opinion towards a product, brand, and the company over time. The main goal of this thesis is to try to replicate the findings presented in Huang (2018) ”The customer knows best: The investment value of consumer opinions.” To achieve this, the data collection process will closely follow the steps outlined by Huang (2018, p.10-16), aiming to validate and critically discuss the aforementioned results. To establish a connection between a company's stock movement and its products on Amazon, it's crucial to gather Amazon Review Data for all product categories. These products should be linked to their respective brands and the publicly traded companies that own them. Given the extensive search required to identify publicly traded companies selling on Amazon.com, and considering the scope of this thesis, I chose to use the list of 346 publicly traded companies on the New York Stock Exchange (NYSE), American Stoch Exchange (AMEX), National Association of Securities Dealers Automated Quotations (NASDAQ) Huang (2018) as a preliminary reference point[4].

In the second step of data gathering, I collect review data from Amazon.com. In light of time constraints, I utilized an online database for collecting Amazon reviews, assembled by Ni et al., 2019 through data-scraping of the product websites. This data includes the product's name, the brand's name, and the Amazon Standard Identification Number (ASIN) - the latter serving as a unique identifier for each distinct product and for the reviews of that product, the review's date, the numerical star rating, and the Amazon Account Identification (ID) - which acts as the primary identifier for each reviewer. For Amazon.co.jp Japan, I build a scraper that gathers all this information from the product webpage. Cumulatively the Database, it encompasses 233.1 million reviews from May 1996 to October 2018.[5] Adopting a procedure similar to Huang (2018), I dropped duplicate reviews posted by the same reviewer account ID on the same day for the same ASIN.

In the third step of data gathering, I connect the brands from the product reviews with the brands under the ownership of the firms included in the sample, thereby facilitating the connection between the review data and the list of

[3]See: https://www.amazon.de/-/en/gp/help/customer/display.html?nodeId=G3UA5WC5S5UUKB5G

[4](See Table OA1 in **Huang2018INTERNETOpinions** Internet Appendix)

[5]This timeframe surpasses the Huang (2018) timeframe by an additional three years, during which Amazon experienced a growth surge of 117% - soaring from $107 billion USD revenues to $233 billion USD.

sample firms. In light of this thesis's time constraints, I look up leading brands per company based on Google Searches. Table 1 presents the number of reviews, products, brands, and firms within the entirety of the sample, as well as by Fama and French 48 industries, linked over the Standard Industry Classifaction (SIC) number of each company. Ultimately, I gather 27 million reviews of 388,176 products, spanning 736 brands from 231 companies in the timeframe 1999 to 2018. Notably, the total number of brands diverges significantly from the count reported by Huang (2018) (1,931 brands), likely owing to my streamlined method of brand assignment. This simplified strategy for brand association led to the exclusion of approximately 100 sample firms during the review linkage process. Despite this, the volume of reviews and products is well above the requisite for this thesis. The three most reviewed industries were Comput-

	Number of Reviews	Number of Products	Number of Brands	Number of firms
Full Sample	27,679,227	388,176	736	231
By Fama and French 48 industires:				
Computers	4,799,232	35,571	35	12
Consumer Goods	2,850,894	42,613	110	24
Apparel	2,696,756	70,044	56	19
Electronic Equipment	2,659,056	17,821	33	12
Rubber and Plastic Products	2,507,529	19,449	24	3
Machinery	1,787,507	38,193	27	9
Recreation	1,411,431	17,950	23	8
Business Supplies	1,197,582	12,667	34	6
Business Services	1,106,286	13,235	52	20
Restaurants, Hotels, Motels	856,000	10,958	11	7
Food Products	785,211	9,321	47	22
Communication	757,680	29,232	21	7
Electrical Equipment	699,099	7,931	28	6
Construction Materials	624,015	11,474	30	6
Defense	531,588	5,660	15	1
Medical Equipment	529,794	8,198	22	7
Automobiles and Trucks	500,325	18,105	5	4
Printing and Publishing	405,951	4,347	14	1
Wholesale	217,135	1,540	17	2
Chemicals	216,036	1,994	12	6
Pharmaceutical Products	145,992	2,999	51	19
Trading	91,530	2,008	5	1
Beer & Liquor	89,205	2,193	16	3
Retail	76,680	1,475	18	13
Measuring and Control Equipment	39,699	992	2	2
Banking	35,370	866	1	1
Candy & Soda	24,492	596	10	3
Agriculture	12,630	206	2	2
Entertainment	8,598	195	7	2
Aircraft	8,538	185	3	2
Petroleum & Natural Gas	7,386	158	5	1

Table 1
Summary statistics on Amazon.com reviews for public firms
The table presented provides a comprehensive summary of statistics of the data-set of Amazon Customer Reviews, collected by Ni et al. (2019) spanning the period from July 1999 to September 2018. In order to be included in this sample, the firm must be publicly traded on either the NYSE, AMEX or Nasdaq. Furthermore, the requisite companies must have corresponding Compustat and CRSP data available on the WRDS platform, along with the condition to have at least ten customer reviews in a month. This table reports the total number of reviews, products, brands, and public firms. It also provides detailed statistics segregated by the Fama and French 48 industries.

ers (4.8 million reviews), Consumer Goods (2.8 million reviews), and Apparel (2.7 million reviews), with Consumer Goods boasting the most extensive array of brands (110 brands).

3.2 Summary Statistics Replication

In the first part of the data summary for further analysis, I intend to replicate the summary statistics for abnormal return by Huang (2018) to provide an overlook of firm characteristics. I create a dataset of firm months with customer reviews covering 1999 to 2018. The final sample includes 19,386 firm months from June 1999 to September 2018. To be included, a firm must have at least ten reviews in a given month. On average, each month consists of 109 firms meeting this criterion. Comparing my timeframe to that of Huang (2018) (July 2004 to December 2015), I find 13,662 firm months, whereas they identified 20,562 firm months. The difference can be attributed to the smaller number of firms (231) in our sample, as a larger number of firms results in more firm months. I further calculate the average star rating for each firm per month based on all customer reviews for their products. To assess changes in customer perception, I compare this monthly average rating to a benchmark. The benchmark represents the average customer rating of the company's products and overall quality over the past 12 months, reflecting consumer expectations. By taking the difference between the current month's average customer rating and the average rating of the past year, I calculate what Huang (2018) refers to as *"abnormal customer ratings"*. This metric is advantageous as it neutralizes biases in customer reviews, remaining unaffected by systematic biases in reviews, provided those biases remain consistent over time.

I divided the data into three time periods to analyze summary statistics. The first period, P1, spans from July 2004 to December 2015. The second period, P2, covers the years 2004 to 2018. Lastly, P3 encompasses the years 1999 to 2018. The purpose of this division was twofold: to validate the findings of Huang (2018) and to examine the results over different periods. Aggregating over a larger number of consumer reviews over a longer timespan in P2 and P3 can reduce noise and provide more accurate information about the products (Huang, 2018, p.21). P1 corresponds to the time frame used by Huang (2018), while P3 extends from the earliest to the latest month of my data. P2 is a combination of Huang's considerations and my dataset. According to Huang (2018), starting in July 2004, Amazon implemented changes to their review system, such as disallowing anonymous reviews and requiring a credit card for posting reviews. These changes were expected to enhance the informativeness of the reviews. Looking at the makeup of my dataset, I find that 14 million reviews are between June 2004 and December 2015, 13 million reviews after 2015, and only 150.000 before June 2004. This tells us two things: One, that the extra review data of P3 is neglectable, making up approximately 0.5% of my total dataset with limited forecast predictability due to insufficient data and the unverified nature of the reviews. Therefore, similar to Huang (2018), I will only be focusing on P1 and P2 but performing every analysis for P3 as well, the results of which can be seen in the Appendix, although not mentioned in the text directly.

Expected results are taken from Huang (2018) Table 2 Panel A. Table 2 shows the summary statistics for P1. In P1, the mean abnormal customer rating for all reviews is around zero (0.019), with an interquartile range of 0.297, similar to the expected values (0.014 for the mean and 0.260 for the interquartile range). In P2, the mean is 0.007, with an interquartile range of 0.260 (See Appendix Table A.1). The average customer ratings show consistent patterns

Table 2
(P1) 2004-2015: Summary Statistics and determinants of abnormal customer ratings
This table reports summary statistics of the determinants of abnormal customer rating for the timeframe P1: 2004 - 2015. The interquartile range represents the difference between the 75th and 25th percentiles. For variable specification refer to Table A.1 of the appendix.

Variable	N	Mean	Std Dev	25th Percentile	Median	75th Percentile
Customer Reviews						
Average customer ratings	12,875	4.003	0.515	3.769	4.097	4.333
Abnormal customer ratings	12,875	0.019	0.374	-0.116	0.028	0.181
# of customer reviews per month	12,875	1,094,175	2,724,370	45,000	189,000	810,000
Firm-level characteristics						
Market cap (millions of dollars)	12,875	29,397,854	67,344,193	1,429,739	4,912,051	16,627,291
Book-to-market	12,875	0.424	0.391	0.221	0.350	0.548
Stock return$_{m-12,m-1}$	12,875	0.123	0.599	-0.130	0.076	0.280
R&D Ratio	12,875	0.021	0.026	0.008	0.016	0.026
Gross profitability	12,875	0.046	0.111	0.012	0.040	0.083
F-score	12,875	5,368	1,637	4,000	6,000	7,000
Dollar volume (millions of dollars)	12,875	4,825,937	13,434,287	250,381	1,140,181	3,577,399
CV of dollar volume	12,875	0.311	0.158	0.210	0.273	0.363
Book leverage	12,875	0.316	0.351	0.085	0.272	0.434
Asset tangibility	12,875	0.164	0.117	0.086	0.136	0.210
# of Analysts	4,610	18,307	16,191	7,000	14,000	24,000
Analyst Revision (percent)	8,976	-0.036	0.900	-0.001	0.000	0.001
Revenue surprise (SURGE)	4,457	-0.118	6.894	-1.785	-0.091	1.519
Earnings surprise (SUE) (percent)	4,100	0.052	0.284	0.001	0.031	0.094

over time: 4.003 (P1) and 4.049 (P2) with interquartile ranges of 0.564 (P1) and 0.507 (P2) (expected: average ratings of 4.1, interquartile range of 0.480). Similarly, the distribution of ratings is mostly centered around 5-star ratings, accounting for approximately 62-64% (see Figure: A.4 in the Appendix). The prevalence of positive reviews can be attributed to three factors. Firstly, products listed on Amazon.com may be perceived to have higher quality, creating a selection effect. Secondly, low-rated products gradually exit the market due to declining sales - a survival effect. Lastly, there may be a systematic bias toward positive evaluations in reviews on Amazon.com. However, abnormal review ratings neutralize time-constant biases, ensuring that our measures remain unaffected (Huang, 2018). In the second phase of replicating summary statistics on abnormal returns, I create firm-level variables related to stock return predictability. These variables are later used to regress abnormal customer ratings on the firm-level factors mentioned earlier. The methodology used in this phase closely follows the approach described by Huang (2018).To obtain the required data, monthly stock return, volume, and price information is acquired from the Center for Research on Stock Prices (CRSP) via the Wharton Research Data Service (WRDS). Quarterly financial statement data from Compustat, as well as analyst numbers and earnings forecast information from the Instituional Broker's Estimation System (I/B/E/S), are also sourced from WRDS. In line with Huang (2018), this study incorporates various firm characteristics likely to be associated with stock return predictability and consumer opinions. These additional characteristics address concerns that customer ratings may merely reflect firms' operating performance, which can already be inferred from financial statements. To mitigate this concern, the study constructs two specific variables: the F-score (Piotroski, 2000) and the gross profitability metric (Novy-Marx, 2013) as cited by Huang (2018). The F-score encompasses nine binary variables that capture financial performance indicators. Gross profitability is determined by dividing income before extraordinary items by the book value of assets. Furthermore, I take into account Research & Development (R&D) expenditures, as these factors can potentially influence consumer opinions through firms' R&D investments (Huang, 2018). Additionally, the level and variability of dollar trading volume, as proposed by Brennan et al. (1998), are considered to control for any potential attention effect.

The summary statistics for my sample of 236 firms in P1, P2 are presented in Table 2 and the Appendix Table A.1, respectively. The average market capitalization of the firms ranges from $28 billion (P1) to $32 billion (P2), significantly higher than the expected $25.72 billion. The disparity could be attributed to the exclusion of lower-capitalization firms in my sample due to the streamlined brand linkage method. Across P1 and P2, the book-to-market ratio remains around 0.4, and the F-Score stands at 5.4, in line with the expected values. Moreover, the monthly trading volume for the entire period is approximately $4.8 billion, with an analyst coverage of around 18,5, and the buy-and-hold returns are approximately 11%-12%, all higher than the expected values by

Huang (2018) (Monthly Trading Volume: \$3.7 billion, buy-and-hold returns: 17%, Analyst Coverage: 16.4). However, this difference could be explained by the smaller size of our firm sample (231), leading to the exclusion of predominantly smaller firms, with on average smaller average monthly trading volume and average analyst coverages. The lower buy-and-hold returns can also be attributed to the omission of smaller firms from the sample, which on average achieve a higher return (Fama and French, 1992), reducing the total average Identical to Huang (2018), I construct two cash flow surprises measures. First, I construct a revenue surprise measure using a Standard Unexpected Revenue Growth Estimator) (SURGE). SURGE of firm *i* in quarter *q* is defined as

$$\text{SURGE}_{i,q} = \frac{\text{REV}_{i,q} - E(\text{REV}_{i,q})}{\sigma_{i_q}} \tag{3.1}$$

where $\text{REV}_{i,q}$ is the revenue per share from the quarterly earnings for firm *i* in quarter *q*, $E(\text{REV}_{i,q})$ is the expected quarterly revenue per share and σ_{i_q} is the total standard deviation of quarterly revenue growth of a firm over the entire time horizon. I also estimate $E(\text{REV}_{i,q})$ under the assumption that the quarterly revenues follow a seasonal random walk with drift (Huang, 2018, p.14):

$$E(\text{REV}_{i,q}) = \text{REV}_{i,q-4} + \frac{1}{8}\sum_{j=1}^{8}(\text{REV}_{i,q-j} - \text{REV}_{i,q-j-4}) \tag{3.2}$$

The revenue surprises (SURGE) for the corresponding periods are -0.118 and 0.072 (P1, P2), respectively. These figures substantially deviate from the expected 0.52. It should be noted that Huang (2018) did not specify the timeframe for calculating the standard deviation of quarterly revenue growth, but the deviation could also result from a difference in sample data. In my study, the standard deviation was calculated across all quarters. Constructing the second measure of cash flow surprises, I calculate the price-scaled forecast error as the Standardized Unexpected Earnings) (SUE) in a manner consistent with Huang (2018). For each quarter, the *SUE* score for each company is extracted from the I/B/E/S on WRDS. I find *SUE* scores of approximately 0.05%, with an interquartile range approximating 0.09, in line with expectations. Due to the time limitations and the defined scope of this study, I have abstained from constructing a measure for hedge fund trading activities.

To ensure that the abnormal customer ratings reflect new information rather than merely capturing lagged stock characteristics, I adopt the approach employed by Huang (2018) and regress abnormal customer ratings on these lagged stock characteristics, including R&D ratio, gross profitability, the level and variation of dollar trading volume, F-Score, book-to-market ratio, past stock returns, market capitalization, book leverage, asset tangibility, and analyst coverage, not including advertisement spending due to the unavailability of such information on Compustat. I also incorporate analyst forecast revisions into the regression models to investigate the potential impact of consumer opinions on certain market participants. Analyst forecast revisions represent the difference between the average current-year Earnings Per Share) (EPS) forecasts during a specific month and the previous month, scaled by the sock price of the previous month. I perform the regression using four different specifications, gradually adding more variables to each regression. However, I encountered problems when adding the R&D ratio due to the limited availability of R&D data on Compustat, leading to a significant reduction in my observation sample when added. To address this issue, I conducted three separate regressions for the timeframe P1 (2004-2015). One regression used the available values of the computed R&D ratio (R1), another excluded the R&D ratio entirely (R2), and a third filled in missing R&D values through time-series interpolation (R3). Table 3 reports the regression results of abnormal customer ratings on lagged stock characteristics, with each column in each regression (R1-R3) adding different sets of variables. The three regressions show minimal variation, primarily affecting the significance levels of single variables. Columns (2) and (3), R1 shows a significant correlation (t-values of 1.84) between Asset Tangibility and one-month-ahead abnormal customer p-values at the 10% level, but I find no significance in R2 and R3 (t-values between 1.2 and 1.5). This discrepancy could be attributed to two reasons: one, the noise in the sample data R1, due to the substantial reduction in the sample size; two, the potential relationship between the R&D ratio and abnormal customer ratings, which may only be captured when using the actual R&D ratio values that have a relationship with Asset Tangibility. The analysis shows that most stock or firm characteristics have limited predictive power for abnormal customer ratings, However, there is a notable distinction for the Coeffiicent of Variation of dollar volume (CV). R1, R2, and R3 indicate significant negative predictability of one-month-ahead abnormal customer ratings, with levels of significance observed between 10% and 5%. This differs from Huang's results, which predict positive significance only in column (4) (significance level 10%). This difference could be due to the correlation between trading volume variation and abnormal customer ratings or the absence of certain variables in my analysis, such as Advertising, Institutional Ownership, and Net buying by hedge funds. In the regressions for period P2, a significant correlation between Asset Tangibility and one-month-ahead abnormal customer ratings is observed, but its statistical significance fluctuates. Additionally, there is a statistically significant correlation at the 1% or 5% significance levels for the variation in trading volume across all regression models. Furthermore, there is a significant correlation between Book-leverage and one-month-ahead abnormal customer ratings. These differences may be attributed to the introduction of new data in the years following Huang's study or the absence of certain variables in my regression models. The rest of the differences between P1, P2, and the expected values are either insignificant in magnitude or statistical relevance. From these observations, it can be inferred that contrary to Huang (2018) findings, abnormal customer ratings may not be entirely independent of all the information embodied by stock and firm characteristics. Particularly, the variation of dollar trading volume in a given month seems to have a highly significant negative correlation with one-month-ahead abnormal customer ratings. Nevertheless, similar to Huang (2018, p.16) conclusions, it is clear that analysts are unable to predict customer opinions with certainty.

4. Empirical Results

This chapter presents the empirical results of my replication of Huang's (2018) study of the effect of customer reviews on stock performance and stock return predictability, the extention on Amamzon reviews in Japan, and the changes in predicability over time

4.1 Abnormal Customer Ratings and Stock return predictability

I employ a calendar time portfolio (Jaffe, 1974; Huang, 2018) to investigate the potential return predictability and investment value associated with abnormal customer ratings on Amazon.com in period P1: July 2004 to December 2015. For each month within this timeframe, the available stocks from my firm-months are classified into tercile portfolios based on their respective abnormal customer ratings. The performance of these portfolios is then tracked over the subsequent month. The bottom tercile portfolio (T1) exhibits a mean abnormal customer rating of -0.311; the middle tercile portfolio (T2) manifests a mean abnormal customer rating of 0.032, and the top tercile portfolio (T3) presents a mean abnormal customer rating of 0.344[6]. In instances where firm months possess only one or two abnormal customer rating values, these values are disregarded, and the respective firm months are not allocated to any portfolio to limit bias towards any portfolio. In terms of portfolio construction, I deploy two weighting strategies that follow the methodology set forth by Huang (2018). The first approach assigns weights based on the number of reviews, while the

[6]See Appendix Table A.5 for more detailed summary Statistics

Table 3

P1: Regressions of one-month ahead abnormal customer ratings on firm characteristics

Regressions of one-month ahead abnormal customer ratings on firm characteristics in period P1: July 2004 - December 2015. R1 shows the regression with actual computed R&D ratio data, R2 excluding R&D ratio data, and R3 using a function to interpolate R&D ratio data in a time series. The t-statistics are provided in parentheses. Stars denote statistical significance: * indicates significance at the 10%level, ** at the 5% level, and *** at the 1% level

	R1				R2				R3			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Log(Market cap)	0.001	0.003	-0.005	-0.004	0.001	0.003	-0.004	-0.002	0.000	0.002	-0.002	-0.001
	(1.147)	(1.790)*	(-0.807)	(-0.464)	(1.147)	(2.131)**	(-0.802)	(-0.336)	(0.354)	(1.304)	(-0.525)	(-0.253)
Book-to-market	-0.005	-0.005	-0.003	0.014	-0.005	-0.004	-0.001	0.003	-0.006	-0.005	-0.003	-0.003
	(-0.600)	(-0.464)	(-0.307)	(0.992)	(-0.600)	(-0.425)	(-0.117)	(0.287)	(-0.743)	(-0.541)	(-0.346)	(-0.265)
Stock return$_{m-12,m-1}$	0.002	0.005	0.008	0.010	0.002	0.002	0.005	0.004	0.004	0.004	0.006	0.006
	(0.358)	(0.716)	(1.192)	(1.026)	(0.358)	(0.636)	(1.046)	(0.704)	(0.629)	(0.636)	(1.046)	(0.704)
Book leverage	0.001	-0.013	-0.013	-0.008	0.001	0.003	0.003	0.002	-0.003	-0.002	-0.003	-0.004
	(0.116)	(-0.921)	(-0.880)	(-0.476)	(0.116)	(0.309)	(0.227)	(0.123)	(-0.272)	(-0.187)	(-0.268)	(-0.293)
Asset tangibility	0.032	0.074	0.074	0.050	0.032	0.036	0.038	0.029	0.044	0.048	0.047	0.038
	(1.025)	(1.849)*	(1.842)*	(1.045)	(1.025)	(1.526)	(1.490)	(1.001)	(1.418)	(1.526)	(1.490)	(1.001)
Log(1+# of analysts)	-0.001	-0.001	-0.005	-0.011	-0.001	-0.002	-0.006	-0.009	0.003	0.002	0.001	-0.001
	(-0.138)	(-0.143)	(-0.585)	(-1.042)	(-0.138)	(-0.447)	(-0.912)	(-1.064)	(0.785)	(0.527)	(0.180)	(-0.236)
R&D Ratio		-0.069	-0.078	-0.010						0.070	0.060	0.095
		(-0.374)	(-0.419)	(-0.046)						(0.51)	(0.441)	(0.577)
Gross profitability		-0.001	-0.011	-0.042		0.024	0.013	-0.016		0.016	0.005	-0.022
		(-0.038)	(-0.308)	(-0.965)		(0.491)	(0.155)	(-0.398)		(0.492)	(0.155)	(-0.555)
F-Score		-0.006	-0.006	-0.006		-0.004	-0.004	-0.007		-0.004	-0.004	-0.006
		(-1.957)*	(-1.931)*	(-1.803)*		(-1.882)*	(-1.772)*	(-2.331)**		(-1.882)*	(-1.772)*	(-2.331)**
Log(CV of dollar volume)			-0.023	-0.034			-0.025	-0.027			-0.026	-0.030
			(-1.836)*	(-2.305)**			(-2.874)***	(-2.677)***			(-2.874)***	(-2.677)***
Log(Dollar volume)			0.007	0.006			0.002	0.002			0.002	0.002
			(1.087)	(0.697)			(0.471)	(0.387)			(0.471)	(0.387)
Analyst Revision				-0.002				-0.003				-0.004
				(-0.485)				(-0.883)				(-0.882)
Number of Observations	12,352	7,156	7,143	5,155	12,352	12,343	12,323	8,768	13,099	13,090	13,054	9,343
Adjusted *R* Squared	0.003	0.003	0.003	0.003	0.003	0.003	0.004	0.003	0.003	0.003	0.003	0.003

second utilizes an equal weighting scheme. The review-weighted strategy is based on the concept that larger crowds can potentially provide more precise information. Moreover, this approach mitigates the risk of the portfolios being predominantly dominated by firms with relatively few reviews, which could amplify the noise in the data (Huang, 2018, p.16). To adjust returns, I employ the Fama-French Carhart four-factor model.

Table 4 presents the four-factor-alpha and factor loadings in P1 from the monthly calendar-time portfolio regression. In line with Huang (2018) I compute the four-factor alpha by regressing monthly portfolio excess returns on the expected monthly returns of risk factors. On average, the review-weighted portfolio outperforms the passive benchmark by approximately 77 basis points per month (significant at the 5% level, t-value 2.41). Similarly, the equally weighted portfolio also outperforms the passive benchmark by around 53 basis points per month (insignificant). This differs from Huang (2018), who finds significance at the 5% level for the equal-weighted top tercile at 55 basis points per month (t-value 2.52). The differecne could be due to the reduced sample size of firms or potential noise from firms with few reviews. Comparing the number of reviews per firm month, my findings indicate an average of approximately 400 more reviews per month compared to Huang (2018) and a 75th percentile at more than double (1100 vs. 450). This suggests that my sample includes fewer firms with few reviews per month and more with more reviews per month. This could lead to different weighings of portfolios in the equal-weighted scheme. Contrary to Huang (2018), my research does not support the claim that stocks with low customer ratings in the bottom tercile (T1) underperform the passive benchmark, but both resutls are statistically insignificant.

In the next step of my analysis, I create Long/Short spread portfolios, buying stocks in the top tercile (T3) of abnormal customer ratings and selling stocks

Table 4
Calendar-time portfolio returns
This table presents the results of calendar-time portfolio regressions. The analysis covers the period P1 (July 2004 to December 2015). To construct the portfolios, I sorted the sample stocks into three terciles based on their abnormal customer rating. The abnormal customer rating is calculated as the difference between the average customer rating in the current month and the average rating in the previous 12 months. The performance of these portfolios was then tracked over the subsequent month. Two weighting schemes were employed: weighting by the number of reviews and equal weighting across stocks. To adjust returns, the Fama-French-Carhart four-factor model was used. The reported alpha estimates and factor loadings were obtained by regressing the monthly portfolio excess returns on the monthly returns from the risk factors. Additionally, a "Long/short" spread portfolio was created by buying the top tercile portfolio and selling the bottom tercile portfolio. The t-statistics are provided in parentheses. Stars denote statistical significance: * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Alpha %	**Market**	**SMB**	**HML**	**UMD**
Panel A: Review weighting					
T1 (low abnormal rating)	0.073%	1.276	0.209	0.140	-0.064
	(0.24)	(15.17)***	(1.42)	(1.01)	(-0.86)
T2	0.233%	1.207	0.495	0.177	-0.154
	(0.86)	(16.44)***	(3.85)***	(1.46)	(-2.38)**
T3 (high abnormal rating)	0.768%	1.231	0.226	0.183	-0.238
	(2.41)**	(14.31)***	(1.50)	(1.29)	(-3.14)***
Long/Short (high - low)	0.695%	-0.047	0.018	0.04349	-0.174
	(1.427)	(-0.37)	(0.07)	(0.20)	(-1.50)
Panel B: Equal weighting					
T1 (low abnormal rating)	0.286%	1.090	0.390	0.065	-0.189
	(1.73)*	(24.39)***	(4.99)***	(0.87)	(-4.80)***
T2	-0.108%	1.210	0.437	0.174	-0.169
	(-0.52)	(21.50)***	(4.44)***	(1.87)*	(-3.41)***
T3 (high abnormal rating)	0.318%	1.092	0.481	0.003	-0.126
	(1.55)	(19.73)***	(4.97)***	(0.03)	(-2.58)**
Long/Short (high - low)	0.033%	0.181	0.091	-0.062	0.063
	(0.14)	(0.03)	(0.84)	(0.60)	(1.17)

in the bottom tercile (T1). I apply this strategy to equally and review-weighted portfolios. Similar to Huang (2018), this Long/Short portfolio outperforms the passive benchmark by 3 to 70 basis points per month without significant exposure to risk factors. However, different to Huang (2018) I do not find significance. This difference could be due to various factors, such as variations in review and firm sample sizes, out-of-sample over-performance of Huang's portfolios, or potential under-performance of mine in that timeframe.

Despite my inability to find a statistically significant alpha for the spread portfolios, I did identify a significant four-factor alpha for the review-weighted portfolio. To shed more light on the reasons behind this predictability of future stock returns based on abnormal customer ratings, I conducted two subsample analyses following the approach by Huang (2018). Firstly, I examined the limits of arbitrage and their impact on investors' ability to use information effectively, as discussed in the studies by Pontiff (1996) and Shleifer and Vishny (1997) cited in Huang (2018). I obtained idiosyncratic historical return volatility to measure arbitrage from the Beta Suite Database on I/B/E/S. Following the methodology outlined by Ang et al. (2006) [7], this idiosyncratic volatility data is calculated as the standard deviation of the residuals from the Fama and French (1992) three-factor model. I divided the sample stocks into two groups based on the median quarterly idiosyncratic volatility. I construct two spread portfolios that buy the high abnormal ratings portfolio (T3) and sell the low abnormal ratings portfolio (T1) in each of the volatility groups and construct one Long/Short portfolio based on high minus low idiosyncratic volatility during P1. I find that in P1 high idiosyncratic volatility yields Fama-French-Carhart four-factor alphas ranging from 11 to 105 monthly basis points, notably higher than the 7.7 to 28 basis points for low idiosyncratic volatility stocks. However, these differences are not statistically significant, deviating from the findings of Huang (2018). Consequently, there is inconclusive evidence regarding the role of arbitrage costs in generating return predictability. To investigate further, I examine the individual grouped tercile returns. In a Long/Short spread (T3 high - T3 low), the top tercile (T3) with high idiosyncratic volatility outperforms the top tercile (T3) with low idiosyncratic volatility by 14 to 56 monthly basis points, which is statistically significant at the 5% level. This suggests that limits to arbitrage (i.e. arbitrage costs) play a role in generating return predictability[8]. One possible explanation for the lack of statistically significant alphas in the idiosyncratic volatility spread portfolios could be the sample effect of my firms. As previously discussed, my firm sample differs slightly from the one used by Huang (2018) in terms of the

[7] Beta Suite Documentation Data Manual, WRDS Research Team, August 2021

[8] See Appnedix Table A.6 for a more detailed breakdown

number and size of firms. My data primarily includes larger firms due to excluding some smaller firms during the data-processing and linking step. According to Ang et al. (2006, p.284), smaller stocks tend to have higher idiosyncratic volatility than larger stocks. Therefore, my sample may exclude even higher idiosyncratic stocks, which may explain the deviation due to a change in the grouping of firms.

Secondly, I focus on the idea that constraints on investor attention and information processing capabilities could potentially slow down the incorporation of customer information into the stock price. Following Hirshleifer and Teoh (2003) and Hong et al. (2000), as cited in Huang (2018), I use analyst coverage and firm size as proxies for investor attention. If investor inattention or limits to investor attention slow down the incorporation of the information embedded in consumer reviews, then the predictability of stock prices should be concentrated among firms with low analyst coverage and small market capitalization. For example, Li (2020) points out that analyst coverage is negatively correlated to equity misevaluation, therefore strengthening market efficiency. I compute the number of financial analysts making forecasts for a sock in a given quarter by retrieving quarterly earnings forecasts of the sample stocks from the I/B/E/S historical database and aggregating the unique Analysts for the given stock in that quarter. For each quarter, I construct the median number of analysts covering a stock, then partition the sample of all stocks into two groups: one higher and one lower than the median. Within each group, I construct the spread portfolios and a Long/Short (low - high analyst coverage) portfolio. Consistent with the findings outlined by Huang (2018), the primary return predictability lies in low analyst-coverage stocks, yielding a monthly alpha of 13 to 130 basis points (5% significance). However, I also find a high analyst-coverage portfolio alpha of 89 monthly basis points (significant at 10%) and, unlike Huang (2018), the equal-weighted portfolio with low analyst coverage does not exhibit a statistically and economically significant alpha. This disparity could be attributed to a smaller sample size and a higher review count. My sample predominantly comprises larger firms, with an average of 18.3 analysts per stock, compared to Huang's 16.4. The high coverage alpha is lower than the low-coverage alpha, aligning with the notion of market efficiency, as the difference in alpha between high and low analyst coverage suggests more effective information integration by analysts. The difference in alpha may also stem from over- or underreaction to abnormal customer rating information. High coverage could lead to overreaction, resulting in temporary price inflation and subsequent correction, as indicated by Ma et al. (2005), thereby yielding a lower alpha. Similarly, for low-coverage stocks, there could be an underreaction to positive news, leading to a slower stock price adjustment and a higher alpha. However, the alpha spread between low and high coverage, ranging from 14 to 41 monthly basis points, is statistically insignificant.

Third, I use Market capitalization testing as a proxy for investor attention is similarly done, creating two groups: small and large. Small cap firms fall below the median market capitalization, and large ones above. Despite the assumption that smaller firms should generate more alpha due to less investor attention, I find that large firms mainly drive predictability with 29 to 110 monthly basis points, contrary to Huang (2018), who finds small firms concentrate the alpha. This difference could be due to variations in the datasets. My sample firms, larger on average, might lead to different portfolio groups based on quarterly median market capitalization. I find a market capitalization 25th percentile of $1.5 billion compared to $1 billion by Huang and a 75th percentile of $18 billion compared to $23 billion by Huang. Therefore, my firm sample might not be small enough to capture the concentration of small firms, and not big enough to capture the full scale of larger firms, potentially skewing the results.[9]. The results show that customer satisfaction scores from Amazon.com economically and statistically significantly predict returns.

4.2 Diminishing Effect of Customer Ratings Post 2015

To investigate the diminishing effect of customer ratings due to the increasing abundance of online reviews, I conduct tercile portfolio regressions on the P2 (2004-2018). Again,the two weighing mechanisms are employed: review weighted and equal-weighted. Returns are adjusted using the Fama-French Carhart four-factor model (Carhart, 1997). To track the change in the impact of customer satisfaction scores, measured as the abnormal customer ratings variable, and assess the potential reduction in mispricing over time, I construct boolean dummy variables: Pre 2015 and Post 2015. These variables replace the constant in the regression, where Pre 2015 is assigned a value of 1 for firm months before December 2015 and 0 thereafter, while Post 2015 is assigned a value of 1 for firm months after December 2015 and 0 before. The cut-off point of December 2015 aligns with the end of the reviewed period in Huang (2018). Table 5 presents the results of the calendar time portfolio regression using the dummy variables. Consistent with previous findings, the review-weighted high abnormal customer ratings tercile (T3) outperforms the passive benchmark by 73 monthly basis points at a 5% significance level.

[9] For a more detail display See Appendix Table A.7

Table 5
Pre and Post 2015 Portfolio Analysis
Tercile portfolio regression in P2: 2004 - 2018 with two boolean dummy variables: Pre 2015 and Post 2015. Pre 2015 is assigned a value of 1 for firm months before December 2015 and 0 thereafter, while Post 2015 is assigned a value of 1 for firm months after December 2015 and 0 before Stars denote statistical significance: * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level

	Pre 2015	Post 2015
Review Weighing		
T1 (low abnormal rating)	0.06%	-0.90%
	(0.21)	(-1.42)
T2	0.30%	0.66%
	(1.12)	(-0.66)
T3 (high abnormal rating)	0.73%	0.23%
	(2.47)**	(0.38)
Long Short (T3-T1)	0.67%	1.13%
	(1.44)	(1.19)
Equal Weighing		
T1 (low abnormal rating)	0.24%	-0.47%
	(1.34)	(-0.47)
T2	-0.05%	0.48%
	(-0.25)	(-0.48)
T3 (high abnormal rating)	0.32%	-0.15%
	(1.59)	(-0.36)
Long Short (T3-T1)	0.08%	0.32%
	(0.36)	(0.69)

The equal-weighted portfolio also outperforms the benchmark, although not significantly, possibly due to data noise caused by smaller variance in review amounts per month affecting portfolio weighing. In contrast, all terciles in the Post 2015 period exhibit economically small or negative alphas without statistical significance. This could be attributed to the insufficiency of post-2015 data, with only three additional years included. However, the review amount between June 2004 and December 2015 is 16 million, while the post-2015 period has 13 million reviews.
Thus, the data should be sufficient to reduce noise and leverage the wisdom of crowds phenomenon (H. Chen et al., 2014). Alternatively, the mispricing effect may have diminished over time, similar to the decline of the Monday Effect. Specialized hedge funds have already incorporated customer satisfaction data into their investment strategies since 2015 (Huang, 2018). With the rapid growth of e-commerce, it is reasonable to assume that more trades are now influenced by customer satisfaction metrics. Notably, the high abnormal customer rating portfolio (T3) still outperforms the Standard & Poor's 500 (S&P 500) Index from December 2015 to December 2018 by 40% (refer to Appendix Figure A.5), indicating a potential positive impact of abnormal customer ratings or a selection bias of firms that trade on Amazon.com Nevertheless, these results suggest a decline in mispricing associated with customer satisfaction but warrant further research.

4.3 Case Study: Japanese Amazon.com Customer Ratings

Expanding the idea and findings of Huang (2018), I employ a novel dataset of 190 thousand reviews from Amazon.co.jp[10]. I retrieve this data from four companies across 2.500 products across the timeframe of Apirl 2008 to June 2023. Table 6 shows the summary statistic for the four companies. My decision on which to include came down to a few factors. For one, I calculated the ten-year daily return volatility of all stocks on the Tokyo Stock Exchange (TYO) using data from the Yahoo Finance API, sorting to acquire only companies in the Retail Trade, Electrical Appliances, Chemicals, and Others Industries, as these are most likely to trade on Amazon.co.jp. From there, I picked those which had readily available brand data on their respective company web pages and moderate 10-year volatility. I focus on return volatility due to the Japanese Stock Market being dominated by low-volatility stocks (Katsura and Shi, 2018; Takkabutr, 2013), which makes it hard to show any relationship, as I showed earlier that abnormal return is concentrated in high-volatility stocks. But, only choosing high-volatility stocks could introduce selection bias, and choosing excessively low-volatility stocks could lead to lower inferences about abnormal returns. The scope of this data is, however, not enough to perform a rigorous

Table 6
Case Study: Public firms trading on Amazon.co.jp
This table reports the summary statistics for the five companies chosen for the timeframe April 2008 to June 2023. Volatility is calculated as the daily return volatility over the past ten years. Average is the average abnormal customer rating over the firm months of the company. All variables are winsorized at the 0.1% and 99.9% levels to minimize the impact of outliers.

Company Name (TYO Ticker)	Volatility	Reviews	Brands	Average	Sector
Fujifilm Holding Coporation (4901)	0.225	66,813	663	0.006	Chemicals
Kyocera Corporation (6971)	0.245	48,257	570	0.002	Electric Appliances
Kao Corporation (4452)	0.184	67,586	1,212	0.060	Chemicals
Unicharm Korporation (8113)	0.188	7,411	103	0.012	Chemicals

statistical analysis. Instead, I perform a correlation analysis. Again, the firm must at least have ten reviews in a given month to be included in the correlation analysis to reduce noise introduced by singular reviews. Figure 1 shows the relationship between the abnormal customer ratings of a given firm and their stock price change. The correlation coefficient ranges from 0.02 to 0.25, with a weighted average correlation coefficient of around 0.14, suggesting at face value only a weak to moderate correlation between consumer opinions and the firm's stock movement.

[10]As a reference, Amazon Japan contributes around 5% to the total revenues of Amazon, compared to 68% from the United States. See https://businessquant.com/amazon-revenue-by-region

However, there are two things two note: One, abnormal customer ratings are just one of a myriad of information sources that affect stock pricing, meaning an average correlation of 0.14 can be considered quite high. Second, it is important to note that correlation does not imply causation, and other factors may be contributing to the observed relationship. Nonetheless, I perform a rudimentary regression with tercile portfolios, categorized into high, normal, and low abnormal customer ratings as I did before. To incorporate the idea that, due to the reduction of information costs, the mispricing of customer satisfaction scores has diminished over time, I similarly employ two boolean dummy variables: Pre- and Post 2015. I find that the review-weighted top tercile with high abnormal customer ratings (T3) outperforms the passive benchmark by 166 monthly basis points Pre 2015, significant at the 10% level, but no significant abnormal returns Post 2015 (See Appendix Table A.8). This result supports my findings from before that mispricing of customer satisfaction has diminished over time and shows that Amazon.co.jp online customer reviews hold investment value.

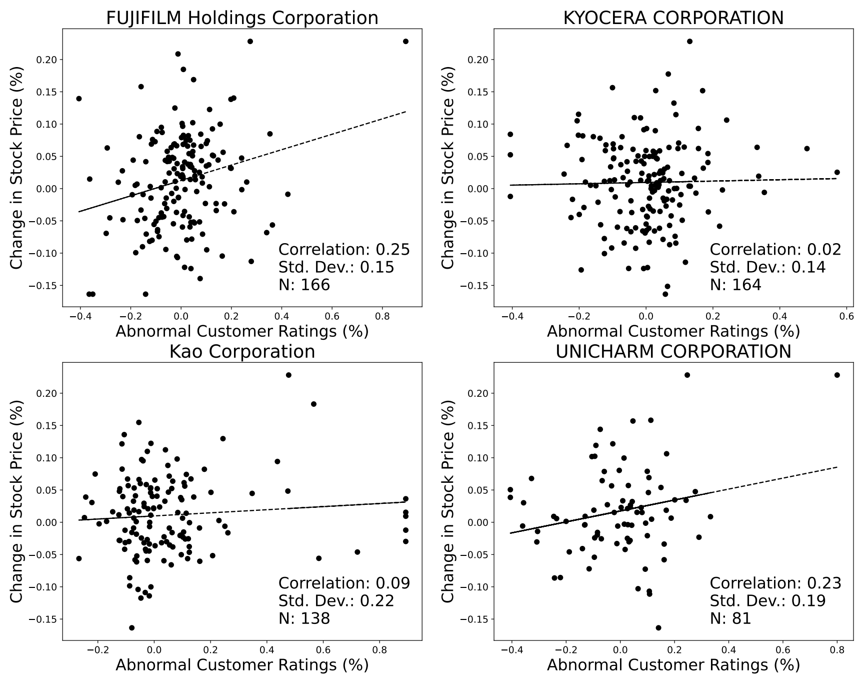


Figure 1: Japan Abnormal Customer Ratings Scatterplot
This figure uses a scatterplot to visualize the relationship between the abnormal customer ratings in a given firm month and the price change of the given stock, calculated as the percent difference between last month's close price and the current firm month's stock price in the timeframe April 2008 to June 2023. *Correlation* is the correlation coefficient between abnormal customer ratings and stock price changes, *Std. Dev* is the standard deviation of the abnormal customer ratings and *N* the number of firm months used to consturct the plot. All values are winsorized at the 0.1% and 99.9% level to minimize the impact of outliers.

5. Conclusion

This paper presents the examination of the idea that online consumer satisfaction scores using customer product reviews from Amazon.com and Amazon.co.jp hold investment value and return predictability. The findings reveal that abnormal customer ratings positively predict subsequent stock returns during the period from June 2004 to December 2015, but point to a disappearance post-2015. A portfolio constructed based on high abnormal customer ratings generates a monthly abnormal return ranging from 32 to 77 basis points, primarily concentrated among stocks characterized by high idiosyncratic volatility, low analyst coverage, and large firm size. These findings support the notion that such stocks may encounter high arbitrage costs and face more limitations to investor attention. Furthermore, a case study analyzing 200,000 Amazon Japan customer reviews between April 2008 and June 2023, covering four companies traded on the TYO, reveals a weighted average correlation coefficient of 0.14 between abnormal customer ratings and stock price changes for a given firm. Before 2015, a review-weighted portfolio of high abnormal customer ratings outperforms the passive benchmark by 166 monthly basis points. However, this outperformance becomes statistically insignificant post-2015. The disappearance of the abnormal customer ratings' predictive power in recent years may be attributed to a decreasing cost of information.

Despite the insightful findings, several limitations should be acknowledged. Firstly, the analysis relies on datasets obtained solely from Amazon.com and Amazon.co.jp, raising concerns about the representativeness of the results. Future research should incorporate data from a more diverse range of online marketplaces and platforms to enhance the generalizability of the findings. Secondly, the sample of firms used to analyze Amazon American customer reviews is biased towards larger companies, potentially impacting results such as the observed concentration of abnormal returns among larger firms or equally weighted portfolios. Finally, the evidence for the Japanese market is relatively limited due to the modest size of the firm and the review data analyzed. In conclusion, this study underscores the potential value of aggregated customer opinions in predicting stock returns. Future research should address the aforementioned limitations by employing larger samples with more recent customer satisfaction scores to investigate the evolving mispricing effect of customer reviews. Additionally, exploring different markets beyond American stocks would contribute to a more comprehensive understanding of the relationship between customer satisfaction and stock pricing.

Appendices

Merging Amazon Review Data with CRSP/Compustat data

Linking Amazon.com review data with CRSP/Compustat data is not a simple task, considering the scope of this study. Amazon.com only provides brand or company names without any additional identifiers. To obtain unique firm identifiers like Compustat's GVKEYS or CRSP's PERMNO/PERMCO keys, I rely on two methods. In the first case, for the main dataset used to replicate the results of Huang (2018), I retrieve a list of companies from the Internet Appendix Table IA-1 (Huang, 2018). This table provides a list of publicly traded firms on NYSE, AMEX, or Nasdaq that have customer reviews on Amazon.com. up to 2015. I choose the apporaoch of first identifying the companies and then their brands due to the time limitations of this paper. I convert the online table into a dataframe using the tabula library in Python. However, these company names also lack unique identifiers. To link the provided company names, I obtain the company name-PERMCO combinations from the *CRSP Tools - Translate to PERMCO/PERMNO* WRDS database for NYSE, AMEX, and NASDAQ between 1998 and 2018. After obtaining the Company Name-PERMCO linkage, I develop an algorithm that compares the company names from both the Internet Appendix table and the company name-PERMCO linkage table, matching the names that have more than 75% overlap. Subsequently, I review the matched dataset and make manual adjustments for individual companies. In cases where a GVKEY for a particular company is not available due to mergers and acquisitions, for example, I obtain the GVKEY for the parent company instead. Finally, I match the PERMCO and GVKEY using the CRSP/Compustat merged database on WRDS.

For the dataset of Amazon.com reviews obtained from Ni et al. (2019), however, only the brand name of the product is provided. Therefore, I create a list of brands for each company from the table using Google Search, AI-Technologies, and Wikipedia. I use this list to filter my review dataset, reducing the total sample size from 233 million to around 27 million reviews, which I can then link with stock performance using the previously obtained GVKEY and PERMCO.

In the second case, for the dataset of Japanese companies, the method is not as straightforward. First, I obtain a list of all 3,880 companies listed on the Tokyo Stock Exchange (東京証券取引所, Tōkyō Shōken Torihikijo, abbreviated TSE/TYO) from the official list of the Japanese Exchange Group (JPX), due to my restricted access to the PACAP Monthly Key Economic Statistics Dataset on WRDS. From there, I retrieve volatility data for each company over the past ten years from Yahoo Finance, sorting them by industry and volatility level. I then select a number of firms to focus on based on the likelihood of trading on Amazon.com and manually search for and obtain subsidiary and brand information

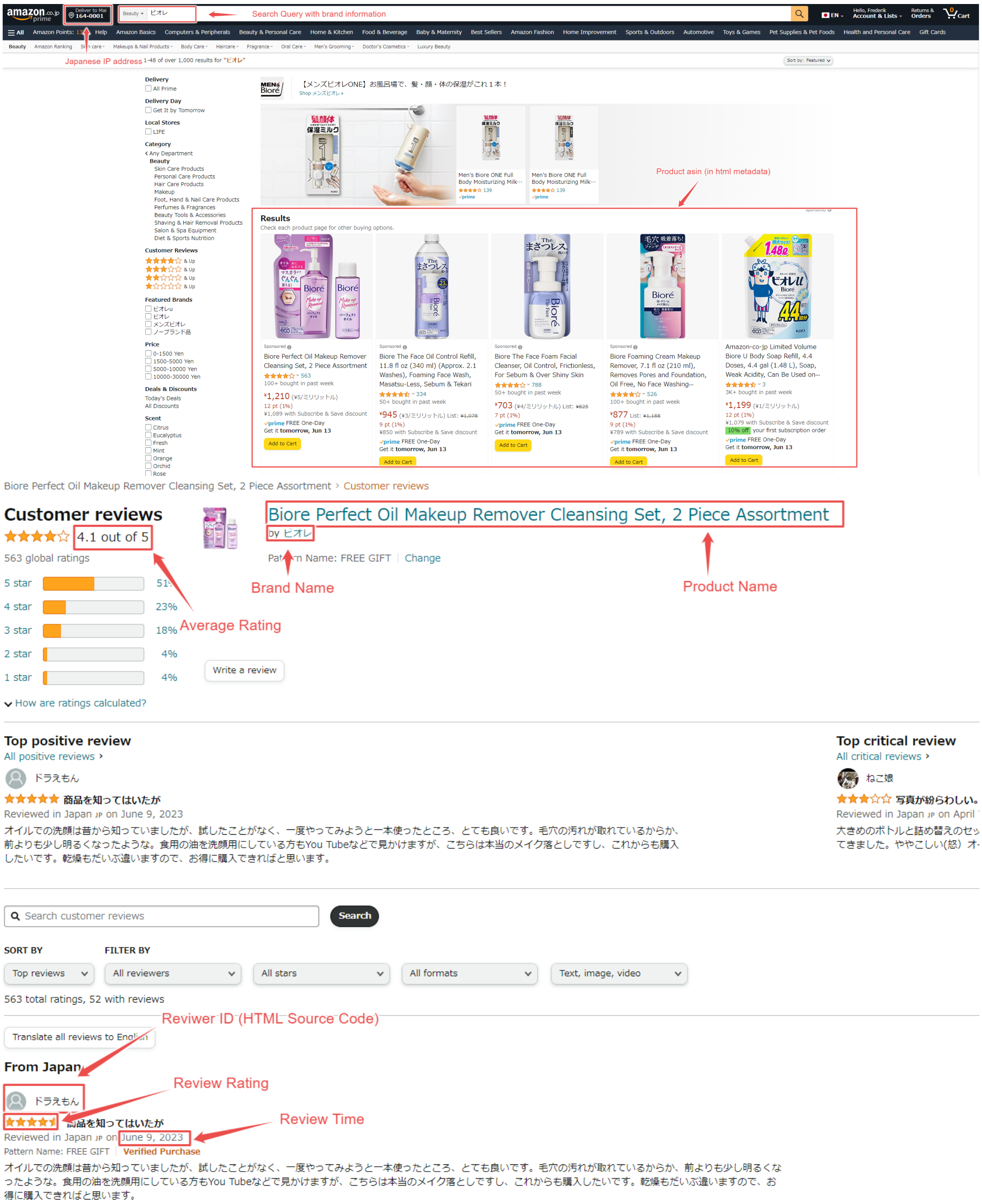


Figure A.1
Sample Webpage on Amazon.com

This figure shows parts of webpages on Amazon.co.jp (Japan) that contain product and review information (Asin Data Webpage ; Review Data Webpage), First, I retrieve product ASIN from the HTML source code, (B082VGJWYS) using specialized search queries (ビオレ), retaining to brands of a selected Japanese company in a specific category, using a Japanese VPN to show products delivered to Japanese Customers (Japanese Postal Code 164-0001). Then, for each asin from an actual product review page, retrieve the following information: the name of the brand (ビオレ), the name of the Product (Biore Perfect Oil Makeup Remover Cleansing Set, 2 Piece Assortment), the numerical star rating (5 starts), and the unique Amazon.com identification number of the reviewer (AF6TGVI2VEKDWITFE32HAOWL6ZSA; displayed only in the HTML source code), and the date of the review (June 9, 2023)

.

from official company websites. I feed the brand, amazon product category and manually identified specific keywords into my Amazon.com review scraper. obtaining the productID (ASIN), brand information, reviewerID, reviewTime, and rating for each review. See Fig A.1 for an example webpage and the scraped data. Acquiring the data proved to be challenging as Amazon.com employs sophisticated scraping protection. Without using proxy servers, which would be too costly, I am limited in the number of asynchronous requests I can send out. I switch between VPNs to counter IP banning by Amazon.co.jp. This significantly restricts the amount of data I can scrape, which is why I opted for a case-study approach instead of rigorous statistical analysis. Ultimately I obtain 700,000 reviews, but after cleaning the reviews for brands of no interest (i.e. those without a publicly traded company or too few reviews), I am left with 200,000 reviews over 2.500 products (asins).

Data availability All data used in this study are publicly available.

Supplementary Analyses

Figure A.2 presents the cumulative outperformance over the S&P 500 Index of the equally weighted and review weighted terciles. A noticeable trend is visible in the review-weighted T3 portfolio, which yields the highest return over the S&P 500 Index, with a cumulative excess return of over 400/%.

(a) Outperformance of Review Weighted Tercile Portfolios above the S&P500 Index: 2004-2015

(b) Outperformance of Equal Weighted Tercile Portfolios above the S&P500 Index: 2004-2015

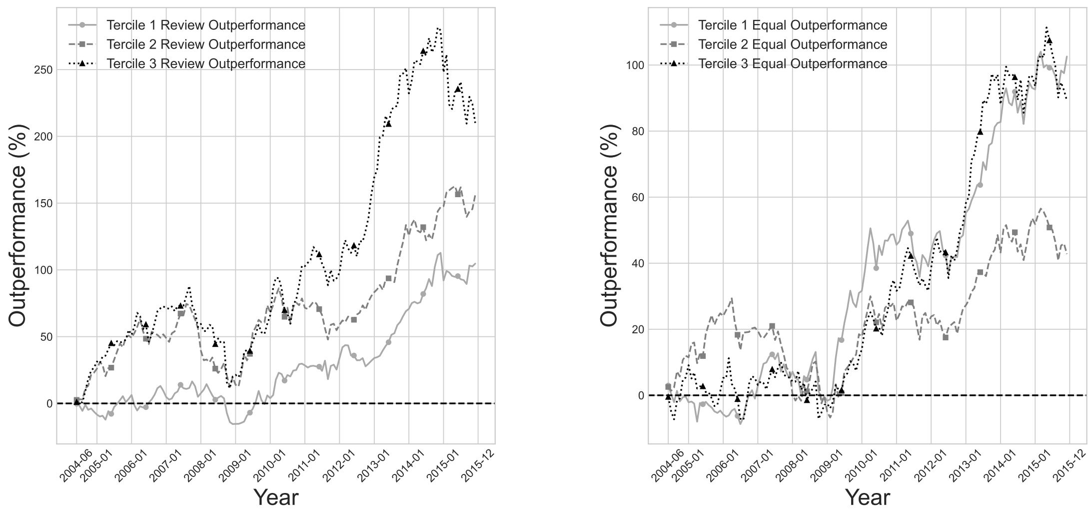
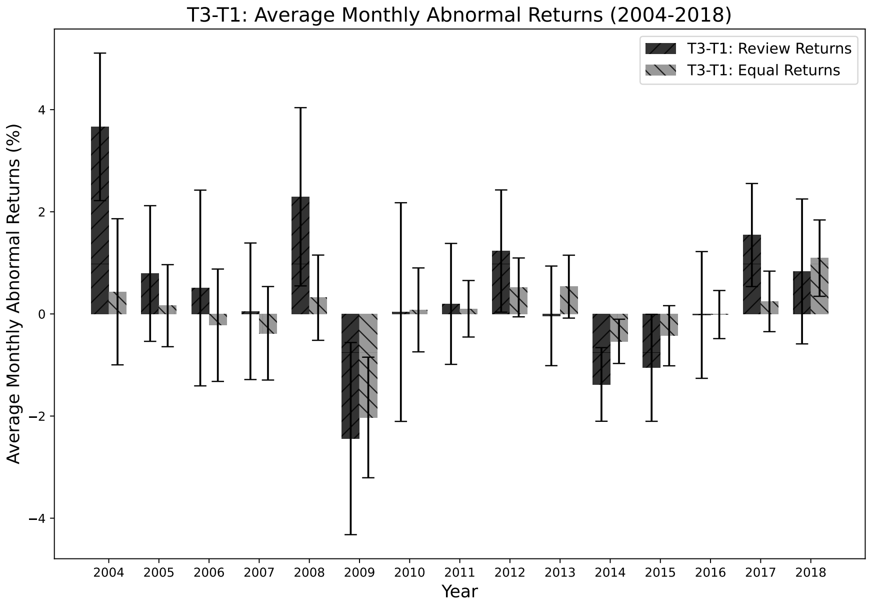


Figure A.2
This figure plots of the outperformance of the equally weighted and review weighted tercile portfolios, calculated as the cumulative return of each tercile above the cumulative return of the S&P 500 Index in the timeframe of June 2004 to December 2015.

I observe a similarity in the returns and graphs of the equal-weighted and review-weighted bottom tercile (T1), both delivering 100% excess cumulative return in the timeframe P1. This could be due to the inherent noise caused by the increased number of reviews per month per stock, making both portfolios similar in weighting. Moreover, the equally weighted top tercile (T3) and bottom tercile (T1) exhibit nearly identical excess cumulative returns. Since the portfolios are noisy due to equal weighting, this similarity in excess returns of specifically the bottom tercile (T1) could be attributed to an attention effect, as proposed by Barber and Odean (2008) or a version of ”value investing” of contrary investors (Odean, 1999). The bottom tercile outperforming the S&P 500 Index might also be due to selection bias. Companies that switched to online retail and the changing market structures have increased returns.

I examine the abnormal review-based strategy’s performance over time, ensuring that the alpha generation is not focused on a particular timeframe, as described in Huang (2018). I calculate the one-month-ahead abnormal return

Figure A.3
Average Monthly Abnormal Returns (2004-2018)
Average monthly alpha performance of a strategy based on reviews. The graph shows the difference in average monthly abnormal returns of portfolios ranked in the top and bottom terciles, according to abnormal customer ratings, on a year-to-year basis. Stocks in the sample are sorted into tercile portfolios based on abnormal customer ratings for each month in P2: July 2004 to December 2015. Abnormal returns for each stock in the terciles are calculated for the following month using the Fama-French-Carhart four-factor model, with factor loadings estimated from the past 60 months of stock returns. The abnormal returns for each tercile portfolio are then computed, both equally weighted and review weighted according to the number of reviews.



for each stock in the tercile portfolios with the Fama-French Carhart four-factor model (Carhart, 1997). I acquire factor loading and factor return data from the Beta Suite Database from WRDS, using a prior 60 months window to estimate factor loading. I then compute the review-weighted and equally-weighted portfolio abnormal return for each tercile, calculated as the realized return over the four-factor expected return. Figure A.3 illustrates the difference in average monthly abnormal returns between the top tercile (T3) and bottom tercile (T1) over time. I observe that most differences in average monthly abnormal returns are centered around 2004 to 2008. Subsequently, these differences fluctuate, dipping into negative alphas in 2014 and 2015, but surprisingly increase again post 2015. This is not, however, due to a stark outperformance of the T3 portfolio in those periods (2017: avg alpha 0.3%m; 2018: -0.4%), but rather due to a significant underperformance of the T1 portfolio (2017: avg alpha -1.1%, 2018: -1.3%). In addition, I do not find that the equally weighted top tercile (T3) constantly yields substantially more alpha than the bottom tercile (T1), which might be again attributable to the increased average monthly reviews introducing noise into the portfolio construction.

Table A.1
(P2) 2004-2018: Summary Statistics and determinants of abnormal customer ratings
This table reports summary statistics of the determinants of abnormal customer rating for the timeframe P2: 2004 - 2018. I base my variable construction on Huang (2018). The average customer rating represents the mean rating received by a company's products within a given month. *Abnormal customer ratings* refers to the deviation of the average rating in a month from the preceding 12-month average. The *number of customer reviews* indicates the average count of reviews for a company's products within a month. *Market cap* represents the firm's market capitalization, calculated by multiplying the number of outstanding shares by the stock price. *Book-to-market* is determined by dividing the book value of common equity by the market value of common equity. *Stock return*$_{m-12,m-1}$ denotes the buy-and-hold stock return over the past 12 months, excluding the most recent month. *R&D ratio* represents the ratio of R&D expenses to book assets. *Gross profitability* is measured as the ratio of income before extraordinary items to the book value of assets. *F-score* is a sum of nine binary variables capturing financial performance signals (Piotroski, 2000). *Dollar volume* signifies the trading volume in dollars during the second-to-last month. *CV of dollar volume* is the coefficient of variation of dollar trading volume calculated over the past 12 months, starting from the second-to-last month. *Book leverage* is the ratio of the book value of total debt to the book value of total assets. *Asset tangibility* is the ratio of net property, plant, and equipment to total assets. All data is taken from the Compustat and CRSP database on WRDS. *# of analysts* represents the number of analysts making forecasts for a stock in a given quarter in the I/B/E/S database. *Analyst revision* indicates the difference between the mean current-year EPS forecasts made by analysts in a given month and the previous month, scaled by the stock price at the end of the previous month. *Revenue surprise (SURGE)* refers to the difference between actual quarterly revenue per share and the expected quarterly revenue per share, scaled by the standard deviation of quarterly revenue growth. *Earnings surprise (SUE)* is the difference between reported quarterly EPS and the median EPS forecast issued by all analysts in the 90-day period leading up to the earnings announcement date taken from I/B/E/S, scaled by the stock price. All variables are winsorized at the 0.1% and 99.9% levels to minimize the impact of outliers.

Variable	N	Mean	Std Dev	25th Percentile	Median	75th Percentile
Customer Reviews						
Average customer ratings	17,059	4.049	0.484	3.842	4.136	4.349
Abnormal customer ratings	17,059	0.007	0.348	-0.112	0.010	0.148
# of customer reviews per month	17,059	1,409,471	3,120,357	60,000	249,000	1,080,000
Firm-level characteristics						
Market cap (millions of dollars)	17,059	32,160,805	75,777,500	1,543,062	5,393,555	18,054,524
Book-to-market	17,059	0.394	0.420	0.197	0.331	0.521
Stock return$_{m-12,m-1}$	17,059	0.114	0.570	-0.124	0.070	0.264
R&D Ratio	17,059	0.021	0.030	0.007	0.015	0.026
Gross profitability	17,059	0.047	0.112	0.012	0.040	0.083
F-score	17,059	5,365	1,651	4,000	6,000	7,000
Dollar volume (millions of dollars)	17,059	4,788,316	12,837,187	294,192	1,246,584	3,755,474
CV of dollar volume	17,059	0.306	0.157	0.204	0.266	0.355
Book leverage	17,059	0.359	0.418	0.119	0.304	0.472
Asset tangibility	17,059	0.166	0.120	0.085	0.135	0.212
# of Analysts	6,046	18,526	15,970	7,000	15,000	24,000
Analyst Revision (percent)	11,908	-0.028	0.797	-0.001	0.000	0.001
Revenue surprise (SURGE)	5,863	0.072	6,794	-1,543	0,062	1,745
Earnings surprise (SUE) (percent)	5,399	0.049	0.268	0.001	0.030	0.088

Table A.2
(P3) 1999-2018: Summary Statistics and determinants of abnormal customer ratings
This table reports summary statistics of the determinants of abnormal customer rating for the timeframe P3: 1999 - 2018.

Variable	N	Mean	Std Dev	25th Percentile	Median	75th Percentile
Customer Reviews						
Average customer ratings	18,245	4.037	0.496	3.823	4.129	4.346
Abnormal customer ratings	18,245	0.002	0.357	-0.121	0.009	0.150
# of customer reviews per month	18,245	1,323,032	3,035,073	51,000	213,000	963,000
Firm-level characteristics						
Market cap (millions of dollars)	18,245	32,830,221	77,418,369	1,491,611	5,232,353	18,006,637
Book-to-market	18,245	0.393	0.415	0.197	0.327	0.518
Stock return$_{m-12,m-1}$	18,245	0.109	0.570	-0.136	0.064	0.261
R&D Ratio	18,245	0.022	0.029	0.007	0.016	0.026
Gross profitability	18,245	0.047	0.111	0.012	0.040	0.083
F-score	18,245	5,375	1,641	4,000	6,000	7,000
Dollar volume (millions of dollars)	18,245	4,790,177	12,760,859	279,785	1,201,544	3,718,417
CV of dollar volume	18,245	0.310	0.163	0.205	0.269	0.359
Book leverage	18,245	0.354	0.419	0.114	0.300	0.469
Asset tangibility	18,245	0.168	0.124	0.085	0.136	0.215
# of Analysts	6,506	18,241	15,887	7,000	15,000	24,000
Analyst Revision (percent)	13,539	-0.028	0.791	-0.001	0.000	0.001
Revenue surprise (SURGE)	6,291	0.121	6.636	-1.497	0.079	1.764
Earnings surprise (SUE) (percent)	5,774	0.051	0.272	0.001	0.031	0.089

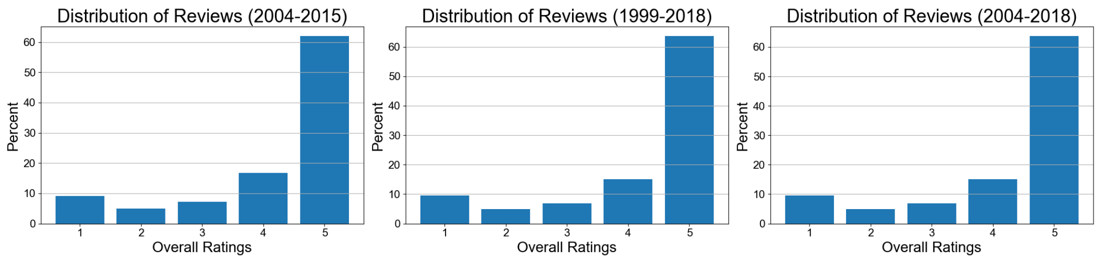


Figure A.4
Distribution of customer product ratings
This figure plots the frequency distribution of customer product ratings on Amazon.com based on the sample of reviews for public firms over the three time periods P1: 2004-2015, P2: 1999-2018, and P3: 2004-2018. The ratings range from one star to five stars.

Table A.3

P2: Regressions of one-month ahead abnormal customer ratings on firm characteristics

Regressions of one-month ahead abnormal customer ratings on firm characteristics in period P2: July 2004 - December 2018. R1 shows the regression with actual computed R&D ratio data, R2 excluding R&D ratio data, and R3 using a function to interpolate R&D ratio data in a time series. The t-statistics are provided in parentheses. Stars denote statistical significance: * indicates significance at the 10%level, ** at the 5% level, and *** at the 1% level

	R1				R2				R3			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Log(Market cap)	0.0009	0.0003	-0.0034	-0.0048	0.0009	0.0017	-0.0023	-0.0025	0.0009	0.0016	-0.0021	-0.0022
	(1.043)	(0.185)	(-0.722)	(-0.782)	(1.043)	(1.581)	(-0.646)	(-0.561)	(1.043)	(1.456)	(-0.604)	(-0.492)
Book-to-market	-0.0053	-0.0001	0.0008	0.0132	-0.0053	-0.0047	-0.0031	0.0022	-0.0053	-0.0047	-0.0031	0.0019
	(-0.828)	(-0.012)	(0.101)	(1.301)	(-0.828)	(-0.715)	(-0.471)	(0.254)	(-0.828)	(-0.724)	(-0.480)	(0.224)
Stock return$_{m-12,m-1}$	-0.0009	0.0028	0.0056	0.0075	-0.0009	-0.0009	0.0017	0.0027	-0.0009	-0.001	0.0017	0.0026
	(-0.182)	(0.482)	(0.951)	(0.888)	(-0.182)	(-0.191)	(0.349)	(0.404)	(-0.182)	(-0.201)	(0.342)	(0.394)
Book leverage	-0.0146	-0.0243	-0.0254	-0.0196	-0.0146	-0.014	-0.0151	-0.0116	-0.0146	-0.0162	-0.0176	-0.015
	(-2.039)**	(-2.632)***	(-2.740)***	(-1.792)*	(-2.039)**	(-1.925)*	(-2.077)**	(-1.402)	(-2.039)**	(-2.138)**	(-2.308)**	(-1.691)*
Asset tangibility	0.0324	0.0654	0.066	0.0516	0.0324	0.0348	0.0364	0.0365	0.0324	0.0372	0.0391	0.0403
	(1.356)	(2.095)**	(2.112)**	(1.373)	(1.356)	(1.423)	(1.484)	(1.232)	(1.356)	(1.514)	(1.586)	(1.350)
Log(1+# of analysts)	-0.0011	0.0028	0.0019	-0.0035	-0.0011	-0.002	-0.0039	-0.0072	-0.0011	-0.002	-0.0037	-0.0071
	(-0.248)	(0.498)	(0.306)	(-0.422)	(-0.248)	(-0.450)	(-0.786)	(-1.119)	(-0.248)	(-0.454)	(-0.744)	(-1.100)
R&D Ratio		0.0935	0.1065	0.1131						0.1062	0.115	0.1348
		(0.703)	(0.795)	(0.704)						(1.017)	(1.094)	(1.075)
Gross profitability		0.0031	-0.0069	-0.0356		0.0118	0.0033	-0.0192		0.0131	0.0043	-0.0191
		(0.110)	(-0.240)	(-1.062)		(0.463)	(0.129)	(-0.635)		(0.510)	(0.168)	(-0.631)
F-Score		-0.0015	-0.0015	-0.0011		-0.0021	-0.0021	-0.003		-0.0021	-0.0021	-0.003
		(-0.646)	(-0.663)	(-0.397)		(-1.211)	(-1.227)	(-1.387)		(-1.196)	(-1.217)	(-1.378)
Log(CV of dollar volume)			-0.0212	-0.0346			-0.0195	-0.0263			-0.0199	-0.0267
			(-2.114)***	(-2.845)***			(-2.642)***	(-2.950)**			(-2.692)***	(-2.990)***
Log(Dollar volume)			0.0023	0.0032			0.003	0.0033			0.0026	0.0028
			(0.421)	(0.45)			(0.631)	(0.524)			(0.631)	(0.525)
Analyst Revision				-0.0031				-0.0035				-0.0034
				(-0.767)				(-0.860)				(-0.858)
Number of Observations	16,527	9,457	9,443	6,794	16,527	16,518	16,496	11,745	16,527	16,518	16,496	11,745
Adj R Squared	0.0006	0.0005	0.0008	0.0008	0.0006	0.0006	0.0009	0.0007	0.0006	0.0006	0.0009	0.0007

Table A.4

P3: Regressions of one-month ahead abnormal customer ratings on firm characteristics

Regressions of one-month ahead abnormal customer ratings on firm characteristics in period P3: July 1999 - December 2018. R1 shows the regression with actual computed R&D ratio data, R2 excluding R&D ratio data, and R3 using a function to interpolate R&D ratio data in a time series. The t-statistics are provided in parentheses. Stars denote statistical significance: * indicates significance at the 10%level, ** at the 5% level, and *** at the 1% level

	R1				R2				R3			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Log(Market cap)	0.000	0.001	-0.006	-0.007	0.000	0.001	-0.004	-0.004	0.000	0.002	-0.004	-0.004
	(0.311)	(0.507)	(-1.338)	(-1.136)	(0.311)	(1.429)	(-1.101)	(-0.844)	(0.311)	(1.510)	(-1.122)	(-0.849)
Book-to-market	-0.002	0.003	0.004	0.016	-0.002	-0.001	0.001	0.006	-0.002	-0.001	0.001	0.006
	(-0.349)	(0.421)	(0.532)	(1.520)	(-0.349)	(-0.181)	(0.071)	(0.737)	(-0.349)	(-0.179)	(0.072)	(0.739)
Stock return$_{m-12,m-1}$	-0.000	0.005	0.008	0.011	-0.000	-0.000	0.003	0.005	-0.000	-0.000	0.003	0.005
	(-0.072)	(0.799)	(1.354)	(1.351)	(-0.072)	(-0.047)	(0.536)	(0.805)	(-0.072)	(-0.042)	(0.537)	(0.806)
Book leverage	-0.011	-0.014	-0.015	-0.011	-0.011	-0.010	-0.012	-0.008	-0.011	-0.009	-0.010	-0.008
	(-1.619)	(-1.546)	(-1.645)	(-0.993)	(-1.619)	(-1.486)	(-1.624)	(-1.029)	(-1.619)	(-1.206)	(-1.340)	(-0.921)
Asset tangibility	0.001	0.024	0.024	0.015	0.001	0.005	0.006	0.006	0.001	0.003	0.004	0.006
	(0.025)	(0.835)	(0.818)	(0.439)	(0.025)	(0.207)	(0.238)	(0.226)	(0.025)	(0.145)	(0.177)	(0.213)
Log(1+# of analysts)	0.001	0.003	-0.000	-0.004	0.001	-0.000	-0.003	-0.006	0.001	-0.000	-0.003	-0.006
	(0.230)	(0.461)	(-0.064)	(-0.492)	(0.230)	(-0.061)	(-0.657)	(-0.958)	(0.230)	(-0.074)	(-0.693)	(-0.958)
R	D Ratio		0.008	0.012	0.048					-0.075	-0.073	-0.015
		(0.057)	(0.093)	(0.296)					(-0.730)	(-0.702)	(-0.118)	
Gross profitability		0.010	0.000	-0.026		0.017	0.009	-0.015		0.016	0.008	-0.015
		(0.366)	(0.001)	(-0.762)		(0.652)	(0.336)	(-0.477)		(0.626)	(0.318)	(-0.477)
F-Score		-0.004	-0.004	-0.004		-0.003	-0.003	-0.004		-0.003	-0.003	-0.004
		(-1.582)	(-1.580)	(-1.336)		(-1.954)	(-1.941)	(-1.336)		(-1.963)	(-1.947)	(-1.988)
Log(CV of dollar volume)			-0.023	-0.038			-0.020	-0.026		-0.019	-0.026	
			(-2.332)**	(-3.194)***			(-2.695)***	(-2.957)***		(-2.661)***	(-2.949)***	
Log(Dollar volume)			0.006	0.006			0.005	0.005		0.005	0.005	
			(1.155)	(0.837)			(1.169)	(0.895)		(1.223)	(0.902)	
Analyst Revision				-0.004				-0.004				-0.004
				(-1.046)				(-1.106)				(-1.106)
Number of Observations	17,803	10,360	10,344	7,408	17,803	17,791	17,767	12,629	17,803	17,791	17,767	12,629
Adj R Squared	-0.000	-0.000	0.000	0.001	-0.000	0.000	0.000	0.001	-0.000	-0.000	0.000	0.001

Table A.5

Summary of Tercile Statistics in P1: June 2004 December 2018

This table provides information on the number of firms, number of firm months, and average rating for each tercile. T1 represents the tercile with low abnormal ratings, T2 represents the middle tercile, and T3 represents the tercile with high abnormal ratings.

	Number of firms	Number of firm months	average rating
T1 (low abnormal rating)	187	138	-0.311
T2	167	138	0.032
T3 (high abnormal rating)	187	138	0.344

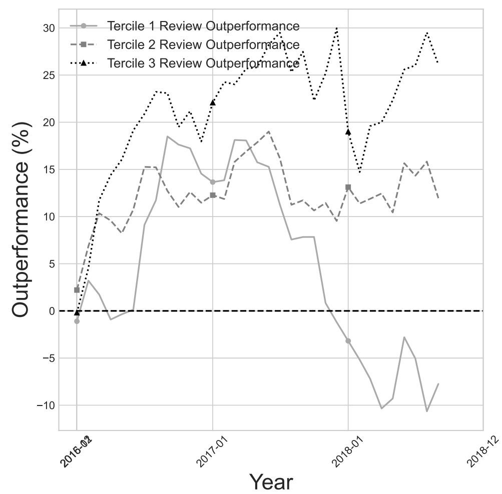
Table A.6
Calendar-time portfolio returns: Idiosyncratic Volatility top terciles
This table presents the monthly calendar-time top tercile (T3) portfolio alphas calculated through the Fama-French Carhart four-factor model. Stocks are first classified into high and low idiosyncratic volatility groups according to the median idiosyncratic volatility in a given quarter and then stocks are classified into tercile portfolios. Alpha estimates are calculated by regressing the monthly returns of the top tercile portfolio against the monthly returns of the Fama-French-Carhart factors. The t-statistics are provided in parentheses. Stars denote statistical significance: * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level

	Review weighting	Equal weighting
Tercile 3 (high abnormal ratings) High idiosyncratic Volatility	0.972%	0.237%
	(1.71)*	(0.66)
Tercile 3 (high abnormal ratings): Low idiosyncratic Volatility	0.558%	0.144%
	(2.1)**	(0.77)
Long/short (T3 High − T3 Low)	0.560%	0.14%
	(2.1)**	(0.55)

Table A.7
Calendar-time portfolio returns: subsample analyses
This table presents the monthly calendar-time spread portfolio alphas calculated through the Fama-French-Carhart four-factor model. In Panel A, stocks are classified into high and low idiosyncratic volatility groups according to the median idiosyncratic volatility. Panels B and C further partition the stocks based on the median number of analyst coverages and the median market capitalization, respectively. For each group, and for each month in timeframe P1: July 2004 to December 2015, a spread portfolio is constructed that buys the top tercile stocks and sells the bottom tercile stocks based on abnormal customer ratings, across the different subsamples. The portfolio's performance is tracked over the next month. Alpha estimates are calculated by regressing the monthly returns of the spread portfolio against the monthly returns of the Fama-French-Carhart factors. The t-statistics are provided in parentheses. Stars denote statistical significance: * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Review weighting	Equal weighting
Panel A: Idiosyncratic Volatility		
Long/Short (T3 − T1) High idiosyncratic Volatility	1.05%	0.11%
	(1.43)	(0.28)
Long/Short (T3 − T1) Low idiosyncratic Volatility	0.28%	0.077%
	(0.75)	(0.37)
Long/short (High − Low)	0.836%	0.123%
	(1.03)	(0.26)
Panel B: Analyst coverage		
Long/Short (T3 − T1) Low analyst coverage	1.306%	0.139%
	(2.01)**	(0.46)
Long/Short (T3 − T1) High analyst coverage	0.892%	-0.002%
	(1.68)*	(-0.1)
Long/short (low − high)	0.414%	0.141%
	(0.35)	(0.55)
Panel C: Market capitalization		
Long/Short (T3 − T1) Small firms	0.034%	-0.029%
	(0.05)	(-0.08)
Long/Short (T3 − T1) Large firms	1.102%	0.298%
	(2.42)**	(1.25)*
Long/short (small − large)	-1.068%	-0.327%
	(-1.46)	(-0.74)

(a) Outperformance of Review Weighted Tercile Portfolios above the S&P500 Index: 2015-2018



(b) Outperformance of Equal Weighted Tercile Portfolios above the S&P500 Index: 2015-2018

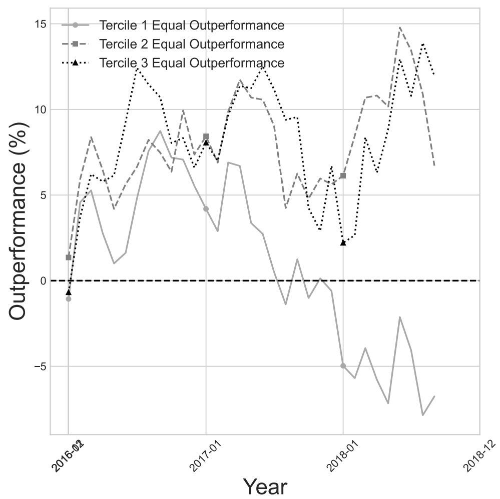


Figure A.5
Outperformance Post 2015
This figure plots the outperformance of the equally weighted and review-weighted tercile portfolios, calculated as the cumulative return of each tercile above the cumulative return of the S&P 500 Index in the timeframe of June 2015 to December 2018. Although not yielding abnormal returns anymore, the top tercile (T3) still outperforms the S&P 500 by around 15-30% and even now see that the bottom tercile (T1) underperforms the S&P 500

Table A.8
Calendar-time portfolio returns: Idiosyncratic Volatility top terciles
This table presents the calendar-time tercile review-weighted portfolio regression covering the Japanese Amazon.com customer reviews from April 2008 to June 2023, employing two dummy variables: Pre 2015 and Post 2015. Pre 2015 is 1 for all firm months in the terciles before December 2015 and 0 thereafter and Post 2015 is 1 for all firm months in the terciles after Deceber 2015, and 0 before. I find that the top tercile Stars denote statistical significance: * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level

	Pre 2015	Post 2015
Tercile 1 (low abnormal ratings)	1.01%	-0.80%
	(1.17)	(-0.84)
Tercile 2	1.38%	0.22%
	(0.14)	(0.20)
Tercile 3 (high abnormal ratings)	1.66%	0.83%
	(1.84)*	(0.844)

Bibliography

Aksoy, Lerzan, Cooil, Bruce, Groening, Christopher, Keiningham, Timothy L, Yalçın, Atakan, Dean, Samuel B, and Richmond, Evelyn R (2008), The Long-Term Stock Market Valuation of Customer Satisfaction, *Journal of Marketing* 72, 105–122.

Anderson, Eugene W, Fornell, Claes, Mazvancheryl, Sanal K, Simon, Carl, Frankel, Rich, Priester, Joe, Batra, Rajeev, Roy, Tirthankar, Ghosh, Mrinal, and Rego, Lopo (2004), Customer Satisfaction and Shareholder Value, *Journal of Marketing* 68, 172–185.

Anderson, Eugene W. (1998), Customer Satisfaction and Word of Mouth, *Journal of Service Research* 1(1), 5–17.

Anderson, Eugene W., Fornell, Claes, and Lehmann, Donald R. (1994), Customer Satisfaction, Market Share, and Profitability: Findings from Sweden, *Journal of Marketing* 58(3), 53.

Ang, Andrew, Hodrick, Robert J., Xing, Yuhang, and Zhang, Xiaoyan (2006), The cross-section of volatility and expected returns, *Journal of Finance* 61(1), 259–299.

Antweiler, Werner and Frank, Murray Z. (2001), Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards, *SSRN Electronic Journal*.

Barber, Brad M. and Odean, Terrance (2008), All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *Review of Financial Studies* 21(2), 785–818.

Bayus, Barry L. (2013), Crowdsourcing new product ideas over time: An analysis of the Dell IdeaStorm community, *Management Science* 59(1), 226–244.

Bolton, R. N., Kannan, P. K., and Bramlett, M. D. (2000), Implications of Loyalty Program Membership and Service Experiences for Customer Retention and Value, *Journal of the Academy of Marketing Science* 28(1), 95–108.

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

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

Chan, Louis K.C., Lakonishok, Josef, and Sougiannis, Theodore (2001), The stock market valuation of research and development expenditures, *Journal of Finance* 56(6), 2431–2456.

Chen, Hailiang, De, Prabuddha, Hu, Yu, and Hwang, Byoung Hyoun (2014), Wisdom of crowds: The value of stock opinions transmitted through social media, *Review of Financial Studies* 27(5), 1367–1403.

Chen, Yubo, Liu, Yong, and Zhang, Jurui (2011), When Do Third-Party Product Reviews Affect Firm Value and What Can Firms Do? The Case of Media Critics and Professional Movie Reviews, *Journal of Marketing* 75, 116–134.

Chintagunta, Pradeep K, Gopinath, Shyam, and Venkataraman, Sriram (2010), The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation Across Local Markets, *Journal of Marketing Science* 29(5), 944–957.

Da, Zhi, Engelberg, Joseph, and Gao, Pengjie (2011), In Search of Attention, *Journal of Finance* 66(5), 1461–1499.

Das, Sanjiv, Martínez-Jerez, Asís, and Tufano, Peter (2005), eInformation: A Clinical Study of Investor Discussion and Sentiment, *Financial Management* 34(3), 103–137.

Dellarocas, Chrysanthos, Zhang, Xiaoquan, and Awad, Neveen F. (2007), Exploring the value of online product reviews in forecasting sales: The case of motion pictures, *Journal of Interactive Marketing* 21(4), 23–45.

Dellavigna, Stefano and Pollet, Joshua M. (2009), Investor Inattention and Friday Earnings Announcements, *The Journal of Finance* 64(2), 709–749.

Edmans, Alex (2011), Does the stock market fully value intangibles? Employee satisfaction and equity prices, *Journal of Financial Economics* 101(3), 621–640.

Fama, Eugene F and French, Kenneth R (1992), The Cross-Section of Expected Stock Returns, *The Journal of Finance* 47(2), 427–465.

Forman, Chris, Ghose, Anindya, and Wiesenfeld, Batia (2008), Examining the Relationship Between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets, *Information Systems Research* 19(3), 291–313.

Fornell, Claes (1992), A National Customer Satisfaction Barometer: The Swedish Experience, *Journal of Marketing* 56(1), 6–21.

Fornell, Claes (2001), The Score of Satisfaction, *Havard Business Review* 79, 120–121.

Fornell, Claes, Johnson, Michael D, Anderson, Eugene W, Cha, Jaesung, and Bryant, Barbara Everitt (1996), The American Customer Satisfaction Index: Nature, Purpose, and Findings, *Journal of Marketing* 60(4), 7–18.

Fornell, Claes, Mithas Sunil, Morgeson, Forest V., and Krishnan, M.S. (2006), Customer Satisfaction and Stock Prices: High Returns, Low Risk, *Journal of Marketing* 70, 3–14.

Fornell, Claes, Morgeson, Forrest V., and Hult, G. Tomas M. (2016), Stock returns on customer satisfaction do beat the market: Gauging the effect of a marketing intangible, *Journal of Marketing* 80(5), 92–107.

Fornell, Claes, Rust, Roland T., and Dekimpe, Marnik G. (2010), The Effect of Customer Satisfaction on Consumer Spending Growth, *Journal of Marketing Research* 47(1), 28–35.

Franzoni, Chiara and Sauermann, Henry (2014), Crowd science: The organization of scientific research in open collaborative projects, *Research Policy* 43(1), 1–20.

Godes, David and Mayzlin, Dina (2002), Using Online Conversations to Study Word of Mouth Communication, *SSRN Electronic Journal*.

Gruca, Thomas S. and Rego, Lopo L. (2005), The impact of cobranding on customer evaluation of brand counterextensions, *Journal of Marketing* 69(3), 1–18.

Heitz, Amanda, Narayanamoorthy, Ganapathi S., and Zekhnini, Morad (2018), Filings of Material Information and the Disappearing Earnings Announcement Premium, *SSRN Electronic Journal*.

Hirshleifer, David, Lim Sonya, Seogyeon, and Teoh, Siew Hong (2009), Driven to Distraction: Extraneous Events and Underreaction to Earnings News, *The Journal of Finance* 64(5), 2289–2325.

Hirshleifer, David and Teoh, Siew Hong (2003), Limited attention, information disclosure, and financial reporting, *Journal of Accounting and Economics* 36(1-3 SPEC. ISS.), 337–386.

Homburg, Christian, Koschate, Nicole, and Hoyer, Wayne D. (2005), Do Satisfied Customers Really Pay More? A Study of the Relationship between Customer Satisfaction and Willingness to Pay, *Journal of Marketing* 69(2), 84–96.

Hong, Harrison, Lim, Terence, and Stein, Jeremy C. (2000), Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *Journal of Finance* 55(1), 265–295.

Hou, Kewei and Moskowitz, Tobias J. (2005), Market Frictions, Price Delay, and the Cross-Section of Expected Returns, *Review of Financial Studies* 18(3), 981–1020.

Huang, Jiekun (2018), The customer knows best: The investment value of consumer opinions, *Journal of Financial Economics* 128(1), 164–182.

Huberman, Gur and Regev, Tomer (2001), Contagious Speculation and a Cure for Cancer: A Nonevent that Made Stock Prices Soar, *The Journal of Finance* 56(1), 387–396.

Ivanov, Vladimir, Joseph, Kissan, and Wintoki, M. Babajide (2013), Disentangling the market value of customer satisfaction: Evidence from market reaction to the unanticipated component of ACSI announcements, *International Journal of Research in Marketing* 30(2), 168–178.

Jacobson, Robert and Mizik, Natalie (2009), *Assessing the Value-Relevance of Customer Satisfaction*, tech. rep.

Jaffe, Jeffrey F (1974), Special Information and Insider Trading, *The Journal of Business* 47(3), 410–428.

Katsikeas, Constantine S., Morgan, Neil A., Leonidou, Leonidas C., and Hult, G. Tomas M. (2016), Assessing performance outcomes in marketing, *Journal of Marketing* 80(2), 1–20.

Katsura, Shinichi and Shi, Jinxia (2018), Anomalies in the Japanese and Chinese Stock Markets: Value Stock Effect, Small-Cap Stock Effect, Volatility Effect.(日中株式市場におけるアノマリー: バリュー株効果, 小型株効果, ボラティリティー効果), *Shokei-gakuso: Journal of Business Studies* 65(1), 1–18.

Kelley, Eric K. and Tetlock, Paul C. (2013), How Wise Are Crowds? Insights from Retail Orders and Stock Returns, *Journal of Finance* 68(3), 1229–1265.

Lazonick, William and O'Sullivan, Mary (2000), Maximizing shareholder value: A new ideology for corporate governance, *Economy and Society* 29(1), 13–35.

Lee, Charles M.C., Ma, Paul, and Wang, Charles C.Y. (2015), Search-based peer firms: Aggregating investor perceptions through internet co-searches, *Journal of Financial Economics* 116(2), 410–431.

Lev, Baruch and Sougiannis, Theodore (1996), The capitalization, amortization, and value-relevance of R&D, *Journal of Accounting and Economics* 21, 107–138.

Li, Keming (2020), Does Information Asymmetry Impede Market Efficiency? Evidence from Analyst Coverage, *Journal of Banking & Finance* 118, 105856.

Luo, Xueming (2009), Quantifying the Long-Term Impact of Negative Word of Mouth on Cash Flows and Stock Prices, *Marketing Science* 28(1), 148–165.

Ma, Yulong, Tang, Alex P, and Hasan, Tanweer (2005), The Stock Price Overreaction Effect: Evidence on Nasdaq Stocks, *Quarterly Journal of Business and Economics* 44(3), 113–127.

Malshe, Ashwin, Colicev, Anatoli, and Mittal, Vikas (2020), How Main Street Drives Wall Street: Customer (Dis)satisfaction, Short Sellers, and Abnormal Returns, *Journal of Marketing Research* 57(6), 1055–1075.

Marquering, Wessel, Nisser, Johan, and Valla, Toni (2006), Disappearing anomalies: a dynamic analysis of the persistence of anomalies, *Applied Financial Economics* 16(4), 291–302.

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

Moe, Wendy W. and Trusov, Michael (2011), The Value of Social Dynamics in Online Product Ratings Forums, *Journal of Marketing Research* 48(3), 444–456.

Ni, Jianmo, Li, Jiacheng, and Mcauley, Julian (2019), *Justifying Recommendations using Distantly-Labeled Reviews and Fine-Grained Aspects*, tech. rep.

Novy-Marx, Robert (2013), The other side of value: The gross profitability premium, *Journal of Financial Economics* 108(1), 1–28.

Odean, Terrance (1999), Do Investors Trade Too Much?, *The American Economic Review*, 1279–1299.

Peng, Lin and Xiong, Wei (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.

Poetz, Marion K. and Schreier, Martin (2012), The value of crowdsourcing: Can users really compete with professionals in generating new product ideas?, *Journal of Product Innovation Management* 29(2), 245–256.

Pontiff, Jeffrey (1996), Costly Arbitrage: Evidence from Closed-End Funds, *The Quarterly Journal of Economics* 111(4), 1135–1151.

Raithel, Sascha, Sarstedt, Marko, Scharf, Sebastian, and Schwaiger, Manfred (2012), On the value relevance of customer satisfaction. Multiple drivers and multiple markets, *Journal of the Academy of Marketing Science* 40(4), 509–525.

Reichheld, Frederick F. and Sasser, W. Earl Jr. (1990), Zero Defections: Quality Comes to Services, *Havard Business Review*.

Shleifer, Andrei and Vishny, Robert W. (1997), The limits of arbitrage, *Journal of Finance* 52(1), 35–55.

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

Srivastava, Rajendra K, Capraro, Anthony J, Deephouse, David L, Roberts, Peter W, Mcmillan, G Steven, College, American, Joshi, Maheshkumar P, Joseph', St, and Crosby, Jack R (1997), The Value of Corporate Reputation: Evidence from Equity Markets, *Corporate Reputation Review* 1(2), 61–68.

Stuart, Jeniffer Ames, Gupta, Sunil, and Lehmann, Donald R (2004), Gupta, Lehmann and Stuart 2004, *Journal of Marketing Reserach* 41, 7–18.

Subrahmanyam, Avanidhar and Titman, Sheridan (1999), The going-public decision and the development of financial markets, *Journal of Finance* 54(3), 1045–1082.

Szakmary, Andrew C. and Kiefer, Dean B. (2004), The Disappearing January/Turn of the Year Effect: Evidence From Stock Index Futures and Cash Markets, *Journal of Futures Markets* 24(8), 755–784.

Takkabutr, Nattapol (2013), "Roles of volatility in Japanese stock market (日本の株式市場における株式リターンボラティリティの役割)", PhD thesis.

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

Tumarkin, Robert and Whitelaw, Robert F. (2001), News or Noise? Internet Postings and Stock Prices, *Financial Analysts Journal* 57(3), 41–51.

Wieseke, Jan, Luo, Xueming, and Homburg, Christian (2010), Custmer Satisfaction , Analyst Stock Recommendations and FIrm Value, *Journal of Marketing Research* 47, 1041–1058.

Wu, Junran, Xu, Ke, and Zhao, Jichang (2020), Online reviews can predict long-term returns of individual stocks, *Journal of Neurocomputing* 417, 406–418.

Zhu, Feng and Zhang, Xiaoquan (Michael) (2010), Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics, *Journal of Marketing* 74(2), 133–148.

Statutory Declaration, Eidesstattliche Erklärung

„Hiermit versichere ich, dass diese Bachelorarbeit / Seminararbeit / Masterarbeit von mir persönlich verfasst ist und dass ich keinerlei fremde Hilfe in Anspruch genommen habe. Ebenso versichere ich, dass diese Arbeit oder Teile daraus weder von mir selbst noch von anderen als Leistungsnachweise andernorts eingereicht wurde. Wörtliche oder sinngemäße Übernahmen aus anderen Schriften und Veröffentlichungen in gedruckter oder elektronischer Form sind gekennzeichnet. Sämtliche Sekundärliteratur und sonstige Quellen sind nachgewiesen und in der Bibliographie aufgeführt. Das Gleiche gilt für graphische Darstellungen und Bilder sowie für alle Internet-Quellen. Ich bin ferner damit einverstanden, dass meine Arbeit zum Zwecke eines Plagiatsabgleichs in elektronischer Form anonymisiert versendet und gespeichert werden kann. Mir ist bekannt, dass von der Korrektur der Arbeit abgesehen werden kann, wenn die Erklärung nicht erteilt wird. “

I hereby declare that the paper presented is my own work and that I have not called upon the help of a third party. In addition, I affirm that neither I nor anybody else has submitted this paper or parts of it to obtain credits elsewhere before. I have clearly marked and acknowledged all quotations or references that have been taken from the works of other. All secondary literature and other sources are marked and listed in the bibliography. The same applies to all charts, diagrams and illustrations as well as to all Internet sources. Moreover, I consent to my paper being electronically stored and sent anonymously in order to be checked for plagiarism. I am aware that the paper cannot be evaluated and may be graded “failed” (“nicht ausreichend”) if the declaration is not made.”

Place, Date	Signature