The customer knows best: The investment value of consumer opinions[†]

Jiekun Huang[a,*]

[a] *University of Illinois at Urbana-Champaign, Gies College of Business, Champaign IL, 61820, USA*

ABSTRACT

This paper investigates whether consumer opinions convey value-relevant information to financial markets. Using a data set of more than 14.5 million customer product reviews on Amazon.com from 2004 through 2015, I find evidence that consumer opinions contain information for stock pricing. A spread portfolio that is long on stocks with high abnormal customer ratings and short on stocks with low abnormal customer ratings delivers an abnormal return of around 55.7 to 73.0 basis points per month. There is no evidence of return reversals in the subsequent year. The return predictability of customer ratings continues to hold after controlling for firm characteristics such as gross profitability, advertising, research and development expenses, and trading volume. Furthermore, abnormal customer ratings positively predict revenues and earnings surprises. These results suggest that consumer opinions contain novel information about firms' fundamentals and stock pricing.

JEL classification: G12, G14, L15.

Keywords: Consumer opinions; Serendipitous information; Stock pricing; Cash flow surprises; Wisdom of crowds.

[†]I am grateful to Avanidhar Subrahmanyam (the referee) for detailed and insightful comments that substantially improved the paper. I thank Heitor Almeida, Lauren Cohen, Don Fullerton, Itay Goldstein, Tim Johnson, Byoung-Hyoun Hwang, Dana Kiku, Eric Kelley, Nolan Miller, Neil Pearson, George Pennacchi, Joshua Pollet, Alexei Tchistyi, Scott Weisbenner, Mao Ye, and seminar participants at University of Illinois at Urbana-Champaign and Cubist Systematic Strategies for comments and helpful discussions. Bitian Qi, Ming Wang, Yue Wang, Zhongnan Xiang, and Xian Zhang provided excellent research assistance. Financial support from the Campus Research Board at the University of Illinois is gratefully acknowledged. I retain responsibility for any remaining errors.

[*]Corresponding author. Tel: +1 217 300 7024; fax: +1 217 244 3102.

E-mail address: huangjk@illinois.edu (J. Huang).

1. Introduction

Customers are important stakeholders of a firm, because the firm's ability to generate cash flows depends in large part on the value created for its customers. As the users of a company's products and services, consumers can possess information about product quality and value, which can have direct implications for the company's future sales and profitability.[1] Subrahmanyam and Titman (1999) coin the term "serendipitous information" for intelligence that investors gather in their everyday activities, such as information about product quality and demand. They argue that such information, while noisy, can provide useful signals of the underlying value of a firm when aggregated. Yet little evidence exists on the information content of consumer opinions for firms' future cash flows and stock returns.

Anecdotal evidence suggests that consumer opinions convey value-relevant information to financial markets.[2] Consider consumers' reviews of TurboTax, a software package made by Intuit Inc. Shortly following the release of the software for the 2003 tax season in November 2002, customers flooded product review websites with complaints about an antipiracy feature, which allowed only the first computer that installed the software to print or electronically file returns. For example, an angry Amazon.com reviewer commented that this policy "makes [TurboTax] an unacceptable product" and called for consumers to boycott it. Two months passed before the news media started to report the story, with the first article appearing on January 31, 2003 (USA Today, 2003). As Fig. 1 shows, the stock price of Intuit Inc. dropped from $27 when the first Amazon.com review on the software was posted on November 27,

[1] The rise of the Internet has greatly facilitated the generation and dissemination of information by ordinary consumers, who not only produce but also read product-related information online. According to Nielsen's 2015 Global Trust in Advertising Survey, 66% of consumers trust online customer reviews, 69% of which say they always or sometimes take action on these opinions, suggesting that these reviews can contain useful information about a company's products and can influence purchase decisions on a large scale.

[2] Anecdotal accounts also suggest that product-related information is of considerable interest to professional portfolio managers. For example, Peter Lynch, the renowned former portfolio manager of Fidelity's Magellan Fund, contends in his popular book *One Up on Wall Street* (1989) that "visiting stores and testing products is one of the critical elements of the analyst's job." A *Harvard Business Review* article comments that "[w]ith the company's retail insights into the browsing, shopping, and purchasing behaviors of its customers, ... Amazon is extraordinarily well-positioned to strategically invest in explicit companies or particular sectors" (Harvard Business Review, 2012).

2002 to $22 when the first news article was published. Perhaps not coincidentally, on March 21, 2003, the company lowered its earnings expectations for the fiscal year because of weaker sales (Wall Street Journal, 2003). Consumers as a group apparently possess useful information about a firm's fundamentals and stock pricing.

[Insert Fig. 1 about here]

In this paper, I investigate whether this pattern holds systematically across a broad sample of firms with online customer reviews. A priori, whether the aggregated opinions of consumer crowds contain new fundamental information and have predictive power for future stock returns is unclear. On the one hand, consumer opinions may fail to provide new information beyond what has been incorporated in the stock price for at least three reasons. First, consumers may lack the incentive to provide truthful information about products. The sharing of information by consumers faces the free-rider problem associated with the private provision of public goods. Second, consumers may lack the expertise to evaluate products. Consumers may make systematic errors in their reviews of a company's products, such as ignoring their own private signals and herding on publicly observed signals (Bikhchandani, Hirshleifer, and Welch, 1992). Also, consumer opinions may be influenced by product advertising and other attention-grabbing promotional activities, resulting in biased reviews.[3] These considerations suggest that the opinions of ordinary consumers are often noisy and subjective, which can render them unreliable for stock price predictions. Third, even if consumer opinions are informative about fundamentals, the information could have already been incorporated into stock prices, which again makes such information useless in predicting future stock price movements.

On the other hand, there are several reasons to posit that the aggregated opinions of consumers contain information for the financial markets. First, consumer opinions not only provide signals about a company's products, but also affect purchase decisions of consumers.

[3]Moreover, some reviews provide false information despite Amazon.com's efforts to weed out suspected fake reviews (Wall Street Journal, 2015).

A large literature in industrial organization suggests that customer perceptions of superior quality can increase firms' future cash flows by enhancing firms' reputation among customers, reducing price elasticities and marketing costs, expanding product offerings, and lowering consumers' informational costs (e.g., Shapiro, 1982, 1983; Allen, 1984; Wernerfelt, 1988; Choi, 1998; Cabral, 2000). The growing popularity of Internet-based consumer opinion platforms can magnify these effects, as consumers not only actively produce product-related information that is readily available to other consumers, but also regularly use such information when making purchase decisions. Therefore, consumer opinions may convey novel information about cash flows. Second, consumer crowds are likely to satisfy the conditions required for the wisdom of crowds to hold true. Each individual consumer has some information about the product in question, and individual consumers can be incentivized to provide truthful information due to private benefits derived from contributing to the provision of a public good, such as peer recognition, ego satisfaction, and joy of giving (e.g., Lerner and Tirole, 2002). Aggregating over a large crowd can ensure that individual consumers' errors cancel out insofar as the errors are not systematically correlated (Subrahmanyam and Titman, 1999). Third, consumer opinions may have predictability for stock returns because of limited investor attention. Because attention is a scarce cognitive resource and investors may consider only a subset of available information due to limits to attention and information processing capacity (see, e.g., Hong and Stein, 1999; Hirshleifer and Teoh, 2003; Peng and Xiong, 2006),it may take some time for the market to fully reflect the information conveyed by consumer opinions, thereby resulting in predictability of stock returns.[4]

In this paper, I investigate the investment value of consumer opinions by making use of a novel data set. I measure consumer opinions using customer product reviews on Amazon.com, which is the largest single source of Internet consumer reviews. Using a sample of more than 14.5 million customer reviews posted by more than 6.4 million reviewers on Amazon.com during the period from 2004 to 2015, I find evidence that consumer opinions

[4] A growing body of empirical literature finds evidence consistent with limited attention among investors (see, among others, Hong, Lim, and Stein, 2000; Huberman and Regev, 2001; Hong, Torous, and Valkanov, 2007; Hou and Moskowitz, 2005; DellaVigna and Pollet, 2007, 2009; Cohen and Frazzini, 2008).

have investment value. Abnormal customer ratings are positively associated with future stock returns, and the economic magnitude is also significant. A spread portfolio that buys stocks with the abnormal consumer rating in the top tercile and sells stocks with that in the bottom tercile delivers a Fama-French-Carhart four-factor alpha of about 55.7 to 73.0 basis points per month. I also find that this result appears concentrated among stocks with high arbitrage costs and more binding limits to investor attention, namely, stocks with high idiosyncratic volatilities, low analyst coverage stocks, and small-cap stocks. Importantly, this return pattern does not reverse in the long run. Furthermore, Fama-MacBeth regressions show that the return predictability of abnormal customer ratings continues to hold after controlling for firm characteristics such as advertising, gross profitability, trading volume, and other predictors of stock returns, suggesting that customer reviews in aggregate provide new information about firms' fundamentals.

Abnormal customer ratings also positively predict revenue surprises and earnings surprises. These results are robust to controlling for other determinants of cash flow surprises such as lagged surprises, advertising, gross profitability, trading volume, and past stock returns. The economic magnitudes are nontrivial. For example, an interquartile increase in abnormal customer ratings is associated with an increase of about 10.0% of the interquartile range of earnings surprises. These findings provide evidence suggesting that consumer opinions contain novel information about future cash flows.

Sophisticated investors, namely, hedge fund managers, seem to make use of the information contained in consumer reviews in their trading decisions. Abnormal customer ratings positively predict net purchases by hedge funds, but do not have predictability for net purchases by non–hedge fund institutions. These results are consistent with the idea that sophisticated investors exploit the return predictability of customer reviews.

My paper makes four main contributions to the literature. First, this study is among the first to test the information content of consumer opinions by distinguishing between an information story (in which customer reviews contain relevant information about fundamentals) and an attention story (in which customer reviews contain no relevant information about fundamentals but cause investors to respond in a naïve fashion). While several marketing studies examine the relation between online reviews and short-run stock returns, their results can be interpreted as consistent with both an information effect and an attention effect.[5] An exception is a contemporaneous marketing study by Fornell, Morgeson, and Hult (2016), which uses survey-based customer satisfaction scores observed at an annual frequency and shows that customer satisfaction scores positively predict stock returns and earnings surprises. My paper discriminates between these two possibilities by examining whether the predictability reverses in the long run and considering the relation between abnormal customer ratings and cash flow surprises. My findings suggest that, consistent with the Subrahmanyam and Titman (1999) proposition about serendipitous information production, customers in aggregate play the role of information producers in the financial market. Second, I construct a comprehensive sample of stocks with customer reviews on Amazon.com, which allows for reliable statistical inferences. The sample contains 346 distinct firms over a period of 12 years, which is much larger than those used by previous studies that examine online customer reviews. Third, by taking into account a number of known predictors in the cross section of stock returns as well as variables that are likely correlated with consumer opinions, my paper sheds light on the nature of the information conveyed by consumer opinions. The findings provide suggestive evidence that customer reviews in aggregate provide new information, which cannot be inferred from traditional sources such as accounting statements. Given the substantial growth of the amount of information produced by large crowds, future studies should investigate how such information can be used in other corporate contexts. Fourth, by exploring whether the predictability is particularly pronounced for stocks with high arbitrage costs and more binding limits to investor attention, this paper illuminates the sources of the stock return predictability of consumer opinions.

The remainder of this paper is organized as follows. Section 2 discusses the related literature. Section 3 describes the data and summary statistics. Section 4 presents the

[5]See, e.g., Chen, Liu, and Zhang (2012), Tirunillai and Tellis (2012), and Luo, Zhang, and Duan (2013).

empirical results, and Section 5 concludes.

2. Literature review

This paper is related to several strands of literature. The first is the finance literature on the influence of product-related information on stock pricing.[6] Subrahmanyam and Titman (1999) posit that such serendipitous information can have large aggregate effects on stock price efficiency. Early empirical studies find evidence that product recall events are associated with significantly negative stock market reactions, suggesting that inferior product quality and, hence, poor customer perceptions negatively affect stock returns (Jarrell and Peltzman, 1985; Barber and Darrough, 1996). More recently, Grullon, Kanatas, and Weston (2004) find that product market advertising increases the breadth of ownership and stock liquidity, suggesting that product advertisements influence investor decisions. Da, Engelberg, and Gao (2011b) show that Internet search volume for firms' products can serve as a leading indicator of a firm's earnings and stock prices.[7] My paper contributes to this literature by focusing on a direct measure of consumers' perceived quality of products. The findings highlight the role of consumers as information producers in financial markets, as suggested by Subrahmanyam and Titman (1999), and the influence of consumers' perceptions of product quality on stock returns.

The second literature this study contributes to is the marketing literature that examines the relation between online product reviews and stock returns (see, e.g., Tirunillai and Tellis, 2012; Chen, Liu, and Zhang, 2012; Luo, Zhang, and Duan, 2013; Luo and Zhang, 2013). For

[6]More broadly, this paper is related to the literature in marketing and accounting that examines the implications of customer satisfaction for firm performance and stock valuation (see, e.g., Ittner and Larcker, 1998; Anderson et al., 2004; Fornell et al., 2006). The literature generally shows that customer satisfaction (measured using customer survey data) is related to improved operating performance and higher firm valuation.

[7]While search volume can help gauge the potential demand for a company's products, whether it captures consumers' opinions about the products is less obvious. For instance, searches for a product can be prompted by both positive and negative news. Customer product ratings, in contrast, are a more direct measure of consumers' perceptions of product quality and value.

example, Tirunillai and Tellis (2012) examine the lead-lag relation between product reviews and stock market variables but find mixed results.[8] These marketing studies generally focus on a relatively small set of firms and do not distinguish between whether the relation is driven by information or investor attention. An important exception is a contemporaneous study by Fornell, Morgeson, and Hult (2016), which is closer to my paper in methodology and sample size and reaches similar conclusions. Using survey-based customer satisfaction scores observed at an annual frequency for a sample of about 300 firms over 15 years, Fornell, Morgeson, and Hult (2016) find that an investment strategy based on customer satisfaction scores delivers an abnormal return of 90 basis points per month. Similar to my study, they find that customer satisfaction positively predicts earnings surprises.

The third literature that this paper is connected to is that on the informational role of large crowds (oftentimes nonprofessional investors) in financial markets. For example, Da, Engelberg, and Gao (2011a), Kelley and Tetlock (2013), Chen, De, Hu, and Hwang (2014), and Lee, Ma, and Wang (2015) find evidence that the collective actions of large groups of financial market participants convey information about future stock returns and cash flows, whereas other studies suggest that some types of crowd activities provide little information about firm fundamentals (see, e.g., Tumarkin and Whitelaw, 2001; Antweiler and Frank, 2004; Das, Martinez-Jerez, and Tufano, 2005; Da, Engelberg, and Gao, 2011a). While some of these studies find evidence that public opinions online contain information for stock pricing, the source of these investors' information remains largely unknown. My paper complements these studies by highlighting product-related opinions of consumers as an important source of information that could be used for stock pricing. Given that consumers are among the most important nonfinancial stakeholders of firms (e.g., Titman, 1984; Maksimovic and Titman, 1991), the findings suggest that the information embedded in consumer opinions is of significant value to the market.[9]

[8] Tirunillai and Tellis (2012) show that numerical ratings (arguably the most salient measure of customers' perception of a company's products) do not have any predictability for subsequent stock returns. They also show that positive tone based on a textual analysis of the review contents does not have stock return predictability, whereas negative tone negatively predicts stock returns.

[9] More broadly, this paper is related to recent studies that examine the value of crowd-sourced opinions in

3. Data and summary statistics

This section describes the data sources and presents summary statistics of the main variables.

3.1. Amazon.com review data

Amazon.com is the largest online retailer in the US, generating $107.01 billion in sales in 2015. Founded in July 1994, the company started letting customers post reviews of its products in 1995. Since then, more than ten million customers have posted more than 200 million reviews on the website, making it the largest single source of Internet consumer reviews. According to Amazon.com's review creation guidelines, "anyone who has purchased items from Amazon.com" can write a product review, and the review "should focus on specific features of the product and [the customer's] experience with it."[10] Thus, customer reviews on Amazon.com likely capture consumers' perceptions of product quality and value. To minimize conflicts of interest, the guidelines prohibit paid reviews and manufacturers' posting of reviews for their own products or negative reviews of competing products. The guidelines also stipulate that the reviews should be about the product, not about the seller, the shipping experience, packaging, or product availability. Customers can rate a product on a scale of one to five stars, with five being the top rating, and enter a text review. All reviews are dated by the time it is first posted, which makes it possible to track consumer opinions over time. Amazon.com maintains all records of products and reviews on its website even when the products are discontinued. Therefore, the reviews are reasonably free of survivorship bias.

I use a three-pronged approach to identify public firms with customer product reviews on Amazon.com. I first retrieve the list of brands from Amazon.com under each product

various settings, such as new product innovations, the funding of start-ups, and scientific research. See, e.g., Poetz and Schreier (2012), Bayus (2013), Mollick (2014), and Franzoni and Sauermann (2014) for evidence on the value of crowdsourcing.

[10] See `http://www.amazon.com/gp/community-help/customer-reviews-guidelines`.

category and identify the companies that own these brands using various sources, including itemMaster.com, Consumer Product Information Database, and Google and Wikipedia searches. For changes of brand ownership due to mergers and acquisitions, I assign the brand to the new owner after the completion of the deal. This approach identifies about 250 public firms. Because not all product categories on Amazon.com provide a comprehensive brand list, firms identified by this approach are limited to those whose brands are listed under a product category. For example, Majesco is not listed as a brand under a category, but Majesco Entertainment Company, a public company that owns the brand, states in its 10-K report that the company has "developed close relationships with a number of retailers, including Amazon."[11] A search of the term "Majesco" on Amazon.com reveals that the company sells products on the e-commerce platform and has customer reviews for its products. I thus complement this approach with a second and a third approach described below that directly searches for brands owned by public firms on Amazon.com.

In my second approach to obtaining product reviews, I identify firms that use Amazon.com to sell their products by searching for the term "Amazon" in 10-K filings of all publicly traded firms in the US. I then check whether the firm sells its products on Amazon.com by searching for the company's brands and products on Amazon.com. In my third approach, I cross-validate by searching on Amazon.com for the brands and products of rivals of the companies identified in the above two approaches. I define rivals as those in the same four-digit Standard Industrial Classification industry. The idea is that if an industry has a firm that sells through Amazon.com, then its rivals could sell through the platform as well. In total, these procedures identify 346 public firms listed on the NYSE, AMEX, and Nasdaq. Table OA.1 of the Online Appendix provides the list of firms included in the sample and the number of reviews for each firm.

To collect the reviews for the sample of public firms, I use a web-crawling program that inputs each brand owned by a public firm as a search term on Amazon.com and outputs all

[11] See http://www.sec.gov/Archives/edgar/data/1076682/000114420414002141/v361719_10k.htm.

reviews for products whose brand name perfectly matches the search term.[12] Fig. 2 shows a sample Amazon.com webpage that contains product and review information. For each product, I retrieve the name of the product, the name of the brand, and the Amazon Standard Identification Number (ASIN). For each review, I collect the name and Amazon account identification (ID) of the reviewer, the date of the review, the numerical star rating, and the review text. The sample of reviews covers the period from July 2004 through December 2015.[13] I remove duplicate reviews posted by the same reviewer account ID on the same day for the same product, which constitute less than 0.01% of the review sample.[14] Table 1 reports the number of reviews, products, brands, and firms for the full sample as well as by Fama and French 12 industries. More than 14.5 million reviews are posted on 269,957 products manufactured by the sample firms. The top three industries in terms of the number of product reviews are business equipment (4.7 million reviews), consumer nondurables (2.6 million reviews), and manufacturing (2.0 million reviews).

[Insert Fig. 2 and Table 1 about here]

3.2. Summary statistics

I construct a panel data set of firm-months with customer reviews. To reduce noise in the data, I require that a firm-month have at least ten reviews to be included in the sample. The final sample has 20,562 firm-months. On average, each month has 150 stocks. For each firm in each month, I compute the simple average star rating of all customer reviews posted

[12] Amazon.com lists almost all product categories by brand (as illustrated in Fig. 2), which enables the program to retrieve the products and reviews by brand. Two exceptions are books and music compact disks (CDs), which are listed under the name of the author or the artist. The program thus does not collect reviews for these two product categories.

[13] The reason for choosing July 2004 as the starting date is twofold. First, Amazon.com in June 2004 disallowed anonymous reviews and introduced a credit card requirement for posting product reviews, which could improve the informativeness of the reviews. Second, the number of reviews for products manufactured by public companies is relatively low before 2004.

[14] Private communications with Amazon.com suggest that a customer can have multiple IDs (accounts) using different credit cards. Unfortunately, the current data I have cannot distinguish between multiple accounts held by the same reviewer using different credit cards. If such duplicate reviews do not contain relevant information (e.g., they can be used to manipulate ratings), they would bias in favor of finding return reversals (which I do not find).

for the firm's products during that month.[15] While individual consumers' opinions can be fraught with errors, averaging them can improve accuracy because errors tend to cancel out across individuals as long as the errors are not systematically correlated. To measure new information conveyed by customer ratings (i.e., surprises in customer ratings), I use average customer ratings in the prior 12 months as benchmarks for consumers' expectations of product quality and value. Averaging over a relatively long time window reduces the influence of transitory fluctuations in ratings and measurement errors. I thus measure abnormal customer ratings as the difference between the average customer rating in a month and that in the prior 12 months. An advantage of this measure is that it differences out time-invariant biases in customer reviews. In other words, even if the reviews are systematically biased, the bias does not affect the measure and the inferences insofar as the level of the bias is constant over time. Abnormal customer ratings can be driven by a number of factors. For example, customer ratings can change when a company introduces new products that are perceived differently by customers than its existing products. Customer ratings changes can also be induced by changes in the quality of products across different batches, changes in the competitive landscape, and changes in consumers' tastes and preferences.

Panel A of Table 2 shows that the mean abnormal customer rating is approximately zero and the interquartile range is 0.26. The average customer rating is 4.10, with an interquartile range of 0.48. The distribution of the level of ratings is concentrated at five stars, with about 64% of the ratings being five-star ratings (see Fig. OA.1 in the Online Appendix). The reviews are predominantly positive for three possible reasons. First, a selection effect could exist in that products sold on Amazon.com may on average have (or be perceived to have) relatively high quality. Second, products that receive low ratings could see their sales on Amazon.com decline and, in some cases, stop being offered, leading to fewer buyers and,

[15] Ideally, one would like to weight the reviews and products by their relative importance for the company's cash flows. However, due to a lack of product-level information (e.g., the fraction of earnings generated by a product), I treat each review equally, assuming that each review captures some aspect of the underlying quality of the firm's products that are common to the products. If some reviews or products are inconsequential to the fundamental value of the firm, this aggregation scheme would introduce noise into the tests and bias against finding significant results.

hence, fewer negative reviews. On the other hand, favorable reviews and product sales could be mutually reinforcing, with favorable reviews attracting more buyers and these buyers in turn posting more favorable reviews to the extent that consumers have correlated opinions. This would shift the distribution of ratings to the right. Third, the reviewers on Amazon.com could be systematically upward-biased. It is worth noting that the use of abnormal ratings (i.e., the average rating in a month minus the average rating during the prior 12 months) differences out time-invariant biases in customer reviews. Thus, the bias does not affect the measure and the inferences insofar as the level of the bias is constant over time.

[Insert Table 2 about here]

To construct other firm-level variables, I obtain stock return and volume data from the Center for Research on Stock Prices (CRSP), financial statement data from Standard and Poor's (S&P) Compustat, and analysts' earnings forecasts from the Institutional Brokers' Estimate System (I/B/E/S). In addition to commonly used firm characteristics such as size and book-to-market, I consider a number of characteristics that are likely to be associated with stock return predictability or consumer opinions. In particular, to mitigate the concern that customer ratings may simply proxy for firms' operating performance that can be inferred from financial statements, I construct the F-score of Piotroski (2000) and gross profitability of Novy-Marx (2013). The F-score is the sum of nine binary variables capturing financial performance signals, and gross profitability is defined as the ratio of income before extraordinary items to book value of assets. I also consider advertising and research and development (R&D) expenditures, because consumer opinions can be influenced by firms' advertising campaigns and R&D investments. Moreover, I consider the level and variation of dollar trading volume of Brennan, Chordia, and Subrahmanyam (1998) and Chordia, Subrahmanyam, and Anshuman (2001) to control for an attention effect.

Panel A of Table 2 reports the summary statistics of these firm characteristics in my sample. For example, the average firm has a market capitalization of $25.72 billion, a book-to-market ratio of 0.42, a buy-and-hold return of 17.1% during the prior 12 months, gross profitability of 0.11, an F-score of 5.3, a monthly trading volume of \$3.7 billion, and 16.4 analysts. Table OA.2 in the Online Appendix compares the characteristics of the firms included in my sample with same-industry firms in the CRSP and Compustat universe. Firms with Amazon.com customer reviews on average tend to have a larger market capitalization, lower book-to-market ratio, higher past stock return, greater advertising expenditure, higher gross profitability and F-score, higher trading volume, and higher analyst coverage than the median same-industry firm in the broader sample. The fact that the firms that have available data on customer reviews are generally larger and have more analyst coverage than those without reviews suggests that the sample selection may bias against finding significant return predictability given that large-cap firms and firms with greater analyst coverage are generally priced more efficiently.

I construct two measures to capture cash flow surprises. The first is revenue surprises. Following Jegadeesh and Livnat (2006), I use standardized unexpected revenue growth estimator ($SURGE$) as a measure of revenue surprises. I define the revenue surprise of firm i in quarter q as

$$SURGE_{i,q} = \frac{REV_{i,q} - E(REV_{i,q})}{\sigma_{i,q}}, \tag{1}$$

where $REV_{i,q}$ is the quarterly revenue per share for firm i in quarter q, $E(REV_{i,q})$ is the expected quarterly revenue per share, and $\sigma_{i,q}$ is the standard deviation of quarterly revenue growth. $E(REV_{i,q})$ is estimated under the assumption that quarterly revenue follows a seasonal random walk with drift, i.e., $E(REV_{i,q}) = REV_{i,q-4} + \frac{1}{8}\sum_{j=1}^{8}(REV_{i,q-j} - REV_{i,q-j-4})$. Panel A of Table 2 shows the summary statistics of revenue surprises. $SURGE$ has a mean of 0.52 and an interquartile range of about 1.31.

The second is earnings surprises. I obtain quarterly earnings forecasts for the sample of stocks from the I/B/E/S historical database. I use price-scaled forecast errors (SUE) as a measure for earnings surprises, defined as the difference between reported quarterly earnings per share (EPS) and the median of the most recent EPS forecasts of all analysts issued during the 90-day period prior to the earnings announcement date scaled by the stock price. Panel A of Table 2 shows that SUE has a mean of 0.077% and an interquartile range of about 0.197 percentage points.

I also construct measures for trading activities by hedge funds. I obtain the list of hedge funds from Gao and Huang (2016), which contains 494 distinct hedge fund managers. I retrieve their quarterly holdings from Thomson Reuters CDA/Spectrum Institutional (13F) Holdings Database. I define net purchases of stock i by hedge funds in quarter q as

$$NetBuy_{i,q}^{HF} = \frac{ShrOwn_{i,q}^{HF}}{ShrOut_{i,q}} - \frac{ShrOwn_{i,q-1}^{HF}}{ShrOut_{i,q-1}}, \tag{2}$$

where $ShrOwn_{i,q}^{HF}$ ($ShrOwn_{i,q-1}^{HF}$) is the number of shares of firm i held by hedge funds in quarter q ($q-1$) and $ShrOut_q$ ($ShrOut_{q-1}$) is firm i's number of shares outstanding in quarter q ($q-1$). To be included in the calculation of the measure, a fund has to be a 13F filer in both quarters q and $q-1$. In other words, increases (decreases) in hedge fund holdings due to entries (exits) of funds are excluded from the calculation. To provide a basis for comparison, I construct a similar measure for all other 13F institutions, i.e., excluding hedge funds. Panel A of Table 2 shows that net buying by hedge funds of the sample firms during 2004 to 2015 has a mean close to zero and an interquartile range of about 1.25 percentage points. Net buying by non–hedge funds has a mean of 0.056% and an interquartile range of about 3.16 percentage points.

Panel B of Table 2 presents regression analysis of the determinants of abnormal customer ratings. I regress abnormal customer ratings on lagged stock characteristics, including advertising, R&D, gross profitability, F-score, the level and variation of dollar trading volume, market capitalization, book-to-market ratio, past stock returns, book leverage, asset tangibility, analyst coverage, and institutional ownership. None of the stock or firm characteristics can reliably predict abnormal customer ratings. This finding suggests that abnormal customer ratings are largely independent of the information contained in stock and firm characteristics, which are observed at lower frequencies.[16] In addition, to test the possibility that some market participants may possess information about consumer opinions, I include analyst forecast revisions and hedge funds' net purchases as explanatory variables in the regressions. I define analyst forecast revisions as the difference between the mean of current-year EPS forecasts made by analysts during a given month and that in the previous month scaled by the stock price at the end of the previous month. The results, reported in the last column of the panel, show that the coefficients on these variables are insignificant, suggesting that neither analysts nor hedge funds can predict consumer opinions.

4. Empirical results

This section presents empirical results on the information content of customer reviews.

4.1. Abnormal customer ratings and stock return predictability

4.1.1. Calendar-time portfolio tests

I use a calendar-time portfolio approach to examine the investment value of abnormal customer ratings. For each month from July 2004 through December 2015, I sort sample stocks into tercile portfolios based on abnormal customer ratings. I then track the performance of the three portfolios over the following month. I employ two weighting schemes across firms, weighting by the number of reviews and equal weighting. The first weighting scheme is motivated by the idea that crowds may provide more precise information when they are larger. This weighting scheme lessens the possibility that the portfolios are dominated by firms with a relatively small number of reviews that may contain more noise. I use the

[16] I use two alternative specifications to address the concern that the dependent variable is bounded, namely, a Tobit specification and a logit transformation of the dependent variable. The results, reported in Table OA.3 in the Online Appendix, show that abnormal customer ratings continue to be largely independent from lagged firm characteristics. Therefore, the lack of significance reported here does not seem to be explained by misspecification problems.

Fama-French-Carhart four-factor model to adjust returns. I compute a four-factor alpha by regressing monthly portfolio excess returns on the monthly returns from the risk factors.

Table 3 reports the alpha and factor loadings from the monthly calendar-time portfolio regressions of the investment strategy for both review weighted and equal weighted portfolios. Stocks with abnormal customer ratings in the top tercile outperform a passive benchmark by about 53 basis points per month, which is significant at the 5% level.[17] Stocks with abnormal customer ratings in the bottom tercile under-perform by 2.4 to 19.8 basis points per month (statistically insignificant). A spread portfolio that buys stocks in the top tercile of abnormal customer ratings and sells stocks in the bottom tercile outperforms passive benchmarks by around 55.7 to 73.0 basis points per month. The spread portfolio generally does not have significant exposure to the risk factors, which is not surprising considering that abnormal customer ratings likely capture firm-specific information.

[Insert Table 3 about here]

The alpha earned by the spread portfolio is based on gross returns without accounting for transaction costs. The trading strategy likely incurs large transaction costs, because it requires relatively frequent trading and some stocks can be expensive to short. For example, the average monthly portfolio turnover rate, i.e., the average of purchase and sales turnover rates, of the spread portfolio is around 50.0%. Considering that the one-way trading costs for US stocks are 16 basis points per dollar trading volume in 2004 (the start of the sample period in this paper) and decrease over the years (see, Table V of French, 2008), the trading strategy is likely to deliver a positive alpha net of trading costs, although the magnitude of the alpha would be smaller.

[17] The tercile portfolios generally have negative loadings on the momentum (UMD) factor. This is driven almost entirely by a three-month period from March to May 2009, which is associated with momentum crashes (Daniel and Moskowitz, 2016). The UMD factor delivers highly negative returns (the average monthly return being −19.48%) during these three months, whereas the average monthly return of the overall market during the same period is highly positive at 8.12%. Therefore, these three months generate a large negative correlation between the UMD factor and the individual stock returns for the average stock in the market. In untabulated analysis, the coefficient on the UMD factor generally becomes insignificant when these three months are removed.

To shed light on the sources of the predictability of future stock returns based on abnormal customer ratings, I conduct two subsample analyses. First, limits to arbitrage may inhibit informed investors' ability to fully and instantaneously capitalize on their information (Pontiff, 1996; Shleifer and Vishny, 1997). Therefore, the return predictability of consumer opinions should be stronger for stocks with more binding limits to arbitrage. I use idiosyncratic stock return volatility as a measure for arbitrage costs. I compute idiosyncratic volatility for a stock-month as the standard deviation of the residuals estimated from the Fama and French (1993) three-factor model on daily data during the prior 90 days. I partition the sample of stocks into two groups based on the median idiosyncratic volatility. I construct spread portfolios for each of the two groups of stocks separately. Panel A of Table 4 reports the abnormal returns of these two spread portfolios. The Fama-French-Carhart four-factor alphas of the spread portfolios consisting of stocks with high idiosyncratic volatilities are around 103.0 to 136.6 basis points per month and are statistical significant at the 5% level, whereas those for low idiosyncratic volatility stocks are small and insignificant. The difference in the abnormal returns between the two spread portfolios is around 89.9 to 138.0 basis points and is significant at the 10% level for both weighting schemes. This finding suggests that arbitrage costs play a role in generating the return predictability.

[Insert Table 4 about here]

Second, limits to investor attention and information processing capacities may delay the incorporation of customer information into the stock price. I use analyst coverage and firm size as proxies for investor attention (Hirshleifer and Teoh, 2003; Hong, Lim, and Stein, 2000). If investor inattention slows the incorporation of the information embedded in consumer reviews, the predictability should be concentrated among firms with low analyst coverage and small firms. I retrieve quarterly earnings forecasts for the sample of stocks from the I/B/E/S historical database and compute the number of financial analysts making forecasts for a stock in a given quarter. I partition the sample of stocks into two groups based on the median number of analysts covering the stock. I construct spread portfolios within each of the two groups of stocks separately. Panel B of Table 4 reports the abnormal returns of these two spread portfolios. The return predictability of consumer reviews comes mainly from stocks with fewer analysts. The spread portfolio consisting of stocks with low analyst coverage delivers a four-factor alpha of 73.4 to 148.0 basis points per month. The spread portfolio consisting of stocks with high analyst coverage has insignificant alphas. The difference is around 34.1 to 129.1 basis points and is significant at the 10% level for the spread portfolio weighted by the number of reviews. I conduct similar subsample tests using market capitalization as the partitioning variable. Small firms are firms with a market cap below the median, and large firms are those with a higher than the median market cap. Panel C of Table 4 shows that the predictability is driven mainly by small firms (although the performance difference is statistically insignificant). These findings are broadly consistent with the notion that limits to investor attention play a role in the observed return predictability.

I conduct several additional tests to assess the robustness of the main results. First, I consider several alternative risk benchmarks to adjust the returns of the portfolios in the calendar-time portfolio tests, including the Fama and French (1993) three-factor model, the Fama and French (2015) five-factor model, and a liquidity-augmented Fama-French-Carhart model. In addition to the Fama and French three factors, the five-factor model includes two new factors, namely a robust minus weak (RMW) factor and a conservative minus aggressive (CMA) factor. These two additional factors capture the performance difference driven by operating profitability and investments, which helps mitigate the concern that the observed alphas on the spread portfolios are driven by high rating firms being more profitable or having a less aggressive investment policy. I construct the liquidity factor IML (illiquid minus liquid) using an algorithm similar to the one in Fama and French (1993) for their SMB (small minus big) and HML (high minus low) factors. I use the Amihud (2002) illiquidity ratio as a proxy for stock liquidity. Moreover, to address the concern that the results could be driven by industry effects, I compute industry–adjusted stock returns by subtracting from each stock return over a period the return of the corresponding industry over the same period (following Moskowitz and Grinblatt, 1999). Panel A of Table 5 shows that the results are qualitatively similar to those obtained using the main specification. For example, the alpha estimates on the spread portfolios range from 51.3 to 78.9 basis points when I use the Fama and French five-factor model to adjust the returns.

[Insert Table 5 about here]

Second, I construct an alternative measure for abnormal customer ratings by using a six-month window prior to the current month to measure the normal level of customer ratings. The abnormal customer rating is thus measured as the average rating in the current month minus that in the prior six months. The results, reported in Panel B of Table 5, show that the four-factor alphas on the spread portfolio continue to be positive and significant at conventional levels.

Third, to evaluate the robustness of the strategy based on customer ratings over time, I examine the performance of the strategy by year. The number of reviews has increased dramatically over time as Amazon.com grows its merchandise offerings and customer base and the review platform becomes more popular among consumers.[18] For example, the average number of reviews posted for a firm-month has increased from 144 in 2004 to $1,671$ in 2015. The increase in the popularity of customer reviews can have two effects on the stock return predictability of consumer opinions over time. First, aggregating over a larger number of consumer reviews can reduce noise and provide more precise information about the products. This suggests that reviews in more recent years should be more informative about stock returns. Second, as consumer reviews become more informative, more arbitrageurs may trade on the information embedded in the reviews. As a result, the information may be incorporated into stock prices more rapidly, giving rise to weaker return predictability of consumer opinions in more recent years.

To examine the performance of the review-based strategy over time, I calculate the one-month-ahead abnormal return for each stock in the tercile portfolios sorted by abnormal

[18] According to the 10-K filings of Amazon.com Inc., net sales revenue has increased from $6.92 billion in 2004 to $107.01 billion in 2015.

customer ratings using the Fama-French-Carhart four-factor model. I use monthly stock returns data during the prior 60 months to estimate factor loadings. I then compute the number of reviews–weighted and equally weighted abnormal returns of each tercile portfolio. Fig. 3 plots the differences in the average monthly abnormal returns between the top and bottom tercile portfolios by year. The outperformance of the top tercile portfolio relative to the bottom tercile portfolio does not appear to be driven by one or two extreme years. For the portfolios weighted by the number of reviews, the top tercile portfolio of abnormal customer ratings consistently outperforms the bottom tercile portfolio in all sample years except 2007 and 2008.[19] Similarly, for the equally weighted portfolios, the top tercile portfolio outperforms the bottom tercile portfolio in all years except 2004 and 2015. I perform the augmented Dickey-Fuller test to examine whether the performance of the strategy is stationary. The result strongly rejects the null that the return differences between the top and bottom tercile portfolios are nonstationary (p-value < 0.001). These results suggest that the two countervailing effects described above largely offset each other.

4.1.2. Long-run stock returns

If consumer opinions contain relevant information about firms' fundamentals, the return predictability of abnormal customer ratings should not reverse subsequently. If, however, abnormal customer ratings contain no relevant information about fundamentals but cause investors to respond in a naïve fashion, the predictability should eventually reverse. I thus conduct the calendar-time portfolio tests over different holding periods. I consider four holding periods: three months (months 2 through 4), six months (months 2 through 7), nine months (months 2 through 10), and 12 months (months 2 through 13) after portfolio formation, all skipping the first month post-formation. I again construct spread portfolios that buy stocks in the top tercile of abnormal customer ratings and sell stocks in the bottom tercile. The portfolios are rebalanced monthly by adding stocks that enter the top or the

[19] The four-factor alpha on the review-weighted spread portfolio in the baseline specification increases from 0.730% to 0.938% when the crisis period, 2007 and 2008, is removed.

bottom tercile and dropping stocks that have reached their holding periods. The results, reported in Table 6, show that the four-factor alphas on the spread portfolios are all positive, but statistically indistinguishable from zero. These results provide evidence that the return predictability of abnormal customer ratings does not reverse in the long run, suggesting that the information embedded in consumer reviews is likely about fundamentals.

[Insert Table 6 about here]

4.1.3. Fama-MacBeth regressions

The results so far show that consumer opinions contain relevant information for stock pricing. A natural question is whether this information is new information. While the results in Panel B of Table 2 show that abnormal customer ratings are largely uncorrelated with lagged accounting variables, the results could simply capture return predictability from contemporaneous accounting variables. I thus conduct Fama-MacBeth regressions to test the return predictability of abnormal customer ratings by explicitly controlling for accounting variables measured contemporaneously to customer ratings and other known predictors in the cross section of stock returns. For each month, I run the following cross-sectional regression:

$$ExcessRet_{i,t+1} = \alpha + \beta_1 AbnRating_{i,t} + \gamma \mathbf{X}_{i,t} + \varepsilon_{i,t}, \tag{3}$$

where $ExcessRet_{i,t+1}$ is the excess return, i.e., raw return in excess of the one-month T-bill rate, of stock i in month $t+1$;[20] $AbnRating_{i,t}$ is the abnormal customer rating of stock i during month t; $\mathbf{X}_{i,t}$ is a vector of common firm characteristics of firm i in month t. I include the following firm characteristics, measured using the most recent quarterly data, as controls: market capitalization, book-to-market ratio, past stock return, the F-score of Piotroski (2000), gross profitability of Novy-Marx (2013), level and variation of dollar trading

[20] For robustness, I repeat the regression with the Fama-French-Carhart four-factor adjusted return as the dependent variable. The results, reported in Table OA.4 in the Online Appendix, are qualitatively similar to those reported here.

volume of Brennan, Chordia, and Subrahmanyam (1998) and Chordia, Subrahmanyam, and Anshuman (2001), and variables that are likely correlated with consumer opinions, including advertising and R&D.

Table 7 reports the time series averages of the cross-sectional regression coefficients. I calculate Fama-MacBeth t-statistics using Newey-West standard errors with four lags. Model 1 shows that the coefficient on abnormal customer ratings is 1.72 and significant at the 5% level when abnormal customer ratings are the only regressor in the regression. The subsequent columns add different sets of control variables. The results show that abnormal customer ratings continue to exhibit stock return predictability after adding these controls, with the coefficient on abnormal ratings largely unchanged across the specifications. The economic magnitude is large as well. For example, the coefficient estimate on abnormal customer ratings in Model 6 indicates that, after controlling for other predictors of stock returns, an interquartile range increase in abnormal customer ratings (about 0.26) is associated with an increase of approximately 34.40 basis points in the one-month-ahead excess returns. These results suggest that the information contained in consumer opinions is likely new information, which cannot be inferred from other sources such as accounting statements. While other accounting and firm characteristics predict returns in the cross section, controlling for these factors does not affect the relation between abnormal customer ratings and future stock returns.

[Insert Table 7 about here]

Table 7 also reveals a number of other interesting findings. Consistent with Piotroski (2000) and Novy-Marx (2013), F-score and gross profitability positively predict subsequent stock returns. Also, consistent with the evidence in Brennan, Chordia, and Subrahmanyam (1998) and Chordia, Subrahmanyam, and Anshuman (2001), the level of dollar trading volume and the variability of dollar trading volume are significant negative predictors of stock returns. Other firm characteristics, including advertising, R&D, market capitalization, book-to-market, and past stock returns, are generally insignificant in predicting future stock returns, which could be specific to the sample employed.[21]

4.2. Abnormal customer ratings and cash flow surprises

To affect stock pricing, consumer opinions must contain novel information about firms' cash flows. I use revenue surprises and earnings surprises to capture new information in cash flows. Because revenues and earnings are released at a quarterly frequency, I compute the abnormal customer rating at a quarterly frequency as the average customer rating during a quarter minus that during the prior four quarters. The final sample for revenue surprises contains $7,283$ observations with non-missing data on relevant variables between July 2004 and December 2015, and that for earnings surprises has $5,503$ observations because calculating analyst forecast dispersion, one of the control variables, requires that the firms be covered by at least two analysts.

Following Tetlock, Saar-Tsechansky, and Macskassy (2008) and Chen, De, Hu, and Hwang (2014), I run the following panel regression:

$$SURPRISE_{i,q} = \alpha + \tau_q + \beta_1 AbnRating_{i,q} + \beta_2 SURPRISE_{i,q-1} + \gamma \mathbf{X}_{i,q-1} + \varepsilon_{i,q}, \tag{4}$$

where $SURPRISE_{i,q}$ is either revenue or earnings surprises for firm i in quarter q (as discussed in Subsection 3.2); $AbnRating_{i,q}$ is the abnormal customer rating during quarter q; $\mathbf{X}_{i,q-1}$ is a vector of firm characteristics of firm i in quarter $q-1$, including advertising, R&D expenses, gross profitability, F-score, the level and variation of dollar trading volume, firm size, book-to-market ratio, and past stock returns; and τ_q is time fixed effects.[22] I also estimate, as an alternative specification, regressions that include firm fixed effects and

[21] For example, the SMB, HML, and UMD factors earn 0.11%, −0.05%, and 0.20% per month (with t-statistics of 0.57, 0.24, and 0.49), respectively, during my sample period.

[22] For robustness, I include contemporaneous advertising and R&D expenses, measured as of quarter q, as controls to mitigate the concern that these two variables drive both abnormal customer ratings and cash flow surprises. The results, reported in the Online Appendix, are little changed from those reported here.

exclude the lagged dependent variable. I cluster standard errors by firm and by quarter (Petersen, 2009). Abnormal customer ratings are measured before the release of quarterly cash flow information, which typically occurs 30–40 days after a quarter-end. If customer ratings contain fundamental information, the coefficient on $AbnRating_{i,q}$ should be significant and positive.

Table 8 presents the regression results with revenue surprises as the dependent variable. Abnormal customer ratings are a significant positive predictor of revenue surprises. The economic magnitude is nontrivial. For example, Model 3 shows that an interquartile increase in abnormal customer ratings is associated with an increase of 8.84 percentage points in revenue surprises, which is about 6.7% of the interquartile range of revenue surprises. This finding suggests that consumer opinions contain novel cash flow information. The predictability of abnormal customer ratings is obtained after controlling for other determinants of revenue surprises. For example, lagged revenue surprises strongly positively predict current revenue surprises. There is also some evidence that gross profitability, F-score, and past stock returns positively predict revenue surprises.

[Insert Table 8 about here]

Turning to earnings surprises, Table 9 shows that abnormal customer ratings significantly positively predict earnings surprises, suggesting that consumer opinions contain novel information about firms' earnings. In terms of economic magnitudes, Model 3 shows that an interquartile increase in abnormal customer ratings is associated with an increase of 0.020 percentage points in price-scaled unexpected earnings (SUE), which is about 10% of the interquartile range of SUE. This finding also suggests that the average analyst does not fully impound the information contained in consumer opinions in her forecasts. Consistent with prior work, Table 9 also shows that lagged earnings surprises positively predict current earnings surprises. Also, some evidence shows that stock returns prior to the release of quarterly earnings positively predict earnings surprises.

[Insert Table 9 about here]

4.3. *Abnormal customer ratings and trading by sophisticated investors*

Because consumer opinions contain value-relevant information about stocks, sophisticated investors may exploit such information in their trading decisions. Hedge fund managers appear to fit the profile of informed traders in the equity market. A growing literature shows that hedge funds outperform other institutional investors, such as mutual funds, suggesting that hedge fund managers possess stock-picking skills (see, among others, Ackermann, McEnally, and Ravenscraft, 1999; Brown, Goetzmann, and Ibbotson, 1999; Brunnermeier and Nagel, 2004). If hedge fund managers exploit information contained in consumer opinions, they should trade in the direction indicated by consumer signals.

To examine whether abnormal customer ratings predict hedge fund trades, I run Fama-MacBeth regressions of net purchases by hedge funds and those by non–hedge funds in quarter q on lagged abnormal customer ratings and controls for stock returns and firm characteristics (following Cohen and Frazzini, 2008). Because hedge fund trades are inferred from quarterly holdings reports, I again calculate the abnormal customer rating at a quarterly frequency as the average customer rating during a quarter minus that during the prior four quarters. For each quarter, I run the following cross-sectional regression:

$$NetBuy_{i,q} = \alpha + \beta_1 AbnRating_{i,q-1} + \beta_2 Ret_{i,q} + \beta_3 Ret_{i,q-1} + \beta_3 Ret_{i,[q-5,q-2]} + \gamma \mathbf{X}_{i,q-1} + \varepsilon_{i,q}, \quad (5)$$

where $NetBuy_{i,q}$ is net purchases either by hedge funds or by non–hedge fund institutions in quarter q; $AbnRating_{i,q-1}$ is abnormal customer ratings in quarter $q-1$; $Ret_{i,q}$, $Ret_{i,q-1}$, and $Ret_{i,[q-5,q-2]}$ are the stock returns in quarter q , quarter $q-1$, and quarters $q-5$ through $q-2$, respectively;[23] $\mathbf{X}_{i,q-1}$ is a vector of common firm characteristics of firm i in quarter $q-1$, including advertising, R&D, gross profitability, F-score, level and variation of dollar

[23] The results are qualitatively similar when $Ret_{i,q}$ is excluded from the regression.

trading volume, firm size, and book-to-market ratio. If hedge funds make use of information contained in customer ratings, the coefficient on $AbnRating_{i,q-1}$ should be significant and positive.

Table 10 reports the time series averages of the cross-sectional regression coefficients. I calculate Fama-MacBeth t-statistics using Newey-West standard errors with four lags. The coefficient on $AbnRating_{i,q-1}$ in the regression of hedge funds' net purchases is positive and significant. In terms of economic magnitudes, an interquartile increase in abnormal customer ratings is associated with an increase of 0.059 percentage points in hedge funds' net purchases, which is about 5% of the interquartile range of net buying by hedge funds. The magnitude is reasonable considering that hedge funds likely face a myriad of information sources and consumer information is just one of them. In contrast, the coefficient on $AbnRating_{i,q-1}$ in the regression of non–hedge funds' net purchases is negative and insignificant. The difference in the coefficient between the two regressions is significant at the 5% level.

[Insert Table 10 about here]

Because some hedge fund managers have better information processing abilities than others, there could be heterogeneity across hedge funds in their propensity to trade on information contained in consumer opinions. I hypothesize that hedge funds with more trading in stocks that have consumer reviews are likely to be better informed. I term these hedge funds "specialized hedge funds". Because customer reviews contain information about firms' cash flows in the near future, it requires informed investors to trade actively. Therefore, abnormal customer ratings should have a stronger predictive power for trades by specialized hedge funds. To test this, I partition hedge funds into two groups based on the median of the trading weight in stocks with Amazon.com reviews, i.e., trading volume in stocks with Amazon.com reviews as a fraction of total trading volume over the last four quarters. I infer trading volume from disclosed quarterly holdings by assuming that hedge funds do not trade intra-quarterly between two consecutive quarterly reports and the changes in holdings during a quarter occur at the end of the quarter. I compute net purchases by the two groups of hedge funds separately. I repeat the regressions in Eq. (5) for net purchases by specialized and less specialized hedge funds separately and report the results in the last two columns of Table 10. The results show that abnormal customer ratings significantly positively predict net buying by specialized hedge funds. On the other hand, an insignificant relation exists between abnormal customer ratings and net purchases by less specialized hedge funds. These results provide evidence that hedge fund managers that specialize in processing information embedded in consumer reviews actively trade on such information. Combined with the findings in Panel B of Table 2, these results suggest that, although hedge fund trades do not predict consumer product reviews, they do respond to information contained in recent product reviews.

5. Conclusions

In this paper, I examine the investment value of consumer opinions. Using a large data set of customer product reviews on Amazon.com, I find that abnormal customer ratings positively predict subsequent stock returns. A spread portfolio that buys stocks with abnormal customer ratings in the top tercile and sells stocks in the bottom tercile generates an abnormal return of about 55.7 to 73.0 basis points per month. The results appear to be concentrated among stocks with high idiosyncratic volatilities, stocks with low analyst coverage, and small-cap stocks, which likely face high arbitrage costs and more binding limits to investor attention. Fama-MacBeth regressions show that the return predictability of customer ratings continues to hold after controlling for firm characteristics such as gross profitability, advertising, and trading volume. I find, consistent with the idea that consumer ratings contain novel cash flow information, that abnormal customer ratings positively predict revenue and earnings surprises and the return predictability does not reverse in the long run. Last, abnormal customer ratings are a significant predictor of net purchases by hedge fund managers, suggesting that sophisticated investors exploit the information contained in consumer opinions. Taken together, these findings provide evidence that the aggregated opinions of consumer crowds contain valuable information about cash flows and stock pricing.

The results in this paper highlight the role of consumers as information producers in financial markets. Compared with traditional information intermediaries such as equity analysts, consumer crowds can provide more timely information on a company's products and cash flows. Given the collective wisdom of consumers, future research should investigate how firms and other stakeholders such as creditors and suppliers can make use of the information conveyed by consumer opinions.

References

Ackermann, C., McEnally, R., Ravenscraft, D., 1999. The performance of hedge funds: Risk, return, and incentives. Journal of Finance 54, 833–874.

Allen, F., 1984. Reputation and product quality. RAND Journal of Economics 15, 311–327.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets 5, 31–56.

Anderson, E. W., Fornell, C., Mazvancheryl, S., 2004. Customer satisfaction and shareholder value. Journal of Marketing 68, 172–185.

Antweiler, W., Frank, M., 2004. Is all that talk just noise? The information content of Internet stock message boards. Journal of Finance 59, 1259–1294.

Barber, B., Darrough, M., 1996. Product reliability and firm value: the experience of American and Japanese automakers, 1973–1992. Journal of Political Economy 104, 1084–1099.

Bayus, B., 2013. Crowdsourcing new product ideas over time: an analysis of the Dell IdeaStorm community. Management Science 59, 226–244.

Bikhchandani, S., Hirshleifer, D., Welch, I., 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of Political Economy 100, 992–1026.

Brennan, M., Chordia, T., Subrahmanyam, A., 1998. Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. Journal of Financial Economics 49, 345–373.

Brown, S., Goetzmann, W., Ibbotson, R., 1999. Off-shore hedge funds: survival and performance. Journal of Business 72, 91–117.

Brunnermeier, M., Nagel, S., 2004. Hedge funds and the technology bubble. Journal of Finance 59, 2013–2040.

Cabral, L., 2000. Stretching firm and brand reputation. RAND Journal of Economics 31, 658-673.

Chen, H., De, P., Hu, Y., Hwang, B., 2014. Wisdom of crowds: the value of stock opinions transmitted through social media. Review of Financial Studies 27, 1367–1403.

Chen, Y., Liu, Y., Zhang, J., 2012. 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, 116–135.

Choi, J., 1998. Brand extension as informational leverage. Review of Economic Studies 65, 655-669.

Chordia, T., Subrahmanyam, A., Anshuman, V., 2001. Trading activity and expected stock returns. Journal of Financial Economics 59, 3–32.

Cohen, L., Frazzini, A., 2008. Economic links and predictable returns. Journal of Finance 63, 1977–2011.

Da, Z., Engelberg, J., Gao, P., 2011a. In search of attention. Journal of Finance 66, 1461–1499.

Da, Z., Engelberg, J., Gao, P., 2011b. In search of fundamentals. Unpublished working paper, University of Notre Dame.

Daniel, K., Moskowitz, T., 2016. Momentum crashes. Journal of Financial Economics 122, 221–247.

Das, S., Martinez-Jerez, A., Tufano, P., 2005. eInformation: a clinical study of investor discussion and sentiment. Financial Management 34, 103–137.

DellaVigna, S., Pollet, J., 2007. Demographics and industry returns. American Economic Review 97, 1667–1702.

DellaVigna, S., Pollet, J., 2009. Investor inattention and Friday earnings announcements. Journal of Finance 64, 709–749.

Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

Fama, E., French, K., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Fornell, C., Mithas, S., Morgenson, F., Krishnan, M., 2006. Customer satisfaction and stock prices: high returns, low risk. Journal of Marketing 70, 1–14.

Fornell, C., Morgeson, F., Hult, T., 2016. Stock returns on customer satisfaction do beat the market: gauging the effect of a marketing intangible. Journal of Marketing 80, 92–107.

Franzoni, C., Sauermann, H., 2014. Crowd science: the organization of scientific research in open collaborative projects. Research Policy 43, 1–20.

French, K., 2008, Presidential address: The cost of active investing. Journal of Finance 63, 1537–1573.

Gao, M., Huang, J., 2016. Capitalizing on Capitol Hill: informed trading by hedge fund managers. Journal of Financial Economics 121, 521–545.

Grullon, G., Kanatas, G., Weston, J., 2004. Advertising, breadth of ownership, and liquidity. Review of Financial Studies 17, 439–461.

Harvard Business Review, 2012. What if Google had a hedge fund? September 28.

Hirshleifer, D., Teoh, S., 2003. Limited attention, information disclosure, and financial reporting. Journal of Accounting and Economics 36, 337–386.

Hong, H., Lim, T., Stein, J., 2000. Bad news travels slowly: size, analyst coverage, and the profitability of momentum strategies. Journal of Finance 55, 265–295.

Hong, H., Stein, J., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. Journal of Finance 54, 2143–2184.

Hong, H., Torous, W., Valkanov, R., 2007. Do industries lead the stock market? Journal of Financial Economics 83, 367–396.

Hou, K., Moskowitz, T., 2005. Market frictions, price delay, and the cross-section of expected returns. Review of Financial Studies 18, 981–1020.

Huberman, G., Regev, T., 2001. Contagious speculation and a cure for cancer: a nonevent that made stock prices soar. Journal of Finance 56, 387–396.

Ittner, C., Larcker, D., 1998. Are nonfinancial measures leading indicators of financial performance? An analysis of customer satisfaction. Journal of Accounting Research 36, 1–35.

Jarrell, G., Peltzman, S., 1985. The impact of product recalls on the wealth of sellers. Journal of Political Economy 93, 512–536.

Jegadeesh, N., Livnat, J., 2006. Revenue surprises and stock returns. Journal of Accounting and Economics 41, 147–171.

Kelley, E., Tetlock, P., 2013. How wise are crowds? Insights from retail orders and stock returns. Journal of Finance 68, 1229–1265.

Lee, C., Ma, P., Wang, C., 2015. Search-based peer firms: Aggregating investor perceptions through Internet co-searches. Journal of Financial Economics 116, 410–431.

Lerner, J., Tirole, J., 2002. Some simple economics of open source. Journal of Industrial Economics 50, 197–234.

Luo, X., Zhang, J., 2013. How do consumer buzz and traffic in social media marketing predict the value of the firm? Journal of Management Information Systems 30, 213–238.

Luo, X., Zhang, J., Duan, W., 2013. Social media and firm equity value. Information Systems Research 24, 146–163.

Lynch, Peter, 1989, One Up on Wall Street. Penguin Books, New York.

Maksimovic, V., Titman, S., 1991. Financial policy and reputation for product quality. Review of Financial Studies 4, 175–200.

Mollick, E., 2014. The dynamics of crowdfunding: determinants of success and failure. Journal of Business Venturing 29, 1–16.

Moskowitz, T., Grinblatt, M., 1999. Do industries explain momentum? Journal of Finance 54, 1249–1290.

Novy-Marx, R., 2013. The other side of value: the gross profitability premium. Journal of Financial Economics 108, 1–28.

Peng, L., Xiong, W., 2006. Investor attention, overconfidence, and category learning. Journal of Financial Economics 80, 563–602.

Petersen, M., 2009. Estimating standard errors in finance panel data sets: comparing approaches. Review of Financial Studies 22, 435–480.

Piotroski, J., 2000. Value investing: the use of historical financial statement information to separate winners from losers. Journal of Accounting Research 38, 1–41.

Poetz, M., Schreier, M., 2012. The value of crowdsourcing: can users really compete with professionals in generating new product ideas? Journal of Product Innovation Management 29, 245–256.

Pontiff, J., 1996. Costly arbitrage: evidence from closed-end funds. Quarterly Journal of Economics 111, 1135–1151.

Shapiro, C., 1982. Consumer information, product quality, and seller reputation. Bell Journal of Economics 13, 20–35.

Shapiro, C., 1983. Premiums for high quality products as returns to reputations. Quarterly Journal of Economics 98, 659–679.

Shleifer, A., Vishny, R., 1997. The limits of arbitrage. Journal of Finance 52, 35–55.

Subrahmanyam, A., Titman, S., 1999. The going-public decision and the development of financial markets. Journal of Finance 54, 1045–1082.

Tetlock, P., Saar-Tsechansky, M., Macskassy, S., 2008. More than words: quantifying language to measure firms' fundamentals. Journal of Finance 63, 1437–1467.

Tirunillai, S., Tellis, G., 2012. Does chatter really matter? Dynamics of user-generated content and stock performance. Marketing Science 31, 198–215.

Titman, S., 1984. The effect of capital structure on a firm's liquidation decision. Journal of Financial Economics 13, 137–151.

Tumarkin, R., Whitelaw, R., 2001. News or noise? Internet postings and stock prices. Financial Analysts Journal 57, 41–51.

USA Today, 2003. TurboTax's theft controls anger some. January 31.

Wall Street Journal, 2003. Intuit shares plunge after profit warning. March 21.

Wall Street Journal, 2015. Amazon sues websites selling product reviews. April 10.

Wernerfelt, B., 1988. Umbrella branding as a signal of new product quality: an example of signalling by posting a bond. RAND Journal of Economics 19, 458–466.

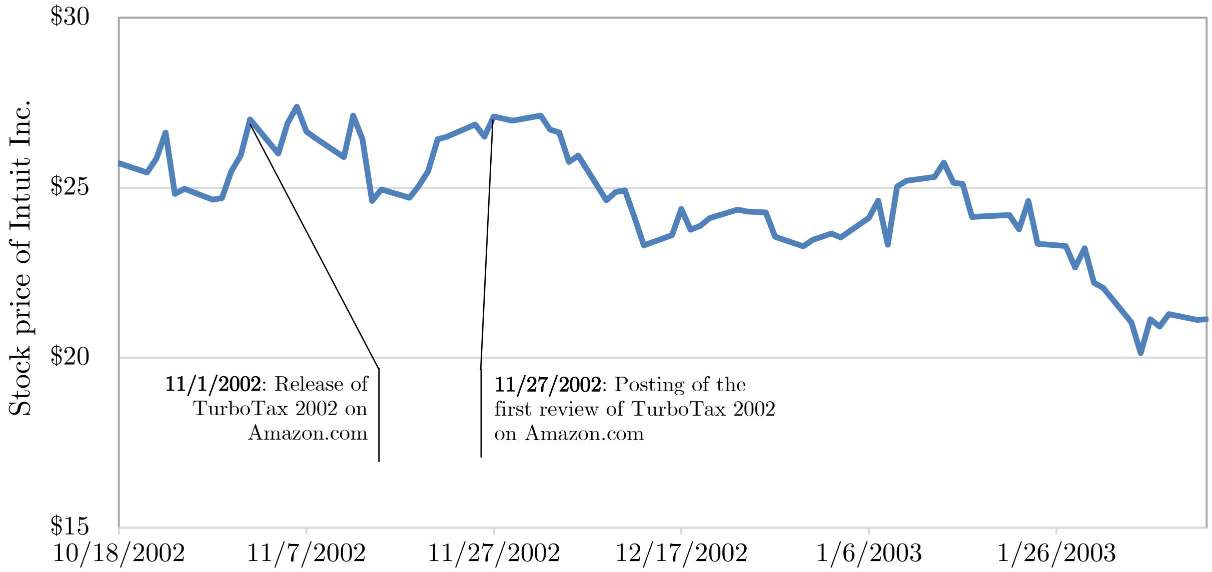

Fig. 1. Intuit Inc. This figure plots the stock price of Intuit Inc., the maker of tax-preparation software TurboTax, from October 2002 to February 2003. The dates when the firm released TurboTax 2002 on Amazon.com and when the first review for the product appeared on Amazon.com are indicated.

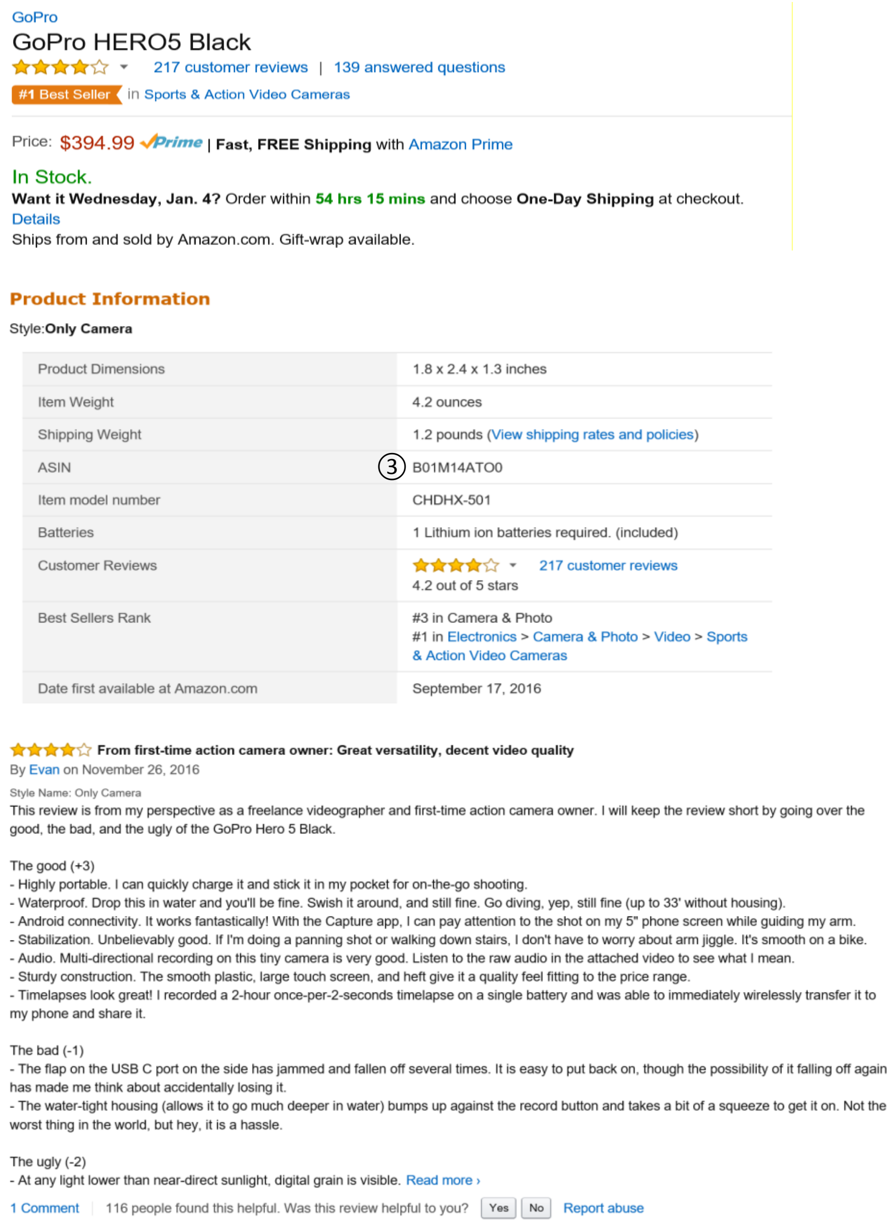

Fig. 2. Sample webpage of Amazon.com. This figure shows parts of a webpage on Amazon.com that contain product and review information (`http://www.amazon.com/GoPro-CHDHX-501-HERO5-Black/dp/B01M14ATO0`). I retrieve the following information from this webpage: the name of the brand (GoPro), the name of the product (GoPro HERO5 Black), the Amazon Standard Identification Number (B01M14ATO0), the numerical star rating (four stars), the review title ("From first-time action camera owner: Great versatility, decent video quality"), the name of the reviewer (Evan) and the Amazon.com account identification of the reviewer (AB4QNDY5QSHKP; displayed only in the HTML source code), the date of the review (November 26, 2016), and the full text of the review.

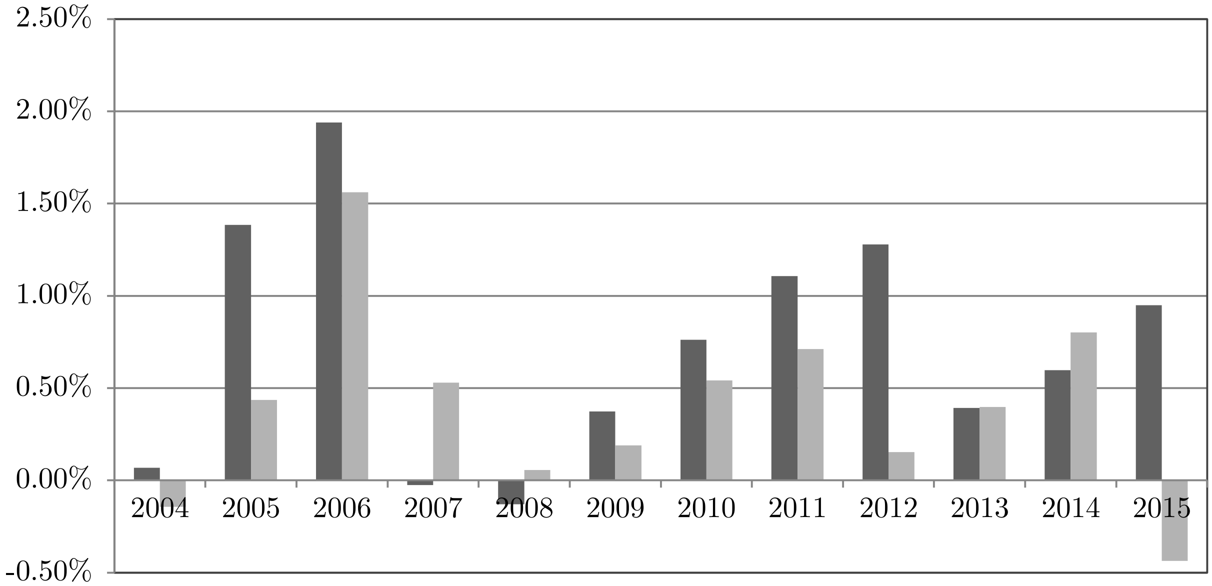

Fig. 3. Performance of the review-based strategy by year. This figure plots the differences in the average monthly abnormal returns between the top and bottom tercile portfolios ranked by abnormal customer ratings by year. For each month from July 2004 through December 2015, I sort sample stocks into tercile portfolios based on the abnormal customer ratings. I calculate the one-month-ahead abnormal returns for each stock in the tercile portfolios using the Fama-French-Carhart four-factor model. I use stock returns in the past 60 months to estimate factor loadings. I then compute the number of reviews–weighted and equally weighted abnormal returns of each tercile portfolio. The dark and grey bars represent the differences in the average monthly abnormal returns between the top and bottom tercile portfolios that are weighted by the number of reviews and equally weighted, respectively.

Table 1
Summary statistics on Amazon.com reviews for public firms

This table reports the summary statistics for the sample of Amazon.com customer reviews for products of public firms from July 2004 through December 2015. To be included in the sample, a firm has to be publicly traded on finance NYSE, AMEX, or Nasdaq, have financial and stock returns data from Compustat and CRSP, and have at least ten customer reviews in a month. I report the number of reviews, products, brands, and public firms for the full sample as well as by Fama and French 12 industries. Firms in the sample come from nine out of 12 Fama and French industries. The industries that do not have firms with Amazon.com reviews are energy, utilities, and financial industries.

	Number of reviews	Number of products	Number of brands	Number of firms
Full sample	14,555,765	269,957	1,931	346
By Fama and French 12 industries				
Consumer non-durables	2,583,822	66,008	798	80
Consumer durables	1,886,991	29,444	94	24
Manufacturing	2,036,257	48,719	170	46
Chemicals	1,248,339	29,779	384	21
Business equipment	4,652,443	47,571	176	68
Telecommunication	787,979	14,823	14	11
Shops	518,004	12,764	112	51
Healthcare	385,722	11,368	153	38
Others	456,208	9,481	30	7

Table 2
Summary statistics and determinants of abnormal customer ratings

This table reports summary statistics and regression analysis of the determinants of abnormal customer ratings. Panel A reports the summary statistics for the sample of firms with customer reviews from July 2004 through December 2015. *Average customer rating* is the average customer rating for a company's products over a given month. *Abnormal customer rating* is the average customer rating in a month minus that in the prior 12 months. *# of customer reviews* is the number of customer reviews for a company's products in a month. *Market cap* is the market capitalization of the firm, calculated as the number of shares outstanding multiplied by the stock price. *Book-to-market* is the book value of common equity divided by the market value of common equity. *Stock return*$_{m-12, m-1}$ is the buy-and-hold stock return during the past 12 months, skipping the most recent month. *Advertising* is the ratio of advertising expenses to sales. *R&D* is the ratio of R&D expenses to book assets. *Gross profitability* is the ratio of income before extraordinary items to book value of assets. *F-score* is the sum of nine binary variables capturing financial performance signals (Piotroski, 2000). *Dollar volume* is the dollar trading volume 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, beginning in 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. *# of analysts* is the number of analysts making forecasts in the I/B/E/S database for a stock in a given quarter. *Institutional ownership* is the number of shares held by institutional investors as a fraction of the number of shares outstanding. *Analyst revision* is the difference between the mean of current-year EPS forecasts made by analysts during a given month and that in the previous month scaled by the stock price at the end of the previous month. *Revenue surprise* (*SURGE*) is the difference between actual quarterly revenue per share and the expected quarterly revenue per share scaled by the standard deviation of quarterly revenue growth, as described in Eq. (1). The expected quarterly revenue is estimated under the assumption that it follows a seasonal random walk with drift. *Earnings surprise* (*SUE*) is the difference between reported quarterly EPS and the median EPS forecast of all analysts issued during the 90-day period prior to the earnings announcement date scaled by the stock price. *Net buying by HFs* (*Net buying by non–HFs*) is the change in the fraction of outstanding shares held by hedge funds (non–hedge fund institutions) between two consecutive quarters, as described in Eq. (2). All variables are winsorized at the 0.1% and 99.9% levels to minimize the effect of outliers. Panel B reports regression analysis of the determinants of one-month-ahead abnormal customer ratings. All regressions in Panel B include time fixed effects and stock fixed effects. Numbers in parentheses are *t*-statistics based on standard errors corrected for heteroskedasticity and clustering at the stock level.

Panel A: Summary statistics

Variable	N	Mean	Standard Deviation	25th percentile	Median	75th percentile
Customer reviews						
Average customer ratings	20,562	4.096	0.444	3.893	4.167	4.377
Abnormal customer ratings	20,562	0.014	0.309	-0.104	0.024	0.153
# of customer reviews	20,562	700.077	2048.420	33.000	109.000	449.000
Firm-level characteristics						
Market cap (millions of dollars)	20,562	25,723.300	56,088.990	997.479	5,533.220	21,868.000
Book-to-market	20,562	0.418	0.588	0.213	0.358	0.589
Stock $\text{return}_{m-12, m-1}$	20,562	0.170	1.034	-0.086	0.123	0.335
Advertising	20,562	0.040	0.053	0.004	0.022	0.054
R&D	20,562	0.009	0.015	0.000	0.000	0.014
Gross profitability	20,562	0.113	0.074	0.070	0.100	0.140
F-score	20,562	5.326	1.892	4.000	5.000	7.000
Dollar volume (millions of dollars)	20,562	3,701.090	11,572.960	117.757	944.920	3,228.530
CV of dollar volume	20,562	0.367	0.240	0.225	0.298	0.425
Book leverage	20,562	0.529	0.254	0.366	0.512	0.659
Asset tangibility	20,562	0.169	0.133	0.080	0.135	0.222
# of analysts	20,562	16.419	13.286	6.000	14.000	24.000
Institutional ownership	20,562	0.627	0.287	0.516	0.689	0.816
Analyst revisions (percent)	9,916	-0.108	2.157	-0.124	0.000	0.104
Cash flow surprises and institutional trades						
Revenue surprise (SURGE)	7,283	0.525	1.742	-0.215	0.292	1.095
Earnings surprise (SUE) (percent)	5,503	0.077	0.632	0.000	0.060	0.197
Net buying by HFs (percent)	7,886	0.009	2.140	-0.616	0.000	0.634
Net buying by non–HFs (percent)	7,886	0.056	4.407	-1.512	0.035	1.644

Panel B: Regression of one-month-ahead abnormal customer ratings on firm characteristics

	Dependent variable: One-month-ahead abnormal customer ratings			
	(1)	(2)	(3)	(4)
Advertising		-0.043	-0.037	0.302
		(0.40)	(0.34)	(1.36)
R&D		0.023	0.031	0.145
		(0.09)	(0.11)	(0.26)
Gross profitability		-0.112	-0.115	-0.035
		(1.34)	(1.37)	(0.42)
F-score		0.000	0.000	-0.000
		(0.10)	(0.11)	(0.05)
Log(Dollar volume)			0.005	0.009
			(1.28)	(0.92)
Log(CV of dollar volume)			0.005	0.018
			(0.66)	(1.84)*
Log(Market cap)	0.005	0.005	0.002	0.002
	(0.77)	(0.84)	(0.23)	(0.19)
Book-to-market	0.002	0.001	0.001	0.003
	(0.26)	(0.14)	(0.17)	(0.14)
Stock return$_{m-12, m-1}$	0.000	0.000	0.000	-0.006
	(0.05)	(0.22)	(0.03)	(0.64)
Book leverage	-0.022	-0.016	-0.017	0.008
	(0.85)	(0.57)	(0.62)	(0.18)
Asset tangibility	-0.020	-0.013	-0.007	-0.096
	(0.22)	(0.14)	(0.07)	(0.69)
Log(1+# of analysts)	-0.011	-0.011	-0.013	-0.021
	(1.45)	(1.45)	(1.51)	(1.43)
Institutional ownership	0.016	0.014	0.007	0.025
	(1.04)	(0.93)	(0.41)	(0.91)
Net buying by HFs				0.000
				(0.16)
Analyst revision				-0.177
				(1.02)
Number of observations	19,603	19,554	19,554	9,900
Adjusted R-squared	0.07	0.07	0.07	0.07

Table 3
Calendar-time portfolio returns

This table reports calendar-time portfolio regression results. For each month from July 2004 through December 2015, I sort sample stocks into tercile portfolios based on the abnormal customer rating, defined as the average customer rating in the month minus that in the prior 12 months. I then track the performance of the three portfolios over the following month. I employ two weighting schemes, weighting by the number of reviews and equal weighting across stocks. I use the Fama-French-Carhart four-factor model to adjust returns. The alpha estimates and factor loadings reported below are obtained by regressing monthly portfolio excess returns on the monthly returns from the risk factors. "Long/short" is a spread portfolio that buys the top tercile portfolio and sells the bottom tercile portfolio. Numbers in parentheses are *t*-statistics. Significance at the 10% (*), 5% (**), or 1% level (***) is indicated.

	α	Market β	SMB β	HML β	UMD β
Panel A: Review weighting					
T1 (low abnormal rating)	-0.198%	1.124	0.482	0.132	-0.308
	(0.74)	(15.69)***	(3.83)***	(1.11)	(4.91)***
T2	-0.014%	1.245	0.246	0.022	-0.188
	(0.06)	(19.90)***	(2.24)**	(0.21)	(3.43)***
T3 (high abnormal rating)	0.532%	1.051	0.179	0.165	-0.134
	(1.98)**	(14.50)***	(1.41)	(1.37)	(2.11)**
Long/Short (high − low)	0.730%	-0.073	-0.303	-0.033	0.174
	(2.17)**	(0.81)	(1.90)*	(0.22)	(2.19)**
Panel B: Equal weighting					
T1 (low abnormal rating)	-0.024%	1.006	0.513	0.003	-0.239
	(0.15)	(22.19)***	(6.44)***	(0.04)	(6.01)***
T2	0.264%	1.081	0.385	0.152	-0.090
	(1.52)	(23.09)***	(4.68)***	(1.96)*	(2.19)**
T3 (high abnormal rating)	0.533%	0.956	0.636	-0.115	-0.361
	(2.52)**	(16.74)***	(6.34)***	(1.21)	(7.20)***
Long/short (high − low)	0.557%	-0.050	0.123	-0.118	-0.121
	(2.66)***	(0.88)	(1.24)	(1.26)	(2.45)**

Table 4
Calendar-time portfolio returns: subsample analyses

This table reports the Fama-French-Carhart four-factor alphas on a monthly calendar-time spread portfolio that buys the top tercile stocks and sells the bottom tercile stocks in abnormal customer ratings for various subsamples. Panel A partitions the sample of stocks into a high idiosyncratic volatility group and a low idiosyncratic volatility group based on the median idiosyncratic volatility. Panels B and C partition the sample of stocks into two groups based on the median number of analysts covering the firm and the median market capitalization, respectively. Within each group of stocks in each month from July 2004 through December 2015, I form a spread portfolio that buys the top tercile stocks and sells the bottom tercile stocks sorted by abnormal customer ratings. I track the performance of the spread portfolio over the following month. The alpha estimates reported below are obtained by regressing monthly returns of the spread portfolio on the monthly returns of the Fama-French-Carhart factors. Numbers in parentheses are *t*-statistics. Significance at the 10% (*), 5% (**), or 1% level (***) is indicated.

	Review weighting	Equal weighting
Panel A: By idiosyncratic volatility		
High idiosyncratic volatility	1.366%	1.030%
	(2.12)**	(2.50)**
Low idiosyncratic volatility	-0.015%	0.131%
	(0.05)	(0.63)
Long/short (high − low)	1.380%	0.899%
	(1.81)*	(1.86)*
Panel B: By analyst coverage		
Low analyst coverage	1.480%	0.734%
	(2.38)**	(2.11)**
High analyst coverage	0.190%	0.393%
	(0.53)	(1.57)
Long/short (low − high)	1.291%	0.341%
	(1.73)*	(0.78)
Panel C: By market capitalization		
Small firms	1.162%	1.181%
	(2.42)**	(2.75)***
Large firms	0.597%	0.412%
	(2.11)**	(1.72)*
Long/short (small − large)	0.565%	0.768%
	(0.97)	(1.52)

Table 5
Robustness checks

This table reports robustness checks of calendar-time portfolio tests. Panel A uses alternative risk benchmarks to adjust returns: the Fama and French (1993) three-factor model, the Fama and French (2015) five-factor model, a liquidity-augmented Fama-French-Carhart model, and industry-adjusted stock returns (following Moskowitz and Grinblatt, 1999). Panel B constructs an alternative measure for abnormal ratings using a six-month window to define the benchmark period. For each specification, I report the alphas on a monthly calendar-time spread portfolio that buys the top tercile stocks and sells the bottom tercile stocks in abnormal customer ratings. Numbers in parentheses are *t*-statistics. Significance at the 10% (*), 5% (**), or 1% level (***) is indicated.

	Review weighting	Equal weighting
Panel A: Alternative risk adjustments		
Using the Fama and French three-factor model	0.787% (2.31)**	0.517% (2.43)**
Using the Fama and French five-factor model	0.789% (2.24)**	0.513% (2.32)**
Using a liquidity-augmented Fama-French-Carhart model	0.771% (2.26)**	0.602% (2.85)***
Using industry-adjusted abnormal returns	0.678% (1.98)*	0.476% (2.32)**
Panel B: Alternative measures for abnormal ratings		
Prior six months as benchmark	0.592% (2.01)**	0.573% (2.03)**

Table 6
Calendar-time portfolio returns over different holding periods

This table reports the performance of calendar-time portfolios over different holding periods. For each month from July 2004 through December 2015, I sort sample stocks into tercile portfolios based on the abnormal customer rating. I then track the performance of the three portfolios over four holding periods: three months (months 2 through 4), six months (months 2 through 7), nine months (months 2 through 10), and 12 months (months 2 through 13) after portfolio formation, all skipping the first month post-formation. The portfolios are rebalanced monthly by adding stocks that enter the top or the bottom tercile and dropping stocks that have reached their holding periods. I employ two weighting schemes, weighting by the number of reviews and equal weighting across stocks. The alpha estimates reported below are obtained by regressing monthly returns of the spread portfolio that buys the top tercile portfolio in abnormal customer ratings and sells the bottom tercile portfolio on the monthly returns of the Fama-French-Carhart factors. Numbers in parentheses are t-statistics. Significance at the 10% (*), 5% (**), or 1% level (***) is indicated.

Holding period	Review weighting	Equal weighting
Months [2, 4]	0.122%	0.143%
	(0.56)	(0.97)
Months [2, 7]	0.132%	0.084%
	(0.74)	(0.75)
Months [2, 10]	0.093%	0.095%
	(0.61)	(0.91)
Months [2, 13]	0.009%	0.110%
	(0.07)	(1.18)

Table 7
Fama-MacBeth regressions

This table reports the coefficient estimates obtained from Fama-MacBeth regressions of one-month-ahead excess stock returns on abnormal customer ratings and other cross-sectional predictors of stock returns described in Eq. (3). For each specification in each month, I run a cross-sectional regression with the monthly excess return (in percent) as the dependent variable. I report the time series averages of the cross-sectional regression coefficients. All variables are defined in Table 2. Numbers in parentheses are Fama-MacBeth t-statistics calculated using Newey-West standard errors with four lags. Significance at the 10% (*), 5% (**), or 1% level (***) is indicated.

	Dependent variable: One-month-ahead excess stock returns (percent)					
	(1)	(2)	(3)	(4)	(5)	(6)
Abnormal customer ratings	1.718	1.681	1.525	1.426	1.580	1.323
	(2.44)**	(2.34)**	(2.27)**	(2.23)**	(2.18)**	(2.08)**
Gross profitability			4.125	3.360		3.295
			(2.22)**	(1.65)		(1.57)
F-score			0.396	0.412		0.391
			(5.02)***	(4.65)***		(4.26)***
Advertising				1.122		0.456
				(0.31)		(0.13)
R&D				-3.672		-3.882
				(0.37)		(0.43)
Log(Dollar volume)					-0.204	-0.175
					(2.06)**	(1.70)*
Log(CV of dollar volume)					-0.868	-0.747
					(2.76)***	(2.40)**
Log(Market cap)		-0.056	-0.087	-0.115	0.054	-0.017
		(0.49)	(0.78)	(1.01)	(0.44)	(0.13)
Book-to-market		0.342	0.829	0.644	0.309	0.561
		(0.69)	(1.53)	(1.08)	(0.61)	(0.88)
Stock return$_{m-12, m-1}$		0.035	-0.331	-0.297	0.303	-0.061
		(0.05)	(0.47)	(0.40)	(0.42)	(0.08)
Number of observations	20,562	20,562	20,562	20,562	20,562	20,562
Average R-squared	0.01	0.07	0.10	0.13	0.11	0.17

Table 8
Abnormal customer ratings and revenue surprises

This table reports ordinary least squares regressions of revenue surprises on abnormal customer ratings and controls described in Eq. (4). The dependent variable is the standardized unexpected revenue growth estimator (SURGE). *Stock return*$_{t-30, t-3}$ and *Stock return*$_{t-365, t-31}$ are the cumulative return during the period from 30 to three days prior to an earnings announcement and that during the period from 365 to 31 days prior to an earnings announcement, respectively. All other variables are defined in Table 2. Numbers in parentheses are *t*-statistics based on standard errors clustered by firm and by quarter. Significance at the 10% (*), 5% (**), or 1% level (***) is indicated.

	Dependent variable: Revenue surprise (SURGE)				
	(1)	(2)	(3)	(4)	(5)
Abnormal customer ratings	0.443	0.440	0.340	0.359	0.360
	(2.16)**	(2.26)**	(2.17)**	(2.16)**	(2.25)**
Lagged dependent variable			1.248		
			(26.63)***		
Log(Market cap)		-0.020	-0.026		0.174
		(0.50)	(1.14)		(1.45)
Book-to-market		-0.100	-0.027		-0.072
		(1.23)	(0.50)		(0.75)
Advertising		-1.873	0.115		-0.133
		(2.59)**	(0.29)		(0.07)
R&D		0.479	1.273		-3.817
		(0.17)	(0.75)		(1.47)
Gross profitability		0.067	-0.012		0.016
		(3.70)***	(0.80)		(0.98)
F-score		0.875	-0.569		2.507
		(1.36)	(1.47)		(3.11)***
Log(Dollar volume)		0.084	0.029		0.061
		(3.00)***	(1.66)*		(1.00)
Log(CV of dollar volume)		-0.125	-0.076		-0.133
		(1.39)	(1.22)		(1.51)
Stock return$_{t-30, t-3}$		0.528	0.121		0.404
		(5.27)***	(1.92)*		(4.12)***
Stock return$_{t-365, t-31}$		0.924	0.677		0.942
		(4.02)***	(3.89)***		(5.29)***
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	No	No	No	Yes	Yes
Number of observations	7,283	7,283	7,283	7,283	7,283
Adjusted R-squared	0.07	0.12	0.41	0.29	0.31

Table 9
Abnormal customer ratings and earnings surprises

This table reports ordinary least squares regressions of earnings surprises on abnormal customer ratings and controls described in Eq. (4). The dependent variable is the standardized unexpected earnings (SUE). *Stock return*$_{t-30, t-3}$ and *Stock return*$_{t-365, t-31}$ are the cumulative return during the period from 30 to 3 days prior to an earnings announcement and that during the period from 365 to 31 days prior to an earnings announcement, respectively. All other variables are defined in Table 2. Numbers in parentheses are *t*-statistics based on standard errors clustered by firm and by quarter. Significance at the 10% (*), 5% (**), or 1% level (***) is indicated.

	Dependent variable: Earnings surprise (SUE)				
	(1)	(2)	(3)	(4)	(5)
Abnormal customer ratings	0.077	0.080	0.076	0.074	0.073
	(2.33)**	(2.40)**	(2.26)**	(2.07)**	(1.92)*
Lagged dependent variable			0.096		
			(2.64)***		
Forecast dispersion		-3.087	-3.387		-2.963
		(0.90)	(1.15)		(1.26)
Log(Market cap)		-0.033	-0.031		-0.078
		(1.88)*	(1.77)*		(1.47)
Book-to-market		0.053	0.039		-0.178
		(0.56)	(0.44)		(1.23)
Advertising		-0.163	-0.112		1.370
		(0.54)	(0.40)		(2.28)**
R&D		2.558	2.284		0.341
		(2.83)***	(2.81)***		(0.16)
Gross profitability		0.009	0.006		0.002
		(1.35)	(0.90)		(0.34)
F-score		0.057	-0.036		0.336
		(0.25)	(0.16)		(0.80)
Log(Dollar volume)		0.044	0.041		0.027
		(2.30)**	(2.11)**		(0.88)
Log(CV of dollar volume)		0.013	0.011		0.034
		(0.48)	(0.45)		(1.08)
Stock return$_{t-30, t-3}$		0.193	0.161		0.150
		(3.99)***	(3.61)***		(2.57)**
Stock return$_{t-365, t-31}$		0.660	0.647		0.632
		(3.34)***	(3.32)***		(3.33)***
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	No	No	No	Yes	Yes
Number of observations	5,503	5,503	5,503	5,503	5,503
Adjusted R-squared	0.02	0.04	0.06	0.11	0.13

Table 10
Abnormal customer ratings and institutional trading

This table reports Fama-MacBeth regressions of institutional trading on lagged abnormal customer ratings and controls described in Eq. (5). The dependent variable in the first and second columns is net buying by hedge funds and that by non-hedge funds in quarter q, respectively. The dependent variable in the third and fourth columns is net buying by specialized hedge funds and that by less specialized hedge funds in quarter q, respectively. Specialized hedge funds are defined as those with a trading weight in stocks with Amazon.com reviews greater than the median, and less specialized hedge funds are those with a below median trading weight. *Stock return*$_q$, *Stock return*$_{q-1}$, and *Stock return*$_{[q-5,\,q-2]}$ are the buy-and-hold stock returns in quarter q, quarter $q-1$, and quarters $q-5$ through $q-2$, respectively. All other variables are defined in Table 2. For each specification in each quarter, I run a separate cross-sectional regression. I report the time series averages of the cross-sectional regression coefficients. Numbers in parentheses are Fama-MacBeth t-statistics calculated using Newey-West standard errors with four lags. Significance at the 10% (*), 5% (**), or 1% level (***) is indicated.

	Net buying by HFs (1)	Net buying by non–HFs (2)	Net buying by specialized HFs (3)	Net buying by less specialized HFs (4)
Abnormal customer ratings	0.227	-0.047	0.183	0.058
	(2.39)**	(0.29)	(2.74)***	(1.07)
Log(Market cap)	-0.017	-0.093	0.036	-0.040
	(0.41)	(1.09)	(0.88)	(3.09)***
Book-to-market	0.151	-0.412	0.114	-0.008
	(1.66)	(2.46)**	(2.24)**	(0.13)
Advertising	-0.338	-1.850	0.035	-0.332
	(0.55)	(1.47)	(0.08)	(0.85)
R&D	-1.469	-13.454	1.712	-1.296
	(0.58)	(2.16)**	(0.96)	(0.91)
Gross profitability	0.086	-0.045	-0.241	0.042
	(0.15)	(0.04)	(0.51)	(0.15)
F-score	-0.005	0.021	-0.012	0.000
	(0.30)	(0.51)	(1.25)	(0.04)
Log(Dollar volume)	0.007	0.058	-0.040	0.032
	(0.19)	(0.88)	(1.06)	(2.83)***
Log(CV of dollar volume)	-0.009	-0.035	0.031	-0.053
	(0.10)	(0.23)	(0.58)	(1.42)
Stock return$_q$	0.141	2.687	-0.135	0.254
	(0.52)	(4.65)***	(0.49)	(1.26)
Stock return$_{q-1}$	0.029	-0.023	-0.183	0.011
	(0.10)	(0.03)	(0.64)	(0.05)
Stock return$_{[q-5,q-2]}$	0.010	0.108	0.049	-0.042
	(0.09)	(0.49)	(0.58)	(0.81)
Number of observations	7,886	7,886	7,886	7,886
Average R-squared	0.13	0.12	0.12	0.12