Review of Quantitative Finance and Accounting (2023) 60:565–606
https://doi.org/10.1007/s11156-022-01104-5

ORIGINAL RESEARCH

The effect of customer concentration on stock sentiment risk

Jian Wang[1,2] · **Yanhuang Huang**[1] · **Hongrui Feng**[3] · **Jun Yang**[4]

Accepted: 31 August 2022 / Published online: 16 October 2022

Abstract
The impact of stock sentiment risk on their returns has been well documented in literature, but exploration into the determinants of stock sentiment risk is lacking. We theorize that concentrated customer bases help mitigate stock sentiment risk. Empirical results based on a large sample from the U.S. market strongly support this hypothesis. Specifically, the mitigating effect takes place through three channels. Companies with high customer concentration tend to have better performance and information quality and attract more long-horizon institutional investors. All these factors contribute to diminishing stock sentiment risk. The results are robust when the endogeneity concern is addressed by investigating the effect of an exogenous shock, or when they are examined with alternative measures of sentiment risk. The negative relationship between customer concentration and stock sentiment risk is ubiquitous but even stronger during the 2008 financial crisis.

Keywords Customer concentration · Sentiment risk · Firm performance · Information quality · Institutional investors

JEL Classification G32 · G30 · G40

1 Introduction

The effect of stock sentiment risk on returns has been well documented and investment strategies have been proposed to capitalize it. Baker and Wurgler (2006, 2007) are among the first to argue that sentiment beta—a stock price's sensitivity to market-wide sentiment change—can provide additional explanatory power on stock price movement. Their findings reveal that low-sentiment-beta stocks significantly outperform high-sentiment-beta stocks when other risk factors are controlled. Massa and

✉ Hongrui Feng
hxf51@psu.edu

[1] School of Business Administration, Northeastern University, Shenyang, China

[2] Institute of Behavioral and Service Operations Management, Northeastern University, Shenyang, China

[3] Black School of Business, Penn State Behrend, Erie, PA, USA

[4] F. C. Manning School of Business Administration, Acadia University, Wolfville, NS, Canada

Yadav (2015) reveal that mutual funds employing the contrarian strategy, which bets against investor sentiment to take advantage of investors' fanaticism and antipathy, consistently yield higher returns than funds employing the catering strategy, which loads on high sentiment-beta stocks to please investors. Zheng et al. (2018) have similar findings in their study on hedge funds.

Despite the compelling evidence for the importance of stock sentiment risk, explorations into its contributing factors have been lacking. Stock sentiment risk is a reflection of not only investor behaviors but also, more deeply, firm characteristics and fundamentals.[1] It has been observed that high sentiment-risk stocks are usually younger, smaller, and more volatile (Baker and Wurgler 2006, 2007; Berger and Turtle 2012). In this study, we argue that, new to literature, customer concentration is an important determinant of stock sentiment risk. Whether a company relies on sales to a few major customers or diffuse customers has far reaching effects on various aspects of corporate strategy, risk taking, and resultant performance. While many extant studies examine the impact of customer concentration on firm profitability measures such as ROA (Balakrishnan et al. 1996; Patatoukas 2012; Irvine et al. 2016; Hui et al. 2019; Cohen and Li 2020), other specific impacts explored include earnings management (Raman and Shahrur 2008), incentive conflicts in inter-firm relationships (Costello 2013), cost of capital (Dhaliwal et al. 2016), and stock liquidity risk (Boscaljon et al. 2021). We reason that the impact of customer concentration on corporate strategy, as revealed in literature, would be reflected on how their stock prices fluctuate with market sentiment. Our main hypothesis is that all else equal, firms with higher customer concentration experience lower stock sentiment risk. Working on a large sample of U.S. firms between 1997 and 2017 with over 300,000 observations, we find that a firm's customer concentration is negatively associated with its stock sentiment risk after controlling various relevant variables. That is, the stock price of a firm with a higher level of customer concentration tends to be less susceptible to market sentiment.

We theorize that customer concentration can mitigate stock sentiment risk for several reasons, which constitute the channels for such an effect. Three hypotheses are proposed and tested for firms' operating and accounting performance, information quality, and long-horizon institutional ownership. All the hypotheses are supported by the empirical results.

We conjecture that firms with high customer concentration tend to deliver better operating and accounting performance (measured by cash flow volatility, sales growth, earnings before interest, tax, depreciation and amortization, or EBITDA, and gross profit) and better information quality (measured by audit-related fees and incidence of restatement of financial statements) (Kalwani and Narayandas 1995; Kumar 1996; Kinney and Wempe 2002). Krishnan et al. (2019) find reduced audit costs caused by customer-base concentration, which suggests information efficiency gains in the audit process. Improved operating and accounting performance and information quality in turn mitigate stock sentiment risk. Such conjectures are

[1] The idea that stocks' systematic risk can be explained by firm fundamentals is prevalent. For example, Campbell et al. (2010) argue that the betas of growth and value stocks are determined by the cash-flow fundamentals of these companies. Ng (2011) links liquidity risk to companies' information quality, and Cao and Petrasek (2014) demonstrate that certain types of institutional ownership lower liquidity risk of stocks.

consistent with Baker and Wurgler (2006) who find that firms with high sentiment risk tend to be new, small, opaque, and uncertain in prospects.

The third channel through which customer concentration mitigates sentiment risk is the attraction of institutional investors, particularly those with long investment horizon. Long-horizon institutional investors are less concerned with short term market sentiment trends and focus more on long-run performance. In general, institutional shareholders can improve the monitoring and information quality of the invested firms, leading to lower sentiment risk. However, the investment horizons of institutional shareholders vary and such variations matter for their impacts (Bushee 2004; Gaspar et al. 2005; Boscaljon et al. 2021). Institutional investors with a long horizon have more meaningful impacts than transient ones. For example, Boscaljon et al. (2021) find that firms with high government customer concentration attract long–horizon institutional investors, who in turn contribute to lower stock liquidity risk.

Our results are robust to alternative measures of stock sentiment risk and inclusion of firm, industry and time fixed effects. While it is challenging to establish a causal relationship based on observational data, an effective method to overcome this challenge is to examine the effect of a quasi-experiment or an exogeneous shock. Therefore, in order to check the causal effect of customer concentration on stock sentiment beta, we examine the impact of tariff jumps, an exogenous shock at industry level. Significant tariff upswings elevate customer concentration for domestic suppliers, and the results show that they lead to lower sentiment risk. Upon passing the parallel trend test, the complete difference-in-difference (DiD) analysis offers similar results.[2] The evidence further confirms that customer concentration plays an important role in containing stock sentiment risk by controlling for the endogeneity issue.

Since sentiment risk is an important systematic risk that investors face especially during market calamity, it is worthwhile to zoom in on the containing effect of customer concentration on sentiment risk during a financial crisis. Our empirical results confirm that its mitigating effect is indeed stronger during the 2008–2009 financial crisis, although such an effect is ubiquitous in the whole sample. Our main results hold consistently with monthly or yearly data. Further, we find that the containing effect of customer concentration on sentiment beta is weaker for companies with foreign sales and younger companies.[3]

This study makes a threefold contribution to improve our understanding of stock sentiment risk. First, despite the well-documented evidence on the systematic impact of sentiment risk on stock and portfolio returns, explorations into its contributing factors have been very limited. It has been noted that stocks with high sentiment risk are typically from companies that are young, small, and more volatile (Baker and Wurgler 2006, 2007; Berger and Turtle 2012). A company's age does not reliably make it more resilient to market sentiment swings, and one may wonder whether high volatility is the cause or consequence of high sentiment sensitivity. This study identifies customer concentration as a contributing factor to stock sentiment risk for the first time in literature, and the empirical analysis provides clear evidence that customer-base concentration can alleviate stock sentiment risk. Second, we identify three channels through which customer concentration impacts stock sentiment risk. We theorize, which is supported by empirical results, that companies with higher

[2] We appreciate a reviewer making these suggestions.

[3] We appreciate the valuable comments by the reviewers that suggest explorations into the impact of data frequency, foreign sales, and firm age.

customer concentration tend to have better performance and information quality, and attract more long-horizon institutional investors. These positive effects help curtail stock sentiment risk. Third, this study provides a valuable bridge between two bodies of literature: those on the effects of customer-base concentration and those on stock sentiment risk. To the best of our knowledge, this study represents the first attempt to probe this connection. Given the findings that sentiment risk is a novel grip for formulating investment strategies, this study should interest investment professionals. In the meantime, corporate managers would also benefit from the new perspective for managing stock sentiment risk.

The rest of this paper is structured as follows. Section 2 surveys the literature and develops our research hypotheses. Section 3 describes the data and methodology. Section 4 provides our main empirical results. Section 5 collects the concluding remarks.

2 Literature review and research hypotheses

2.1 Literature review

This study is uniquely positioned at the junction of two groups of literature. The first group is on stock sentiment beta. Baker and Wurgler (2006, 2007) suggest stock prices are affected by market-wide sentiment that creates cross-sectional variation in stock returns. Using proxies to measure the degree of investment sentiment, they find that returns are relatively low for stocks with high sensitivity to sentiment, which typically possess the following characteristics: young, fast growing, volatile, un-profitable, and non-dividend-paying. Berger and Turtle (2012) find confirmatory evidence that investor sentiment sensitivity increases with multiple measure of opacity: sentiment-prone stocks tend to be small, young, volatile, and composed of relatively intangible assets. Massa and Yadav (2015) find further evidence that low-sentiment stocks outperform high-sentiment stocks by examining portfolios of mutual funds. Some more recent studies demonstrate that similar patterns exist in individual stocks as well (Antoniou et al. 2016; Yang and Hu 2021). However, explorations on the contributing factor to stock sentiment beta has been lacking in literature. This study helps to fill this gap by identifying a specific factor, customer-base concentration, and its effecting mechanisms.

The second group of related literature is on the effects of customer concentration. Prior studies suggest that major customer concentration could contribute to operational, business or financial risk (e.g., Ma et al. 2020). Dhaliwal et al. (2016) find evidence that client business risk increases with higher customer concentration because suppliers are at risk of losing a key customer and have higher cash flow risk reflected in higher costs of capital. Studies related to auditing provide evidence that dependent suppliers may try to influence major customers' perceptions through opportunistic earnings management and risky tax planning strategies (Raman and Shahrur 2008; Huang et al. 2016). On the other hand, major customer concentration may allow suppliers to have more focused business activities because of more transparent information sharing. Kalwani and Narayandas (1995), Kumar (1996) and Kinney and Wempe (2002) find evidence of lower operating complexity through information sharing along the supply chain. Such improved efficiency and information transparency are attractive to institutional investors and make stock prices less affected by dynamic market sentiment. Krishnan et al. (2019) provides evidence on improved audit efficiency

and quality due to concentrated customer bases. Suppliers with more concentrated customer bases are less likely to experience material restatements of previously audited financial statements.

There has been mixed evidence regarding the impact of customer concentration on firm profitability. Patatoukas (2012) challenges the conventional wisdom that customer concentration impairs firm profitability and reports a positive relation between the two. Yet Hui et al. (2019) provides contradictory evidence. Irvine et al. (2016) introduces a dynamic relationship life-cycle hypothesis that the relation between custom-base concentration and profitability is negative in the early years of the relationship, but becomes positive as the relationship matures. They provide strong confirmation of Patatoukas' (2012) conclusion on the value relevance of major customer relationship.

Another related body of research is regarding the role of institutional investors on stock market risks. Despite evidence on the importance of institutional trading or ownership (Brown and Cliff 2004; Pirinsky and Wang 2006), there has been no consensus regarding the impact of institutional investors in capital markets. On one hand, some studies argue that institutional investors are more rational and thus play a positive role in stabilizing stock prices, strengthening corporate supervision and improving market efficiency (An and Zhang 2013; Crane et al. 2019). On the one hand, other research has shown that institutional investors often prefer high-risk assets and adopt momentum-driven trading strategies (Falkenstein 1996; Qureshi et al. 2017) and demonstrate strong herding behaviors (Choi and Skiba 2015), therefore aggravating market volatility. To reconcile the conflicting findings, it is important to recognize that institutional investors are not a homogeneous group. Legal differences aside, a key feature that captures institutional investors' trading behaviors and potential impact is their investment horizon. Institutional investors can be classified from "transient" investors to "dedicated" investors, based on their investment horizon, with many others in between (Bushee 2004). Such classifications can provide unique insights into the impact of institutional investors in various contexts, from corporate control (Gaspar et al. 2005), stock liquidity risk (Cao and Petrasek 2014), to the effect of financial disclosure quality on stock volatility (Bushee and Noe 2000) or holding costs of trading strategies (Bushee et al. 2019). Generally, long-horizon institutional investors contribute to lower return volatility and liquidity risk, when other factors are controlled. Therefore, we conjecture that long-horizon institutional investors also contribute to mitigating stock sentiment risk.

2.2 Hypotheses

Based on the literature and discussions above, we conjecture that, on balance, customer concentration has positive effects on suppliers by reducing demand uncertainty and facilitating relationship-specific investments. These effects make these companies' stock less susceptible to market sentiment swings. Therefore, we have this main hypothesis:

Hypothesis 1 All else being equal, firms with higher customer concentration have lower sentiment risk.

Furthermore, we theorize that customer concentration mitigates stock sentiment risk via multiple channels (mechanisms). Companies with concentrated customer base tend to have better performance and more comprehensible information disclosure. Such companies are

more likely to attract long-term institutional investors. All these effects help uphold investors' confidence so that their stock price is less susceptive to market sentiment. We formally state these hypotheses as follows:

Hypothesis 2a All other things being equal, firms with higher customer concentration have better operating and accounting performance, which in turn contributes to lower sentiment risk.

Hypothesis 2b All other things being equal, firms with higher customer concentration have better information quality, which in turn contributes to lower sentiment risk.

Hypothesis 2c All other things being equal, firms with higher customer concentration attract more long-horizon institutional investors, which in turn contributes to lower sentiment risk.

3 Data and methodology

3.1 Data

Our final sample includes 341,069 firm-month observations with fiscal years ending from 1997 to 2017. We begin with all common stocks publicly listed on NYSE, NASDAQ, or AMEX. We exclude stocks with a share price greater than $1000 or lower than $5 at the end of each month to mitigate the impact of thin trading and illiquidity (e.g., Pastor and Stambaugh 2003).[4] We obtain audit fee and restatement data from Audit Analytics. Such data is only available between 2000 and 2017. Excluded are observations from the financial services and utilities sectors (SIC codes 4900–4999 and 6000–6999) and those in American Depositary Share (ADR) programs. Financial service and utility companies are subject to hefty regulations, and some utilities firms are owned or controlled by states (Krishnan et al. 2019). The wholesale and retail markets in utility industries are subject to a unique set of rules and have experienced drastic changes in the past decades (Delmas and Tokat 2005; Kelly and Moody 2005), which makes measures of their customer concentration incommensurable with other industries'. Due to the distinct nature of their business, the separation of financial and operating activities in financial service firms is artificial. As a result, financial ratios for financial companies are incomparable to those of nonfinancial companies. Observations with missing information on required variables are also excluded. Stock market data are extracted from CRSP, and accounting data are obtained from COMPUSTAT. The level of customer concentration is calculated annually. Sentiment betas are computed monthly. There are approximately 1700 observations per monthly cross section. We winsorize non-categorical variables at the top and bottom one percentile of each cross-section to mitigate the impact of extreme observations. The breakdown of final sample firms is shown in Table 1. On average, 62.35% of the sample

[4] To formally test the potential impact of thin trading, we follow Miller et al. (1994) to adjust stock returns for thin trading. Then we rerun the main regressions with the adjusted sentiment beta, whose results are consistent with those in the paper and are thus unreported.

Table 1 Breakdown of sample firms

Year	NYSE (%)	NASDAQ (%)	AMEX (%)
Panel A: breakdown by stock exchanges			
1997	61.37	33.59	5.04
1998	61.05	33.49	5.46
1999	60.95	34.76	4.29
2000	61.34	33.23	5.43
2001	63.72	31.65	4.63
2002	62.84	32.78	4.38
2003	61.44	33.31	5.25
2004	63.15	32.04	4.81
2005	62.38	31.99	5.63
2006	61.46	32.09	6.45
2007	60.24	33.17	6.59
2008	60.93	33.62	5.45
2009	63.04	31.81	5.15
2010	63.65	32.97	3.38
2011	62.81	31.73	5.46
2012	63.49	33.02	3.49
2013	62.96	31.48	5.56
2014	63.44	30.96	5.60
2015	61.29	32.03	6.68
2016	63.14	31.90	4.96
2017	63.27	31.25	5.48
Average	62.35	32.52	5.20
Industry (based on SIC codes)	1997–2003	2004–2010	2011–2017
Panel B: breakdown by industries			
Consumer non-durables	113	120	117
Consumer durables	870	869	884
Manufacturing	68	75	81
Energy (oil, gas, and coal extraction and products)	71	64	72
Chemicals and allied products	43	56	41
Business equipment	65	73	62
Telephone and television transmission	37	39	34
Wholesale, retail, and some services	328	326	341
Healthcare, medical equipment, and drugs	67	63	58
Other	19	19	16

Panel A reports the proportions of sample stocks listed on NYSE, NASDAQ, and AMEX, in each sample year

Panel B reports the average number of sample firms (rounded to nearest integer) in each industry during three subsample periods, 1997–2003, 2004–2010, and 2011–2017, respectively. We classify firms into 10 industries (excluding financial services and utilities) following the industry classification system developed by Fama and French (1997). We identify the industry based on a firm's primary standard industrial classification (SIC) code as reported in COMPUSTAT, and make refinements of two-digit classifications when there are sufficient numbers of observations to form more narrowly defined groups

firms are listed on NYSE, 32.52% on NASDAQ, and 5.20% on AMEX. Sample firms are from ten major industries excluding financial services and utilities.

Since our sample period is after the significant expansion of the COMPUSTAT database in 1978, our sample includes firms that were later delisted and is thus largely free from delisting and survivorship bias. There are 365 delistings in our final sample, and 78.84% of sample firms survive the whole study period. A delisted firm remains in the sample until the month when it is removed from the data file. We follow Gaunt (2004) to address missing stock prices in their delisting months.[5] To attend to backfilled data bias, we do not include firms until they have been in COMPUSTAT for 2 years because COMPUSTAT rarely includes more than 2 years of historical data when it adds firms (Banz and Breen 1986; Fama and French 1993).

3.2 Variables

Customer concentration (*concus*): as the primary explanatory variable in this study, the measure on a firm's customer-base concentration follows Patatoukas (2012), represented by the notation *concus*. It is an application of the Herfindahl–Hirschman index, which captures both the number of major customers and the importance of each major customer in the sales revenue. Specifically, the customer concentration of firm i in year t is defined as

$$concus_{i,t} = \sum_{j=1}^{n} \left(\frac{Sales\ to\ Customer_{i,j,t}}{Total\ sales_{i,t}} \right)^2$$

Statement of Financial Accounting Standards (SFAS) No. 131 and Regulation S-K of the SEC require all public companies to disclose the identities of their major customers representing more than 10% of their total sales. By definition, *concus* is bounded between 0 and 1: it is equal to 1 if the firm earns all of its revenue from a single customer; as the customer base diversifies until no major customer has more than 10% of sales, it equals 0. This measure of customer concentration has been adopted in numerous publications including Irvine et al. (2016), Dhaliwal et al. (2016), Krishnan et al. (2019), and Cohen and Li (2020).

Stock sentiment risk: our primary stock sentiment risk measure is the beta from the following Carhart four-factor regression plus the sentiment factor:

$$r_{i,t} = \alpha_i + \beta_{i,t}^{M} MKT_t + \beta_{i,t}^{S} SMB_t + \beta_{i,t}^{H} HML_t + \beta_{i,t}^{U} UMD_t + \beta_{i,t}^{ST} BWS_t + \vartheta_{i,t}$$

where $r_{i,t}$ is the monthly excess return for stock i in month t, and *MKT*, *SMB*, *HML*, and *UMD* are the factors from Fama and French (1993) and Carhart (1997). *BWS* is the Baker–Wurgler (BW) Sentiment index. In the rest of the paper, the notation β^{ST4} represents the sentiment beta from the above Carhart 4-factor model and serves as the primary measure of stock sentiment risk. For robustness checks, sentiment beta is also estimated from Fama–French 3-factor model and CAPM model plus the sentiment factor, represented by β^{ST3} and β^{ST1} respectively.

[5] For a robustness test on this issue, we assign delisting-month returns following Shumway (1997) and Shumway and Warther (1999). We also employ the Heckman (1979) model to test the possible impact of delisting bias on our main results. The untabulated results are consistent with those reported in the paper and thus refute concerns for delisting and survivorship bias.

Institutional investor turnover (*InstTO*): to measure the institutional investors' investing horizon with a company, we adopt the institutional investor turnover (*InstTO*) proposed by Gaspar et al. (2005) as an inverse indicator. If the institutional investors of a company are mostly short-term oriented, they buy and sell their investments frequently (thus *InstTO* is high), while a low *InstTO* indicates that the institutional shareholders, on average, have a relatively long investment horizon with a company (i.e., they hold shares for a considerable length of time).

Effecting channel measures: as discussed above, we hypothesize that customer concentration affects stock sentiment risk via three channels: operating and accounting performance, information quality, and institutional investors. Four performance measures are adopted: cash flow volatility (*cfvol*), earnings before interest, tax, depreciation and amortization scaled by market value (*ebitda*), gross profitability scaled by total assets (*gp*), and sales growth ($\Delta sale$). Three measures are adopted for information quality: audit fees (*audit_fee*), audit and related fees (*total_fee*), and an indicator variable for restatement of previously audited financial statements (*restatement*). Since not all institutional investors are the same, following Gaspar et al. (2005) and Boscaljon et al. (2021), among others, the institutional investor turnover (*InstTO*) is adopted as an inverse measure of their investment horizon. If a company has low institutional investor turnover, it means that these institutions have long investment horizons with this company.

Control variables: stock sentiment risk and firm performance are affected by a variety of firm characteristics besides customer base concentration. In the context of study, they serve as control variables. We follow the guidance in literature when choosing them (e.g., Chen et al. 2001; Petersen 2009; Gow et al. 2010; Patatoukas 2012). Various aspects of companies' characteristics—age, capital structure, profitability, balance sheet, liquidity, and ownership structure—are considered.

Description of all main variables is summarized in Appendix A.

3.3 Methodology

The main empirical methodology employed throughout the paper is the panel regression. Various combinations of firm, industry, and year fixed effects are controlled to demonstrate the results' robustness. To test the main hypothesis, Hypothesis H1, the following regression is employed:

$$b\_csi4_{i,t} = \beta_0 + \beta_1 concus_{i,t-1} + \beta_2 controls_{i,t-1} + \varepsilon_{i,t}$$

The lead-lag design of the regressions follows Patatoukas (2012), Irvine et al. (2016) and Cohen and Li (2020), which intends to alleviate the endogeneity concern. To test the channels through which customer concentration affects stock sentiment risk, two-stage least squares (2SLS) regressions are conducted. In the first stage, a channel measure is regressed on *concus* along with control variables:

$$channel_{i,t} = \gamma_0 + \gamma_1 concus_{i,t-1} + \gamma_2 controls_{i,t-1} + \eta_{i,t}$$

In the second stage, the stock sentiment beta is regressed on the fitted value of the channel from the first stage:

$$b\_csi4_{i,t} = \delta_0 + \delta_1 \widehat{channel}_{i,t-1} + \delta_2 controls_{i,t-1} + \zeta_{i,t}$$

Table 2 Summary statistics

Variable	Mean	SD	p25	p50	p75	obs
β^{ST4}	− 0.006	0.184	− 0.064	− 0.003	0.053	341,069
β^{ST1}	− 0.008	0.180	− 0.066	− 0.005	0.044	361,248
β^{ST3}	− 0.009	0.179	− 0.063	− 0.003	0.050	356,400
β^{CCI}	0.001	0.164	− 0.056	0.009	0.076	247,324
Concus	0.191	0.177	0.048	0.126	0.293	341,069
cfvol	0.607	1.045	0.314	1.320	3.018	326,704
Δsale	0.079	0.021	− 0.051	0.074	0.104	341,069
gp	0.507	0.913	0.194	0.560	0.691	341,069
Ebitda	0.033	0.115	0.034	0.217	0.589	341,069
Audit_fee	15.571	33.214	2.461	6.374	14.246	168,017
Total_fee	19.852	36.063	2.874	7.257	16.973	168,017
Restatement	0.113	0.317	0.000	0.000	0.000	168,017
InstTO	0.206	0.064	0.165	0.199	0.236	341,069
Leverage	0.207	0.793	0.017	0.209	0.354	341,069
mtb	2.401	2.507	1.121	1.556	2.540	341,069
Cash	0.284	0.246	0.062	0.233	0.447	341,069
ppe	0.200	0.229	0.042	0.103	0.260	341,069
cf_na	0.307	1.023	0.047	0.216	0.519	341,069
nwc_na	0.032	0.658	− 0.056	0.074	0.253	341,069
Capex	0.056	0.074	0.017	0.037	0.071	341,069
rd	0.059	0.690	0.000	0.007	0.084	341,069
Log_at	1.705	0.194	0.542	1.101	1.725	341,069
Inst_ratio	0.198	0.264	0.086	0.223	0.761	341,069
Inage	3.428	0.411	3.178	3.367	3.714	341,069
Soft	0.690	0.308	0.394	0.703	0.831	341,069
Δcsale	0.247	0.417	− 0.024	0.109	0.307	341,069
Roa	0.029	0.163	− 0.009	0.059	0.109	341,069
Δroa	0.001	0.106	− 0.038	0.003	0.034	330,176
sg	0.129	0.364	− 0.056	0.079	0.246	341,069
Loss	0.250	0.433	0.000	0.000	0.000	341,069
Debtmaturity	0.382	0.177	0.124	0.257	0.441	350,952
Foreign	0.603	0.412	0.000	1.000	1.000	341,069

This table presents the summary statistics of sample firms for the period from 1997 to 2017. All continuous variables are winsorized at the 1st and 99th percentiles

Three channels are investigated, including operating and accounting performance, information quality, and institutional investor horizon. Accordingly, in each category, multiple measures are usually adopted as discussed above.

4 Empirical results

4.1 Summary statistics

Table 2 reports the summary statistics for the main variables. The measure on customer concentration, *concus*, has a mean of 0.191 and median 0.126, comparable with the data in Patatoukas (2012) and Boscaljon et al. (2021). A standard deviation of 0.177 indicates significant variations in customer concentration amongst companies: many have a very dispersed customer base, while some others have major customers. The means for the one-factor-based to four-factor-based sentiment beta are between − 0.006 and − 0.009, with sufficient variations amongst companies (the standard deviations are all around 0.18). The gap between the mean and median of most control variables is relatively small, indicating a balanced distribution.

Table 3 reports the Pearson correlation coefficients of sentiment betas, customer concentration, and control variables. As expected, the three versions of sentiment risk are highly correlated. *Concus* is negatively correlated with all three sentiment betas (β^{ST4}, β^{ST3}, β^{ST1}) at 1% level, indicating that firms with a concentrated customer base usually have a low sentiment beta. Among the four performance measures, *cfvol* is positively correlated with sentiment betas, while the other three (*gp, Δsale, ebitda*) are negatively correlated with sentiment betas. Low sentiment risk is thus correlated with strong operating and accounting performance and low cashflow uncertainty. The three measures on information quality—*audit_fee*, *total_fee*, and *restatement*—are all positively correlated with sentiment betas and negatively correlated with customer concentration. Institutional investor turnover (*InstTO*), an inverse indicator of their investment horizon, is positively correlated with sentiment betas. Long-horizon institutional ownership is therefore correlated with low sentiment risk. These said, all the observations are simple correlations between pairs of variables when other factors are not controlled. To check the potential effect of multicollinearity on regression results, the variance inflation factor (VIF) is calculated for each variable. The VIFs are mostly less than 2, with the maximum of 2.27, which is well below the generally agreed threshold of concern (i.e., 2.5 or higher).

4.2 Effect of customer concentration on stock sentiment risk

We begin our analysis by examining how a concentrated customer base may impact stock sentiment risk using a panel regression approach. Panel A in Table 4 shows that stock sentiment risk is negatively linked to the level of customer concentration. In all three model specifications, this relation is not only statistically strong at 1% level but also economically significant. For example, in the model with firm and month fixed effects (i.e., column (2)), a one-standard-deviation increase in *concus* is associated with a 0.00053, or 8.85%, reduction, on average, in sentiment beta. Stocks of companies with concentrated customer base are less prone to be affected by market-wide sentiment. Among control variables, *leverage* is positively related to stock sentiment risk, consistent with prior literature (Chen et al. 2001, 2019; Hutton et al. 2009; Kim et al. 2011, 2014). *Mtb*, *cash*, and *Ins_Ratio* are negatively related to stock sentiment risk. Consistent with findings in literature (Fee and Thomas 2004; Bhattacharyya and Nain 2011), stocks of companies with good growth potential, more liquid assets, or higher institutional ownership tend to have lower sentiment risk. In models (1)–(3), various combinations of firm, industry, and month fixed effects

Table 3 Correlation coefficient matrix

	β^{ST4}	β^{ST1}	β^{ST3}	β^{CCI}	concus	cfvol	Δsale	gp	ebitda	audit_fee	total_fee	restatement	instTO
β^{ST4}	1												
β^{ST1}	0.885***	1											
β^{ST3}	0.940***	0.946***	1										
β^{CCI}	0.742***	0.718***	0.706***	1									
concus	− 0.003***	− 0.07***	− 0.001***	− 0.007***	1								
cfvol	0.007***	0.030***	0.005***	0.027***	− 0.014***	1							
Δsale	− 0.004***	− 0.005***	− 0.001***	− 0.014***	0.065***	− 0.124***	1						
gp	− 0.009***	− 0.021***	− 0.007***	− 0.004***	0.001***	− 0.277***	0.245***	1					
ebitda	− 0.010***	− 0.004***	− 0.010***	− 0.018***	0.082***	− 0.128***	0.279***	0.042***	1				
audit_fee	0.002***	0.007***	0.002***	0.006***	− 0.074***	0.669***	− 0.103***	− 0.732***	− 0.642***	1			
total_fee	0.001***	0.008***	0.002***	0.011***	− 0.074***	0.670***	− 0.100***	− 0.757***	− 0.658***	0.955***	1		
restatement	0.007***	0.005***	0.006***	0.012***	− 0.076***	0.045***	− 0.005***	− 0.040***	− 0.038***	0.047***	0.041***	1	
instTO	0.607***	0.619***	0.583***	0.412***	− 0.034***	0.003***	− 0.001***	− 0.003***	− 0.002***	0.004***	0.004***	0.003***	1
leverage	− 0.001***	− 0.007***	− 0.005***	− 0.036***	− 0.007	− 0.006***	− 0.014*	− 0.012***	− 0.005***	0.009***	0.012***	0.004***	0.028***
mtb	− 0.001***	0.022***	0.013***	0.023***	− 0.000	− 0.012***	− 0.000***	− 0.012***	− 0.014*	− 0.036***	− 0.032***	0.003**	0.061***
cash	− 0.008***	0.008***	0.002***	− 0.019***	− 0.001	0.006***	− 0.028***	− 0.033***	− 0.044***	− 0.075*	− 0.075	− 0.006	0.166***
ppe	− 0.021***	− 0.034***	− 0.026***	− 0.009***	0.005	− 0.022***	0.001**	− 0.020**	0.034***	− 0.070	− 0.072*	0.001*	− 0.025***
cf_na	− 0.005***	− 0.012***	− 0.012***	− 0.233***	− 0.002	0.050***	0.022***	0.062***	0.066***	0.065	0.065	− 0.007	− 0.053***
nwc_na	0.014*	0.005***	0.006***	0.004***	− 0.003	− 0.039***	− 0.006***	− 0.031***	− 0.033***	− 0.025*	− 0.026	0.003	− 0.084***
capex	− 0.014*	− 0.025***	− 0.014**	− 0.048**	− 0.005	− 0.011***	0.006***	− 0.012***	0.022***	− 0.073	− 0.073	− 0.008	0.077***
rd	0.017**	0.017**	0.019**	0.033**	− 0.000	− 0.003**	− 0.003**	− 0.007*	− 0.007*	− 0.010**	− 0.009**	0.000	0.001***
log_at	− 0.023***	− 0.016**	− 0.029***	− 0.017***	− 0.011	0.490***	0.139***	0.501***	0.487***	0.582***	0.576***	− 0.055***	− 0.059***
inst_ratio	− 0.005**	− 0.001	− 0.002	− 0.004**	0.009***	− 0.088***	− 0.024***	− 0.089***	− 0.087***	− 0.113***	− 0.111***	0.024***	− 0.040***
lnage	− 0.013***	− 0.012***	− 0.011***	0.014***	− 0.044***	0.128***	0.033***	0.157***	0.139***	0.189***	0.202***	0.037***	− 0.205***
soft	0.006	− 0.008	0.001	0.042*	0.001	0.013**	− 0.007	− 0.003	− 0.004	0.002	0.004	0.008***	− 0.029
Δcsale	− 0.007*	− 0.006	− 0.005	− 0.013***	− 0.000	0.003	0.036*	0.024	0.017*	0.000	0.008	− 0.008***	0.010
roa	− 0.014**	0.021	− 0.005**	− 0.003**	0.000	− 0.098**	− 0.017	0.016	0.004	− 0.052**	0.013**	− 0.008**	− 0.231

Table 3 (continued)

	β^{ST4}	β^{ST1}	β^{ST3}	β^{CCI}	concus	cfvol	Δsale	gp	ebitda	audit_fee	total_fee	restatement	instTO
Δroa	− 0.003*	0.004	0.001	0.032	− 0.001	− 0.001**	0.008**	0.003	0.006	− 0.001	0.000	0.005***	0.018**
sg	− 0.004	− 0.004	− 0.004	− 0.056	0.000	− 0.002	0.001	− 0.001*	− 0.000	− 0.002***	− 0.001***	− 0.003***	− 0.001
loss	0.021	0.019**	0.028**	− 0.033	0.016**	− 0.103	− 0.067	− 0.137	− 0.142	− 0.166***	− 0.164***	0.015***	0.080**
debtmaturity	0.000	− 0.004*	− 0.001	0.026	− 0.000	− 0.006	− 0.003**	− 0.006*	− 0.006*	− 0.006***	− 0.005***	− 0.008***	0.005
foreign	− 0.010**	− 0.005	− 0.004	0.011*	0.003	− 0.031	− 0.005	− 0.011**	− 0.030*	0.019**	0.018**	0.022	− 0.088**

This table reports the Pearson correlation coefficients of sentiment beta measures, customer concentration, and other variables. All continuous variables are winsorized at the 1st and 99th percentiles

Table 4 Effect of customer concentration on sentiment beta

	Panel A: level values			Panel B: change values		
	(1)	(2)	(3)	(4)	(5)	(6)
	β^{ST4}	β^{ST4}	β^{ST4}	$\Delta\beta^{ST4}$	$\Delta\beta^{ST4}$	$\Delta\beta^{ST4}$
Concus	− 0.002***	− 0.003***	− 0.001***			
	(− 4.094)	(− 4.859)	(− 4.578)			
Δconcus				− 0.002***	− 0.003***	− 0.005***
				(− 4.017)	(− 4.080)	(− 3.457)
Leverage	0.004**	0.005***	0.004***	− 0.003***	− 0.003***	0.001***
	(2.445)	(2.793)	(3.261)	(− 3.778)	(− 2.862)	(3.600)
mtb	− 0.001***	− 0.001***	− 0.001***	0.001**	0.001**	0.000
	(− 11.707)	(− 12.295)	(− 10.514)	(2.489)	(2.423)	(0.879)
Cash	− 0.006*	− 0.009***	− 0.013***	0.006***	0.004***	− 0.004
	(− 1.907)	(− 2.591)	(− 5.952)	(3.911)	(3.572)	(− 1.130)
ppe	0.007	0.010***	− 0.016***	− 0.026**	− 0.026**	− 0.006
	(1.327)	(3.901)	(− 5.782)	(− 2.441)	(− 2.474)	(− 1.253)
cf_na	− 0.011***	− 0.012***	− 0.012***	− 0.001	− 0.002***	0.001
	(− 3.496)	(− 3.593)	(− 2.885)	(− 0.399)	(− 3.325)	(0.693)
nwc_na	0.001**	0.002**	− 0.002	0.001***	0.001***	− 0.001
	(2.232)	(2.481)	(− 0.451)	(2.957)	(2.926)	(− 1.201)
Capex	− 0.001	0.009	− 0.027***	− 0.016***	− 0.005***	0.009
	(− 0.093)	(1.537)	(− 5.263)	(− 3.288)	(− 3.378)	(0.940)
rd	0.001***	0.001***	− 0.003***	0.001	0.002	0.001**
	(6.241)	(6.090)	(− 7.775)	(0.657)	(0.683)	(2.329)
Log_at	0.004***	− 0.001	0.000	− 0.000	− 0.002	− 0.001**
	(5.316)	(− 1.173)	(0.315)	(− 0.331)	(− 1.473)	(− 2.225)
Inst_ratio	− 0.052***	− 0.045***	− 0.018***	− 0.032	− 0.025	− 0.021
	(− 3.990)	(− 3.477)	(− 1.498)	(− 1.280)	(− 0.983)	(− 1.006)
Inage	− 0.032***	− 0.038***	− 0.033**	− 0.038**	− 0.031***	− 0.034***
	(− 2.846)	(− 6.937)	(− 1.978)	(− 2.135)	(− 3.228)	(− 3.945)
_cons	− 0.024***	− 0.023***	− 0.006	0.024***	0.017***	0.070***
	(− 6.460)	(− 6.319)	(− 1.110)	(6.218)	(4.592)	(12.301)
Fixed firm	Yes	Yes	No	Yes	Yes	No
Fixed month	No	Yes	Yes	No	Yes	Yes
Fixed industry	No	No	Yes	No	No	Yes
N	341,069	341,069	341,069	130,926	130,926	130,926
Adi_R^2	0.509	0.515	0.502	0.405	0.406	0.414

This table presents the panel regression estimation results where sentimental beta (four factors) is regressed on 1-year lagged customer concentration and control variables including leverage, mtb, cash, ppe, cf_na, nwc_na, capex, rd, log_at, inst_ratio and lnage. Models (1) and (4) include firm fixed effects, models (2) and (5) include both firm and month fixed effects, and models (3) and (6) include both month and industry fixed effects. The dependent variable and the independent variable (concus) in models (1)–(3) are level values, while those in models (4)–(6) are change values. The dependent variable, independent variable and all control variables are defined in Appendix A. The t-Statistics are in parentheses below parameter estimates, and they are based on clustered standard errors at year and industry levels. ***, **, and * indicate significance at 1%, 5%, 10% levels, respectively

are controlled. The main result regarding the effect of customer concentration on stock sentiment risk is highly consistent and robust.

Panel B of Table 4 shows the regression results with change values, which intend to alleviate the endogeneity concern that the effect of customer concentration may be driven by correlated omitted variables. Regression analyses based on level values are inherently vulnerable to correlated omitted variables. To mitigate this concern, we conduct a multivariate changes analysis by regressing changes in stock sentiment risk on the corresponding changes in customer concentration while other factors are controlled. Estimation results in Panel B are consistent with those in Panel A, confirming that an increase in customer concentration can lead to a decrease in stock sentiment risk.

All panel regressions throughout the paper include firm or industry fixed effects and time fixed effects. Firm or industry fixed effects control for unobservable time-invariant firm-level or industry-level differences, and time (month or year) fixed effects account for common time-variant factors, such as macroeconomic and market conditions. With such controls, for a firm that experiences a change in customer concentration, other firms in that period serve as control firms (Cohen and Li, 2020). Intuitively, this approach is similar to the Difference-in-Difference design as in Bertrand and Mullainathan (2003) and Huang et al. (2016).

An important issue is whether the effect of customer concentration on sentiment beta could be nonlinear. We investigate this possibility in two ways. First, we add a quadratic term of *concus* (i.e., $concus^2$) to our main regressions. The results, as shown in Table 5, indicate that the $Concus^2$ is insignificant while *concus* remains highly significant. Such a finding suggests that the relationship between *concus* is sentiment beta is largely linear. Second, we check whether piecewise regressions are warranted. Instead of splitting the sample somewhat arbitrarily, we adopt the threshold regressions for panel data proposed by Hansen (1999, 2000). The basic idea of threshold regressions is to let the data decide whether change points (thresholds) exist that separate the different regimes. The threshold regression is equipped to detect and explore the potentially nonlinear effect of an explaining variable on the explained variable. As indicated in Table 6, all the LM tests for single, double, or triple thresholds yield insignificant results in different fixed-effects panel regression models. Therefore, they suggest no significant threshold effects, thus confirming the generally linear relationship between customer concentration and stock sentiment risk.

4.3 Effect mechanisms

In this section, we investigate the channels through which customer concentration may impact stock sentiment risk.

4.3.1 Operating and accounting performance

We first consider the possibility that our results are driven by firm performance associated with concentrated customer base, which results in lower sentiment risk. Major customers can influence firms' cash flows, business models, and liquidity management (e.g., Patatoukas 2012; Dhaliwal et al. 2016). We conjecture that a more concentrated customer base may decrease a firm's operating risk, which leads to better operating and accounting performance and thus contributes to lower stock sentiment risk. Two-stage least squares regressions are employed in this investigation. In the first stage, we regress the performance measures (*cfvol, ebitda, gp, Δsale*) on customer concentration along with the control

Table 5 Check for non-linear effect of customer concentration on sentiment beta

	(1) β^{ST4}	(2) β^{ST4}	(3) β^{ST4}
$Concus^2$	0.000	0.000	0.000
	(0.784)	(0.722)	(0.648)
Concus	− 0.001***	− 0.001***	− 0.001***
	(− 2.909)	(− 3.019)	(− 3.045)
Leverage	0.004**	0.005***	0.004***
	(2.443)	(2.789)	(3.268)
mtb	− 0.001***	− 0.001***	− 0.002***
	(− 3.707)	(− 3.295)	(− 3.510)
Cash	− 0.006*	− 0.009***	− 0.013***
	(− 1.910)	(− 2.599)	(− 5.987)
ppe	0.007	0.010*	− 0.016***
	(1.322)	(1.887)	(− 5.796)
cf_na	− 0.007	− 0.009	− 0.006
	(− 0.497)	(− 0.596)	(− 0.884)
nwc_na	0.002	0.003	− 0.005
	(0.133)	(0.482)	(− 0.437)
Capex	− 0.001	0.009	− 0.027***
	(− 0.092)	(1.538)	(− 5.259)
rd	0.001***	0.001***	− 0.001***
	(6.242)	(6.092)	(− 7.784)
log_at	0.004***	− 0.001	0.002
	(5.311)	(− 1.173)	(0.395)
inst_ratio	− 0.052***	− 0.045***	− 0.018
	(− 3.989)	(− 3.478)	(− 1.495)
Inage	− 0.033***	− 0.033***	− 0.031***
	(− 4.745)	(− 4.872)	(− 4.565)
_cons	− 0.021***	0.002	0.006
	(− 3.361)	(0.495)	(0.974)
Fixed firm	Yes	Yes	No
Fixed month	No	Yes	Yes
Fixed industry	No	No	Yes
N	341,069	341,069	341,069
Adi_R^2	0.107	0.101	0.107

This table presents the estimation results of panel regressions where sentimental beta (four factors) is regressed on 1-year lagged customer concentration, the square of customer concentration and control variables including leverage, mtb, cash, ppe, cf_na, nwc_na, capex, rd, log_at, inst_ratio and lnage. Model (1) includes firm fixed effects, model (2) includes both firm and month fixed effects, and model (3) includes both month and industry fixed effects. The dependent variable, independent variable and all control variables are defined in Appendix A. The t-Statistics are in parentheses below parameter estimates, and they are based on clustered standard errors at year and industry levels. ***, **, and * indicate significance at 1%, 5%, 10% levels, respectively

Table 6 Tests for thresholds in potential non-linear effect of customer concentration on sentiment beta

	Firm-fixed			Firm-fixed and month-fixed			Month-fixed and industry-fixed		
	Threshold value	*P* value	Confidence intervals	Threshold value	*P* value	Confidence intervals	Threshold value	*P* value	Confidence intervals
Single threshold	0.012	0.7471	(0.0107, 0.0125)	0.009	0.8525	(0.0072, 0.0115)	0.018	0.9061	(0.0115, 0.0224)
Double threshold (Th-21)	0.031	0.6832	(0.0279, 0.0363)	0.022	0.7362	(0.0186, 0.0264)	0.026	0.7726	(0.0199, 0.0303.)
Double threshold (Th-22)	0.107	0.8387	(0.1024, 0.1138)	0.115	0.7691	(0.1078, 0.1213)	0.119	0.7537	(0.1136, 0.1247)
Triple threshold (TH-31)	0.011	0.9023	(0.0083, 0.0134)	0.013	0.6214	(0.0075, 0.0209)	0.015	0.8910	(0.0093, 0.0167)
Triple threshold (TH-32)	0.134	0.6912	(0.1389, 0.1410)	0.149	0.7648	(0.1388, 0.1527)	0.139	0.7934	(0.1304, 0.1472)
Triple threshold (TH-33)	0.227	0.7086	(0.2185, 0.2337)	0.218	0.9035	(0.2097, 0.2213)	0.222	0.8227	(0.2135, 0.2328)

(1) The threshold estimates refer to the level of *concus*. (2) Th-21 and Th-22 are the first and second threshold estimated in the double-threshold model; Th-31, Th-32, Th-33 are the first, second and third threshold estimated in the triple-threshold model, respectively. (3) P-value from LM test represents the significance of the threshold value, and there is no threshold effect if p-value > 0.1. (4) The industry-fixed and month-fixed effects are controlled with industry dummies and month dummies, respectively. (5) Reported confidence intervals are at 95% confidence level

Table 7 Mechanism one: operating and accounting performance

	Panel A				Panel B			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	cfvol	ebitda	gp	Δsale	β^{ST4}	β^{ST4}	β^{ST4}	β^{ST4}
Autolag	0.076***	0.086***	0.054***	0.037***				
	(5.215)	(3.996)	(3.189)	(4.741)				
Concus	− 0.011***	0.003***	0.006***	0.009***				
	(− 4.056)	(4.134)	(3.472)	(4.236)				
Pre-cfvol					0.024***			
					(3.709)			
Pre-ebitda						− 0.013***		
						(− 3.196)		
Pre-gp							− 0.007***	
							(− 3.863)	
Pre-Δsale								− 0.015***
								(− 3.763)
Leverage	0.440	− 0.054***	− 0.734	− 0.653***	0.003	0.004**	0.004**	0.003*
	(0.691)	(− 3.494)	(− 0.902)	(− 3.506)	(1.561)	(1.979)	(1.981)	(1.931)
mtb	− 0.002***	− 0.008***	− 0.008***	− 0.005***	− 0.002***	− 0.001***	− 0.003***	− 0.002***
	(− 3.405)	(− 3.925)	(− 3.617)	(− 3.145)	(− 8.984)	(− 5.194)	(− 5.187)	(− 5.187)
Cash	− 0.031***	− 0.035***	− 0.034***	− 0.010***	− 0.009**	− 0.010***	− 0.010***	− 0.010***
	(− 3.441)	(− 3.932)	(− 4.111)	(− 3.909)	(− 2.573)	(− 3.015)	(− 3.022)	(− 3.112)
ppe	− 0.008**	− 0.065	− 0.006	− 0.039***	0.009*	0.008	0.008	0.007
	(− 2.372)	(− 1.497)	(− 1.413)	(− 3.166)	(1.801)	(1.493)	(1.474)	(1.285)
cf_na	0.033***	0.116***	0.118	0.119	0.012***	0.011***	0.011***	0.011***
	(3.495)	(3.051)	(0.767)	(0.284)	(9.128)	(5.744)	(5.737)	(5.738)
nwc_na	− 0.122	− 0.453	− 0.610***	− 0.440***	− 0.002***	− 0.001*	− 0.001*	− 0.001*
	(− 0.669)	(− 1.216)	(− 3.172)	(− 3.310)	(− 4.301)	(− 1.816)	(− 1.810)	(− 1.775)

Table 7 (continued)

	Panel A				Panel B			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	cfvol	ebitda	gp	Δsale	β^{ST4}	β^{ST4}	β^{ST4}	β^{ST4}
Capex	− 0.032***	− 0.046***	− 0.051***	− 0.058***	0.001***	0.001***	0.002***	0.001***
	(− 3.526)	(− 5.106)	(− 5.725)	(− 6.521)	(5.630)	(5.207)	(5.207)	(5.216)
rd	0.001	− 0.004	− 0.003	0.019	0.004***	0.004***	0.004***	0.004***
	(0.129)	(− 0.274)	(− 0.157)	(0.360)	(5.256)	(5.680)	(5.682)	(5.237)
Log_at	0.214***	0.600***	0.342***	0.074***	− 0.051***	− 0.051***	− 0.051***	− 0.052***
	(3.501)	(3.637)	(3.535)	(4.185)	(− 3.941)	(− 3.926)	(− 3.928)	(− 3.977)
Inst_ratio	0.074	0.096	0.062	− 0.102	− 0.024***	− 0.022**	− 0.019**	− 0.012
	(0.728)	(0.359)	(0.354)	(− 0.084)	(− 2.686)	(− 2.483)	(− 2.132)	(− 1.370)
Inage	− 0.008***	− 0.011***	− 0.013***	− 0.009***	− 0.028***	− 0.029***	− 0.029***	− 0.030***
	(− 4.243)	(− 4.517)	(− 4.746)	(− 5.063)	(− 3.277)	(− 3.463)	(− 3.497)	(− 3.185)
_cons	− 0.039***	− 0.054***	− 0.045***	− 0.044***	− 0.018**	− 0.023***	− 0.017**	− 0.027***
	(− 4.408)	(− 6.034)	(− 5.021)	(− 4.908)	(− 2.046)	(− 2.656)	(− 1.973)	(− 3.101)
Fixed firm	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed year	Yes	Yes	Yes	Yes	No	No	No	No
Fixed month	No	No	No	No	Yes	Yes	Yes	Yes
N	326,704	341,069	341,069	341,069	326,704	341,069	341,069	341,069
Adi_R^2	0.501	0.436	0.524	0.470	0.537	0.481	0.516	0.503

This table reports the two-stage regression estimation results of Mechanism #1, Operating Performance. Panel A reports the first-stage panel estimation results where the operating performance measures, ebitda, gp, Δsale, or cfvol, are regressed on 1-year lagged concus and control variables including leverage, mtb, cash, ppe, cf_na, nwc_na, capex, rd, log_at, inst_ratio and lnage. The dependent variable in model (1) is cfvol, in model (2) is ebitda, in model (3) is gp, and in model (4) is ΔSale. All models in Panel A include firm and year fixed effects. Panel B reports the second-stage estimation results where sentimental beta (four-factor) is regressed on the concus-predicted operating performance, Pre-cfvol, Pre-ebitda, Pre-gp, or Pre-Δsale, and control variables as above. In models (5)–(8), the dependent variable is sentimental beta (four-factor). To obtain concus-predicted operating performance, we regress operating performance measures on concus and controls (as shown in Panel A), and then use the fitted value in the regressions. All models in Panel B include firm and month fixed effects. The dependent variables, independent variables and all control variables are defined in Appendix A. The t-statistics are in parentheses below parameter estimates, and they are based on robust standard errors clustered at year and industry levels. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively

variables in Table 4. In the second stage, we regress the sentiment beta on the predicted performance measure and control variables.

Panel A of Table 7 reports the results of the first stage regressions where the dependent variables are *cfvol*, *ebitda*, *gp*, or Δ*Sale*, and the main independent variable is lagged *concus*. The coefficients of *concus* are significant at 1% level for all four performance measures: *concus* positively predicts firms' *ebitda*, *gp* and Δ*Sale*, and negatively predicts *cfvol*. The higher the customer concentration, the lower the cash flow volatility, the stronger the earnings, gross profitability, and sales growth.

We next estimate the impact of the firms' operating and accounting performance driven by *concus* on stock sentiment risk. The *concus*-predicted *cfvol*, *ebitda*, *gp* and Δ*sale* are fitted values from the regressions in Panel A, and are attached with the prefix pre- to indicate predicted values. We then run regressions of β^{ST4} on *pre-cfvol*, *pre-ebitda*, *pre-gp*, or *pre-Δsale*, whose results are reported in panel B of Table 7. The coefficients of *pre-cfvol*, *pre-ebitda*, *pre-gp*, and *pre-Δsale* are all significant at 1% level, positive for *pre-cfvol* and negative for the others. The results in Table 7 confirm that a concentrated customer base generally benefits the firm's accounting performance and cash flow volatility, which in turn helps attenuate its stock sentiment risk. Hypothesis H2a is thus supported.

4.3.2 Information quality

We now investigate information quality as a potential channel through which customer concentration affects the stock sentiment risk. Three indicators of information quality are considered. Audit fees consist of all fees necessary to perform audit, while total fees are the sum of audit and audit-related fees. Fee data is available only for post-2000 from Audit Analytics. *Restatement* is an indicator variable for the restatement of previously audited financial statements. It indicates failure to detect misstatements in originally audited financial statements and, therefore, provides an objective and visible indicator of audit and information quality (Kinney et al. 2004; Lobo and Zhao 2013; DeFond and Zhang 2014). Companies with concentrated customer bases are easier to audit, thus having lower audit and related fees, and less likely to have restatements. Other things being equal, the unambiguous and apprehensible information in financial statements is convenient for investors to digest and has less room for suspicion and speculation. Stocks of these companies are thus less apt to be affected by market sentiment.

Same as above, we use 2SLS regressions to assess this channel. In the first stage, we regress the information quality measures on customer concentration together with control variables following Krishnan et al. (2019). In the second stage, we regress sentiment beta on the predicted information quality measure with controls.

Panel A of Table 8 shows the estimation results of panel regressions where the dependent variables are *lnaudit_fee*, *lntotal_fee* or *restatement*, and the main independent variable is lagged *concus*. The coefficients of *concus* are significant at the 1% level for all three indicators of information quality. The negative coefficients of *concus* in the regressions of *lnaudit_fee* and *lntotal_fee* mean that suppliers with more concentrated customers spend less on audit and related fees, which indicates faster speed and higher auditing efficiency. Financial statements from these companies are easier for investors to digest. Moreover, restatement is less likely to occur for such companies. Overall, concentrated customers contribute to higher information quality.

We next estimate the impact of the firms' information quality driven by the *concus* on stock sentiment risk. The *concus*-driven information quality measures are the fitted values

Table 8 Mechanism two: information quality

	Panel A			Panel B		
	(1)	(2)	(3)	(4)	(5)	(6)
	lnaudit_fee	lntotal_fee	Restatement	β^{ST4}	β^{ST4}	β^{ST4}
Autolag	0.046	0.045	0.033**			
	(0.920)	(0.267)	(2.500)			
Concus	− 0.008***	− 0.021***	− 0.107***			
	(− 3.642)	(− 3.781)	(− 3.775)			
Pre-lnaudit_fee				0.001***		
				(3.570)		
Pre-lntotal_fee					0.003***	
					(3.651)	
Pre-restatement						0.021***
						(4.540)
Leverage	0.017***	0.015***	− 0.000	− 0.029***	− 0.030***	− 0.029***
	(4.336)	(3.698)	(− 0.002)	(− 3.237)	(− 3.398)	(− 3.219)
mtb	− 0.001***	− 0.000***	0.000	− 0.035***	− 0.035***	− 0.034***
	(− 5.149)	(− 3.094)	(0.047)	(− 3.852)	(− 3.893)	(− 3.724)
Log_at	0.069***	0.072***	− 0.059***	0.011***	0.010***	0.010***
	(4.273)	(4.527)	(− 4.760)	(8.900)	(8.100)	(8.554)
rd	− 0.000	− 0.000	− 0.001	0.001***	0.002***	0.001***
	(− 0.760)	(− 0.449)	(− 0.783)	(5.172)	(5.177)	(5.142)
ppe	0.268***	0.085	− 2.192	0.021***	0.027***	0.026***
	(2.838)	(0.847)	(− 0.668)	(2.830)	(3.613)	(3.411)
Debtmaturity	0.000	0.000	− 0.001			
	(1.004)	(0.858)	(− 1.190)			
roa	− 0.022***	− 0.027***	− 0.178			
	(− 7.586)	(− 8.645)	(− 1.636)			
Loss	0.005***	0.003*	− 0.082			
	(2.758)	(1.790)	(− 1.364)			
sg	− 0.000	− 0.000	− 0.028			
	(− 0.371)	(− 1.205)	(− 1.637)			
Foreign	0.004	0.001				
	(1.288)	(0.507)				
Soft			0.492**			
			(2.570)			
Δcsale			− 0.004*			
			(− 1.958)			
Δroa			0.104			
			(1.161)			
Cash				0.005	0.007	0.009*
				(0.987)	(1.294)	(1.790)
Capex				0.025**	0.025**	0.024**
				(2.413)	(2.412)	(2.290)
cf_na				0.010***	0.010***	0.010***
				(9.461)	(9.516)	(9.466)

Table 8 (continued)

	Panel A			Panel B		
	(1)	(2)	(3)	(4)	(5)	(6)
	lnaudit_fee	lntotal_fee	Restatement	β^{ST4}	β^{ST4}	β^{ST4}
nwc_na				− 0.025	− 0.028**	− 0.016
				(− 1.445)	(− 2.306)	(− 0.898)
Inst_ratio				0.005***	0.005***	0.005***
				(4.549)	(4.530)	(4.565)
Inage				− 0.031***	− 0.031***	− 0.030***
				(− 5.733)	(− 5.847)	(− 5.587)
_cons	− 0.262***	− 0.156***	− 5.040***	− 0.040***	− 0.038***	− 0.037***
	(− 28.092)	(− 15.895)	(− 23.818)	(− 3.786)	(− 3.629)	(− 3.426)
Fixed firm	Yes	Yes	Yes	Yes	Yes	Yes
Fixed year	Yes	Yes	Yes	No	No	No
Fixed month	No	No	No	Yes	Yes	Yes
N	168,017	168,017	168,017	168,017	168,017	168,017
Adi_R^2	0.566	0.503	0.627	0.497	0.526	0.510

This table reports the two-stage regression estimation results of Mechanism #2, Information Quality. Panel A reports the first-stage estimation results. Column (1) and (2) report panel estimation results where the information quality measure, lnaudit_fee or lntotal_fee, is regressed on lagged concus and control variables including leverage, mtb, log_at, rd, ppe, debtmaturity, roa, loss, sg, and foreign. Column (3) reports the logistic regression result where restatement is regressed on lagged concus and control variables including leverage, mtb, log_at, rd, ppe, debtmaturity, roa, loss, sg, soft, Δcsale, and Δroa. All models in Panel A include firm and year fixed effects. Panel B reports the second-stage panel estimation results where sentimental beta (four-factor) is regressed on the concus-predicted information quality, Pre-lnaudit_fee, Pre-lntotal_fee, or Pre-restatement, and control variables. In models (4)–(6), the dependent variable is sentimental beta (four-factor). To obtain concus-predicted information quality, we regress information quality measures on concus and controls, and then use the fitted value in the regressions. All models in Panel B include firm and month fixed effects. The dependent variables, independent variables and all control variables are defined in Appendix A. The t-statistics are in parentheses below parameter estimates, and they are based on robust standard errors clustered at year and industry levels. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively

from the first step of 2SLS. In panel B of Table 8, the coefficients are positive for *pre-lnaudit_fee*, *pre-lntotal_fee* and *pre-restatement*, all significant at the 1% level. Better information quality, therefore, contributes to lower stock sentiment risk.

Together, the results in Table 8 suggest that concentrated customer base helps to improve companies' information quality and, consequently, lower stock sentiment risk. Hypothesis 2b is thus supported.

4.3.3 Long-horizon institutional investors

We then test the third channel by examining the effect of customer concentration on stock sentiment risk via firms' institutional ownership turnover using 2SLS panel regressions. In the first stage, we regress *instTO* on customer concentration. To address the concern that institutional ownership turnover may also be driven by other firm characteristics, the complete set of control variables in Table 4 is included. Fixed firm and fixed time effects are controlled in the regressions. In the second stage, we regress sentiment beta on the predicted *instTO*.

Table 9 Mechanism three: institutional ownership turnover

	(1) instTO	(2) β^{ST4}	(3) inst_ratio
Autolag	0.041***		
	(3.853)		
Concus	− 0.007***		0.009***
	(− 3.164)		(3.428)
Pre-instTO		0.058***	
		(6.080)	
Leverage	− 0.001***	0.007**	0.029***
	(− 4.145)	(2.381)	(3.224)
mtb	− 0.001**	− 0.003***	0.004***
	(− 2.536)	(− 4.237)	(2.960)
Cash	− 0.003*	− 0.001	0.006
	(− 1.870)	(− 0.340)	(0.676)
ppe	− 0.001***	0.006	− 0.003***
	(− 4.213)	(0.980)	(− 2.958)
cf_na	0.009***	0.011***	− 0.015***
	(5.518)	(4.777)	(− 3.788)
nwc_na	− 0.002***	0.001	− 0.006**
	(− 3.871)	(0.961)	(− 2.504)
Capex	− 0.024**	− 0.019**	− 0.027***
	(− 2.338)	(− 2.253)	(− 2.996)
rd	0.002***	0.001***	− 0.001***
	(4.337)	(5.279)	(− 4.310)
Log_at	− 0.001	0.003***	− 0.002***
	(− 1.035)	(3.785)	(− 4.302)
Inst_ratio	0.001	− 0.078***	
	(1.354)	(− 5.253)	
Inage	− 0.013***	− 0.033***	− 0.051***
	(− 3.902)	(− 2.997)	(− 5.749)
_cons	− 0.015***	− 0.045	− 0.024***
	(− 3.043)	(− 0.125)	(− 4.853)
Fixed firm	Yes	Yes	Yes
Fixed year	Yes	No	Yes
Fixed month	No	Yes	No
N	341,069	341,069	341,069
Adi_R^2	0.408	0.418	0.440

This table reports the two-stage regression results of Mechanism #3, Institutional Ownership Turnover. Column (1) reports the first-stage panel estimation results where the institutional ownership turnover measure, instTO, is regressed on 1-year lagged concus and control variables including leverage, mtb, cash, ppe, cf_na, nwc_na, capex, rd, log_at, inst_ratio and lnage. Column (2) reports the second-stage estimation results where sentimental beta (four-factor) is regressed on the concus-predicted Institutional Ownership Turnover, Pre-instTO, and control variables as above. To obtain concus-predicted Institutional Ownership Turnover, we regress instTO on concus and controls (as shown in column (1)), and then use the fitted value in the regression. Column (3) reports estimation results where Inst_Ratio is regressed on 1-year lagged concus and control variables. Models (1) and (3) include firm and year fixed effects, and model (2) includes firm and month fixed effects. The dependent variables, independent variables and all control variables are defined in Appendix A. The t-statistics are in parentheses below parameter estimates, and they are based on robust standard errors clustered at year and industry levels. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively

Since institutional ownership turnover, *InstTO*, can be understood as an inverse indicator of their investment horizon, the negative coefficient of *concus* in model (1) of Table 9 suggests that higher customer concentration is associated with longer term institutional ownership. Reported in model (3) is the estimation result where institutional ownership percentage, *Inst_Ratio*, is the explained variable. It shows that when other factors are controlled, firms with concentrated customer base tend to have higher institutional ownership. Overall, these results demonstrate that companies with concentrated customer base tend to have more institutional investors, and these institutions are inclined to stay longer.

Next, we estimate the impact of the firms' institutional ownership turnover driven by *concus* on the stock sentiment beta. In model (2), the coefficient of *pre-InstTO* is positive at 1% level, confirming that the long-horizon institutional investors, partially attracted by highly concentrated customers, help attenuate firms' stock sentiment risk. It is important to note that the *Inst_Ratio* is significantly negative at 1% level. A large institutional investor base helps lower stock sentiment risk. Overall, institutional investors, particularly those with a long investment horizon, make a company's stock price less susceptible to market sentiment. Hypothesis H2c is therefore supported.

4.4 Robustness checks

4.4.1 Alternative measures of sentiment risk

All of the results thus far employ the measure of stock sentiment risk estimated from the Carhart four-factor model. To check the robustness, we also use sentiment risk measures estimated from CAPM-type single-factor model (denoted as β^{ST1}) and Fama–French three-factor model (denoted as β^{ST3}) and report the results in Tables 10 and 11. The estimated coefficients of *concus* in the regressions with alternative sentiment betas are invariably negative at 1% in all model specifications, consistent with Table 4 and thereby demonstrating their robustness.

Its construction decides that the BW sentiment index is comprehensive, and, more importantly, independent from common macroeconomic conditions. Our approach for estimating stock sentiment beta follows a very large body of well-established research (Baker and Wurgler, 2006; Berger and Turtle, 2012; Massa and Yadav, 2015). To further check the robustness of our results, we estimate an alternative measure of sentiment beta (β^{CCI}) following Keiber and Samyschew (2019). Specifically, the same set of risk factors are adopted, including inflation rate (CIFR), industrial production (CIP), short-term interest rate (CSIR), oil price (COIL), and market return (MR). The investor sentiment measure adopted is Consumer Confidence Index (CCI) for the U.S. market. We re-estimate our main model with β^{CCI} as the explained variable, whose untabulated results are highly consistent with Table 4, thus demonstrating the robustness of our central thesis—a higher concentration in customer base contributes to a lower sentiment beta.

Another potential issue is mismatch in frequency between the explained variable (sentiment beta) and the explaining variable (customer concentration) and control variables. Since stock return data is available at much higher frequency than accounting data, their presence in the same regression is a common practice in asset pricing literature since Fama and French (1992). Nevertheless, to further check the robustness of our results, we also conduct the main regression at the yearly frequency by calculating stock sentiment beta in two ways. First, the annual stock sentiment beta is directly estimated in a 5-year

Table 10 Effect of customer concentration on sentiment beta (single-factor)

	Panel A level values			Panel B change values		
	(1)	(2)	(3)	(4)	(5)	(6)
	β^{ST1}	β^{ST1}	β^{ST1}	$\Delta\beta^{ST1}$	$\Delta\beta^{ST1}$	$\Delta\beta^{ST1}$
Concus	− 0.001***	− 0.004***	− 0.002***			
	(− 3.086)	(− 3.330)	(− 3.273)			
Δconcus				− 0.003***	− 0.004***	− 0.006***
				(− 2.928)	(− 3.058)	(− 3.029)
Leverage	0.006***	0.012***	− 0.004***	− 0.006	− 0.004	0.007**
	(3.333)	(6.956)	(− 2.976)	(− 1.618)	(− 1.123)	(1.994)
mtb	− 0.000***	− 0.000***	− 0.000**	0.000*	0.000**	0.000
	(− 8.580)	(− 8.591)	(− 2.409)	(1.666)	(2.129)	(0.078)
Cash	− 0.011***	− 0.015***	− 0.011***	− 0.005	− 0.012*	0.003
	(− 3.377)	(− 4.731)	(− 5.137)	(− 0.652)	(− 1.691)	(0.444)
ppe	0.009***	0.011**	− 0.016***	− 0.037***	− 0.041***	− 0.010
	(3.117)	(2.135)	(− 6.064)	(− 3.284)	(− 3.628)	(− 1.278)
cf_na	0.012	0.011***	0.011***	− 0.001***	− 0.003***	0.003***
	(1.192)	(3.714)	(3.369)	(− 3.669)	(− 3.373)	(3.174)
nwc_na	− 0.001***	− 0.003***	− 0.002***	0.001	0.002***	0.003***
	(− 3.123)	(− 3.200)	(− 4.938)	(0.071)	(3.254)	(3.314)
Capex	− 0.008	0.005	− 0.006***	− 0.023*	0.003	− 0.008
	(− 1.459)	(0.885)	(− 7.344)	(− 1.714)	(0.258)	(− 0.520)
rd	0.013***	0.011***	− 0.015***	0.004***	0.009***	0.013***
	(7.427)	(7.392)	(− 3.837)	(3.235)	(3.115)	(3.263)
log_at	0.008***	0.006***	0.007***	− 0.001***	− 0.002	− 0.001
	(11.085)	(10.205)	(5.838)	(− 3.635)	(− 1.212)	(− 1.164)
inst_ratio	− 0.020	− 0.017	0.024**	− 0.008	0.008	0.014
	(− 1.617)	(− 1.373)	(2.106)	(− 0.303)	(0.283)	(0.428)
lnage	− 0.029***	− 0.027***	− 0.025**	− 0.021	− 0.028***	− 0.022
	(− 2.719)	(− 8.919)	(− 2.200)	(− 0.344)	(− 3.353)	(− 0.882)
_cons	− 0.013*	− 0.014**	0.053***	− 0.010***	− 0.006*	0.009*
	(− 1.865)	(− 1.990)	(8.250)	(− 2.855)	(− 1.829)	(1.671)
Fixed firm	Yes	Yes	No	Yes	Yes	No
Fixed month	No	Yes	Yes	No	Yes	Yes
Fixed industry	No	No	Yes	No	No	Yes
N	361,248	361,248	361,248	150,323	150,323	150,323
Adi_R^2	0.496	0.509	0.494	0.405	0.403	0.409

This table presents the estimation results of panel regressions where sentimental beta (single-factor) is regressed on 1-year lagged customer concentration and control variables including leverage, mtb, cash, ppe, cf_na, nwc_na, capex, rd, log_at, inst_ratio and lnage. Models (1) and (4) include firm fixed effects, models (2) and (5) include both firm and month fixed effects, and models (3) and (6) include both month and industry fixed effects. The dependent variable and the independent variable (concus) in models (1)–(3) are level values, while those in models (4)–(6) are change values. The dependent variable, independent variable and all control variables are defined in Appendix A. The t-statistics are in parentheses below parameter estimates, and they are based on robust standard errors clustered at year and industry levels. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively

Table 11 Effect of customer concentration on sentiment beta (three-factor)

	Panel A level values			Panel B change values		
	(1)	(2)	(3)	(4)	(5)	(6)
	β^{ST3}	β^{ST3}	β^{ST3}	$\Delta\beta^{ST3}$	$\Delta\beta^{ST3}$	$\Delta\beta^{ST3}$
Concus	− 0.004***	− 0.002***	− 0.001***			
	(− 4.010)	(− 4.379)	(− 3.430)			
Δconcus				− 0.007***	− 0.006***	− 0.005***
				(− 3.516)	(− 3.633)	(− 3.223)
Leverage	0.002	0.004*	− 0.002	− 0.004***	− 0.002***	0.001
	(0.860)	(1.816)	(− 0.291)	(− 3.025)	(− 3.648)	(0.092)
mtb	− 0.001***	− 0.001***	− 0.001***	0.001**	0.002**	0.002
	(− 9.189)	(− 9.672)	(− 7.438)	(2.322)	(2.281)	(1.260)
Cash	− 0.005	− 0.009**	− 0.013***	− 0.002	− 0.006	− 0.007*
	(− 1.247)	(− 2.383)	(− 5.649)	(− 0.291)	(− 0.962)	(− 1.944)
ppe	0.006	0.010*	− 0.017***	− 0.026**	− 0.027***	− 0.002***
	(1.070)	(1.695)	(− 5.647)	(− 2.536)	(− 2.596)	(− 3.422)
cf_na	0.013	0.009***	0.008	− 0.012	− 0.001***	0.001
	(1.010)	(3.816)	(0.353)	(− 0.565)	(− 3.603)	(1.609)
nwc_na	− 0.001*	− 0.001	− 0.001**	− 0.003	− 0.003	− 0.002*
	(− 1.776)	(− 1.056)	(− 2.314)	(− 0.236)	(− 0.027)	(− 1.742)
Capex	0.005	0.020***	− 0.024***	− 0.031**	− 0.016	− 0.020**
	(0.787)	(3.129)	(− 4.307)	(− 2.510)	(− 1.298)	(− 2.040)
rd	0.003***	0.001***	− 0.007***	0.005***	0.005***	0.004*
	(5.949)	(5.818)	(− 3.697)	(3.090)	(3.160)	(1.670)
Log_at	0.005***	− 0.000	0.001***	− 0.002	− 0.002	− 0.001***
	(6.218)	(− 0.203)	(2.636)	(− 1.254)	(− 1.300)	(− 3.110)
Inst_ratio	− 0.022	− 0.012	0.000	− 0.010	− 0.003	0.009
	(− 1.506)	(− 0.811)	(0.000)	(− 0.395)	(− 0.130)	(0.454)
lnage	− 0.020***	− 0.026***	− 0.022**	0.023***	− 0.026***	0.021
	(− 2.616)	(− 5.252)	(− 1.964)	(5.102)	(− 4.094)	(0.395)
_cons	− 0.014**	− 0.019**	− 0.029***	0.018***	0.010***	0.010***
	(− 1.989)	(− 2.400)	(− 4.191)	(4.326)	(2.577)	(2.577)
Fixed firm	Yes	Yes	No	Yes	Yes	No
Fixed month	No	Yes	Yes	No	Yes	Yes
Fixed industry	No	No	Yes	No	No	Yes
N	356,400	356,400	356,400	143,037	143,037	143,037
Adi_R^2	0.413	0.401	0.415	0.317	0.333	0.326

This table presents results of panel regressions where sentimental beta (three-factor) is regressed on 1-year lagged customer concentration and control variables including leverage, mtb, cash, ppe, cf_na, nwc_na, capex, rd, log_at, inst_ratio and lnage. Models (1) and (4) include firm fixed effects, models (2) and (5) include both firm and month fixed effects, and models (3) and (6) include both month and industry fixed effects. The dependent variable and the independent variable (concus) in models (1)–(3) are level values, while those in models (4)–(6) are change values. The dependent variable, independent variable and control variables are defined in Appendix A. The t-statistics are in parentheses below parameter estimates, and they are based on robust standard errors clustered at year and industry levels. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively

rolling window. Second, the average of 12 sentiment betas from each month is perceived as the sentiment beta for the year. Their results are highly consistent with Table 4 and thus are unreported.

4.4.2 Exogeneous shock and difference-in-difference (DiD) analyses

To further address the endogeneity concern regarding the relation between customer concentration and stock sentiment risk, we examine the effect of an exogenous shock in the spirit of a DiD model (Deng et al. 2020; Huang et al. 2016). The idea is to introduce an exogeneous shock that directly affects customer concentration and examine its impact on stock sentiment risk.

A natural exogenous industry-wide shock is unexpected change of industry level import tariffs (e.g., Feenstra and Romalis 2014; Fresard 2010). Import tariff jumps make procurements from foreign suppliers more expensive and therefore cause many customers to switch to domestic suppliers. Thus, the level of customer concentration would increase for domestic companies (Hoberg and Phillips 2010, 2016; Hoberg et al. 2014). Specifically, we use the American import tariff rates obtained from Census Bureau imports data to identify cases where a firm's major customers are affected by an exogenous shock at the industry level. We set *tariff* equal to 1 if the firm-year is after the tariff jump year (i.e., when the tariff growth rate is 20% or higher), and 0 otherwise. The validity of tariff jumps as an exogenous shock for examining of the causal effect of customer concentration on sentiment beta is supported by Table 12. Upon a tariff jump, customer concentration of affected companies indeed increases as some customers switch to domestic suppliers from foreign suppliers. Thus, we can examine its effect by comparing the before-after changes in stock sentiment risk induced by elevated import tariffs (treatment firms) and those unaffected by a tariff jump (control firms). Our model identification is a staggered DiD with multiple treatment groups and multiple time periods, similar to the specification in Bertrand and Mullainathan (2003) and Imbens and Wooldridge (2009). We regress sentiment beta on the dummy variable *tariff* and control variables (i.e., replacing *concus* with *tariff* in Table 4).

We find in Table 13 that the coefficient of *tariff* is significantly negative at 1% level with all control variables included, supporting our expectation that stock sentiment risk decreases after import tariff jumps that raise domestic customer concentration. The magnitude of the DiD estimators suggests that the sentiment beta for the tariff-affected firms decreases between 0.023 and 0.029, compared to similar firms not affected by the tariff jump.

We also conduct the complete DiD regression. The parallel path assumption is tested following Bertrand and Mullainathan (2003) and Serfling (2016), in view of the staggered nature of our experiment group. As shown in Table 14, the interaction terms corresponding to the years before treatment are all insignificant, suggesting that firms in the control group (unaffected by tariff jumps) and the experiment group (affected by tariff jump) are similar prior to the treatment. The parallel assumption is therefore validated and the application of the DiD model is justified.

In essence, the exogenous shock analysis in Table 13 is an abridged DiD analysis, where *tariff* is equivalent to treatment × time in complete DiD models. In the DiD regression results reported in Table 15, the coefficient of the interaction term, treatment × time, is significantly negative, suggesting that the treatment group experiences a significant decrease in sentiment beta due to the tariff jump compared to the control group.

Table 12 Validity of tariff jump as exogenous shock

	(1) concus	(2) concus	(3) concus
Autolag	− 0.107***	− 0.104***	− 0.105***
	(− 13.477)	(− 13.125)	(− 12.783)
Tariff	0.248***	0.239***	0.256***
	(4.014)	(4.582)	(4.211)
Leverage	− 0.095*	− 0.100**	− 0.095
	(− 1.862)	(− 1.963)	(− 0.159)
mtb	0.002	0.003	− 0.007
	(0.524)	(0.150)	(− 0.768)
Cash	− 0.667***	− 0.660***	− 0.671***
	(− 6.871)	(− 6.816)	(− 4.012)
ppe	− 0.531***	− 0.627***	− 0.719***
	(− 3.391)	(− 4.034)	(− 3.348)
cf_na	0.012	0.009	0.011
	(1.142)	(0.845)	(0.840)
nwc_na	− 0.029	− 0.036*	− 0.357*
	(− 1.463)	(− 1.816)	(− 1.905)
Capex	0.208	0.059	− 0.127*
	(1.140)	(0.327)	(− 1.678)
rd	0.002	0.001	0.001
	(0.657)	(0.806)	(0.078)
log_at	− 0.029***	− 0.024***	− 0.095**
	(− 4.625)	(− 3.313)	(− 2.129)
inst_ratio	− 0.134	− 0.248	− 0.199
	(− 0.356)	(− 0.659)	(− 0.858)
lnage	0.271***	0.245***	0.287***
	(3.003)	(2.809)	(2.669)
_cons	0.972***	0.738***	− 0.936
	(6.837)	(5.646)	(− 0.556)
Fixed firm	Yes	Yes	No
Fixed month	No	Yes	Yes
Fixed industry	No	No	Yes
N	341,069	341,069	341,069
Adi_R2	0.318	0.320	0.314

This table presents estimation results of panel regressions where concus is regressed on tariff (a dummy variable) and 1-year lagged control variables including leverage, mtb, cash, ppe, cf_na, nwc_na, capex, rd, log_at, inst_ratio and lnage. Specifically, the tariff level of each year is compared to the previous year. If the growth rate exceeds 20%, the value of tariff for that year and subsequent years is assigned to be 1, and 0 for the previous years. Model (1) includes firm fixed effects, model (2) includes both firm and month fixed effects, and model (3) includes both month and industry fixed effects. The dependent variable and all control variables are defined in Appendix A. The t-statistics are in parentheses below parameter estimates, and they are based on robust standard errors clustered at year and industry levels. ***, **, and * indicate significance at 1%, 5% and 10% levels, respectively

Table 13 Effect of exogenous shock in tariff on sentiment beta

	(1) β^{ST4}	(2) β^{ST4}	(3) β^{ST4}
Tariff	− 0.029***	− 0.026***	− 0.023***
	(− 9.503)	(− 8.908)	(− 8.308)
Leverage	0.004**	0.005***	0.004***
	(2.329)	(2.688)	(3.300)
mtb	− 0.002***	− 0.013***	− 0.003***
	(− 3.711)	(− 3.415)	(− 3.564)
Cash	− 0.006*	− 0.008**	− 0.013***
	(− 1.693)	(− 2.249)	(− 5.948)
ppe	0.005	0.008***	− 0.016***
	(1.022)	(3.470)	(− 5.859)
cf_na	− 0.011***	− 0.012***	− 0.014
	(− 3.508)	(− 3.655)	(− 0.905)
nwc_na	0.009***	0.021	− 0.004
	(3.071)	(0.334)	(− 0.453)
Capex	− 0.002	0.006	− 0.028***
	(− 0.344)	(0.966)	(− 5.418)
rd	0.034***	0.021***	− 0.017***
	(6.233)	(6.122)	(− 7.730)
log_at	0.004***	0.002	0.031
	(5.689)	(0.076)	(0.281)
inst_ratio	− 0.052***	− 0.046***	− 0.017
	(− 3.988)	(− 3.524)	(− 1.457)
lnage	− 0.028***	− 0.032**	− 0.032**
	(− 4.224)	(− 2.007)	(− 2.012)
_cons	0.063***	0.062***	0.039***
	(9.625)	(9.423)	(6.001)
Fixed firm	Yes	Yes	No
Fixed month	No	Yes	Yes
Fixed industry	No	No	Yes
N	341,069	341,069	341,069
Adi_R^2	0.607	0.601	0.614

This table presents panel regressions estimation results where sentimental beta (four factors) is regressed on the dummy variable for an exogenous shock to tariff, and 1-year lagged control variables including leverage, mtb, cash, ppe, cf_na, nwc_na, capex, rd, log_at, inst_ratio and lnage. Specifically, the tariff level of each year is compared to the previous year. If the growth rate exceeds 20%, the value of tariff for that year and subsequent years is assigned to be 1, and 0 for the previous years. Model (1) includes firm fixed effects, model (2) includes both firm and month fixed effects, and model (3) includes both month and industry fixed effects. The dependent variable and all control variables are defined in Appendix A. The t-statistics are in parentheses below parameter estimates, and they are based on robust standard errors clustered at year and industry levels. ***, **, and * indicate significance at 1%, 5% and 10% levels, respectively

Table 14 Parallel trend test

	(1) β^{ST4}	(2) β^{ST4}	(3) β^{ST4}
Before5	0.005	0.002	0.004
	(0.317)	(0.077)	(0.191)
Before4	− 0.001	− 0.006	− 0.003
	(− 0.045)	(− 0.329)	(− 0.129)
Before3	0.007	0.005	0.011
	(0.398)	(0.324)	(0.622)
Before2	0.001	− 0.008	0.003
	(0.075)	(− 0.445)	(0.150)
Before1	− 0.004	− 0.001	− 0.001
	(− 0.212)	(− 0.011)	(− 0.028)
Treat × time	− 0.056***	− 0.054***	− 0.060***
	(− 4.471)	(− 4.855)	(− 4.363)
After1	− 0.068***	− 0.075***	− 0.073***
	(− 4.347)	(− 4.463)	(− 4.060)
After2	− 0.072***	− 0.080***	− 0.085***
	(− 4.271)	(− 4.272)	(− 3.749)
After3	− 0.067***	− 0.064***	− 0.062***
	(− 3.451)	(− 3.519)	(− 3.288)
After4	− 0.021**	− 0.023***	− 0.019**
	(− 2.232)	(− 2.767)	(− 2.251)
After5	0.018	0.025	0.027
	(0.934)	(1.208)	(1.491)
Leverage	− 0.024***	− 0.023***	0.004
	(− 4.189)	(− 4.058)	(1.155)
mtb	0.001	0.002	− 0.013***
	(1.083)	(1.155)	(− 3.620)
Cash	0.014*	0.015*	− 0.031***
	(1.731)	(1.913)	(− 6.165)
ppe	0.095***	0.094***	− 0.005
	(6.453)	(6.361)	(− 0.694)
cf_na	0.004***	0.003***	− 0.001
	(3.029)	(2.840)	(− 1.107)
nwc_na	− 0.011***	− 0.010***	0.002
	(− 3.173)	(− 2.968)	(0.900)
Capex	0.003	− 0.001	− 0.055***
	(0.184)	(− 0.076)	(− 3.607)
rd	0.001***	0.001***	0.001
	(4.216)	(4.436)	(0.908)
log_at	− 0.002	0.002	− 0.003***
	(− 1.606)	(1.321)	(− 6.897)
inst_ratio	− 0.139***	− 0.142***	− 0.159***
	(− 4.064)	(− 4.157)	(− 5.309)
lnage	− 0.029***	− 0.015***	− 0.016***
	(− 3.427.)	(− 3.414.)	(− 3.007)

Table 14 (continued)

	(1) β^{ST4}	(2) β^{ST4}	(3) β^{ST4}
_cons	0.012	0.019	0.012
	(0.751)	(1.040)	(0.653)
Fixed firm	Yes	Yes	No
Fixed month	No	Yes	Yes
Fixed industry	No	No	Yes
N	141,168	141,168	141,168
Adi_R^2	0.298	0.320	0.318

This table reports panel estimation results of parallel trend test where sentimental beta (four factors) is regressed on DID interaction terms and 1-year lagged control variables including leverage, mtb, cash, ppe, cf_na, nwc_na, capex, rd, log_at, inst_ratio and lnage. A DID term is the interaction of treatment and time. Both treatment and time are dummy variables. For example, *Before1*, is product of the year dummy for the year before the treatment (i.e., tariff jump) and the treatment dummy; similarly, *After1*, is product of the year dummy for the year after the treatment and the treatment dummy. We definite treatment as a substantial jump in tariff level. If the growth rate of the tariff level exceeds 20% in one industry, the treatment values of all firms in this industry are assigned to 1, or 0 otherwise; If the tariff increases more than 20%, the time dummies for that year and subsequent years are assigned to 1, and 0 for the previous years. Model (1) includes firm fixed effects, model (2) includes both firm and month fixed effects, and model (3) includes both month and industry fixed effects. The dependent variable and all control variables are defined in Appendix A. The t-statistics are in parentheses below parameter estimates, and they are based on robust standard errors clustered at year and industry levels. ***, **, and * indicate significance at 1%, 5% and 10% levels, respectively

Change regressions are another way to examine the causal relationship between customer concentration and stock sentiment beta. We adopt a research design similar to Chava and Purnanandam (2010). Essentially this methodology uses the companies before the tariff jump as the control group and is powerful in detecting causal relations. Since the parallel path test in Table 14 shows that the impact of tariff jumps lasts for 4 years, the change of all variables (with the exception of firm age) is calculated as the difference between the fourth year post tariff jump and the year of the tariff jump. The untabulated change results are statistically significant and directionally similar to those in Table 4, thus rendering supports for our causal proposition.

In summary, the results of an exogenous shock test and complete DiD regressions suggest that increased customer concentration due to higher tariffs at an industry level triggers a decrease in stock sentiment risk. Such findings support the view demonstrated in Table 4 that sentiment beta is influenced by the level of customer concentration.

4.4.3 Impact of financial crisis and globalization

Risk factors are often more impactful during a crisis. For example, Lou and Sadka (2011) document that stock liquidity risk is more prominent for firms during the 2008 financial crisis. To investigate the potential difference in the impact of customer concentration on

Table 15 Multivariate difference-in-difference regressions

	(1) β^{ST4}	(2) β^{ST4}	(3) β^{ST4}
Treatment	− 0.004**	− 0.007***	− 0.005**
	(− 2.325)	(− 2.652)	(− 2.438)
Time	0.032	− 0.017*	− 0.029**
	(0.429)	(− 1.924)	(− 2.289)
Treatment × time	− 0.020***	− 0.016***	− 0.012***
	(− 5.777)	(− 5.065)	(− 4.056)
Leverage	0.003	0.003	0.007***
	(1.182)	(1.410)	(3.842)
mtb	− 0.001***	− 0.001***	− 0.001***
	(− 2.924)	(− 3.838)	(− 3.948)
Cash	− 0.001	− 0.001	− 0.013***
	(− 0.196)	(− 0.054)	(− 4.134)
ppe	0.082***	0.091***	0.009*
	(8.611)	(9.651)	(1.790)
cf_na	0.009***	0.009**	0.010
	(2.647)	(2.562)	(1.487)
nwc_na	− 0.003	− 0.001	− 0.002**
	(− 0.464)	(− 0.877)	(− 2.184)
Capex	0.020*	0.028***	− 0.027***
	(1.868)	(2.660)	(− 2.782)
rd	0.001***	0.002***	0.001***
	(3.014)	(3.872)	(3.671)
log_at	− 0.003**	− 0.006***	− 0.002***
	(− 2.574)	(− 3.582)	(− 3.013)
inst_ratio	− 0.031	− 0.030	− 0.044**
	(− 1.515)	(− 1.465)	(− 2.388)
lnage	− 0.030***	− 0.032***	− 0.031***
	(− 3.079)	(− 3.262)	(− 3.011)
_cons	− 0.005	0.023***	− 0.039***
	(− 0.553)	(3.573)	(− 4.080)
Fixed firm	Yes	Yes	No
Fixed month	No	Yes	Yes
Fixed industry	No	No	Yes
N	341,069	341,069	341,069
Adi_R^2	0.391	0.307	0.400

This table presents multivariate DID estimation results of panel regressions where sentimental beta (four factors) is regressed on treatment, time, treatment × time and 1-year lagged control variables including leverage, mtb, cash, ppe, cf_na, nwc_na, capex, rd, log_at, inst_ratio and lnage. Both treatment and time are dummy variables. Specifically, we definite treatment as a substantial jump in tariff level. If the growth rate of the tariff level exceeds 20% in one industry, the treatment values of all firms in this industry are assigned to 1, or 0 otherwise; If the tariff increases more than 20%, the time dummies for that year and subsequent years are assigned to 1, and 0 for the previous years. Model (1) includes firm fixed effects, model (2) includes both firm and month fixed effects, and model (3) includes both month and industry fixed effects. The dependent variable and all control variables are defined in Appendix A. The t-statistics are in parentheses below parameter estimates, and they are based on robust standard errors clustered at year and industry levels. ***, **, and * indicate significance at 1%, 5% and 10% levels, respectively.

Table 16 Effect of customer concentration on sentiment beta: crisis versus non-crisis periods

Panel A: crisis period (2008–2009)			
	(1)	(2)	(3)
	β^{ST4}	β^{ST4}	β^{ST4}
Concus	− 0.006***	− 0.004***	− 0.002***
	(− 3.799)	(− 4.691)	(− 4.105)
Leverage	0.080***	0.083***	0.011***
	(3.028)	(3.438)	(4.391)
mtb	− 0.005***	− 0.005***	− 0.001***
	(− 3.577)	(− 3.886)	(− 3.037)
Cash	0.011	− 0.025**	− 0.033***
	(0.952)	(− 2.179)	(− 8.001)
ppe	0.089***	0.035	0.023***
	(3.808)	(1.525)	(4.337)
cf_na	− 0.009***	− 0.007***	− 0.008***
	(− 3.142)	(− 3.306)	(− 3.672)
nwc_na	− 0.008***	− 0.007***	0.001
	(− 2.652)	(− 2.594)	(0.405)
Capex	− 0.048***	0.022	− 0.040***
	(− 2.604)	(1.281)	(− 3.471)
rd	0.002***	0.002***	− 0.001
	(8.484)	(8.841)	(− 0.407)
log_at	0.015***	0.014***	− 0.005***
	(3.453)	(3.244)	(− 4.087)
inst_ratio	− 0.082**	− 0.077**	− 0.081***
	(− 2.569)	(− 2.425)	(− 3.186)
lnage	− 0.039***	− 0.035***	0.033
	(− 3.028)	(− 4.966)	(1.627)
_cons	− 0.020***	− 0.016***	− 0.017***
	(− 6.594)	(− 5.398)	(− 5.585)
Fixed firm	Yes	Yes	No
Fixed month	No	Yes	Yes
Fixed industry	No	No	Yes
N	39,627	39,627	39,627
Adi_R^2	0.440	0.432	0.494
Panel B: non-crisis period			
	(4)	(5)	(6)
	β^{ST4}	β^{ST4}	β^{ST4}
Concus	− 0.006***	− 0.006***	− 0.007***
	(− 3.028)	(− 4.056)	(− 3.684)
Leverage	0.003	0.003	0.003**
	(1.351)	(1.605)	(2.132)
mtb	− 0.000***	− 0.000***	− 0.000***
	(− 10.085)	(− 9.982)	(− 8.859)

Table 16 (continued)

Panel B: non-crisis period			
	(4)	(5)	(6)
	β^{ST4}	β^{ST4}	β^{ST4}
Cash	− 0.004	− 0.006***	− 0.012***
	(− 1.185)	(− 3.583)	(− 5.074)
ppe	0.005	0.009	− 0.021***
	(0.917)	(1.594)	(− 6.932)
cf_na	0.012	0.013	0.013
	(0.863)	(1.001)	(1.211)
nwc_na	− 0.001	− 0.002***	− 0.001
	(− 0.838)	(− 3.608)	(− 1.628)
Capex	− 0.011	− 0.001	− 0.027***
	(− 1.643)	(− 0.165)	(− 4.798)
rd	0.001***	0.001***	− 0.001***
	(6.782)	(6.438)	(− 8.113)
log_at	0.004***	− 0.001*	0.001***
	(4.556)	(− 1.917)	(3.341)
inst_ratio	− 0.054***	− 0.045***	− 0.015
	(− 3.849)	(− 3.203)	(− 1.167)
lnage	− 0.032***	− 0.031**	− 0.028**
	(− 5.134)	(− 2.090)	(− 2.218)
_cons	− 0.032***	− 0.035***	− 0.013*
	(− 5.066)	(− 5.559)	(− 1.796)
Fixed firm	Yes	Yes	No
Fixed month	No	Yes	Yes
Fixed industry	No	No	Yes
N	301,442	301,442	301,442
Adi_R^2	0.310	0.387	0.269

This table presents the panel regression estimation results where sentimental beta (four factors) is regressed on 1-year lagged customer concentration and control variables including leverage, mtb, cash, ppe, cf_na, nwc_na, capex, rd, log_at, inst_ratio and lnage, in two sub-periods, the financial crisis period (2008–2009) and the non-crisis period (the sample period excluding 2008 and 2009). Panel A represents the crisis period while Panel B represents the non-crisis period. Models (1) and (4) include firm fixed effects, models (2) and (5) include both firm and month fixed effects, and models (3) and (6) include both month and industry fixed effects. The dependent variable, independent variable and all control variables are defined in Appendix A. The t-statistics are in parentheses below parameter estimates, and they are based on robust standard errors clustered at year and industry levels. ***, **, and * indicate significance at 1%, 5% and 10% levels, respectively

stock sentiment risk during financial crisis, we divide our sample into two subsamples: 2008–2009 crisis period, and non-crisis periods (other years). Table 16 reports panel regression results in both subsamples, with various combinations of fixed firm, year, and industry effects being controlled. In both panels, the effect of customer concentration is always negative at 1% level, consistent with Table 4. These results confirm that the impact of customer concentration on sentiment risk is ubiquitous rather than crisis driven. The

existence of major customers always plays a meaningful role in reducing a firm's stock sentiment risk. However, on average, the magnitudes of estimated coefficients for *concus* during the crisis period are nearly tripled compared to those in the full sample (Table 4). The effect of customer concentration is indeed elevated during crisis, suggesting that concentrated customer bases have a more prominent anchoring effect that somewhat calms down investors during market catastrophe.

Technological changes such as developments in social media and ecommerce have broadened customers' procurement opportunities. We investigate whether the effect of customer concentration has changed since the 2008 global financial crisis by conducting regressions in two subsamples: 1997–2007 and 2008–2017. The untabulated results are similar to Table 4. The comparison between the two subsamples shows that the effect of customer concentration is somewhat stronger in more recent years. It appears that as the markets become more contestable facilitated by technologies, a concentrated customer base is more important in curbing stock sentiment risk.

4.4.4 Impact of foreign sales

Retention of international customers may be more challenging given their global orientation. In contrast, domestic customers may perceive a narrower set of buying opportunities. To investigate the potential impact of foreign sales, we incorporate the dummy variable of foreign sales, *foreign*, in our models. As shown in Table 17, *foreign* per se is insignificant, but its interaction term with *concus* is significantly positive. The overall effect of *concus* on sentiment beta is still negative, which is consistent with our main results, but this containing effect is somewhat weaker for companies with foreign sales. Because foreign customers naturally have a more global orientation, they are more likely to shift to other suppliers than domestic customers when opportunities arise. Therefore, high customer concentration with a foreign component tends to have a weaker impact on stock sentiment beta.

4.4.5 Impact of firm age

Customers in firms of different age may have different characteristics. To investigate the contextual impact of firm age on the relationship between customer concentration and sentiment beta, we conduct subsample regressions in two ways. All firms are sorted by age, and regressions are conducted in two contrasting subsamples: top half (or 1/3) versus bottom half (or 1/3). In untabulated results, we find that the impact of customer concentration on stock sentiment risk is greater and more significant for older firms. Large customers are more entrenched when present in older firms, because of loyalty, search costs, or inertia. In contrast, customers in younger firms are less determinate. Therefore, customer concentration has a more pronounced impact on sentiment beta in mature companies.

5 Conclusions

This study is motivated by the increasing evidence that stock sentiment risk is a novel grip for formulating investment strategies. However, no other contributing factors to sentiment risk have been identified beyond the observations that companies whose stocks have high sentiment risk tend to be young, small, and uncertain. We theorize that highly concentrated customer bases would help mitigate stock sentiment risk (Hypothesis 1). Furthermore, we

Table 17 Effect of foreign sales

	(1) β^{ST4}	(2) β^{ST4}	(3) β^{ST4}
Concus	− 0.003***	− 0.004***	− 0.003**
	(− 2.819)	(− 3.095)	(− 2.264)
Foreign	− 0.002	− 0.001	− 0.002
	(− 1.369)	(− 1.002)	(− 0.731)
Concus × foreign	0.001***	0.001***	0.002***
	(3.792)	(3.032)	(3.834)
Leverage	0.005*	0.006**	0.006***
	(1.704)	(2.208)	(2.980)
mtb	− 0.001***	− 0.001***	− 0.001***
	(− 3.277)	(− 3.748)	(− 3.489)
Cash	0.004	0.001	− 0.010***
	(0.912)	(0.144)	(− 4.080)
ppe	0.004	0.008	− 0.025***
	(0.610)	(1.228)	(− 7.779)
cf_na	− 0.011***	− 0.010***	− 0.011***
	(− 3.669)	(− 3.817)	(− 3.631)
nwc_na	− 0.001	− 0.002	− 0.002***
	(− 0.506)	(− 0.368)	(− 2.973)
Capex	− 0.004	0.010	− 0.043***
	(− 0.542)	(1.431)	(− 6.925)
rd	0.001***	0.001***	− 0.001***
	(6.970)	(6.824)	(− 7.066)
log_at	0.004***	− 0.002***	− 0.001***
	(4.145)	(− 3.060)	(− 4.734)
inst_ratio	− 0.083***	− 0.073***	− 0.067***
	(− 5.625)	(− 4.942)	(− 4.854)
lnage	− 0.037***	− 0.035***	− 0.041***
	(− 3.395)	(− 3.253)	(− 3.762)
_cons	0.003	0.010*	0.050***
	(0.351)	(1.817)	(5.976)
Fixed firm	Yes	Yes	No
Fixed month	No	Yes	Yes
Fixed industry	No	No	Yes
N	341,069	341,069	341,069
Adi_R^2	0.328	0.360	0.301

This table presents the panel estimation results where sentimental beta (four factors) is regressed on 1-year lagged customer concentration, foreign, concus × foreign and control variables including leverage, mtb, cash, ppe, cf_na, nwc_na, capex, rd, log_at, inst_ratio and lnage. Foreign is a dummy variable that takes the value of 1 when the firm reports foreign income or foreign income taxes, or 0 otherwise. Model (1) includes firm fixed effects, model (2) includes firm and month fixed effects, and model (3) includes month and industry fixed effects. The dependent variable, independent variable and all control variables are defined in Appendix A. The t-Statistics are in parentheses below parameter estimates, and they are based on clustered standard errors at year and industry levels. ***, **, and * indicate significance at 1%, 5%, 10% levels, respectively.

conjecture that this containing effect takes effect via three channels: improved performance and information quality, and more long-term oriented institutional investors (Hypothesis 2a-2c). Empirical analyses of a large sample from the U.S. over two decades provide strong support to all these hypotheses. Stock sentiment risk is negatively linked to the level of customer concentration at 1% level across different model specifications. Economically, a one-standard-deviation increase in customer concentration is associated with an 8.85% reduction in sentiment beta when firm and month fixed effects are controlled. As for the effect mechanisms, companies with higher customer concentration have lower cash flow volatility and stronger earnings, gross profitability, and sales growth. Companies with concentrated customer bases have unambiguous and apprehensible information, thus having lower audit related fees, and are less likely to have restatements. Further, such firms tend to have higher institutional ownership, especially by those with a long investment horizon. All these characteristics—strong financial and operating performance, quality and transparent information, and high institutional shareholding—make a company's stock price less susceptible to market sentiment.

To mitigate the endogeneity concern regarding the relationship between customer concentration and stock sentiment beta, we examine the effect of an exogenous shock. When import tariff jumps unexpectedly, the level of customer concentration rises for domestic companies as procurement from foreign suppliers becomes more expensive. By comparing the changes in stock sentiment risk induced by elevated import tariffs and those unaffected by a tariff jump in a staggered DiD model, we learn that stock sentiment risk indeed decreases after import tariff jumps that raise domestic customer concentration. Such findings strengthen our main thesis that sentiment beta is influenced by the level of customer concentration.

Our results are robust to alternative measures of sentiment risk, or additional controls of foreign sales or firm age. Although the mitigating effect of customer concentration on stock sentiment risk is ubiquitous, it is stronger during the 2008 financial crisis.

Our paper contributes to literature by providing new evidence on the positive effects of concentrated customer bases. This is of particular interest when it is linked to lower stock sentiment risk, which contributes to improved stock returns. Therefore, the findings offer investment professionals a novel perspective for formulating their investment strategies. This study would also interest corporate executives for managing stock sentiment risk by leveraging the newly identified contributing factor, customer concentration.

Several research questions are worth pursuing in future studies. For one, would the identities of major customers—governments or corporations—cause a distinct effect of customer concentration on suppliers' stock sentiment risk? Further, should data permit, it would be interesting to explore how major customers' geographical location, history with a supplier, or financial strength may impact the relationship between customer concentration and stock sentiment risk.

Appendix A: description of main variables

Variable name	Description
β^{ST4}	Sentiment beta of stocks obtained from Carhart four-factor model, following Baker and Wurgler (2006)
β^{ST1}	Sentiment beta of stocks obtained from market single-factor model (CAPM)

Variable name	Description
β^{ST3}	Sentiment beta of stocks obtained from Fama–French there-factor model
β^{CCI}	Sentiment beta of stocks based on Consumer Confidence Index (CCI), following Keiber and Samyschew (2019)
Concus	HHI ratio of Customer Concentration as defined in Patatoukas (2012)
InstTO	Institutional investor turnover as proposed by Gaspar et al. (2005)
Audit_fee	Audit fees in $US thousands from Audit Analytics
Total_fee	The sum of audit and audit-related fees in $US thousands from Audit Analytics[a]
Restatement	An indicator variable for restatements of previously audited financial statements[b]
Leverage	The ratio of total debt divided by total asset
mtb	The ratio of the market value of equity to the book value of equity measured at the beginning of the fiscal year
ppe	The ratio of net PP&E to the total asset
Cash	The ratio of cash and all securities readily transferable to cash as listed in the Current Asset section to the total assets
cf_na	The ratio of the Operating Income Before Depreciation (OIBDP) to non-current assets
nwc_na	The working capital minus cash and short-term investments, divided by the non-current assets
Capex	The ratio of capital expenditures to non-current assets
rd	Research and development expenditure divided by total assets
Log_at	The natural log of total assets
inst_ratio	Institutional ownership, defined as the equity held by institutional investors at the end of the last quarter in fiscal year t
lnage	The natural log of the number of years since IPO
roa	The return-on-assets calculated as the ratio of operating income after depreciation divided by total assets
sg	The annual percentage of growth rate in sales
Loss	An indicator variable for negative operating income after depreciation[c]
Debtmaturity	The ratio of short-term debt to total debt
Foreign	An indicator variable for foreign income or foreign income taxes[d]
Δcsale	The year-over-year growth in cash sales, where cash sales is calculated as sales minus the change in accounts receivables
Δroa	The year-over-year change in return-on-assets
Soft	The ratio of total assets minus net PP&E minus cash and cash equivalents divided by total assets
gp	The firm's gross profitability as revenues minus cost of goods sold, scaled by total assets following Novy-Marx (2013)
Ebitda	Earnings before interest, tax, depreciation and amortization scaled by the market value of equity
Δsale	Sales growth from year t to year t + 1
cfvol	The annualized standard deviation of cash flow in 5 years

[a]Audit-related fees are assurance and related services that traditionally are performed by the independent accountant and include employee benefit plan audits, due diligence related to mergers and acquisitions, accounting consultations and audits in connection with acquisitions, internal control reviews, attest services that are not required by statute or regulation, and consultation concerning financial accounting and reporting standards

[b]Indicator variable = 1 for financial statement restatements not attributed to clericala errors identified by Audit Analytics, 0 otherwise

[c]Indicator variable = 1 if operating income after depreciation is negative, 0 otherwise.

[d]Indicator variable = 1 if the firm reports foreign income or foreign income taxes, 0 otherwise.

Acknowledgements We thank Ramesh Rao for helpful comments.

Funding Wang and Huang acknowledge financial support from the National Natural Science Foundation of China (Grant Numbers 71571038 and 71971048), the Fundamental Research Funds for Central Universities in China (Grant Number N2006010) and LiaoNing Revitalization Talents Program in China (Grant Number XLYC1907015).

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

An H, Zhang T (2013) Stock price synchronicity, crash risk, and institutional investors. J Corp Financ 21:1–15. https://doi.org/10.1016/j.jcorpfin.2013.01.001

Antoniou C, Doukas JA, Subrahmanyam A (2016) Investor sentiment, beta, and the cost of equity capital. Manag Sci 62(2):347–367. https://doi.org/10.1287/mnsc.2014.2101

Baker M, Wurgler J (2006) Investor sentiment and the cross-section of stock returns. J Financ 61:1645–1680. https://doi.org/10.1111/j.1540-6261.2006.00885.x

Baker M, Wurgler J (2007) Investor sentiment in the stock market. J Econ Perspect 2:129–215. https://doi.org/10.1257/jep.21.2.129

Balakrishnan R, Linsmeier TJ, Venkatachalam M (1996) Financial benefits from JIT adoption: effects of customer concentration and cost structure. Account Rev 71(2):183–205

Banz RW, Breen WJ (1986) Sample-dependent results using accounting and market data: some evidence. J Financ 41(4):779–793. https://doi.org/10.1111/j.1540-6261.1986.tb04548.x

Berger D, Turtle HJ (2012) Cross-sectional performance and investor sentiment in a multiple risk factor model. J Bank Financ 36(4):1107–1121. https://doi.org/10.1016/j.jbankfin.2011.11.001

Bertrand M, Mullainathan S (2003) Enjoying the quiet life? Corporate governance and managerial preferences. J Polit Econ 111(5):1043–1075. https://doi.org/10.1086/376950

Bhattacharyya S, Nain A (2011) Horizontal acquisition and buying power: a product market analysis. J Financ Econ 99(2):97–115. https://doi.org/10.1016/j.jfineco.2010.08.007

Boscaljon B, Feng H, Jia Y, Sun Q (2021) Government customers, institutional investment horizons, and liquidity risk. Rev Quant Financ Acc 56:273–296. https://doi.org/10.1007/s11156-020-00894-w

Brown GW, Cliff MT (2004) Investor sentiment and the near-term stock market. J Empir Financ 11(1):1–27. https://doi.org/10.1016/j.jempfin.2002.12.001

Bushee BJ (2004) Identifying and attracting the "right" investors: evidence on the behavior of institutional investors. J Appl Corp Financ 16(4):28–35. https://doi.org/10.1111/j.1745-6622.2004.00005.x

Bushee BJ, Goodman TH, Sunder SV (2019) Financial reporting quality, investment horizon, and institutional investor trading strategies. Account Rev 94(3):87–112. https://doi.org/10.2308/accr-52202

Bushee BJ, Noe CF (2000) Corporate disclosure practices, institutional investors, and stock return volatility. J Account Res. https://doi.org/10.2307/2672914

Campbell JY, Polk C, Vuolteenaho T (2010) Growth or glamour? Fundamentals and systematic risk in stock returns. Rev Financ Stud 23(1):305–344. https://doi.org/10.1093/rfs/hhp029

Cao C, Petrasek L (2014) Liquidity risk and institutional ownership. J Financ Mark 21:76–97. https://doi.org/10.1016/j.finmar.2014.05.001

Carhart MM (1997) On persistence in mutual fund performance. J Financ 52(1):57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

Chava S, Purnanandam A (2010) CEOs versus CFOs: incentives and corporate policies. J Financ Econ 97(2):263–278. https://doi.org/10.1016/j.jfineco.2010.03.018

Chen J, Hong H, Stein J (2001) Forecasting crashes: trading volume, past returns, and conditional skewness in stock prices. J Financ Econ 61(2):345–381. https://doi.org/10.1016/S0304-405X(01)00066-6

Chen J, Tong J, Wang W, Zhang F (2019) The economic consequences of labor unionization: evidence from stock price crash risk. J Bus Ethics 157(3):775–796. https://doi.org/10.1007/s10551-017-3686-0

Choi N, Skiba H (2015) Institutional herding in international markets. J Bank Financ 55:246–259. https://doi.org/10.1016/j.jbankfin.2015.02.002
Cohen DA, Li B (2020) Customer-base concentration, investment, and profitability: the US government as a major customer. Account Rev 95(1):101–131. https://doi.org/10.2308/accr-52490
Costello AM (2013) Mitigating incentive conflicts in inter-firm relationships: evidence from long-term supply contracts. J Account Econ 56(1):19–39. https://doi.org/10.1016/j.jacceco.2013.02.001
Crane AD, Koch A, Michenaud S (2019) Institutional investor cliques and governance. J Financ Econ 133(1):175–197. https://doi.org/10.1016/j.jfineco.2018.11.012
DeFond M, Zhang J (2014) A review of archival auditing research. J Account Econ 58(2/3):275–326. https://doi.org/10.1016/j.jacceco.2014.09.002
Delmas M, Tokat Y (2005) Deregulation, governance structures, and efficiency: the US electric utility sector. Strateg Manag J 26(5):441–460. https://doi.org/10.1002/smj.456
Deng X, Gao L, Kim JB (2020) Short-sale constraints and stock price crash risk: causal evidence from a natural experiment. J Corp Financ 60:101498. https://doi.org/10.1016/j.jcorpfin.2019.101498
Dhaliwal D, Judd JS, Serfling M, Shaikh S (2016) Customer concentration risk and the cost of equity capital. J Account Econ 61(1):23–48. https://doi.org/10.1016/j.jacceco.2015.03.005
Falkenstein EG (1996) Preferences for stock characteristics as revealed by mutual fund portfolio holdings. J Financ 51(1):111–135. https://doi.org/10.1111/j.1540-6261.1996.tb05204.x
Fama EF, French KR (1992) The cross-section of expected stock returns. J Financ 47(2):427–465. https://doi.org/10.2307/2329112
Fama EF, French KR (1993) Common risk factors in the returns on stocks and bonds. J Financ Econ 33(1):3–56. https://doi.org/10.1016/0304-405X(93)90023-5
Fama EF, French KR (1997) Industry costs of equity. J Financ Econ 43(2):153–193. https://doi.org/10.1016/S0304-405X(96)00896-3
Fee E, Thomas S (2004) Sources of gains in horizontal mergers: evidence from customer, supplier, and rival firms. J Financ Econ 74(1):423–460. https://doi.org/10.1016/j.jfineco.2003.10.002
Feenstra RC, Romalis J (2014) International prices and endogenous quality. Q J Econ 129(2):477–527. https://doi.org/10.1093/qje/qju001
Fresard L (2010) Financial strength and product market behavior: the real effects of corporate cash holdings. J Financ 65(3):1097–1122. https://doi.org/10.1111/j.1540-6261.2010.01562.x
Gaspar JM, Massa M, Matos P (2005) Shareholder investment horizons and the market for corporate control. J Financ Econ 76(1):135–165. https://doi.org/10.1016/j.jfineco.2004.10.002
Gaunt C (2004) Size and book to market effects and the Fama French three factor asset pricing model: evidence from the Australian stock market. Account Financ 44(1):27–44. https://doi.org/10.1111/j.1467-629x.2004.00100.x
Gow ID, Ormazabal G, Taylor DJ (2010) Correcting for cross-sectional and time-series dependence in accounting research. Account Rev 85(2):483–512. https://doi.org/10.2308/accr.2010.85.2.483
Hansen BE (1999) Threshold effects in non-dynamic panels: estimation, testing, and inference. J Econom 93(2):345–368. https://doi.org/10.1016/S0304-4076(99)00025-1
Hansen BE (2000) Sample splitting and threshold estimation. Econometrica 68(3):575–603. https://doi.org/10.1111/1468-0262.00124
Heckman JJ (1979) Sample selection bias as a specification error. Econometrica 47(1):153–161. https://doi.org/10.2307/1912352
Hoberg G, Phillips G (2010) Real and financial industry booms and busts. J Financ 65(1):45–86. https://doi.org/10.1111/j.1540-6261.2009.01523.x
Hoberg G, Phillips G (2016) Text-based network industries and endogenous product differentiation. J Polit Econ 124(5):1423–1465. https://doi.org/10.1086/688176
Hoberg G, Phillips G, Prabhala N (2014) Product market threats, payouts, and financial flexibility. J Financ 69(1):293–324. https://doi.org/10.1111/jofi.12050
Huang H, Lobo G, Wang C, Xie H (2016) Customer concentration and corporate tax avoidance. J Bank Financ 72:184–200. https://doi.org/10.1016/j.jbankfin.2016.07.018
Hui KW, Liang C, Yeung PE (2019) The effect of major customer concentration on firm profitability: competitive or collaborative? Rev Account Stud 24(1):189–229. https://doi.org/10.1007/s11142-018-9469-8
Hutton AP, Marcus AJ, Tehranian H (2009) Opaque financial reports, R2, and crash risk. J Financ Econ 94(1):67–86. https://doi.org/10.1016/j.jfineco.2008.10.003
Imbens GW, Wooldridge JM (2009) Recent developments in the econometrics of program evaluation. J Econ Lit 47(1):5–86. https://doi.org/10.1257/jel.47.1.5
Irvine PJ, Park SS, Yıldızhan Ç (2016) Customer-base concentration, profitability, and the relationship life cycle. Account Rev 91(3):883–906. https://doi.org/10.2308/accr-51246

Kalwani MU, Narayandas N (1995) Long-term manufacturer-supplier relationships: Do they pay off for supplier firms? J Mark 59(1):1–16. https://doi.org/10.1177/002224299505900101
Keiber KL, Samyschew H (2019) The pricing of sentiment risk in European stock markets. Eur J Financ 25(3):279–302. https://doi.org/10.1080/1351847X.2018.1521340
Kelly S, Moody D (2005) Wholesale electric restructuring: was 2004 the "tipping point"? Electr J 18(2):11–18. https://doi.org/10.1016/j.tej.2005.02.001
Kim Y, Li H, Li S (2014) Corporate social responsibility and stock price crash risk. J Bank Financ 43(2):1–13. https://doi.org/10.1016/j.jbankfin.2014.02.013
Kim J-B, Li Y, Zhang L (2011) CFOs versus CEOs: equity incentives and crashes. J Financ Econ 101(2):713–730. https://doi.org/10.1016/j.jfineco.2011.03.013
Kinney WR, Palmrose Z-V, Scholz S (2004) Auditor independence, no-audit services, and restatements: Was the U.S. government right? J Account Res 42(3):561–588. https://doi.org/10.1111/j.1475-679X.2004.t01-1-00141.x
Kinney MR, Wempe WF (2002) Further evidence on the extent and origins of JIT's profitability effects. Account Rev 77(1):203–225. https://doi.org/10.2308/accr.2002.77.1.203
Krishnan GV, Patatoukas PN, Wang AY (2019) Customer-base concentration: implications for audit pricing and quality. J Manag Account Res 31(1):129–152. https://doi.org/10.2308/jmar-52040
Kumar N (1996) The power of trust in manufacturer–retailer relationships. Harv Bus Rev 74(November):92–106
Lobo GJ, Zhao Y (2013) Relation between audit effort and financial report misstatements: evidence from quarterly and annual restatements. Account Rev 88(4):1385–1412. https://doi.org/10.2308/accr-50440
Lou X, Sadka R (2011) Liquidity level or liquidity risk? Evidence from the financial crisis. Financ Anal J 67(3):51–62. https://doi.org/10.2469/faj.v67.n3.5
Ma X, Wang W, Wu J, Zhang W (2020) Corporate customer concentration and stock price crash risk. J Bank Financ 119:105903. https://doi.org/10.1016/j.jbankfin.2020.105903
Massa M, Yadav V (2015) Investor sentiment and mutual fund strategies. J Financ Quant Anal 50(4):699–727. https://doi.org/10.1017/S0022109015000253
Miller MH, Muthuswamy J, Whaley RE (1994) Mean reversion of Standard and Poor's 500 index basis changes: arbitrage-induced or statistical illusion? J Financ 49(2):479–513. https://doi.org/10.1111/j.1540-6261.1994.tb05149.x
Ng J (2011) The effect of information quality on liquidity risk. J Account Econ 52(2–3):126–143. https://doi.org/10.1016/j.jacceco.2011.03.004
Novy-Marx R (2013) The other side of value: The gross profitability premium. J Financ Econ 108(1), 1–28.
Pastor L, Stambaugh RF (2003) Liquidity risk and expected stock returns. J Polit Econ 111(3):642–685. https://doi.org/10.1086/374184
Patatoukas PN (2012) Customer-base concentration: implications for firm performance and capital markets. Account Rev 87(2):363–392. https://doi.org/10.2308/accr-10198
Petersen MA (2009) Estimating standard errors in finance panel data sets: comparing approaches. Rev Financ Stud 22(1):435–480. https://doi.org/10.1093/rfs/hhn053
Pirinsky C, Wang Q (2006) Does corporate headquarters location matter for stock returns? J Financ 61(4):1991–2015. https://doi.org/10.1111/j.1540-6261.2006.00895.x
Qureshi F, Kutan AM, Ismail I, Gee CS (2017) Mutual funds and stock market volatility: an empirical analysis of Asian emerging markets. Emerg Mark Rev 31:176–192. https://doi.org/10.1016/j.ememar.2017.05.005
Raman K, Shahrur H (2008) Relationship-specific investments and earnings management: evidence on corporate suppliers and customers. Account Rev 83(4):1041–1081. https://doi.org/10.2308/accr.2008.83.4.1041
Serfling M (2016) Firing costs and capital structure decisions. J Financ 71(5):2239–2286. https://doi.org/10.1111/jofi.12403
Shumway T (1997) The delisting bias in CRSP data. J Financ 52(1):327–340. https://doi.org/10.1111/j.1540-6261.1997.tb03818.x
Shumway T, Warther VA (1999) The delisting bias in CRSP's Nasdaq data and its implications for the size effect. J Financ 54(6):2361–2379. https://doi.org/10.1111/0022-1082.00192
Yang C, Hu X (2021) Individual stock sentiment beta and stock returns. N Am J Econ Financ 55:101306. https://doi.org/10.1016/j.najef.2020.101306
Zheng Y, Osmer E, Zhang R (2018) Sentiment hedging: how hedge funds adjust their exposure to market sentiment. J Bank Financ 88:147–160. https://doi.org/10.1016/j.jbankfin.2017.11.016