Investor Sentiment and Stock Return Comovement: the Role of Information and Innovation

Haohan Ren[1]

Current Version: November 2017

[1] Haohan Ren is from the Department of Finance, CUHK Business School, The Chinese University of Hong Kong, Shatin, Hong Kong. Author's contact information: haohan@baf.cuhk.edu.hk, (852) 39431824. I am grateful for valuable comments from Hendrik Bessembinder, Jay Cao, Ling Cen, Kalok Chan, Darwin Choi, Si Cheng, Zhi Da, Zhenyu Gao, Allaudeen Hameed, Wenxi Jiang, Yizhou Xiao, Bohui Zhang, Hua Zhang, and Wenrui Zhang.

Investor Sentiment and Stock Return Comovement: the Role of Information and Innovation

Abstract

I find that stock return comovement following positive investor sentiment is lower than that following negative investor sentiment. Further analysis suggests that this difference is associated with higher firm-specific information production and innovation output following positive sentiment. Specifically, following positive investor sentiment, the media and financial analysts produce more firm-specific information, short sellers and institutional investors conduct more informed trading, and firms produce more innovations. Various cross-sectional tests confirm that the difference in information production and innovation output indeed contributes to the difference in comovement between positive sentiment periods and negative sentiment periods. Overall, my results shed light on the nontrivial role of information producers and innovation generators in shaping the relation between sentiment and comovement.

Keywords: Investor sentiment; Information production; Innovation; Stock return comovement
JEL Code: G12; G14

1. Introduction

The traditional theory based on frictionless markets argues that stock return comovement should be determined by the comovement of the stocks' fundamental values. However, empirical evidence suggests that the correlation of firms' fundamentals only partially explains stock return comovement and that systematic noise trading is a plausible alternative explanation (Morck, Yeung, and Yu 2000). As an important driving force of systematic noise trading, investor sentiment is thus linked to stock return comovement (Barberis, Shleifer, and Wurgler, 2005; Kumar and Lee, 2006; Morck, Yeung, and Yu 2013). Previous literature on how sentiment affects stock return comovement naturally focuses on the role of noise traders' irrational behaviors. However, other market participants (e.g. informed arbitrageurs and firm CEOs) also respond actively to investor sentiment (De Long et al., 1990; Baker and Wurgler, 2000, 2002) and their effects on stock return comovement can be non-trivial. To draw a complete picture of the relation between investor sentiment and stock return comovement, it is important to consider the role of other market participants. In this paper, we directly examine whether and how other market participants' responses to investor sentiment affect stock return comovement.

Motivated by Veldkamp (2006a, 2006b) and Morck, Yeung, and Yu (2013) (MYY), we focus on three types of market participants: information suppliers, informed arbitrageurs and creative firms. Specifically, the information suppliers we consider are the media and analysts and the informed arbitrageurs are short sellers and institutional investors. We hypothesize that following high sentiment periods, 1) information suppliers would produce more information (Hypothesis 1); 2) informed arbitrageurs would conduct more informed arbitrage that incorporates more information into stock prices (Hypothesis 2); and 3) creative firms would produce more innovations that elevate fundamental volatility (Hypothesis 3). Both higher information production and higher firm-specific fundamental volatility are found to be associated with lower stock return comovement (MYY, 2000; Campbell et al., 2001; Durnev et al., 2003; Jin and Myers, 2006; Chun et al., 2008; Hutton, Marcus, and Tehranian, 2009; Irvine and Pontiff 2009; Dang, Moshirian, and Zhang, 2015). We thus predict a negative relation between lagged investor sentiment and stock return comovement. Our three hypotheses on how the three types of market participants react to investor sentiment are crucial to our analysis and are built on the following theoretical and empirical findings of the previous literature.

First, Veldkamp (2006a, 2006b) show that demand for information about assets increases with the value of the assets. According to her model, the value of information about assets depends on the total payoff variance that is determined by both risk and the value of assets at risk. Intuitively, for a given amount of risk, investors want to know more about high-valued assets because they have more value at risk in those assets. Since stocks are overvalued following high sentiment periods and undervalued following low sentiment periods, we conjecture that following high sentiment periods, more information will be produced because of increased demand for information, which leads to lower stock return comovement.

Second, MYY (2013) argue that when investors are overly pessimistic, stocks are underpriced and the cost of equity capital is likely to be high, which deters arbitrage by informed investors. However, when investors are overly optimistic, stocks are overpriced, and the cost of equity capital is likely to be low. With a lower cost of equity capital, informed arbitrageurs can conduct more informed arbitrage. Because informed arbitrage is an important channel through which firm-specific information is incorporated into stock prices, stock return comovement following high sentiment periods is expected to be lower than that following low sentiment periods.

Third, MYY (2013) also conjecture that the lower cost of equity capital following high sentiment periods not only increases informed arbitrage but also enables creative firms to produce more innovations that elevate firm-specific fundamental volatility. The elevated fundamental volatility eventually reduces stock return comovement. Consistent with the other two market participants, creative firms' response to investor sentiment also predicts a negative relation between lagged sentiment and stock return comovement.

We first test our prediction on the relation between lagged investor sentiment and stock return comovement. We compute the annual investor sentiment index as the average of the monthly Baker and Wurgler (2006) sentiment index over the year. We follow MYY (2000) to compute the R^2s of the regressions of weekly stock returns on CRSP value-weighted returns and use the logistic-transformed R^2s as our stock return comovement measure. We estimate both stock-level regressions and market-level regressions. We find that R-square is negatively related to lagged investor sentiment, suggesting that stock return comovement following high sentiment periods is significantly lower than that following low sentiment periods. The results are also economically significant. A one-standard deviation increase in investor sentiment is associated with a decrease of 18.8% of the standard deviation in the logistic-transformed R-squares. Our results are robust to an alternative measure of investor sentiment—a binary measure of the sentiment index following Stambaugh, Yuan, and Yu (2012) and an alternative measure of stock return comovement proposed by Piotroski and Roulstone (2004). We also find that the negative relation between lagged investor sentiment and stock return comovement holds in five other major stock markets (United Kingdom, Germany, France, Canada and Japan). To ease the concern that our results are merely capturing the countercyclical pattern of stock return comovement documented by Brockman, Liebenberg, and Schutte (2010) (BLS), we add the sentiment index to the main regressions of BLS, and we find that lagged investor sentiment is still negatively and significantly related to stock return comovement.

We next examine whether our findings are associated with the above three hypotheses derived from Veldkamp (2006a, 2006b) and MYY (2013). To investigate the first hypothesis that information suppliers produce more information following high sentiment periods, we follow Dang, Moshirian, and Zhang (2015) to construct a news commonality measure and a news coverage measure. A lower level of news commonality is associated with higher firm-specific news production and a higher level of news coverage is associated with higher total news production. We find a negative relation between lagged investor sentiment and news commonality and a positive relation between lagged investor sentiment and news coverage, suggesting that more firm-specific news and total news are produced following high sentiment periods. In terms of economic significance, a one-standard-deviation increase in sentiment index is associated with a decrease of 3.1% of the standard deviation in news commonality.

Having documented an increase in information production following high sentiment periods, we then provide evidence that information suppliers do contribute to the effect of investor sentiment on stock return comovement. The media and analysts are traditionally regarded as important information suppliers (Piotroski and Roulstone, 2004; Veldkamp, 2006a; Chan and Hameed, 2006; Hameed et al., 2015). We find that the negative relation between lagged investor sentiment and stock return comovement is more pronounced in the stocks that have higher media coverage and higher analyst coverage, which is consistent with our hypothesis that information suppliers' reactions to sentiment can have a significant impact on stock return comovement.

The second and third hypotheses rely on one intuitive assumption: high investor sentiment leads to low cost of equity capital. We first analyze this assumption before we examine the two hypotheses. Specifically, we follow Hail and Leuz (2006) to compute five measures of cost of equity capital and investigate their relations with our investor sentiment measure. We find that all five measures of cost of equity capital are negatively correlated with investor sentiment at the 1% level. Economically, a one-standard deviation increase in sentiment index is associated with a decrease of 6.4% of the standard deviation in the cost of capital. Our findings confirm that high investor sentiment is associated with a lower cost of equity capital.

We then examine our hypothesis about the informed arbitrageurs' responses to investor sentiment. We consider short sellers and institutional investors as informed arbitrageurs. We find that both stocks' short interest and institutional investors' trading intensity increase after investor sentiment increases, indicating that higher investor sentiment is associated with more informed arbitrage. In addition, the negative relation between stock return comovement and lagged investor sentiment is more pronounced in stocks with lower cost of short selling and stocks with higher institutional ownership. These stocks are exactly the ones that are more likely to be traded by short sellers and institutional investors. Overall, our findings suggest that informed arbitrageurs' reactions to investor sentiment have significant explanatory power for our main findings.

We finally investigate the third hypothesis. If the lower cost of equity capital induced by high sentiment enables creative firms to produce more innovations that reduce stock return comovement, we should find a) a positive relation between investor sentiment and innovation output and b) a more significant reduction in stock return comovement for stocks whose values are sensitive to innovations.

We obtain the innovation data from Kogan et al. (2017). We use the number of patents, the adjusted number of citations and the value of patents as our innovation output measures and find that the level of innovation output is positively associated with investor sentiment. An increase of one standard deviation in the sentiment index coincides with an increase of 2.7% of the standard deviation in the value of patents scaled by book assets. We also find that the effect is more pronounced in financially constrained firms. A recent paper by Dang and Xu (2016) finds similar results. Finally, we follow Hsu, Tian, and Xu (2014) to compute a measure of high-tech intensiveness and use this measure to identify the stocks whose values are sensitive to innovations. We find that the negative relation between lagged investor sentiment and stock return comovement is more pronounced in high-tech stocks, suggesting that creative innovations are an important channel through which investor sentiment affects stock return comovement.

Our paper contributes to two strands of the literature. We first contribute to the stock return comovement literature by providing additional empirical evidence on the relation between investor sentiment and stock return comovement. Prior studies mainly focus on the role of sentimental noise traders in generating stock return comovement but neglect the fact that other market participant also actively react to investor sentiment. For example, Barberis, Shleifer, and Wurgler (2005) propose “friction-based” and “sentiment-based” theories to explain excessive comovement. Kumar and Lee (2006) find that correlated retail trading, which is likely to be driven by sentiment, is positively related to stock return comovement for stocks with high retail concentration. In these studies, investor sentiment is linked to stock return comovement through the noise traders’ irrational behaviors. Our results and the well-documented explanations based on noise traders’ irrationality are not mutually exclusive. We offer an additional channel through which investor sentiment affects stock return comovement. We also contribute to the literature on the effects of bubbles on financial markets (e.g., Ventura 2012; Martin and Ventura, 2012). Our finding that high investor sentiment lowers the cost of equity capital for creative innovators and thus leads to creative innovations is consistent with the theoretical models in which bubbles reduce inefficient investment and increase efficient ones (Martin and Ventura 2012).

2. Data and sample construction

We describe how we construct our key variables in this section. These key variables include investor sentiment index, stock return comovement, implied cost of equity capital and innovation output. We obtain investor sentiment index from Baker and Wurgler (2006) and we construct stock return comovement using stock returns from CRSP. We exclude stocks whose prices are below 2 USD. We follow Hail and Leuz (2006) to use financial variables from Compustat and analyst forecast data from I\B\E\S to compute our implied cost of equity measures. Our innovation output data are from Kogan et al. (2017). The sample period of our main analysis is between 1981 and 2014 because the earnings forecasts of two-periods to five-periods ahead and the long-term earnings growth forecasts used to compute implied cost of equity capital are available from 1981.

2.1 Investor sentiment index

Our investor sentiment measure is from Baker and Wurgler (2006). Their sentiment index is based on five potential sentiment proxies including the dividend premium, the closed-end fund discount, the number and first-day returns on IPOs, and the equity share in new issues. They provide monthly sentiment index from July 1965 to December 2014. We compute the average of monthly sentiment index of each year as our annual sentiment measure ($Sentiment_t$). We also follow the literature (e.g. Stambaugh, Yu, and Yuan, 2012) to compute the sample median of the annual sentiment index and create a dummy variable, $Senti\_high_t$, which equals 1 if the annual sentiment index of year t is above the sample median and 0 otherwise. We use $Sentiment_t$ in our main analyses and $Senti\_high_t$ in robustness tests.

2.2 Stock return comovement

We construct two stock return comovement measures. We first follow MYY (2000) to regress weekly stock return on CRSP value-weighted market return for each stock in each year. Specifically, we estimate the following regression for each stock in each year

$$R_{i,t}=\alpha+\beta*MktRet_{i,t}+\varepsilon_{i,t} \qquad (1)$$

where $R_{i,t}$ is the weekly (Wednesday to Wednesday) return of firm i in week t and $MktRet_{i,t}$ is the CRSP value-weighted weekly return in week t. We include only common stocks (share code of 10 or 11) and we winsorize weekly returns at 99%. We require a stock to have at least 30 weekly returns within a year and we drop stocks with prices lower than 2. We then apply a logistic transformations to the R-squares of Eq.(1)

$$Ln\_Rsq_{i,T}=ln(R^2_{i,T}/(1-R^2_{i,T})) \qquad (2)$$

where $R^2_{i,T}$ is the R-square of Eq.(1) for stock i in year T and $Ln\_Rsq_{i,T}$ is our return comovement measure based on MYY (2000). For robustness check, we also follow Piotroski and Roulstone (2004) to construct another stock return comovement measure. We add lagged market return and industry return to Eq.(1) to estimate the R-squares

$$R_{i,t}=\alpha+\beta_1*MktRet_{i,t}+\beta_2*MktRet_{i,t-1}+\beta_3*IndRet_{i,t}+\beta_4*IndRet_{i,t-1}+\varepsilon_{i,t} \qquad (3)$$

where $IndRet_{i,t}$ is the value-weighted return of the industry which firm i belongs to in week t. Our industry classification is based on two-digit SIC. We also exclude the return of stock i when calculating $IndRet_{i,t}$. Similarly, our second measure of stock return comovement is defined as

$$Ln\_Rsq\_Ind_{i,T}=ln(Ind\_R^2_{i,T}/(1-Ind\_R^2_{i,T})) \qquad (4)$$

where $Ind\_R^2_{i,T}$ is the R-square of Eq.(3) for stock i in year T and $Ln\_Rsq\_Ind_{i,T}$ is our return comovement measure based on Piotroski and Roulstone (2004).

2.3 Implied cost of equity capital

We follow Hail and Leuz (2006) to use the average of four measures of ICOC proposed by the prior literature as our main ICOC measure. Specifically, the four measures are *r_GLS* from Gebhardt, Lee, and Swaminathan (2001), *r_CT* from Claus and Thomas (2001), *r_OJN* from Ohlson and Juettner-Nauroth (2005) and *r_PEG* from Easton (2004), respectively. Four different models are employed to compute these four measures of ICOC. Each measure is essentially the internal rate of return that equates the current stock price to the present value of expected future residual incomes or abnormal earnings. Since all of these measures are widely used in prior literature (e.g. Hail and Leuz (2006); Chen et al. (2010); Cao et al. (2015); Cao et al. (2017)) and it is not clear which measure is a best proxy for ICOC, we also report results using each of the four measures. We strictly follow Hail and Leuz (2006) to compute these measures and please refer to Hail and Leuz (2006) and Appendix C for a detailed description of the construction procedures.

2.4 Innovation output

We obtain patent data from Kogan et al. (2017). Their patent data are from Google Patents and they also complement the citations extracted from Google data with the hand-collected reference data of Nicholas (2008). Please refer to the Online Appendix of Kogan et al. (2017) for a detailed description of the data. Their patent database includes 1,928,123 patents that can be matched to firms in CRSP from 1926 to 2010 and 27% of these patents are not included in the NBER Patent data.

We merge the patent data with CRSP/Compustat merged data and we adopt the data requirements in Kogan et al. (2017). Specifically, we exclude firms with missing values for book assets and SIC codes. We also remove firms in industries that have no patents during our sample period. Because there is a significant drop in numbers of patents applied for between 2005 and 2006 due to the truncation problem, our sample period is specified as 1981 to 2005. We exclude firms in financial and utility industries (SIC codes 6000 to 6799 and SIC codes 4900 to 4949).

We follow Hall, Jaffe, and Trajtenberg (2001, 2005) and Kogan et al. (2017) to construct three innovation output measures. The first measure of innovation output is the number of patents applied for by a firm in a specific year ($Patent_{i,t}$). The second measure of innovation output is the adjusted number of citations of all patents applied for by a firm in a specific year ($Cites_{i,t}$). The raw number of citations is subject to truncation bias: patents continue to receive citations over long periods but we only observe the citations up to the end of the sample period. The truncation bias is a more serious concern for patents granted in the later years of the sample period because these patents have less time to receive citations. To address the truncation bias concern, we follow Hall, Jaffe, and Trajtenberg (2005) to scale the raw number of citations of each patent by the average number of citations of all patents applied for in the same year and in the same technology class. The third measure of innovation output is the value of all patents applied for by a firm in a specific year ($Patent\_Val_{i,t}$). Kogan et al. (2017) provide this measure in their patent data. To compute the value of a given patent, they first estimate the stock return around the patent issuance date that is attributable to the value of the patent. They then define the value of the patent as the product of the estimate of the stock return due to the value of the patent times the market capitalization of the firm that is issued the patent. Please refer to Kogan et al. (2017) for a detailed description of the method they employ to compute the value of each patent. We follow Kogan et al. (2017) to scale all the three innovation output measures by book assets. In robustness tests, we also follow the literature to use unscaled innovation output measures.

2.5 Other control variables

We also include a set of firm-level and market-level control variables in our analyses. For analyses on stock return comovement, we follow Chan and Hameed (2006) to include firm size, turnover rate and stock return volatility. To address the concern that other market characteristics may drive both investor sentiment and stock return comovement, we also include market total value, market return, market return volatility and NBER recession indicator. For analyses on innovation output, we follow prior research on innovation (e.g. Chang et al. 2015; Cornaggia et al. 2015) to control book assets, return on assets, R&D expenditure scaled by book assets, firm age, sales growth and market-to-book ratio. The definitions of these variables are in Appendix B. The summary statistics of these variables are reported in Table 1.

[Insert Table 1 Here]

3. Stock comovement and investor sentiment

We examine the relation between stock return comovement and lagged investor sentiment in this section. We first report results of our baseline model. We then perform several robustness tests using alternative key variables, more control variables including business cycle measures, extended sample periods, and international data.

3.1 Baseline Model

Our baseline model is specified as

$$Ln\_Rsq_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\beta_2*Size_{i,t}+\beta_3*Turnover_{i,t}+\beta_4*RetStd_{i,t}$$

$$+\beta_5*MktVal_t+\beta_6*MktRet_t+\beta_7*MktStd_t+\beta_8*NBER\_Recession_t+\varepsilon_{i,t} \qquad (5)$$

where $Ln\_Rsq_{i,t}$ is the logistic-transformed R-square defined by Eq. (2) for stock i in year t. We also follow Piotroski and Roulstone (2004) to use an alternative stock comovement measure, $Ln\_Rsq\_ind_{i,t}$ defined by Eq. (4), as our dependent variable. $Sentiment_{t-1}$ is the average of the monthly BW sentiment index of year t-1. We include firm fixed effect and cluster standard errors at firm level in the regressions. To be consistent with the sample periods of the following analyses that use data from I\B\E\S, our sample period is set as 1981 to 2014. In an additional robustness test, we also report results using the full sample period of the sentiment index which is 1965 to 2014.

The results are presented in Table 2. The dependent variable in Column (1) and Column (2) is $Ln\_Rsq_{i,t}$ defined by Eq. (2) and the dependent variable in Column (3) and (4) is $Ln\_Rsq\_Ind_{i,t}$ defined by Eq. (4). Our key variable of interest is the $Sentiment_{t-1}$. We find that the coefficients on the $Sentiment_{t-1}$ are negative and highly significant at 1% level in all four regression models (t-statistics ranges from -8.65 to -23.11). The negative relation is also economically significant. For example, a coefficient of -0.117 on $Sentiment_{t-1}$ indicates that a one-standard-deviation increase in the sentiment index is associated with a decrease of 5.9% of the standard deviation in the logistic-transformed R-squares.

The coefficients on the control variables are generally consistent with findings in prior literature. For instance, firm size is positive correlated with stock comovement because big firms have larger weights in total market returns. Turnover rate is positively correlated with stock comovement because more actively traded stocks tend to react to information in a more timely and synchronous manner (Chan and Hameed (2006)).

We conduct several robustness tests. Bris, Goetzmann, and Zhu (2007) find that stocks' R^2s following negative market returns are higher than that following positive market returns. To address the concern that our finding may be explained by their results, we include past two years' stock returns and market returns as control variables and repeat the analysis. The results are reported in Table A1. The negative relation between lagged investor sentiment and stock return comovement remains unchanged. In addition, the Baker and Wurgler sentiment index is available from 1965 so we perform an unreported test using data from 1965 to 2014. We find quite similar results.

[Insert Table 2 here]

3.2 Market-level analyses

In addition to the stock level analyses in section 3.1, we also conduct market level analyses in this section. We first compute the market level stock comovement for each year from 1981 to 2014 and then examine the relation between market level stock comovement and investor sentiment. Specifically, we estimate the following model:

$$Ln\_MktRsq_t=\alpha+\beta_1*Sentiment_{t-1}+\beta_2*MktVal_t+\beta_3*MktRet_t+\beta_4*MktStd_t$$
$$+\beta_5*NBER\_Recession_t+\varepsilon_t \quad (6)$$

where Ln_MktRsq_t is the market stock comovement in year t. We follow MYY (2000) to define Ln_MktRsq_t as the logistic-transformed average of R^2s of all stocks in year t:

$$MktR^2_t=\frac{\sum_i R^2_{i,t}*SST_{i,t}}{\sum_i SST_{i,t}} \quad (7)$$

$$Ln\_MktRsq_t=ln(MktR^2_t/(1-MktR^2_t)) \quad (8)$$

where R^2_{it} is the R-square of Eq. (1) for stock i in year t and SST_{it} is the sum of squared total return variations of stock i in year t. We also use simple average of R^2 of all stocks to compute $MktR^2_t$.

[Insert Table 3 here]

The results are presented in Table 3. We first use MYY's R^2 to compute market stock comovement. We use simple average R^2 in Model (1) and SST-weighted average in Model (2). We then replace MYY's R^2 with PR's R^2 and report the results in Model (3) and (4). Although we only have 34 observations for the regressions, we find a negative relation between the market stock comovement and the sentiment index. The t-statistics of the coefficients on $Sentiment_{t-1}$ ranges from -1.84 to -2.07. Given the limited number of observations, we think the negative relation can be considered as statistically significant. In terms of economic significance, as shown in Model (4), a one-standard-deviation increase in the sentiment index is related to a decrease of 18.8% of the standard deviation in the logistic-transformed market R-squares. The findings in the market level analyses are consistent with the findings in the stock level analyses.

3.3 Analyses based on a binary measure of sentiment

The key sentiment variable used in the above analyses is the level of the BW sentiment index. In this section, we follow Stambaugh, Yu, and Yuan (2012) to use a binary measure of sentiment, $Senti\_high_{t-1}$, to conduct robustness tests. $Senti\_high_{t-1}$ is a dummy variable which equals 1 if the sentiment index of year t-1 is above the sample median of sentiment index and 0 otherwise. We simply replace $Sentiment_{t-1}$ with $Senti\_high_{t-1}$ in Eq. (5) and Eq. (6) and then repeat the analyses.

Table A2 reports results of the stock level regressions (Eq. (5)) and Table A3 reports the results of the market level regressions (Eq. (6)). We find that current period stock return comovement is negatively and significantly correlated with the last period binary sentiment measure at both stock level and market level. In terms of the statistical significance of the stock level regressions, the coefficients on $Senti\_high_{t-1}$ across all four models of Table A2 are highly significant at the 1% level (t-statistics range from -2.93 to -16.64). Despite the limited number of observations for the market level regressions, the coefficients on $Senti\_high_{t-1}$ in three out of four models of Table A3 are significant at the 10% level (t-statistics range from -1.77 to -1.90). The results are also economically significant. For example, the coefficient on $Senti\_high_{t-1}$ in Model (2) of Table A2 indicates that the difference in stock return comovement between low sentiment periods and high sentiment periods is 7.5% percent of standard deviation in the logistic-transformed R-square. For market level regressions, as shown in Model (2) of Table A3, the difference in stock return comovement between low sentiment periods and high sentiment periods is 35.3% of the standard deviation in market average stock return comovement. The significantly negative relation between current period stock comovement and last period investor sentiment still holds after we replace the level of sentiment index ($Sentiment_{t-1}$) with the binary sentiment measure ($Senti\_high_{t-1}$).

3.4 International evidence

Baker, Wurgler, and Yuan (2012) construct investor sentiment indices for other five countries. These countries are United Kingdom, France, Canada, Germany and Japan. With these sentiment indices, we examine whether the negative relation between investor sentiment and stock return comovement still holds in the international markets. Specifically, we first compute the R-squares of stocks in major stocks exchanges in each of the five countries and then we estimate Eq. (5). Following Chan and Hameed (2006), we do not include industry returns when estimating R-squares because industry returns are problematic in international markets. We report the results in Table 4. We use the level of sentiment index in Model (1) and (2) and use the binary sentiment index in Model (3) and (4). The coefficients on the sentiment measures are always negative and significant at the 1% level, which is consistent with our findings based on the U.S. sample.

[Insert Table 4 Here]

3.5 Business cycle

Brockman, Liebenberg, and Schutte (2010) find that stock return comovement is countercyclical because of higher information production during boom periods. A natural concern is that the negative relation between sentiment and stock return comovement is merely a reflection of the relation between sentiment and business cycle. To address this concern, we first examine the correlation between our sentiment index and their measure of business cycle[2]. Because their analysis is at quarterly frequency, we compute the quarterly sentiment index as the average of monthly sentiment index over the quarter. The correlation between quarterly sentiment and quarterly GDP growth is only 0.1 with a p-value of 0.2. This is not surprising because the Baker and Wurgler (2006) sentiment index has been orthogonalized to several macroeconomic conditions when they construct this measure.

Second, we include the quarterly sentiment index in the regressions of Table 4 of BLS (2010) and re-estimate these regressions using the U.S. market data. To be consistent with their settings, the dependent variable is the measure of stock return comovement used in BLS (2010) which follows Campbell et al. (2001). We also use the same control variables as their paper. The results are reported in Table 5. Model 2 and Model 3 show that the negative relation between sentiment

[2] BLS (2010) use three business cycle variables including two indicators of expansions and GDP growth. The following two tests are based on GDP growth. We use the other two indicators in unreported tests and find similar results.

and stock return comovement is still significant even after controlling for GDP growth which is a main indicator of business cycle in BLS (2010). The economic significance is also non-trivial. The R-squares of Model 1 increase by 83% after we include the sentiment index, suggesting that investor sentiment captures another significant portion of variations of stock return comovement that cannot be explained by business cycle. In addition, a one-standard-deviation increase in investor sentiment (GDP growth) is associated with a decrease of 8.88% (10.25%) of the standard deviation in stock return comovement.

[Insert Table 5 Here]

4. The effect of information suppliers

The findings in section 3 are consistent with our main prediction built on the three hypotheses on how different types of market participants, i.e., information suppliers, informed arbitrageurs and creative firms, respond to investor sentiment. Recall that the three hypotheses are:

Hypothesis 1: Following high investor sentiment, information suppliers would produce more information;

Hypothesis 2: Following high investor sentiment, informed arbitrageurs would conduct more informed arbitrage that incorporates more information into stock prices;

Hypothesis 3: Following high investor sentiment, creative firms would produce more innovations that elevate fundamental volatility.

We examine Hypothesis 1 in this section and the other two hypotheses in the following sections. If Hypothesis 1 is a valid explanation for our main findings, we should find: a) more firm-specific information in the whole market is produced following high sentiment periods and b) the negative relation between lagged investor sentiment and stock return comovement concentrates on the stocks that are widely covered by information suppliers. We follow the literature to consider the media and analysts as traditional information suppliers (e.g. Veldkamp, 2006a; Hameed et al., 2015). We first examine the relation between lagged investor sentiment and firm-specific news production and then test the effects of the media coverage and the analyst coverage on the negative relation between lagged investor sentiment and stock returns comovement.

4.1 Firm-specific news production

We provide evidence that firm-specific information increases with investor sentiment in this section. Measuring firm-specific information production is difficult. We use the news commonality measure ($NewsR^2_{i,t}$) constructed by Dang, Moshirian and Zhang (2015) as our firm-specific information production measure. $NewsR^2_{i,t}$ is the R-square of the regression of the weekly news score (*ESS*) for firm i on the weekly market news score (*MktESS*) in year t. A high $NewsR^2_{i,t}$ means that the information about firm i is highly related to the market information, suggesting less firm-specific information being produced. The news score (*ESS*) provided by RavenPack basically indicates how firm-specific news events are categorized and rated as having a positive or negative effect on stock prices. It ranges from 0 to 100 (0 means most negative, 50 means neutral and 100 means most positive). We examine the relation between $NewsR^2$ and our two sentiment measures (*senti_high* and *sentiment*) and report the results in Table 6. We find the coefficients on the sentiment measures are all negative and significant at the 1% level. Economically, a one-standard-deviation increase in sentiment index is associated with a decrease of 3.1% of the standard deviation in news commonality. The negative relation between sentiment and news R-square indicates that more firm-specific information is produced following high sentiment periods. Although stock return comovement is more related to firm-specific news production, we also analyze how investor sentiment affects total news production. We construct an annual measure of news coverage for each firm ($NewsCover_{i,t}$) and study its relation with investor sentiment. The result is reported in Column 3 of Table 6. The positive and significant coefficient on investor sentiment suggests that total news production is also higher following high sentiment periods.

[Insert Table 6 Here]

4.2 The effect of media coverage

The media are without doubt important information suppliers in financial markets. They have incentives to produce more information when the demand for information increases. If information generators play an important role in driving the negative relation between lagged investor sentiment and stock return comovement, we expect the negative relation to be more pronounced in stocks widely covered by the media.

We first construct a measure of news coverage ($NewsCover_{i,t}$) for each stock each year using the data from RavenPack. We then estimate the following regression model:

$$Ln\_Rsq\_Ind_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\beta_2*Sentiment_{t-1}*NewsCover_{i,t-1}+$$

$$\beta_3*NewsCover_{i,t-1}+\beta_4*X_{i,t}+\varepsilon_{i,t} \qquad (9)$$

where $NewsCover_{i,t-1}$ is the annual average of the daily number of times that stock i is mentioned by the media in year t-1. To be included in our sample, observations' relevance scores and novelty scores must be 100. Our focus is β_2. We report the results in Table 7. Consistent with our prediction, β_2s are always negative and significant at the 1% level. Economically, a one-standard-deviation increase in news coverage is associated with an increase of 37.6% of the sentiment sensitivity of stock return comovement. The media are a nontrivial factor that is driving our main finding.

[Insert Table 7 Here]

4.3 The effect of analyst coverage

Analysts are also important information generators in the stock market. Stocks followed by more analysts are more transparent and suffer less from information asymmetry problem (e.g. Brennana and Subrahmanyam (1995); Hong, Lim, and Stein (2000)). Following the above arguments in section 4.2, we expect to find a significant effect of analyst coverage on the negative relation between lagged investor sentiment and stock return comovement.

Our analyst coverage data are from I\B\E\S. We estimate the following regression:

$$Ln\_Rsq\_Ind_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\beta_2*Sentiment_{t-1}*AnalystCover_{i,t-1}+ \beta_3*AnalystCover_{i,t-1}+\beta_4*X_{i,t}+\varepsilon_{i,t} \quad (10)$$

where $AnalystCover_{i,t-1}$ is our measure of analyst coverage. We follow He and Tian (2013) to construct two measures of analyst coverage. The first one is the natural logarithm of 1 plus the average of monthly number of analyst estimates for stock i in year t (*AnalystCover1*). The second one is the natural logarithm of 1 plus the number of unique analysts who make at least 1 earnings forecast for stock i in year t (*AnalystCover2*). The results are presented in Table 8. The negative and significant coefficients on the interaction term suggest that the negative relation between investor sentiment and stock return comovement concentrate on stocks with high analyst coverage. Overall, this finding and the one in section 4.2 are consistent with our first hypothesis about information suppliers.

[Insert Table 8 Here]

5. Implied cost of equity capital (ICOC) and investor sentiment

So far, our evidence supports the Hypothesis 1 about information suppliers we propose in section 1. Before we move on to test the Hypothesis 2 about informed arbitrageurs and the Hypothesis 3 about creative firms, we examine one intuitive assumption that both Hypothesis 2 and Hypothesis 3 rely on. The assumption is that high investor sentiment leads to low cost of equity capital. We provide evidence for the assumption in this section by analyzing the relation between lagged investor sentiment and cost of equity capital.

We follow Hail and Leuz (2006) to construct a measure of ICOC (*r_avg*) and estimate the following model

$$r\_avg_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\gamma*X_{i,t}+\varepsilon_{i,t} \qquad (11)$$

where $r\_avg_{i,t}$ is the average of four measures of ICOC of stock i in year t. The four measures of ICOC are *r_GLS* from Gebhardt, Lee, and Swaminathan (2001), *r_CT* from Claus and Thomas (2001), *r_OJN* from Ohlson and Juettner-Nauroth (2005) and *r_PEG* from Easton (2004). A detailed description of these measures is provided in section 2.3 and Appendix C. $X_{i,t}$ are the same set of control variables used in Eq. (5).

[Insert Table 9 Here]

As shown in Column (1) of Panel A of Table 9, *r_avg* is negatively correlated with *Sentiment* with a t-value of -16.53 and a coefficient of -0.003. The result is also economically significant. A one-standard-deviation increase in sentiment index is associated with a decrease of 6.4% of the standard deviation in the cost of equity capital. The dependent variables in Column (2) to Column (5) are r_PEG, r_CT, r_OJN and r_GLS, repectively. We find quite similar results: the coefficients on *Sentiment* are -0.003 for the Model using r_PEG (t-stat=-9.59), -0.003 for the

Model using r_CT (t-stat=-18.01), -0.005 for the Model using r_OJN (t-stat=-14.00) and -0.002 for the Model using r_GLS (t-stat=-12.37), respectively.

We also perform a market level analysis where we use the value-weighted average of ICOC across all stocks as the dependent variable:

$$ICOC_t=\alpha_i+\beta_1*Sentiment_{t-1}+\gamma*X_t+\varepsilon_t \qquad (12)$$

where $ICOC_t$ is the value-weighted average of *r_AVG/r_PEG/r_CT/r_OJN/r_GLS* across all stocks in year t. X_t are the same set of market level control variables used in Eq. (6). Panel B of Table 9 presents the results. The sentiment index is negatively correlated with the market cost of equity capital ($ICOC_t$) in all five models. The coefficient on $Sentiment_{t-1}$ is significant at the 5% level (t-stat=-2.20) for the model based on *r_CT* and marginally significant (t-stat=-1.64) for the model based on *r_avg*, repectively. The small number of observations and the potential measurement errors in each individual measure of ICOC may limit our ability to find significant results for the other three Models. Economically, a coefficient of -0.004 in Model (1) indicates that a one-standard-deviation increase in sentiment index is related with a decrease of 11.5% of the standard deviation in cost of equity capital.

For robustness check (unreported), we replace the level of sentiment index (*sentiment*) with the binary sentiment measure (*senti_high*) in Eq. (11) and Eq. (12) and repeat the analyses. The results remain unchanged. Therefore, evidence in this section suggests that the overpricing induced by high sentiment indeed leads to lower cost of equity capital for firms.

6. The effect of informed arbitrageurs

In this section, we investigate Hypothesis 2 which is about how informed arbitrageurs' responses to investor sentiment affect stock return comovement. We consider short sellers and institutional investors as informed arbitrageurs (Boehmer, Jones, and Zhang (2008); Diether, Lee, and Werner (2009); Puckett and Yan, 2011; Chen, Jegadeesh, and Wermers, 2000).

6.1 Short sellers

Evidence in prior literature suggests that short-sellers are mainly sophisticated investors (e.g. Asquith, Pathak, and Ritter (2005); Boehmer, Jones, and Zhang (2008); Diether, Lee, and Werner (2009); Engelberg, Reed, and Ringgenberg (2012)). For example, Boehmer, Jones, and Zhang (2008) find that 75% short sales are executed by institutional investors and Diether, Lee and Werner (2009) document a robust negative relation between short volume and future stock returns. We can reasonably assume that firm-specific information is incorporated into stock prices through short sales. If informed arbitrageurs contribute to our main findings in section 3, short sellers, as one important group of arbitrageurs, would be likely to play a role. Specifically, we predict that a) there would be more short selling activities following high sentiment periods and b) the negative relation between lagged investor sentiment and stock return comovement would be more pronounced in stocks with lower cost of short selling because short sellers are less constraint to generate firm-specific information through their trades for these stocks.

To test these two predictions, we estimate the following two regressions, respectively:

$$ShortIntensity_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\beta_2*X_{i,t}+\varepsilon_{i,t} \qquad (13)$$

$$Ln\_Rsq\_Ind_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\beta_2*Sentiment_{t-1}*ShortCost_{i,t-1}+\beta_3*ShortCost_{i,t-1}+\beta_4*X_{i,t}+\varepsilon_{i,t} \qquad (14)$$

where Eq. (13) is used to examine the first prediction and Eq. (14) is for the second prediction. $ShortIntensity_{i,t}$ is the intensity of short selling activity. We use both the change of the natural logarithm of a stock's short interest ratio and the change of the natural logarithm of a stock's lendable value as the short selling intensity measures[3]. The results of Eq. (13) are reported in Panel A of Table 10. Consistent with our prediction, the change of short interest (lendable value) is positively (negatively) associated with lagged investor sentiment, suggesting that more shares are sold short and fewer shares are available for lending following high sentiment periods. Short sellers are more active after high investor sentiment.

Panel B of Table 10 reports the results of Eq. (14). $ShortCost_{i,t-1}$ represents our measure of cost of short selling. We construct two measures of cost of short selling based on stocks' loan fees and stocks' lendable value, respectively. The first one is $LoanFee\_high_{i,t}$ which equals 1 if the average loan fee of stock i in year t is above the sample median and 0 otherwise. The second one is the $Lendable\_high_{i,t}$ which equals 1 if the average lendable value of stock i in year t is above the sample median and 0 otherwise. The lendable value of a stock is scaled by the market value of that stock. Our focus is β_2 and we expect it to be positive for the regressions using $LoanFee\_high_{i,t-1}$ and negative for the regressions using $Lendable\_high_{i,t-1}$. Consistent with our prediction, the coefficients on the interaction terms are always positive and significant at the 1% level in column (1) and (2) where we use $LoanFee\_high_{i,t-1}$ and negative and significant in column (3) and (4) where we use $Lendable\_high_{i,t-1}$. In unreported tests, we also use the binary sentiment measure and find quite similar results. The effect is also economically significant: the coefficient on $Sentiment_{t-1}$ for stocks with low cost of short selling is -0.328, while it is -0.174 for stocks with high cost of short selling, indicating a 47% difference in the sentiment sensitivity of stock return comovement. The results suggest that the difference in stock comovement between high sentiment periods and low sentiment periods concentrates on stocks with low cost of short selling. Overall, the short sellers' response to investor sentiment supports Hypothesis 2 in section 1 regarding the behavior of informed arbitrageurs.

[3] The lendable value of a stock is scaled by the market value of that stock

[Insert Table 10 Here]

6.2 Institutional investors

Another group of informed arbitrageurs we investigate is institutional investors. Compared to retail investors, institutional investors are often referred to as more rational and informed investors (e.g. Puckett and Yan, 2011; Chen, Jegadeesh, and Wermers, 2000; Kacperczyk, Sialm, and Zheng, 2005). Their trades are likely to convey more firm-specific information to the market. If informed arbitrageurs like institutional investors play an important role in shaping the negative relation between lagged investor sentiment and stock return comovement, we conjecture that a) institutional investors are more active following high sentiment periods and b) the negative relation should be more pronounced in stocks with higher institutional ownership.

We collect the institutional ownership data from Thomson Reuters 13f Institutional Holding database. We estimate the following two regressions to examine the above two conjectures, respectively:

$$InstIntensity_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\beta_2*X_{i,t}+\varepsilon_{i,t} \quad (15)$$

$$Ln\_Rsq\_Ind_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\beta_2*Sentiment_{t-1}*InstHold_{i,t-1}+\beta_3*InstHold_{i,t-1}+\beta_4*X_{i,t}+\varepsilon_{i,t} \quad (16)$$

where Eq. (15) is for the first conjecture and Eq. (16) is for the second conjecture. Our measure of institutional trades intensity, $InstIntensity_{i,t}$, is the absolute value of the change in natural logarithm of institutional holdings of stock i in year t. This measure follows the spirit of Piotroski and Roulstone (2004). Higher $InstIntensity_{i,t}$ is likely to be associated with larger net changes in institutional holdings and more active institutional trades. The results of Eq. (15) are presented in Panel A of Table 11. The coefficient on lagged investor sentiment is positive and significant at the 1% level. Institutional investors are trading more actively following high sentiment periods, which confirms our first conjecture.

We present the results of Eq. (16) in Panel B of Table 11. $InstHold_{i,t-1}$ is the value of stock i held by institutions scaled by the market value of stock i in year t-1. As shown in the table, β_2s are negative and significant at the 1% level in all four models. Economically, a one-standard-deviation increase in $InstHold_{i,t-1}$ is associated with an increase of 95.2% in the sentiment sensitivity of stock return comovement. Again, these findings are largely consistent with Hypothesis 2 in section 1.

[Insert Table 11 Here]

7. The effect of creative firms

Finally, we analyze Hypothesis 3 in this section. We examine whether the low cost of equity capital induced by high sentiment enables firms to accelerate creative innovations that increase firm-specific event intensity and eventually decreases stock return comovement.

7.1 Innovation output and investor sentiment

If high investor sentiment fosters creative destructions, we expect that the innovation output following high sentiment periods should be higher than that following low sentiment periods. We rely on the following regression to test our prediction

$$Innovation_{i,t} = \alpha_i + \beta_1 * Sentiment_{t-1} + \gamma * X_{i,t} + \varepsilon_{i,t} \quad (17)$$

where $Innovation_{i,t}$ is the innovation output of stock i in year t. We use the number of patents ($Patents_{i,t}$), the adjusted number of citations ($Cites_{i,t}$) and the value of the patents ($Patent\_Value_{i,t}$) as our innovation output measures. We follow Kogan et al (2017) to scale the three measures by book assets. $X_{i,t}$ is a set of control variables including natural logarithm of book assets, return on assets, R&D expenditures, natural logarithm of firm age, sales growth rate and market-to-book ratio. We expect β_1 to be positive and significant. The results are reported in Table 12.

[Insert Table 12 Here]

As shown in Table 12, β_1s in all specifications are positive and significant at the 1% level. Economically, a β_1 of 0.613 in Model (6) suggests that a one-standard-deviation increase in sentiment index is associated with an increase of 2.7% of the standard deviation in the value of patents scaled by book assets. The results here are consistent with our prediction that high sentiment enables firms to produce more innovations. In an additional unreported test, we replace the level of sentiment with the binary sentiment measure and re-estimate Eq. (17) and we find quite similar results.

Theoretically, a decrease in cost of capital should have a larger impact on financially constrained firms (Almeida, Campello, and Weisbach, 2004). If investor sentiment affects innovation output through the cost of equity capital channel, we should find that the effect documented in Table 12 is more pronounced in financially constrained firms. We use three measures to identify financially constrained firms: $High\_HP_{i,t}$, $Low\_payout_{i,t}$ and $No\_divid_{i,t}$. $High\_HP_{i,t}$ is a dummy variable which equals 1 if the Hadlock-Pierce index of firm i is above the sample median HP in year t and 0 otherwise. $Low\_payout_{i,t}$ is a dummy variable which equals 1 if the payout ratio of firm i is below the sample median payout ratio in year t and 0 otherwise. $No\_divid_{i,t}$ is a dummy variable which equals 1 if firm i does not pay dividend in year t and 0 otherwise. We interact our measures of financial constraints with the sentiment variable in Eq. (17) and re-estimate the model. The results are presented in Table 13. The coefficients on the interaction terms are positive and significant at the 1% level in all the specifications. The results suggest that the positive relation between sentiment and innovation output concentrates on firms that are financially constrained, which is consistent with our prediction.

[Insert Table 13 Here]

7.2 The effect of high-tech intensiveness

After showing a difference in innovation output between high sentiment periods and low sentiment periods, we next provide evidence of the direct impact of innovations on the relation between lagged investor sentiment and stock return comovement. Hypothesis 3 implies that the negative relation would be more pronounced in firms whose values are sensitive to innovations. We follow Hsu, Tian, and Xu (2014) to use a measure of high-tech intensiveness to identify these firms. Our regression is specified as:

$$Ln\_Rsq\_Ind_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\beta_2*Sentiment_{t-1}*High\text{-}Tech_i+ \beta_3*High\text{-}Tech_i+\beta_4*X_{i,t}+\varepsilon_{i,t} \quad (18)$$

where $High\text{-}Tech_i$ is a dummy variable which equals 1 if the high-tech intensiveness of the industry that firm i belongs to is above the sample median and 0 otherwise. Industry j's high-tech intensiveness is computed in the following way. Every year t, we use all firms in industry j to regress the natural logarithm of the ratio of market value to total assets on R&D expense scaled by total assets over the past five years. The coefficient on R&D expense scaled by total assets is the temporary high-tech intensiveness of industry j in year t. We take the time series median of industry j's temporary high-tech intensiveness as the high-tech intensiveness of industry j. The results of the regressions are reported in Table 14. The negative coefficients on the interaction terms suggest that the negative relation between investor sentiment and stock return comovement concentrates on stocks that are sensitive to innovations. Economically, the difference in the sentiment sensitivity of stock return comovement between high-tech firms and low-tech firms is 15%. We argue that this is evidence showing that creative innovations are an important channel through which investor sentiment affects stock return comovement.

[Insert Table 14 Here]

8. Conclusion

I find a negative relation between lagged investor sentiment and stock return comovement. I investigate how market participants other than noise traders contribute to this negative relation. I show that both firm-specific news production and innovation output are higher following high sentiment periods. Moreover, the negative relation is more pronounced in the stocks that have higher news coverage, higher analyst coverage, lower cost of short selling, higher institutional ownership, and higher high-tech intensiveness. Overall, my evidence suggests a nontrivial role of other market participants in shaping the relation between sentiment and comovement, which confirms the hypotheses derived from Veldkamp (2006a, 2006b) and Morck, Yeung, and Yu (2013).

Reference

Asquith, Paul, Parag A. Pathak, and Jay R. Ritter. "Short interest, institutional ownership, and stock returns." *Journal of Financial Economics* 78, no. 2 (2005): 243-276.

Baker, Malcolm, and Jeffrey Wurgler. "The equity share in new issues and aggregate stock returns." *Journal of Finance* 55, no. 5 (2000): 2219-2257.

Baker, Malcolm, and Jeffrey Wurgler. "Market timing and capital structure." *Journal of Finance* 57, no. 1 (2002): 1-32.

Baker, Malcolm, and Jeffrey Wurgler. "Investor sentiment and the cross- section of stock returns." *Journal of Finance* 61, no. 4 (2006): 1645-1680.

Baker, Malcolm, Jeffrey Wurgler, and Yu Yuan. "Global, local, and contagious investor sentiment." *Journal of Financial Economics* 104, no. 2 (2012): 272-287.

Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler. "Comovement." *Journal of Financial Economics* 75, no. 2 (2005): 283-317.

Boehmer, Ekkehart, Charles M. Jones, and Xiaoyan Zhang. "Which shorts are informed?" *Journal of Finance* 63, no. 2 (2008): 491-527.

Brennan, Michael J., and Avanidhar Subrahmanyam. "Investment analysis and price formation in securities markets." *Journal of Financial Economics*38, no. 3 (1995): 361-381.

Bris, Arturo, William N. Goetzmann, and Ning Zhu. "Efficiency and the bear: Short sales and markets around the world." *Journal of Finance* 62, no. 3 (2007): 1029-1079.

Brockman, Paul, Ivonne Liebenberg, and Maria Schutte. "Comovement, information production, and the business cycle." *Journal of Financial Economics* 97, no. 1 (2010): 107-129.

Campbell, John Y., Martin Lettau, Burton G. Malkiel, and Yexiao Xu. "Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk." *Journal of Finance* 56, no. 1 (2001): 1-43.

Cao, Ying, James N. Myers, Linda A. Myers, and Thomas C. Omer. "Company reputation and the cost of equity capital." *Review of Accounting Studies* 20, no. 1 (2015): 42-81.

Cao, Ying, Linda A. Myers, Albert Tsang, and Yong George Yang. "Management forecasts and the cost of equity capital: international evidence." *Review of Accounting Studies* 22, no. 2 (2017): 791-838.

Chan, Kalok, and Allaudeen Hameed. "Stock price synchronicity and analyst coverage in emerging markets." *Journal of Financial Economics* 80, no. 1 (2006): 115-147.

Chen, Hsiu-Lang, Narasimhan Jegadeesh, and Russ Wermers. "The value of active mutual fund management: An examination of the stockholdings and trades of fund managers." *Journal of Financial and quantitative Analysis* 35, no. 3 (2000): 343-368.

Chen, Zhihong, Dan S. Dhaliwal, and Hong Xie. "Regulation fair disclosure and the cost of equity capital." *Review of Accounting Studies* 15, no. 1 (2010): 106-144.

Chun, Hyunbae, Jung-Wook Kim, Randall Morck, and Bernard Yeung. "Creative destruction and firm-specific performance heterogeneity." *Journal of Financial Economics* 89, no. 1 (2008): 109-135.

Claus, James, and Jacob Thomas. "Equity premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stock markets." *Journal of Finance* 56, no. 5 (2001): 1629-1666.

Dang, Tung Lam, Fariborz Moshirian, and Bohui Zhang. "Commonality in news around the world." *Journal of Financial Economics* 116, no. 1 (2015): 82-110.

Dang, Tri Vi, and Zhaoxia Xu. "Market Sentiment and Innovation Activities." Working Paper (2016).

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. "Noise trader risk in financial markets." *Journal of Political Economy* 98, no. 4 (1990): 703-738.

Diether, Karl B., Kuan-Hui Lee, and Ingrid M. Werner. "Short-sale strategies and return predictability." *Review of Financial Studies* 22, no. 2 (2009): 575-607.

Durnev, Artyom, Randall Morck, Bernard Yeung, and Paul Zarowin. "Does greater firm- specific return variation mean more or less informed stock pricing?." *Journal of Accounting Research* 41, no. 5 (2003): 797-836.

Easton, Peter D. "PE ratios, PEG ratios, and estimating the implied expected rate of return on equity capital." *The Accounting Review* 79, no. 1 (2004): 73-95.

Engelberg, Joseph E., Adam V. Reed, and Matthew C. Ringgenberg. "How are shorts informed?: Short sellers, news, and information processing." *Journal of Financial Economics* 105, no. 2 (2012): 260-278.

Gebhardt, William R., Charles Lee, and Bhaskaran Swaminathan. "Toward an implied cost of capital." *Journal of Accounting Research* 39, no. 1 (2001): 135-176.

Hail, Luzi, and Christian Leuz. "International differences in the cost of equity capital: Do legal institutions and securities regulation matter?." *Journal of Accounting Research* 44, no. 3 (2006): 485-531.

Hall, Bronwyn H., and Rosemarie Ham Ziedonis. "The patent paradox revisited: an empirical study of patenting in the US semiconductor industry, 1979-1995." *RAND Journal of Economics* (2001): 101-128.

Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg. "Market value and patent citations." *RAND Journal of Economics* (2005): 16-38.

Hameed, Allaudeen, Randall Morck, Jianfeng Shen, and Bernard Yeung. "Information, analysts, and stock return comovement." *Review of Financial Studies* 28, no. 11 (2015): 3153-3187.

He, Jie Jack, and Xuan Tian. "The dark side of analyst coverage: The case of innovation." *Journal of*

Financial Economics 109, no. 3 (2013): 856-878.

Hong, Harrison, Terence Lim, and Jeremy C. Stein. "Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies." *Journal of Finance* 55, no. 1 (2000): 265-295.

Hsu, Po-Hsuan, Xuan Tian, and Yan Xu. "Financial development and innovation: Cross-country evidence." *Journal of Financial Economics* 112, no. 1 (2014): 116-135.

Hutton, Amy P., Alan J. Marcus, and Hassan Tehranian. "Opaque financial reports, R 2, and crash risk." *Journal of Financial Economics* 94, no. 1 (2009): 67-86.

Irvine, Paul J., and Jeffrey Pontiff. "Idiosyncratic return volatility, cash flows, and product market competition." *Review of Financial Studies* 22, no. 3 (2008): 1149-1177.

Jin, Li, and Stewart C. Myers. "R^2 around the world: New theory and new tests." *Journal of Financial Economics* 79, no. 2 (2006): 257-292.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng. "On the industry concentration of actively managed equity mutual funds." *Journal of Finance* 60, no. 4 (2005): 1983-2011.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman. "Technological innovation, resource allocation, and growth." *The Quarterly Journal of Economics* 132, no. 2 (2017): 665-712.

Kumar, Alok, and Charles Lee. "Retail investor sentiment and return comovements." *Journal of Finance* 61, no. 5 (2006): 2451-2486.

Martin, Alberto, and Jaume Ventura. "Economic growth with bubbles." *The American Economic Review* 102, no. 6 (2012): 3033-3058.

Morck, Randall, Bernard Yeung, and Wayne Yu. "The information content of stock markets: why do emerging markets have synchronous stock price movements?." *Journal of Financial Economics* 58, no. 1 (2000): 215-260.

Morck, Randall, Bernard Yeung, and Wayne Yu. R-squared and the economy. No. w19017. *National Bureau of Economic Research*, 2013.

Ohlson, James A., and Beate E. Juettner-Nauroth. "Expected EPS and EPS growth as determinants of value." *Review of Accounting Studies* 10, no. 2 (2005): 349-365.

Piotroski, Joseph D., and Darren T. Roulstone. "The influence of analysts, institutional investors, and insiders on the incorporation of market, industry, and firm-specific information into stock prices." *The Accounting Review* 79, no. 4 (2004): 1119-1151.

Puckett, Andy, and Xuemin Sterling Yan. "The interim trading skills of institutional investors." *The Journal of Finance* 66, no. 2 (2011): 601-633.

Schumpeter J. 1911. Theorie der Wirtschaftlichen Entwicklung. Leipzig, Ger.: Dunker & Humblot

Stambaugh, Robert F., Jianfeng Yu, and Yu Yuan. "The short of it: Investor sentiment and anomalies." *Journal of Financial Economics* 104, no. 2 (2012): 288-302.

Veldkamp, Laura L. "Media frenzies in markets for financial information." *The American Economic Review* 96, no. 3 (2006): 577-601.

Veldkamp, Laura L. "Information markets and the comovement of asset prices." *The Review of Economic Studies* 73, no. 3 (2006): 823-845.

Ventura, Jaume. "Bubbles and capital flows." *Journal of Economic Theory* 147, no. 2 (2012): 738-758.

Table 1 Summary statistics

This table presents the summary statistics of the variables used in this paper. *Ln_Rsq* is the logistic-transformed R-square following MYY (2000). *Ln_ME* is the natural logarithm of year-end market capitalization. *Turnover* is the average monthly turnover rate in a year. *RetVol* is the volatility of daily stock returns in a given year. *r_CT* is a measure of cost of equity capital (COC) from Claus and Thomas (2001), *r_OJN* is a measure of COC from Ohlson and Juettner-Nauroth (2005), *r_GLS* is a measure of COC from Gebhardt, Lee, and Swaminathan (2001), *r_PEG* is a measure of COC from Easton (2004). *r_AVG* is the average of these four measures. *Patents_value* is the value of all patents applied for by a firm in a specific year. *Cites* is the adjusted number of citations of all patents applied for by a firm in a specific year. *Patents* is the number of patents applied for by a firm in a specific year. $Senti\_high_t$ is a dummy variable which equals 1 if the sentiment index of year t is above the sample median and 0 otherwise. *Sentiment* is the average of monthly Baker and Wurgler sentiment index in a given year. *MktVal* is the natural logarithm of total capitalization across all stocks in the market. *MktRet* is the annual CRSP value-weighted index return. *MktRetVol* is the volatility of daily index returns within a year.

Panel A: R2 Analysis								
Variable	N	Mean	Std	Min	P25	P50	P75	Max
Ln_Rsq_Ind	146921	-1.479	1.136	-7.889	-2.226	-1.471	-0.699	2.568
Ln_Rsq	148144	-2.706	2.157	-24.273	-3.648	-2.261	-1.247	2.028
LN_NewRsq	53798	-4.674	2.240	-23.951	-5.727	-4.203	-3.115	-0.165
Size	146921	5.208	2.082	-2.096	3.666	5.048	6.611	13.374
Turnover	146921	0.088	0.136	0.000	0.021	0.048	0.106	8.440
RetVol	146921	0.033	0.019	0.001	0.019	0.028	0.041	0.530
Panel B: Cost of Equity Capital Analysis								
r_AVG	50807	0.114	0.027	0.061	0.094	0.109	0.130	0.224
r_CT	50807	0.097	0.024	0.048	0.080	0.093	0.110	0.175
r_OJN	50807	0.135	0.041	0.068	0.105	0.125	0.156	0.289
r_GLS	50807	0.097	0.024	0.051	0.080	0.094	0.112	0.174
Panel C: Innovation Analysis								
Patents_value	107681	3.703	13.158	0.000	0.000	0.000	0.190	87.360
Cites	107681	1.848	6.845	0.000	0.000	0.000	0.000	48.260
Patents	107681	1.935	6.389	0.000	0.000	0.000	0.234	43.469
Panel D: Market-level Variables								
Sentiment	34	0.318	0.575	-0.625	-0.064	0.206	0.641	2.152
Senti_high	34	0.500	0.508	0	0	0.5	1	1
MktVal	34	15.828	0.953	14.069	14.907	16.205	16.650	17.182
MktRet	34	0.123	0.171	-0.382	0.018	0.157	0.253	0.357
MktRetVol	34	0.010	0.004	0.005	0.007	0.008	0.012	0.025

Table 2 Stock comovement and sentiment: stock-level

This table presents the results of the regression of stock return comovement on investor sentiment. We estimate the following model:

$$Ln\_Rsq_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\beta_2*X_{i,t}+\varepsilon_{i,t}$$

where $Ln\_Rsq_{i,t}$ is the logistic-transformed R-square defined by Eq. (2) for stock i in year t. $Sentiment_{t-1}$ is the average of the monthly BW sentiment index of year t-1. $Ln\_Rsq\_Ind_{i,t}$ is the logistic-transformed R-square defined by Eq. (4) for stock i in year t. Other control variables are defined in Appendix B. The sample period is from 1981 to 2014. In Model (1) and Model (2), we use $Ln\_Rsq_{i,t}$ of MYY (2000) as the dependent variable. In Model (3) and Model (4), we use $Ln\_Rsq\_Ind_{i,t}$ of Piotroski and Roulstone (2004) as the dependent variable . All the regressions include firm fixed effect and standard errors are clustered at firm level. t-statistics are reported in parentheses.

VARIABLES	(1) *Ln_Rsq*	(2) *Ln_Rsq*	(3) *Ln_Rsq_Ind*	(4) *Ln_Rsq_Ind*
Sentiment	**-0.095*****	**-0.222*****	**-0.042*****	**-0.117*****
	(-9.95)	**(-23.11)**	**(-8.65)**	**(-24.77)**
Size		0.527***		0.292***
		(53.38)		(57.75)
Turnover		0.907***		0.508***
		(8.40)		(8.78)
Retstd		-5.713***		-0.973***
		(-10.40)		(-4.00)
Mktval		-0.495***		-0.199***
		(-36.39)		(-28.47)
Mktret		-0.022		-0.085***
		(-0.64)		(-5.26)
Mktretvol		120.163***		72.435***
		(71.99)		(85.86)
NBER_Rec		-0.215***		-0.085***
		(-12.19)		(-9.69)
Constant	-2.687***	1.432***	-1.473***	-0.506***
	(-796.44)	(7.31)	(-855.52)	(-4.97)
Observations	151,265	148,144	150,004	146,921
R-squared	0.358	0.428	0.439	0.532

Table 3 Stock comovement and sentiment: market-level

This table presents the results of the regression of stock return comovement on investor sentiment at the market level. We estimate the following model:

$$Ln\_MktRsq_t=\alpha+\beta_1*Sentiment_{t-1}+\beta 2*X_t+\varepsilon_t$$

where Ln_MktRsq_t is the logistic-transformed market R-square defined by Eq. (8) in year t. $Sentiment_{t-1}$ is the average of the monthly BW sentiment index of year t-1. $Ln\_MktRsq\_Ind_t$ is the logistic-transformed market R-square defined by Piotroski and Roulstone (2004) in year t. Other control variables are defined in Appendix B. The sample period is from 1981 to 2014. We use simple average in Model (1) and SST-weighted average R^2 in Model (2). We then replace MYY's R^2 with Piotroski and Roulstone (2004)'s R^2 and report the results in Model (3) and (4). t-statistics are reported in parentheses.

VARIABLES	(1) *Ln_MktRsq*	(2) *Ln_MktRsq*	(3) *Ln_MktRsq_Ind*	(4) *Ln_MktRsq_Ind*
Sentiment	**-0.254****	**-0.232****	**-0.162***	**-0.139***
	(-2.08)	**(-2.07)**	**(-2.01)**	**(-1.84)**
Mktval	0.075	0.062	0.087*	0.077*
	(1.06)	(1.01)	(1.80)	(1.83)
Mktret	-0.014	-0.222	-0.080	-0.228
	(-0.03)	(-0.51)	(-0.26)	(-0.79)
Mktretvol	73.432**	85.204***	54.846***	63.576***
	(2.67)	(3.60)	(2.96)	(3.98)
NBER_Rec	0.004	0.111	0.025	0.109
	(0.01)	(0.45)	(0.12)	(0.66)
Constant	-3.649***	-3.679***	-3.052***	-3.082***
	(-3.40)	(-3.84)	(-4.29)	(-4.86)
Observations	34	34	34	34
R-squared	0.458	0.609	0.569	0.700

Table 4 Stock comovement and sentiment in international markets

This table presents the results of the regression of stock return comovement on investor sentiment across five countries. These five countries are United Kingdom, Canada, Germany, France and Japan. We estimate the following model:

$$Ln\_Rsq_{i,j,t}=\alpha_i+\beta_1*Sentiment_{j,t-1}+\beta_2*X_{i,j,t}+\varepsilon_{i,j,t}$$

where $Ln\_Rsq_{i,j,t}$ is the logistic-transformed R-square defined by Eq. (2) for stock i of country j in year t. $Senti\_high_{j,t-1}$ is a dummy variable which equals 1 if the sentiment index of country j in year t-1 is above the sample median of the sentiment index of country j and 0 otherwise. $Sentiment_{j,t-1}$ is the average of the monthly sentiment index of country j in year t-1. Other control variables are defined in Appendix B. The sample period is from 1981 to 2014. In Model (1) and Model (2), we use $Senti\_high_{j,t-1}$ as the sentiment index. In Model (3) and Model (4), we use $Sentiment_{j,t-1}$. All the regressions include firm fixed effect and standard errors are clustered at firm level. t-statistics are reported in parentheses.

	(1)	(2)	(3)	(4)
VARIABLES	*Ln_Rsq*	*Ln_Rsq*	*Ln_Rsq*	*Ln_Rsq*
Sentiment	**-0.031*****	**-0.057*****	**-0.095*****	**-0.083*****
	(-4.32)	**(-7.99)**	**(-6.00)**	**(-5.85)**
Size		0.420***		0.422***
		(38.83)		(39.06)
Turnover		-0.000**		-0.000**
		(-2.06)		(-2.08)
Retstd		4.518***		4.390***
		(7.89)		(7.66)
Mktval		-0.209***		-0.223***
		(-11.52)		(-12.66)
Mktret		-0.676***		-0.652***
		(-20.95)		(-20.56)
Mktretvol		37.764***		37.788***
		(34.28)		(34.30)
Constant	-2.927***	-3.127***	-2.884***	-2.898***
	(-13,784.54)	(-13.62)	(-387.90)	(-13.11)
Observations	117,706	117,706	117,706	117,706
R-squared	0.320	0.359	0.321	0.359

Table 5 The effect of business cycle

This table presents the results of the regression of stock return comovement on a business cycle variable, investor sentiment and several control variables.

$$Ln\_Rsq_t\_Campbell=\alpha_i+\beta_1*GDP\_Growth_{t-1}+\beta_2*Sentiment_{t-1}+\beta_3*X_t+\varepsilon_t$$

Where *$Ln\_Rsq_t\_Campbell$* is measure of stock return comovement following Campbell et al. (2010). *GDP_Growth_{t-1}* is the change of natural logarithms of quarterly real GDP in quarter t-1. *$Sentiment_{t-1}$* is the average of monthly Baker and Wurgler (2006) sentiment index over quarter t-1. *$NumofStk_t$* is the logarithm of the number of stocks in the U.S. market in quarterly t. *$Delist\_Pct_t$* is the proportion of stocks that delisted in quarter t. *HHI_Stk_t* and *HHI_Ind_t* are the Herfindahl index of the entire market based on firm-level sales and industry-level sales in quarter t, respectively. All the regressions include firm fixed effect and standard errors are clustered at firm level. t-statistics are reported in parentheses.

	(1)	(2)	(3)
VARIABLES	*R2*	*R2*	*R2*
GDP_Growth	-13.123***	-12.113***	-4.524***
	(-4.19)	(-4.15)	(-2.74)
Sentiment		**-0.140*** **	**-0.045** **
		(-2.78)	**(-2.06)**
NumofStk			-1.113***
			(-7.07)
Delist_Pct			0.289***
			-7.93
HHI_Stk			-182.325***
			(-6.33)
HHI_Ind			6.105
			-1.45
Constant	-0.915***	-0.877***	8.509***
	(-26.88)	(-27.80)	-5.52
Observations	137	137	137
R-squared	0.092	0.168	0.76

Table 6 News production and sentiment

This table presents the relation between news production and investor sentiment. We estimate the following model

$$NewsR^2_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\beta_2*X_{i,t}+\varepsilon_{i,t}$$

Where $NewsR^2_{i,t}$ is the news commonality measure defined by Dang, Moshirian and Zhang (2015) for stock i in year t. $NewsCover_{i,t}$ is the annual average of the daily number of times that stock i is mentioned by the media in year t-1. Other variables are the same as those in Table 2. The sample period is from 2000 to 2014. All regressions include firm fixed effect and standard errors are clustered at firm level. t-statistics are reported in parentheses.

	(1)	(2)	(3)
VARIABLES	*NewsR²*	*NewsR²*	*NewsCover*
Sentiment	**-0.047*** **	**-0.121*** **	**0.016*** **
	(-3.00)	**(-6.02)**	**(13.22)**
Size		0.107***	0.030***
		(5.51)	(4.92)
Turnover		-0.174***	0.063***
		(-2.70)	(3.04)
Retstd		4.246***	0.022
		(4.63)	(0.16)
Mktval		-0.447***	0.827***
		(-5.96)	(50.41)
Mktret		-0.409***	0.070***
		(-5.09)	(14.66)
Mktretvol		-15.165***	30.557***
		(-3.53)	(39.60)
NBER_Rec		-0.179***	-0.076***
		(-4.90)	(-22.52)
Constant	-4.665***	2.275*	-14.031***
	(-1,611.01)	(1.80)	(-48.18)
Observations	54,068	53,798	48,986
R-squared	0.156	0.159	0.759

Table 7 The effect of media coverage

This table presents the effect of analyst coverage on the relation between stock return comovement and investor sentiment. We estimate the following model:

$$Ln\_Rsq\_Ind_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\beta_2*Sentiment_{t-1}*NewsCover_{i,t-1}+\beta_3*NewsCover_{i,t-1}+\beta_4*X_{i,t}+\varepsilon_{i,t}$$

where $Ln\_Rsq\_Ind_{i,t}$ is the logistic-transformed R-square defined by Eq. (4) for stock i in year t. $Sentiment_{t-1}$ is the average of the monthly BW sentiment index of year t-1. $NewsCover_{i,t-1}$ is the annual average of the daily number of times that stock i is mentioned by the media in year t-1. Other control variables are defined in Appendix B. The sample period is from 2000 to 2014. All the regressions include firm fixed effect and standard errors are clustered at firm level. t-statistics are reported in parentheses.

VARIABLES	(1) *Ln_Rsq_Ind*	(2) *Ln_Rsq_Ind*
Sentiment	-0.038***	-0.134***
	(-4.81)	(-16.48)
*Sentiment *NewsCover*	**-0.084*** **	**-0.129*** **
	(-3.22)	**(-5.38)**
NewsCover	0.190***	-0.080***
	(7.42)	(-4.76)
Size		0.275***
		(30.93)
Turnover		-0.151***
		(-5.00)
Retstd		4.272***
		(8.14)
Mktval		0.344***
		(10.61)
Mktret		0.012
		(0.46)
Mktretvol		88.036***
		(50.02)
NBER_Rec		-0.393***
		(-27.73)
Constant	-1.139***	-9.491***
	(-148.79)	(-17.32)
Observations	52,992	52,638
R-squared	0.570	0.637

Table 8 The effect of analyst coverage

This table presents the effect of analyst coverage on the relation between stock return comovement and investor sentiment. We estimate the following model:

$$Ln\_Rsq\_Ind_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\beta_2*Sentiment_{t-1}*AnalystCover_{i,t}+\beta_3*AnalystCover_{i,t}+\beta_4*X_{i,t}+\varepsilon_{i,t}$$

where $Ln\_Rsq\_Ind_{i,t}$ is the logistic-transformed R-square defined by Eq. (4) for stock i in year t. $Sentiment_{t-1}$ is the average of the monthly BW sentiment index of year t-1. We use two measures of analyst coverage. The first one is the natural logarithm of 1 plus the average of monthly number of estimates for stock i in year t (*AnalystCover1*). The second one is the natural logarithm of 1 plus the number of unique analysts who make at least 1 earnings forecast for stock i in year t (*AnalystCover2*). Other control variables are defined in Appendix B. The sample period is from 1981 to 2014. In Model (1) and Model (2), we use $AnalystCover1_{i,t}$ as our measure of loan fee. In Model (3) and Model (4), we use $AnalystCover2_{i,t}$. All the regressions include firm fixed effect and standard errors are clustered at firm level. t-statistics are reported in parentheses.

VARIABLES	(1) *Ln_Rsq_Ind*	(2) *Ln_Rsq_Ind*	(3) *Ln_Rsq_Ind*	(4) *Ln_Rsq_Ind*
Sentiment	-0.020***	-0.106***	-0.013**	-0.100***
	(-3.04)	(-16.73)	(-1.99)	(-15.73)
*Sentiment*AnalystCover*	**-0.014*** **	**-0.011*** **	**-0.020**** **	**-0.016**** **
	(-2.89)	**(-2.47)**	**(-4.75)**	**(-4.18)**
AnalystCover	0.249***	0.076***	0.221***	0.082***
	(25.10)	(9.26)	(28.31)	(12.57)
Size		0.281***		0.278***
		(56.19)		(55.76)
Turnover		0.490***		0.481***
		(8.75)		(8.73)
Retstd		-1.005***		-0.996***
		(-4.15)		(-4.12)
Mktval		-0.200***		-0.203***
		(-28.86)		(-29.29)
Mktret		-0.086***		-0.089***
		(-5.34)		(-5.53)
Mktretvol		71.958***		71.441***
		(85.03)		(84.29)
NBER_Rec		-0.082***		-0.072***
		(-9.41)		(-8.12)
Constant	-1.694***	-0.488***	-1.694***	-0.442***
	(-191.02)	(-4.84)	(-215.42)	(-4.39)
Observations	149,964	146,881	149,964	146,881
R-squared	0.448	0.533	0.450	0.533

Table 9 Implied cost of equity capital (ICOC) and sentiment

This table presents the results of the regressions of five measures of ICOC on investor sentiment index. We estimate the following model

$$r\_AVG_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\gamma*X_{i,t}+\varepsilon_{i,t}$$

where $r\_AVG_{i,t}$ is the average of four measures of ICOC of stock i in year t. The four measures of ICOC are *r_GLS* from Gebhardt, Lee, and Swaminathan (2001), *r_CT* from Claus and Thomas (2001), *r_OJN* from Ohlson and Juettner-Nauroth (2005) and *r_PEG* from Easton (2004). A detailed description of these measures is provided in section 2.3 and Appendix C. $X_{i,t}$ are the same set of control variables used in Eq. (5). Stock-level regressions include firm fixed effect and standard errors are clustered at firm level. t-statistics are reported in parentheses.

Panel A: Stock-level regressions

	(1)	(2)	(3)	(4)	(5)
VARIABLES	*r_AVG*	*r_PEG*	*r_CT*	*r_OJN*	*r_GLS*
Sentiment	**-0.003*****	**-0.003*** **	**-0.003*** **	**-0.005*** **	**-0.002*** **
	(-16.53)	**(-9.59)**	**(-18.01)**	**(-14.00)**	**(-12.37)**
Size	-0.012***	-0.017***	-0.004***	-0.017***	-0.012***
	(-43.01)	(-39.32)	(-14.46)	(-37.49)	(-39.68)
Turnover	0.011***	0.010***	0.013***	0.012***	0.008***
	(4.40)	(3.17)	(5.06)	(3.47)	(4.04)
Retstd	0.138***	0.263***	0.088***	0.283***	-0.083***
	(7.50)	(8.96)	(4.95)	(9.22)	(-5.23)
Mktval	-0.005***	-0.003***	-0.003***	-0.006***	-0.010***
	(-12.69)	(-4.60)	(-6.57)	(-8.57)	(-24.19)
Mktret	0.004***	0.006***	-0.003***	0.003***	0.010***
	(5.93)	(5.25)	(-4.36)	(2.88)	(17.72)
Mktretvol	0.013	-0.221***	0.172***	-0.200***	0.300***
	(0.33)	(-3.62)	(4.83)	(-3.11)	(8.95)
NBER_Rec	0.003***	0.005***	0.001**	0.003***	0.003***
	(7.05)	(7.62)	(2.28)	(4.14)	(9.10)
Constant	0.275***	0.281***	0.158***	0.335***	0.325***
	(45.74)	(29.85)	(28.28)	(33.39)	(55.15)
Observations	50,807	50,807	50,807	50,807	50,807
R-squared	0.617	0.550	0.543	0.543	0.656

Panel B: Market-level regressions					
	(1)	(2)	(3)	(4)	(5)
VARIABLES	*r_AVG*	*r_PEG*	*r_CT*	*r_OJN*	*r_GLS*
Sentiment	**-0.004**	**-0.003**	**-0.006****	**-0.004**	**-0.004**
	(-1.64)	**(-0.92)**	**(-2.20)**	**(-1.51)**	**(-1.30)**
Mktval	-0.020***	-0.023***	-0.012***	-0.025***	-0.021***
	(-14.46)	(-12.56)	(-6.74)	(-13.89)	(-14.46)
Mktret	0.010	0.016	-0.004	0.016	0.012
	(1.18)	(1.56)	(-0.39)	(1.51)	(1.32)
Mktretvol	-0.147	-0.231	-0.225	-0.059	-0.073
	(-0.39)	(-0.57)	(-0.44)	(-0.16)	(-0.18)
NBER_Rec	0.008	0.009*	0.012	0.006	0.006
	(1.54)	(1.77)	(1.42)	(1.18)	(1.37)
Constant	0.431***	0.485***	0.283***	0.533***	0.423***
	(19.18)	(16.03)	(10.45)	(17.28)	(18.57)
Observations	34	34	34	34	34
R-squared	0.899	0.896	0.685	0.921	0.892

Table 10 The effect of short selling

This table presents the effect of investor sentiment on short selling activity and the effect of cost of short selling on the relation between stock return comovement and lagged investor sentiment. We estimate the following two models:

$$ShortIntensity_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\beta_2*X_{i,t}+\varepsilon_{i,t}$$

$$Ln\_Rsq\_Ind_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\beta_2*Sentiment_{t-1}*ShortCost_{i,t-1}+\beta_3*ShortCost_{i,t-1}+\beta_4*X_{i,t}+\varepsilon_{i,t}$$

Panel A reports the results of the first model. $ShortIntensity_{i,t}$ is the change of the natural logarithm of a stock's short interest ratio in Column 1 and 2 of Panel A and the change of the natural logarithm of a stock's lendable value in Column 3 and 4. The lendable value of a stock is scaled by the market value of that stock. $Sentiment_{t-1}$ is the average of the monthly BW sentiment index of year t-1. Panel B reports the results of the second model. $Ln\_Rsq\_Ind_{i,t}$ is the logistic-transformed R-square defined by Eq. (4) for stock i in year t. We use two measures of cost of short selling. The first one is $LoanFee\_high_{i,t-1}$ which equals 1 if the average loan fee of stock i in year t-1 is above the sample median and 0 otherwise. The second one is the $Lendable\_high_{i,t-1}$ which equals 1 if the average lendable value of stock i in year t-1 is above the sample median and 0 otherwise. Other control variables are defined in Appendix B. The sample period is from 2002 to 2014. In Column (1) and Column (2), we use $LoanFee\_high_{i,t}$ as our measure of loan fee. In Column (3) and Column (4), we use $Lendable\_high_{i,t}$. All the regressions include firm fixed effect and standard errors are clustered at firm level. t-statistics are reported in parentheses.

Panel A: Shorting activity and sentiment

	(1)	(2)	(3)	(4)
VARIABLES	*ShortInterest_diff*	*ShortInterest_diff*	*Lendable_diff*	*Lendable_diff*
Sentiment	0.153***	0.565***	-0.170***	-0.208***
	(9.06)	(20.44)	(-13.58)	(-11.89)
Size		0.100***		0.059***
		(6.11)		(5.57)
Turnover		0.910***		0.151**
		(8.68)		(1.99)
Retstd		4.696***		-1.182
		(5.25)		(-1.62)
Mktval		-2.798***		-3.085***
		(-54.16)		(-100.92)
Mktret		-0.081		-1.801***
		(-1.01)		(-40.10)
Mktretvol		-77.177***		-225.247***
		(-19.15)		(-79.20)
NBER_Rec		-0.850***		0.969***
		(-16.17)		(32.75)
Constant	0.308***	47.232***	0.498***	54.290***
	(597.98)	(55.48)	(1,485.49)	(105.28)
Observations	36,415	36,184	35,584	35,358
R-squared	0.142	0.290	0.220	0.572

Panel B: The effect short selling				
	(1)	(2)	(3)	(4)
VARIABLES	*Ln_Rsq_Ind*	*Ln_Rsq_Ind*	*Ln_Rsq_Ind*	*Ln_Rsq_Ind*
Sentiment	-0.299***	-0.328***	-0.065***	-0.126***
	(-21.59)	(-19.63)	(-2.62)	(-4.97)
*Sentiment*ShortCost*	**0.271*** **	**0.154*** **	**-0.219*** **	**-0.212*** **
	(9.69)	**(5.83)**	**(-7.70)**	**(-7.76)**
ShortCost	-0.157***	-0.047***	0.259***	0.105***
	(-11.64)	(-3.79)	(15.20)	(6.68)
Size		0.190***		0.190***
		(18.40)		(18.47)
Turnover		-0.074**		-0.088***
		(-2.30)		(-2.69)
Retstd		0.885*		0.982**
		(1.80)		(2.00)
Mktval		0.229***		0.200***
		(6.80)		(5.85)
Mktret		0.429***		0.440***
		(11.73)		(12.15)
Mktretvol		103.010***		103.373***
		(43.53)		(44.67)
NBER_Rec		-0.556***		-0.579***
		(-20.46)		(-21.44)
Constant	-0.923***	-7.115***	-1.144***	-6.713***
	(-192.35)	(-12.55)	(-104.36)	(-11.76)
Observations	39,490	39,229	39,469	39,208
R-squared	0.590	0.660	0.591	0.661

Table 11 The effect of institutional holding

This table presents the effect of investor sentiment on institutional trades and the effect of analyst coverage on the relation between stock return comovement and lagged investor sentiment. We estimate the following two models:

$$InstIntensity_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\beta_2*X_{i,t}+\varepsilon_{i,t}$$

$$Ln\_Rsq\_Ind_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\beta_2*Sentiment_{t-1}*InstHold_{i,t-1}+\beta_3*InstHold_{i,t-1}+\beta_4*X_{i,t}+\varepsilon_{i,t}$$

Panel A reports the results of the first model. $InstIntensity_{i,t}$ is the absolute value of the change in natural logarithm of institutional holdings of stock i in year t. $Sentiment_{t-1}$ is average of the monthly BW sentiment index within year t-1. Panel B reports the results of the second model. $Ln\_Rsq\_Ind_{i,t}$ is the logistic-transformed R-square defined by Eq. (4) for stock i in year t. $InstHold_{i,t-1}$ is the value of stock i held by institutions scaled by the market value of stock i in year t-1. Other control variables are defined in Appendix B. The sample period is from 1981 to 2014. All the regressions include firm fixed effect and standard errors are clustered at firm level. t-statistics are reported in parentheses.

Panel A: Institutional trades and investor sentiment

	(1)	(2)
VARIABLES	*InstIntensity*	*InstIntensity*
Sentiment	**0.043*** **	**0.027*** **
	(15.14)	**(9.11)**
Size		-0.016***
		(-5.12)
Turnover		0.103***
		(3.53)
Retstd		-0.043
		(-0.24)
Mktval		-0.085***
		(-22.00)
Mktret		0.063***
		(6.17)
Mktretvol		2.114***
		(3.98)
NBER_Rec		-0.026***
		(-4.31)
Constant	0.279***	1.688***
	(270.75)	(31.72)
Observations	107,279	106,527
R-squared	0.335	0.346

Panel B: The effect of institutional trades		
	(1)	(2)
VARIABLES	*Ln_Rsq_Ind*	*Ln_Rsq_Ind*
Sentiment	0.011	-0.050***
	(1.27)	(-6.17)
*Sentiment*InstHold*	**-0.086*** **	**-0.167*** **
	(-5.12)	**(-10.28)**
InstHold	0.835***	0.383***
	(32.98)	(19.22)
Size		0.275***
		(53.12)
Turnover		0.498***
		(8.59)
Retstd		-0.675***
		(-2.66)
Mktval		-0.224***
		(-31.54)
Mktret		-0.097***
		(-5.92)
Mktretvol		71.533***
		(83.80)
NBER_Rec		-0.086***
		(-9.75)
Constant	-1.774***	-0.170*
	(-186.79)	(-1.66)
Observations	146,755	144,755
R-squared	0.457	0.538

Table 12 Innovation output and investor sentiment

The table presents the results of the regressions of innovation output on investor sentiment. We estimate the following model

$$Innovation_{i,t} = \alpha_i + \beta_1 * Sentiment_{t-1} + \gamma * X_{i,t} + \varepsilon_{i,t}$$

where $Innovation_{i,t}$ is the innovation output of stock i in year t. We use the number of patents ($Patent_{i,t}$), the adjusted number of citations ($Cites_{i,t}$) and the value of the patents ($Patent\_Value_{i,t}$) as our innovation output measures. All the three measures are scaled by book assets. $X_{i,t}$ is a set of control variables including natural logarithm of book assets, return on assets, R&D expenditures, natural logarithm of firm age, sales growth rate and market-to-book ratio. All variables are defined in Appendix B. The sample period is from 1981 to 2005. All regressions include firm fixed effect and standard errors are clustered at firm level. t-statistics are reported in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	*Cites*	*Cites*	*Patents*	*Patents*	*Patent_value*	*Patent_value*
Sentiment	**0.235*** **	**0.202*** **	**0.227*** **	**0.175*** **	**0.538*** **	**0.613*** **
	(7.66)	**(6.31)**	**(8.65)**	**(6.78)**	**(12.44)**	**(13.18)**
Ln_AT		-0.355***		-0.479***		0.492***
		(-5.71)		(-8.02)		(3.28)
ROA		0.536***		0.193		1.162***
		(2.97)		(1.11)		(3.89)
R&D		14.769***		16.892***		28.244***
		(10.71)		(13.67)		(12.75)
Ln_Age		-0.449***		0.002		0.327***
		(-6.07)		(0.03)		(2.86)
Salegrow		-0.111**		-0.064		-0.224***
		(-2.04)		(-1.30)		(-2.79)
MB		0.159***		0.155***		1.096***
		(4.48)		(5.07)		(14.53)
Constant	1.780***	3.400***	1.874***	2.915***	3.336***	-2.760***
	(131.38)	(11.71)	(161.56)	(10.77)	(174.83)	(-3.49)
Observations	123,054	107,681	123,054	107,681	123,054	107,681
R-squared	0.514	0.534	0.557	0.584	0.631	0.666

Table 13 Innovation output and sentiment: subsample test

The table reports the effect of sentiment on innovation output in financially constraint firms. We estimate the following model

$$Innovation_{i,t}= \alpha_i+\beta_1*Senti_{t-1}+\beta_1*Sentiment_{t-1}*Constraint_{i,t-1}+Constraint_{i,t-1}+\gamma*X_{i,t}+\varepsilon_{i,t}$$

where $Innovation_{i,t}$ is the innovation output of stock i in year t. We use both the number of patents ($Patent_{i,t}$) and the adjusted number of citations ($Cites_{i,t}$) as our innovation output measures. We use three measures to identify financially constrained firms: $High\_HP_{i,t}$, $Low\_payout_{i,t}$ *and* $No\_divid_{i,t}$. $High\_HP_{i,t}$ is a dummy variable which equals 1 if the Hadlock-Pierce index of firm i is above the sample median HP in year t and 0 otherwise. $Low\_payout_{i,t}$ is a dummy variable which equals 1 if the payout ratio of firm i is below the sample median payout ratio in year t and 0 otherwise. $No\_divid_{i,t}$ is a dummy variable which equals 1 if firm i does not pay dividend in year t and 0 otherwise. $X_{i,t}$ are the same set of control variables used in Table 12. All regressions include firm fixed effect and standard errors are clustered at firm level. t-statistics are reported in parentheses.

Panel A: High_HP

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	*Cites*	*Cites*	*Patents*	*Patents*	*Patent_value*	*Patent_value*
Senti	0.235***	0.086***	0.227***	0.073***	0.538***	0.477***
	(7.66)	(3.94)	(8.65)	(4.62)	(12.44)	(9.07)
Senti High_HP*		**0.263*** **		**0.232*** **		**0.312*** **
		(3.70)		**(4.05)**		**(3.27)**
High_HP		0.208		0.170		0.627**
		(1.62)		(1.47)		(2.53)
Ln_AT		-0.306***		-0.438***		0.610***
		(-4.90)		(-7.15)		(3.67)
ROA		0.535***		0.193		1.140***
		(2.97)		(1.11)		(3.82)
R&D		14.733***		16.860***		28.204***
		(10.69)		(13.65)		(12.74)
Ln_Age		-0.416***		0.030		0.400***
		(-5.49)		(0.43)		(3.36)
Salegrow		-0.110**		-0.063		-0.221***
		(-2.03)		(-1.29)		(-2.76)
MB		0.162***		0.158***		1.101***
		(4.56)		(5.15)		(14.58)
Constant	1.780***	3.010***	1.874***	2.589***	3.336***	-3.739***
	(131.38)	(9.32)	(161.56)	(8.21)	(174.83)	(-3.90)
Observations	123,054	107,681	123,054	107,681	123,054	107,681
R-squared	0.514	0.534	0.557	0.585	0.631	0.666

Panel B: Low_Payout						
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	*Cites*	*Cites*	*Patents*	*Patents*	*Patent_value*	*Patent_value*
Sentiment	0.235***	-0.039*	0.227***	-0.018	0.538***	0.272***
	(7.66)	(-1.67)	(8.65)	(-0.77)	(12.44)	(4.07)
Sentiment Low_Payout*		**0.334*** **		**0.269*** **		**0.476*** **
		(6.61)		**(6.19)**		**(5.18)**
Low_Payout		-0.085		-0.135**		-0.292**
		(-1.23)		(-2.12)		(-2.03)
Ln_AT		-0.356***		-0.481***		0.486***
		(-5.74)		(-8.08)		(3.24)
ROA		0.562***		0.204		1.170***
		(3.07)		(1.15)		(3.89)
R&D		14.770***		16.874***		28.202***
		(10.71)		(13.65)		(12.71)
Ln_Age		-0.447***		0.005		0.333***
		(-6.04)		(0.07)		(2.91)
Salegrow		-0.110**		-0.063		-0.222***
		(-2.03)		(-1.28)		(-2.76)
MB		0.162***		0.157***		1.100***
		(4.54)		(5.11)		(14.53)
Constant	1.780***	3.455***	1.874***	3.013***	3.336***	-2.549***
	(131.38)	(11.84)	(161.56)	(11.04)	(174.83)	(-3.18)
Observations	123,054	107,624	123,054	107,624	123,054	107,624
R-squared	0.514	0.534	0.557	0.585	0.631	0.666

Panel C: No_Divid						
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	*Cites*	*Cites*	*Patents*	*Patents*	*Patent_value*	*Patent_value*
Sentiment	0.235***	-0.027	0.227***	-0.004	0.538***	0.278***
	(7.66)	(-1.25)	(8.65)	(-0.18)	(12.44)	(4.69)
Sentiment No_Divid*		**0.337*** **		**0.264*** **		**0.491*** **
		(6.57)		**(6.17)**		**(5.65)**
No_Divid		-0.153*		-0.209***		-0.244
		(-1.70)		(-2.60)		(-1.42)
Ln_AT		-0.358***		-0.485***		0.487***
		(-5.77)		(-8.13)		(3.23)
ROA		0.553***		0.203		1.186***
		(3.06)		(1.16)		(3.97)
R&D		14.737***		16.851***		28.193***
		(10.69)		(13.64)		(12.73)
Ln_Age		-0.445***		0.007		0.333***
		(-6.01)		(0.11)		(2.91)
Salegrow		-0.110**		-0.062		-0.222***
		(-2.03)		(-1.27)		(-2.77)
MB		0.162***		0.157***		1.100***
		(4.55)		(5.12)		(14.55)
Constant	1.780***	3.502***	1.874***	3.071***	3.336***	-2.593***
	(131.38)	(11.77)	(161.56)	(11.04)	(174.83)	(-3.20)
Observations	123,054	107,679	123,054	107,679	123,054	107,679
R-squared	0.514	0.534	0.557	0.585	0.631	0.666

Table 14 The effect of high-tech intensiveness

This table presents the effect of analyst coverage on the relation between stock return comovement and investor sentiment. We estimate the following model:

$$Ln\_Rsq\_Ind_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\beta_2*Sentiment_{t-1}*High\_Tech_i+\beta_3*High\_Tech_i+\beta_4*X_{i,t}+\varepsilon_{i,t}$$

where $Ln\_Rsq\_Ind_{i,t}$ is the logistic-transformed R-square defined by Eq. (4) for stock i in year t. $Sentiment_{t-1}$ is the average of the monthly BW sentiment index of year t-1. $High\text{-}Tech_i$ is a dummy variable which equals 1 if the high-tech intensiveness of the industry that firm i belongs to is above the sample median and 0 otherwise. Industry j's high-tech intensiveness is computed in the following way. Every year t, we use all firms in industry j to regress the natural logarithm of the ratio of market value to total assets on R&D expense scaled by total assets over the past five years. The coefficient on R&D expense scaled by total assets is the temporary high-tech intensiveness of industry j in year t. We take the time series median of industry j's temporary high-tech intensiveness as the high-tech intensiveness of industry j. Other control variables are defined in Appendix B. The sample period is from 1981 to 2005. All the regressions include firm fixed effect and standard errors are clustered at firm level. t-statistics are reported in parentheses.

VARIABLES	(1) *Ln_Rsq_Ind*	(2) *Ln_Rsq_Ind*
Sentiment	-0.034***	-0.109***
	(-4.00)	(-13.56)
*Sentiment*High_Tech*	**-0.028****	**-0.019***
	(-2.50)	**(-1.88)**
High_Tech	-0.073**	-0.021
	(-2.31)	(-0.88)
Size		0.283***
		(52.05)
Turnover		0.462***
		(8.10)
Retstd		-1.356***
		(-4.99)
Mktval		-0.222***
		(-30.25)
Mktret		-0.114***
		(-6.29)
Mktretvol		73.406***
		(76.99)
NBER_Rec		-0.066***
		(-6.75)
Constant	-1.384***	-0.058
	(-72.92)	(-0.54)
Observations	116,660	114,667
R-squared	0.432	0.528

Table A1 Stock return comovement and investor sentiment: more controls

This table presents the results of the regression of stock return comovement on investor sentiment. We estimate the following model:

$$Ln\_Rsq_{i,t}=\alpha_i+\beta_1*Sentiment_{t-1}+\beta_2*X_{i,t}+\varepsilon_{i,t}$$

Sktret, *l1_Stkret*, *l2_Stkret* is the annual return of stock i in year t, t-1 and t-2, respectively. *Mktret*, *l1_Mtkret*, *l2_Mtkret* is the annual market return in year t, t-1 and t-2, respectively. Other variables are defined in Table 3. All the regressions include firm fixed effect and standard errors are clustered at firm level. t-statistics are reported in parentheses.

	(1)	(2)
VARIABLES	*Ln_Rsq*	*Ln_Rsq_Ind*
Sentiment	**-0.266*** **	**-0.154*** **
	(-21.60)	**(-25.76)**
Size	0.544***	0.321***
	(34.99)	(34.57)
Stkret	-0.132***	-0.104***
	(-4.53)	(-4.81)
l1_Stkret	0.089***	0.032***
	(8.28)	(5.08)
l2_Stkret	0.085***	0.034***
	(9.68)	(6.80)
Turnover	0.780***	0.471***
	(5.63)	(7.76)
Retstd	-4.320***	0.085
	(-6.10)	(0.26)
Mktval	-0.476***	-0.202***
	(-26.89)	(-20.44)
Mktret	-0.099**	-0.141***
	(-1.96)	(-5.29)
l1_ Mktret	-0.580***	-0.383***
	(-12.86)	(-17.01)
l2_ Mktret	-0.774***	-0.384***
	(-23.97)	(-22.80)
Mktretvol	122.633***	75.173***
	(62.64)	(74.16)
NBER_Rec	-0.520***	-0.298***
	(-20.58)	(-23.06)
Constant	1.182***	-0.534***
	(4.92)	(-4.13)
Observations	113,478	112,221
R-squared	0.447	0.560

Table A2 Stock comovement and binary sentiment measure: stock level

This table presents the results of the regression of stock return comovement on investor sentiment. We estimate the following model:

$$Ln\_Rsq_{i,t}=\alpha_i+\beta_1*Senti\_high_{t-1}+\beta_2*X_{i,t}+\varepsilon_{i,t}$$

where $Ln\_Rsq_{i,t}$ is the logistic-transformed R-square defined by Eq. (2) for stock i in year t. $Senti\_high_{t-1}$ is a dummy variable which equals 1 if the sentiment index of year t-1 is above the sample median of the sentiment index and 0 otherwise. $Ln\_Rsq\_Ind_{i,t}$ is the logistic-transformed R-square defined by Eq. (4) for stock i in year t. Other control variables are defined in Appendix B. The sample period is from 1981 to 2014. In Model (1) and Model (2), we use $Ln\_Rsq_{i,t}$ of MYY (2000) as the dependent variable. In Model (3) and Model (4), we use $Ln\_Rsq\_Ind_{i,t}$ of Piotroski and Roulstone (2004) as the dependent variable . All the regressions include firm fixed effect and standard errors are clustered at firm level. t-statistics are reported in parentheses.

	(1)	(2)	(3)	(4)
VARIABLES	*Ln_Rsq*	*Ln_Rsq*	*Ln_Rsq_Ind*	*Ln_Rsq_Ind*
Senti_high	**-0.032*** **	**-0.132*** **	**-0.028*** **	**-0.085*** **
	(-2.93)	**(-12.02)**	**(-5.37)**	**(-16.64)**
Size		0.526***		0.292***
		(52.61)		(56.96)
Turnover		0.996***		0.551***
		(8.77)		(9.08)
Retstd		-5.829***		-1.049***
		(-10.51)		(-4.28)
Mktval		-0.478***		-0.193***
		(-35.05)		(-27.44)
Mktret		0.020		-0.064***
		(0.58)		(-4.00)
Mktretvol		117.087***		70.868***
		(70.42)		(84.32)
NBER_Rec		-0.213***		-0.079***
		(-11.94)		(-8.96)
Constant	-2.703***	1.188***	-1.473***	-0.596***
	(-457.19)	(6.02)	(-517.08)	(-5.82)
Observations	151,265	148,144	150,004	146,921
R-squared	0.358	0.426	0.439	0.531

Table A3 Stock comovement and binary sentiment measure: market-level

This table presents the results of the regression of stock return comovement on investor sentiment at the market level. We estimate the following model:

$$Ln\_MktRsq_t=\alpha+\beta_1*Senti\_high_{t-1}+\beta 2*X_t+\varepsilon_t$$

where Ln_MktRsq_t is the logistic-transformed market R-square defined by Eq. (8) in year t. $Senti\_high_{t-1}$ is a dummy variable which equals 1 if the sentiment index of year t-1 is above the sample median of sentiment index and 0 otherwise. $Ln\_MktRsq\_Ind_t$ is the logistic-transformed market R-square defined by Piotroski and Roulstone (2004) in year t. Other control variables are defined in Appendix B. The sample period is from 1981 to 2014. We use simple average in Model (1) and SST-weighted average R^2 in Model (2). We then replace MYY's R^2 with Piotroski and Roulstone (2004)'s R^2 and report the results in Model (3) and (4). t-statistics are reported in parentheses.

	(1)	(2)	(3)	(4)
VARIABLES	*Ln_MktRsq*	*Ln_MktRsq*	*Ln_MktRsq_Ind*	*Ln_MktRsq_Ind*
Senti_high	**-0.246***	**-0.243***	**-0.163***	**-0.150***
	(-1.77)	**(-1.90)**	**(-1.84)**	**(-1.83)**
Mktval	0.083	0.067	0.091*	0.079*
	(1.22)	(1.07)	(1.97)	(1.89)
Mktret	0.021	-0.193	-0.059	-0.212
	(0.04)	(-0.47)	(-0.20)	(-0.78)
Mktretvol	70.979**	82.994***	53.293**	62.266***
	(2.48)	(3.39)	(2.76)	(3.77)
NBER_Rec	0.033	0.141	0.045	0.129
	(0.11)	(0.57)	(0.22)	(0.76)
Constant	-3.719***	-3.698***	-3.081***	-3.081***
	(-3.73)	(-3.88)	(-4.60)	(-4.92)
Observations	34	34	34	34
R-squared	0.442	0.602	0.559	0.697

Appendix B: Variable definitions

Variables	Definitions	Source
Sentiment and R-squares		
Ln_Rsq	The logistic-transformed R-square of the regressions of Eq.(1)	CRSP
Ln_Rsq_Ind	The logistic-transformed R-square of the regressions of Eq.(3)	CRSP
Ln_MktRsq	The logistic-transformed market average R-square defined by Eq. (7)	CRSP
Senti_high	A dummy variable which equals 1 if the annual Baker and Wurgler (2006) sentiment index is above the sample median and 0 otherwise	BW (2006)
Sentiment	The average of monthly Baker and Wurgler (2006) sentiment index within a year	BW (2006)
Ln_ME	The natural logarithm of the year-end market capitalization (shrout*prc)	CRSP
Turnover	The average of monthly turnover rate within a year	CRSP
RetVol	The volatility of daily stock returns in a year	CRSP
Mktval	The year-end total market value (*totval*) for the whole market	CRSP
Mktret	The annual value-weighted market return	CRSP
Mktretvol	The volatility of daily value-weighted market returns within a year	CRSP
NBER_Rec	NBER based recession indicator	NBER
Sentiment and cost of equity capital (COC)		
r_GLS	COC defined by Gebhardt, Lee, and Swaminathan (2001)	Compustat, IBES
r_CT	COC defined by Claus and Thomas (2001)	Compustat, IBES
r_OJN	COC defined by Ohlson and Juettner-Nauroth (2005)	Compustat, IBES
r_PEG	COC defined by Easton (2004)	Compustat, IBES
r_AVG	The simple average of r_GLS, r_CT, r_OJN and r_PEG	Compustat, IBES
Sentiment and innovation output		
Patent	The number of patents a firm applies for scaled by total asset in a year	Kogan, Papanikolaou, Seru, and Stoffman (2017)
Cites	The adjusted number of citations of all patents a firm applies for scaled by total assets in a year	KPSS (2017)
Ln_AT	The natural logarithm of total assets (*at*)	Compustat
ROA	Income before extraordinary items plus ieferred income taxes minus preferred dividends *(ib+txdi-dvp)/at*	Compustat
R&D	R&D expense scaled by total assets *(xrd/at)*	Compustat
Ln_Age	The natural logarithm of the age of a firm	CRSP
Salegrow	The annual growth rate of sales	Compustat
MB	Market-to-book ratio *((at-ceq)+(csho*prcc_f))/at*	Compustat

High_HP	A dummy variable which equals 1 if the HP index (Hadlock and Pierce (2006)) of a firm is above the sample median and 0 otherwise	Compustat
Low_Payout	A dummy variable which equals 1 if the payout ratio *(dvpsx_f/epspx)* of a firm is below the sample median and 0 otherwise	Compustat
No_divid	A dummy variable which equals 1 if a firm does not pay dividend in a year and 0 otherwise	Compustat
Sentiment and news commonality		
NewsR^2	The logistic-transformed R-squares of regressions of weekly news sentiment scores of a firm on weekly market news sentiment scores in a year	RavenPack

Appendix C: Construction of implied cost of equity capital

We follow Hail and Leuz (2006) to construct four measures of COC and use their simple average as our main measure of COC. We provide detailed descriptions of the four models used to construct the four measures in this section. The variables used in the models are defined as follows:

> p_t: the stock price of a firm at time t
>
> eps_t: the expected future earnings per share of a firm for year t
>
> k: the average of payout ratios over the three years before year t
>
> d_t: the expected dividend per share for year t, measured as last-year dividend per share
>
> g_l_t: long-term earnings growth forecast from IBES at time t
>
> g_t: the economic growth rate, measured as the annualized median of one-year-ahead monthly inflation rates
>
> bv_t: the book value of equity per share from Compustat for year t. For the years beyond the initial year, we estimate book value of equity per share as $bv_{t+1}=bv_t+eps_{t+1}-eps_{t+1}*k$

To be included in our sample, an observation should have nonmissing values on p_t, eps_{t+1}, eps_{t+2} and either eps_{t+3} through eps_{t+5} or an estimate of long-term earnings growth. We also require $eps_{t+2} > eps_{t+1}$. To ensure the financial information is publicly available at the time of estimating COC, we measure stock prices and analyst forecasts as of month +6 subsequent the fiscal year end. Therefore, the one-period-ahead earnings forecast of a firm, eps_{t+1}, is six months prior to its fiscal year-end.

1. Easton (2004) Model

$$p_t = (eps_{t+2} + r\_PEG \times d_{t+1} - eps_{t+1}) / r\_PEG^2$$

2. Claus and Thomas (2001) Model

$$p_t = bv_t + \sum_{\tau=1}^{5} \frac{eps_{t+\tau} - r\_CT \times bv_{t+\tau-1}}{(1+r\_CT)^{\tau}} + \frac{(eps_{t+5} - r\_CT \times bv_{t+4})(1+g_t)}{(1+r\_CT)^5 (r\_CT - g_t)}$$

Where bv_t is from Compustat and $bv_{t+\tau}$ is estimated as $bv_{t+\tau-1}+eps_{t+\tau}- eps_{t+\tau}\times k$. If eps_{t+3} through eps_{t+5} are missing, then we use the following equation to estimate them: $eps_{t+\tau}=eps_{t+\tau-1}\times(1+g\_l_{t+\tau-1})$.

3. Ohlson and Juettner-Nauroth (2005) Model

$$p_t = \frac{eps_{t+1}}{r\_OJN} + \frac{eps_{t+2} - eps_{t+1} - r\_OJN(eps_{t+1} - d_{t+1})}{r\_OJN(r\_OJN - g_t)}$$

4. Gebhardt, Lee,and Swaminathan (2001) Model

$$p_t = bv_t + \sum_{\tau=1}^{3} \frac{eps_{t+\tau} - r\_GLS \times bv_{t+\tau-1}}{(1+r\_GLS)^{\tau}} + \sum_{\tau=4}^{11} \frac{(\hat{ROE}_{t+\tau} - r\_GLS) \times bv_{t+\tau-1}}{(1+r\_GLS)^{\tau}} + \frac{(\hat{ROE}_{t+12} - r\_GLS) \times bv_{t+11}}{(1+r\_GLS)^{11} \times r\_GLS}$$

We use the expected future earnings per share for the initial three years. Beyond the initial three years, the forecasted return on equity is assumed to be decreasing linearly over the next nine years to the sector-specific median return on equity over the past five years. If the sector-specific median return on equity is negative, we replace it with the country-year median return on equity.