What explains the dispersion effect? Evidence from institutional ownership

Hwang, Chuan Yang; Wong, Kit Pong; Yi, Long

Published in:
Pacific Basin Finance Journal

DOI:
10.1016/j.pacfin.2021.101698

Published: 01/02/2022

Document Version:
Peer reviewed version

What Explains the Dispersion Effect?
Evidence from Institutional Ownership

Chuan-Yang Hwang
Nanyang Business School
Nanyang Technological University

Kit Pong Wong
School of Economics and Finance
The University of Hong Kong

Long Yi
Finance and Decision Sciences
Hong Kong Baptist University

This paper conducts a joint test of two plausible explanations (difference-in-opinion vs. analyst self-censoring) for why stocks with higher dispersion in analysts' earnings forecasts earn lower subsequent returns (the dispersion effect). We exploit exogenous variations in institutional ownership generated by the annual index reconstitution to address the endogeneity concern of institutional ownership. We find results strongly suggest that analyst self-censoring rather than the more popular difference-in-opinion story is the more plausible explanation for the dispersion effect, at least in a sample where the endogeneity bias of institutional ownership is minimized.

1. Introduction

The dispersion effect refers to the intriguing anomaly in asset pricing studies first documented in Diether, Malloy, and Scherbina (henceforth DMS) (2002)[1] in which stocks that have higher dispersion in analysts' earnings forecasts – and hence are presumably riskier – earn much lower future returns. DMS (2002) propose two plausible explanations to account for the dispersion effect: (i) the difference-in-opinion explanation, and (ii) the self-censoring explanation. The first explanation postulates that forecast dispersion is a proxy for different opinions among investors. Short-sale constraints prohibit pessimistic investors from trading, so stock prices reflect only the optimistic views (Miller, 1977). This induces a greater upward bias in the prices of stocks with higher forecast dispersion, which in turn results in lower future returns. The second explanation postulates that the incentive structure of analysts induces them to self-censor unfavorable earnings forecasts to please managers (McNichols and O'Brien, 1997; Scherbina, 2008). DMS (2002) show that when analysts' forecasts are more dispersed, there are more pessimistic forecasts to self-censor. Thus, there is a greater upward bias in the consensus earnings forecasts of stocks with higher dispersion in analysts' forecasts. If investors do not take the upward bias into account when making their investment decisions, they tend to overvalue stocks with higher forecast dispersion, which, in turn, leads to lower future returns for these stocks.

The literature thus far has emphasized the difference-in-opinion explanation for two reasons. First, the seminal work of DMS (2002) focuses on this story even though the evidence in their paper also supports the self-censoring explanation. Second, it is difficult to obtain data that contain analysts' incentives, while it is relatively easy to obtain proxies for short-sale constraints, which

[1] Richardson, Tuna, and Wysocki (2010) show that DMS (2002) ranks third among the most frequently cited research papers on financial anomalies published since 2000.

is crucial to the tests of the difference-in-opinion explanation. The empirical tests of the difference-in-opinion explanation have yielded mixed results. For example, Nagel (2005) uses institutional ownership as a proxy for short-sale constraint. He finds that the dispersion effect is more pronounced in a subsample of stocks with low institutional ownership, supporting the difference-in-opinion story. Boehme, Danielsen, and Sorescu (2006) also find that stocks that are subject to high short-sale constraints and high forecast dispersion are more likely to be overvalued. However, Avramov et al. (2009) show that proxies for short-sale costs do not capture the dispersion effect. Scherbina (2008) is one of the few researchers to focus on testing the self-censoring explanation. She estimates the extent of self-censoring based on the proportion of analysts who stop revising their annual earnings forecasts and finds that this measure predicts negative earnings surprises and lower future returns. In an international setting, Hwang and Li (2018) find the dispersion effect to be stronger in countries where the demand for analysts' services is stronger, which they interpret as supporting the self-censoring explanation.

This paper aims to disentangle the difference-in-opinion explanation from the self-censoring explanation using institutional ownership in the U.S. market.[2] The idea is based on two salient features of institutional ownership. First, institutional ownership is commonly used as a proxy for short-sale constraints in the literature (Nagel, 2005; Asquith, Pathak, and Ritter, 2005). Stocks with high institutional ownership are less subject to short-sale constraints, as institutional investors are important suppliers of shares for short selling. Thus, stocks with *higher* institutional ownership will exhibit a *weaker* dispersion effect if the underlying reason for the dispersion effect is

[2] Other explanations for the dispersion effect include that of Johnson (2004), who argues that forecast dispersion is a proxy for idiosyncratic parameter risk. Because equity is a call option on firm's assets, in the presence of leverage, expected return decreases with dispersion. Avramov et al. (2009) find that leverage is not relevant for the dispersion effect and argue that the dispersion effect is a manifestation of the negative relationship between default risk and return.

difference-in-opinion. Second, institutional ownership is used as a proxy for demand for analysts' services (Frankel, Kothari, and Weber, 2006), which affects analysts' incentives to self-censor. Like Scherbina (2005), Hwang and Li (2018) model a situation in which analysts receive unfavorable signals about a firm's upcoming earnings, and the incentive structure of analysts induces them either to adjust their forecasts upward at a cost of losing reputation, or to issue no forecasts (i.e., self-censor). They show that when analysts face greater demand for their services, they find the reputational cost of adding an optimistic bias to be higher, which makes them more likely to self-censor unfavorable forecasts. As self-censoring induces a greater upward bias to the consensus forecast than adding an optimistic bias does, stocks with greater demand for analysts' services will have a stronger dispersion effect. Since the demand for analysts' services increases with institutional ownership, stocks with *higher* institutional ownership will exhibit a *stronger* dispersion effect if analyst self-censoring explains the dispersion effect.

We start with a general sample of stocks in the Center for Research in Securities Prices database (CRSP) and confirm the findings in Nagel (2005) and Boehme, Danielsen, and Sorescu (2006) that the dispersion effect is stronger among stocks with lower institutional ownership in the general sample. These results seem to support the popular difference-in-opinion explanation in the literature, as low institutional ownership is associated with more binding short-sale constraints. Institutional ownership, however, can be endogenous in nature. As institutional investors are generally considered more informed, stocks with bad news and hence low future returns would attract fewer institutional investors and hence have low institutional ownership. This effect would be more significant in stocks with greater information asymmetry or uncertainty[3]. Given that

[3] Bushee and Goodman (2007) find that informed trading by institutional investors is more pronounced towards small stocks. Small stocks are more likely to have high information asymmetry as the authors noted that large stocks are with abundant public information. In a cross-country study, Maffett (2012) provides evidence that institutional

dispersion is often considered as a proxy for information asymmetry or uncertainty (Abarbanell, Lanen, and Verrecchia, 1995; Barron et al., 1998), one would expect high dispersion stocks with low institutional ownership to be dominated by those with bad news and low future returns. As it turns out, this group of stocks is mainly responsible for the dispersion effect among low institutional ownership stocks. Therefore, the results we obtained by comparing dispersion effects among stocks with high and low institutional ownership may be spurious, as they could simply arise from comparing stocks with good and bad news in conjunction with the information advantage of institutional investors regarding stocks with high information asymmetry, which has nothing to do with the two explanations we try to disentangle.

To address this problem, following Boone and White (2015) and Crane, Michenaud, and Weston (2016), we utilize the annual Russell 1000/2000 (a.k.a. Russell 3000) index reconstitution to obtain a sample of stocks with variations in institutional ownership that are not the active choice of institutional investors, and with minimum variation in other characteristics.[4] Stocks in the Russell sample differ in institutional ownership as a result of different index weights. We then compare the dispersion effect of stocks in the bottom of the Russell 1000 index that have low institutional ownership with that of stocks in the top of the Russell 2000 index that have high institutional ownership. The results are the exact opposite of those we found using the general sample. In the portfolio strategy, we find a significant dispersion effect for stocks with higher institutional ownership, but no dispersion effect for stocks with lower institutional ownership.

investors' information advantage is stronger in opaque firms with high information asymmetry. In studying the informational role of local institutional investors, Baik, Kang, and Kim (2010) document the information advantage of local institutional investors particularly strong towards stocks with high information asymmetry.

[4] Note that we are not doing a regression discontinuity study. Rather, we rely on Russell index reconstitution to construct two samples of firms that mainly differ in institutional ownership as a result of an event that is exogenous to the choice of institutional investors. In particular, our study only requires the difference in institutional ownership not being the active choice of institutional investors based on return information and dispersion.

Similarly, in regressions, dispersion negatively predicts returns only when institutional ownership is high. These results suggest that after we address the potential inference problem caused by the endogenous nature of institutional ownership, the dispersion effect is actually stronger for stocks with *higher* institutional ownership. Given the difference-in-opinion explanation would predict a stronger dispersion effect for stocks who have more binding short-sale constraint (lower institutional ownership) while analyst self-censoring explanation predicts a stronger dispersion effect for stocks whose demand for analyst service is high (institutional ownership is high), these findings clearly refute the difference-in-opinion in favor of the analyst self-censoring story for the explanation of the dispersion effect.

One may have concern about whether the Russell Index reconstitution indeed affects the demand for analyst service --- a key requirement in the self-censoring story, if the institutional ownership variation induced by the reconstitution is mainly due to passive investors. Fortunately, this concern is unwarranted. Boone and White (2015) show that quasi-indexers (Bushee, 1998), who account for a major part of the institutional ownership variation around the index threshold, despite of being passive in nature, have demand for information including those provided by analysts for reasons that we will explain in later section.

Since the underlying source of the dispersion effect under the self-censoring explanation is the greater bias in analysts' forecasts associated with high dispersion firms (positive dispersion-bias relationship), our earlier discussions also imply that the positive dispersion-bias relationship will be stronger in stocks with *higher* institutional ownership. Again, the potentially endogenous nature of institutional ownership can bias the test inference. The information advantage of institutional investors makes them less likely to hold stocks with unfavorable prospects, thus stocks with low institutional ownership are more likely to have unfavorable prospects, especially for high dispersion stocks (i.e., stocks with high information asymmetry). This – combined with analysts' incentive to self-censor *only* unfavorable forecasts – should lead us to find the dispersion-bias relationship to be *stronger* among stocks with *low* institutional ownership. This is indeed the case. In the general sample, the coefficient on dispersion in a regression where bias is the dependent variable is 0.008 using low institutional ownership stocks, while that for stocks with high institutional ownership is lower at 0.003. This finding indicates a stronger dispersion-bias relationship among stocks with low institutional ownership that might have led us to reject the self-censoring explanation. However, once we avoid the potential endogeneity problem of institutional ownership, we find the opposite. In the Russell sample, the coefficient in the regression with low institutional ownership stocks is 0.003, while that for stocks with high institutional ownership is *higher* at 0.012 and statistically much more significant. These results provide further support to the self-censoring explanation, for the correlation between dispersion-bias and institutional ownership is a prediction unique to the self-censoring explanation; the difference-in-opinion explanation has no prediction on dispersion-bias relationship.

We contribute to four strands of literature. First, we contribute to research on the dispersion effect by showing that using institutional ownership as the proxy for short-sale constraint (e.g., Nagel, 2005; Asquith, Pathak, and Ritter, 2005) can yield misleading results that support the difference-in-opinion hypothesis, because institutional ownership may be endogenous. Using the Russell sample, which is devoid of the endogeneity concern, we find opposite results that support the self-censoring explanation. We are quick to add that our results in no way refute the argument that institutional ownership is a good proxy of short-sale constraint, which in turn prevents the overpricing from being arbitraged away. However, our results do suggest that concern about the possible endogeneity of institutional ownership is warranted at least in the tests of dispersion effect, and that the less popular self-censoring story deserves more attention than it has received. This is consistent with the recent international study of Hwang and Li (2018). They conclude that the self-censoring explanation is more plausible because the dispersion effect is stronger in countries with greater demand for analysts' services.

Second, we contribute to the literature that uses the Russell index reconstitution as a source of exogenous variation of institutional investors and studies its impacts on firm behaviors. Most of this literature focuses on issues in corporate finance. For example, Boone and White (2015) use it to identify the impact of institutional investors on firm transparency; Crane, Michenaud, and Weston (2016) on pay out policy; Appel, Gormley, and Keim (2016a) on corporate governance; Chen, Dong and Lin (2019) on CSR activities; Chang, Lin, and Ma (2015) on mergers and acquisitions. Our study highlights the importance of taking care of the endogenous nature of institutional ownership in the study of asset pricing anomalies.

Third, we contribute to the literature on analysts' incentives. McNichols and O'Brien (1997) find that analysts are reluctant to issue unfavorable earnings forecasts because of the fear of jeopardizing the investment banking business. O'Brien, McNichols, and Lin (2005) document that investment banking ties reduce the speed with which analysts convey unfavorable news. Ljungqvist et al. (2007) find that analysts' recommendation relative to the consensus is positively associated with investment banking relationships, but institutional investors can moderate the positive effect. Like O'Brien, McNichols, and Lin (2005), we also show that the incentive structure of analysts is significant enough to distort information production in financial markets. However, in contrast to the moderating effect of institutional ownership documented by Ljungqvist et al. (2007), we find that analysts' incentives can be exacerbated by institutional ownership as predicted by Hwang and Li (2018).

Lastly, ours contributes to the literature on asset pricing anomalies broadly. Hou, Xue, and Zhang (2015) use an investment approach and find most anomalies can be explained by their q-factor when one focuses on the broad cross section, i.e., disregards small stocks. In a similar vein, in studying most recent documented asset pricing anomalies, Harvey, Liu, and Zhu (2016) show that an anomaly should have t-statistics higher than 3. Ours contributes to this broad literature by disentangling the underlying mechanism of a documented anomaly.

2. Data

We obtain analysts' forecast data from the Institutional Brokers Estimate System (I/B/E/S). Dispersion (*DISP*) in analysts' earnings forecasts is calculated each month as the ratio of the standard deviation of analysts' current fiscal-year annual earnings-per-share forecasts to the absolute value of the mean forecast. Analysts' forecasts are adjusted historically for stock splits in the standard issue of I/B/E/S data, which renders these data unsuitable for the analysis of forecast dispersion. Following DMS (2002), we thus use the raw (unadjusted) data reported in the I/B/E/S Summary History file. Forecast bias (*BIAS*) is defined as the difference between analysts' consensus earnings-per-share forecast in the current month and the corresponding actual earnings-per-share announced in the future, scaled by current month stock price.

We have two samples in this study. The general sample covers all NYSE, AMEX, and NASDAQ stocks from 1984 to 2006.[5] The general sample is divided into five groups by institutional ownership. The second sample consists of 100 stocks in the bottom of the Russell 1000 index and another 100 stocks in the top of the Russell 2000 index each year. We obtain

[5] This is to ensure that our sample period is the same as in the Russell sample, which is unsuitable for our type of research after 2006 due to policy changes that limit the movement of stocks across indexes (Cf. Boone and White (2015) and Crane, Michenaud, and Weston (2016).)

Russell 1000/2000 index constituents and index-weight information from Russell Investment for 1984 to 2006. Stocks in the bottom of the Russell 1000 index have lower institutional ownership than those in the top of the Russell 2000 because index assignment leads to different index weights.

We obtain quarterly institutional ownership data from Thomson Reuters, which keeps track of the 13-F filings of professional money managers (institutional investors). Institutional investors with investment discretion of $100 million or more are required to file Form 13-F with the SEC within 45 days of the end of each calendar quarter to disclose the number of company shares they hold. These institutional investors include investment advisers, banks, insurance companies, broker-dealers, pension funds, etc. Institutional ownership is calculated as the ratio of the total shares held by institutional investors to total shares outstanding. We use the latest quarterly filing to calculate institutional ownership for the general sample. For example, we use the 13-F filing of previous December to calculate the institutional ownership for January, February and March of this year. While for the Russell data, institutional ownership is determined by index membership. Whenever we need to report the institutional ownership for the Russell sample for a given year, we use the average of the institutional ownership based on the four quarterly filings from September to June next year because Russell reconstitution takes effect each June.

We obtain monthly stock return data for NYSE, AMEX, and NASDAQ stocks from the Center for Research in Securities Prices (CRSP) database, and we trim 1% of the return on each tail to reduce the impact of data error on our results. Stock-level characteristics are obtained from the CRSP-COMPUSTAT merged database.

Table 1 lists the summary statistics for main variables in this study for both the general sample and the Russell sample, separated by institutional ownership[6]. In the general sample, the means of dispersion, bias and analyst coverage are much larger when institutional ownership is low. The differences in the mean are statistically significant, which suggests that institutional ownership can be endogenous. After we control for the endogeneity of institutional ownership in the Russell sample, these differences become insignificant. These patterns are consistent with that in Panel A, Table 4 of Boone and White (2015), who show that for a bandwidth of 100, there is no significant differences in analyst coverage and forecast dispersion six months after index reconstitution.

[Insert Table 1 Here]

3. The Dispersion Effect: The Role of Institutional Investors

3.1 Hypothesis Development

Institutional ownership is commonly used as a proxy for short-sale constraint in the literature (Nagel, 2005; Asquith, Pathak, and Ritter, 2005). Stocks with higher institutional ownership are less subject to short-sale constraints as institutional investors are important suppliers of shares to borrow for short selling. Since short-sale constraints are the driving force that explains the dispersion effect in the difference-in-opinion story, one would expect the dispersion effect to be *weaker* among stocks with *higher* institutional ownership.

Institutional ownership can also serve as a proxy for the demand for analysts' services (Frankel, Kothari, and Weber, 2006). In the model of Hwang and Li (2018), when analysts receive unfavorable signals about a firm's upcoming earnings, the incentive structure of analysts induces

[6] Detailed definitions of variables can be found in Appendix A.

them to either adjust their forecasts upward at a cost of losing reputation or choose to issue no forecasts (i.e., self-censor) at another cost that has nothing to do with loss of reputation. As analysts' reputational concerns increase with the demand for their services (Barniv, Myring, and Thomas, 2005), Hwang and Li (2018) show that analysts whose services are in greater demand find that the cost of reputational loss from adding an optimistic bias is higher, and therefore they choose to self-censor more. They further show that although both self-censoring and adding optimistic bias would induce upward bias that increases with analysts' forecast dispersion (the positive dispersion-bias relationship), the effect coming from the former is stronger. Thus, for stocks with *higher* institutional ownership, there would be a *stronger* positive dispersion-bias relationship and hence a stronger dispersion effect. Note that the difference-in-opinion explanation has no prediction regarding the dispersion-bias relationship. We therefore establish the following two competing hypotheses in which institutional ownership exerts precisely the opposite influence on the dispersion effect.

Hypothesis 1 (Difference-In-Opinion): If the difference-in-opinion explanation drives the dispersion effect, stocks with *high (low)* institutional ownership will exhibit a *weaker (stronger)* dispersion effect, and there is no prediction regarding the dispersion-bias relationship.

Hypothesis 2 (Analyst Self-censoring): If the self-censoring explanation drives the dispersion effect, stocks with *high (low)* institutional ownership will exhibit a *stronger (weaker)* dispersion effect. In addition, there will be a positive dispersion-bias relationship, which is stronger (weaker) among high (low) institutional ownership stocks.

Note that the predictions of Hypothesis 1 and Hypothesis 2 are exact opposites of each other. Thus, these two hypotheses are mutually exclusive: if Hypothesis 1 is supported by the data, Hypothesis 2 would automatically be rejected and vice versa.

3.2 Institutional Ownership and the Dispersion Effect: General Sample

We first investigate whether and how institutional ownership plays a role in determining the strength of the dispersion effect using the general sample.

3.2.1 Portfolio Strategy

At the end of each month, we sort stocks (with a price of over five dollars) in the general sample into quintiles (D1 to D5) based on that month's analysts' forecast dispersion. D1 contains stocks with the lowest dispersion, while D5 consists of stocks with the highest dispersion. The portfolios are rebalanced each month. At the same time, we sort stocks independently into five quintiles (I1 to I5) each month based on institutional ownership. I1 is the group of stocks with the lowest institutional ownership, and I5 is the group of stocks with the highest institutional ownership. We end up with 25 (5X5) portfolios each month. We label our portfolios as I*D*. For example, I1D5 is a portfolio of stocks with the lowest institutional ownership (I1) and the highest dispersion in analysts' forecasts (D5). The dispersion effect of group I* is captured by a hedging portfolio that holds a long position in I*D1 and a short position in I*D5. Monthly portfolio return is calculated as the equal-weighted average of the returns of portfolio stocks.

Before presenting the portfolio results, we first verify that institutional ownership is indeed associated with less binding short-sale constraints. Short-sale constraint is defined as the supply of shares available for shorting times 1000 scaled by the total number of shares outstanding for each stock in each month.[7] The short-sale data are available from June 2002 onwards. A higher value means less binding short-sale constraints. For each group formed by institutional ownership (I1 to I5), we calculate the monthly median level of short-sale constraints and report the time-series average of the median. Table 2 reports the results. We see that short-sale constraints are most binding for stocks in I1 and least binding for those in I5. The difference in short-sale constraints is statistically significant at 1% with a t-stat of 4.36. We also present the average of institutional ownership. Not surprisingly, the difference in institutional ownership is large in magnitude and significant between stocks in I1 and I5.

[Insert Table 2 Here]

We next present our 5X5 portfolio strategy results. Table 3 reports the results. In Panel A of the table, we first verify the dispersion effect using all stocks in the general sample. Consistent with DMS (2002), we find a strong dispersion effect, represented as D1-D5>0. Relative to stocks with low dispersion (D1), high dispersion stocks (D5) earn significantly lower returns whether they are measured as raw returns or as risk-adjusted returns (alpha) from the CAPM, the Fama-French three-factor model (FF3), or the Carhart four-factor model (FF4). Panel D presents the average number of stocks in each portfolio. Panel B presents our main results. We see that the dispersion effect is much stronger in stocks with low institutional ownership. The hedging portfolio monthly return decreases monotonically from 1.41% for I1 (long in I1D1, short in I1D5) to the insignificant 0.53% for I5 (long in I5D1, short in I5D5) in terms of raw returns. The results are similar after adjustment by factor models including CAPM, FF3, and FF4. These results are consistent with the findings of Nagel (2005) and Boehme, Danielsen, and Sorescu (2006). It is worth pointing out

[7] We are grateful to Chi-Shen Wei for sharing the data.

that these results are mainly driven by the much lower returns of high dispersion stocks with low institutional ownership relative to those of high dispersion stocks with high institutional ownership. For example, measured in FF3 alpha, the former is lowered than the latter by 1.09% (0.50%+0.59%), which is close to the difference in the dispersion effect between low institutional stocks and high institutional stocks, 0.83% (1.39%-0.56%).

As stated above, Hypothesis 1 and Hypothesis 2 are mutually exclusive when it comes to the test of dispersion effect; thus, we focus our tests on Hypothesis 1, which predicts that I1D1-I1D5>I5D1-I5D5. To this end, we set up a test with the null H_0: I1D1-I1D5 $\leq$ I5D1-I5D5 and the alternative H_A: I1D1-I1D5>I5D1-I5D5, so that when the null is rejected, we can accept the alternative and conclude that the data support Hypothesis 1 (and reject Hypothesis 2). Note that since the alternative is specified as a strict inequality, this is a one-tailed test (cf. Mann 2010, pp. 387). Test results reported in Panel C reveal that the null is rejected with a p-value of <0.01 for all return measures, indicating that Hypothesis 1 is supported by the data; hence difference-in-opinion is more plausible than self-censoring in explaining the dispersion effect in the general sample. This is again consistent with Nagel (2005) and Boehme, Danielsen, and Sorescu (2006).

[Insert Table 3 Here]

3.2.2 Regression Approach

In addition to the portfolio approach presented above, we utilize panel regressions[8] to check the impact of dispersion on stock returns when controlling for other stock-level characteristics. The

[8] We run panel regression instead of the Fama-MacBeth regression for the following reason: 1) in the Russell sample, the number of the stock cross-section is small, 2) Peterson (2009) shows that with sufficient time clusters, estimation in panel setting with standard errors clustered by time performs equally well. We estimate standard error clustered by both time and stock to further control for stock effect, if any.

following regression model is performed on stocks in the general sample with the lowest level of institutional ownership (I1) and on stocks with the highest level of institutional ownership (I5).

$$RET_{i,t+1} = \beta_0 + \beta_1 DISP_{i,t} + \beta_2 LOGMV_{i,t} + \beta_3 LOGBM_{i,t} + \beta_4 MOM_{i,t} + \beta_5 ROE_{i,t} + \beta_6 INV_{i,t} + \beta_7 COV_{i,t} + \mu_{t+1} + \varepsilon_{i,t+1}. \quad (1)$$

$RET_{i,t+1}$ is the return of stock *i* in month *t+1*. $DISP_{i,t}$ is the dispersion in analysts' forecasts at month *t*. Control variables, including log of market capitalization, log of book-to-market ratio, momentum, return on equity, investment, and analyst coverage, are defined in appendix A. We include month fixed effects in the model as well. Standard errors are double clustered at the stock and month level.

Panel A of Table 4 reports the regression results separately for subsamples consisting of stocks with low and high institutional ownership.

[Insert Table 4 Here]

The coefficient on *DISP* (β_1), reported in Column (1) is significantly negative. This suggests a significant dispersion effect after we control for the impact of size, book-to-market, momentum, return on equity, investment, and analyst coverage in the sample of stocks with low institutional ownership. In contrast, the coefficient estimate on *DISP* reported in Column (2) is not significantly different from zero, which suggests that there is no significant dispersion effect in the high IO sample. To test if the dispersion effect in the high IO sample is stronger than that in the low IO sample (i.e., to test Hypothesis 1), we specific the null as H_0: Low IO $\beta_1 \geq$ High IO β_1 and the alternative as H_A: Low IO $\beta_1 <$ High IO β_1 for a one-sided test in the regression setting. The test result reported in Panel C of Table 4 indicates that the null can be rejected with a p value of

0.021 and we can accept the alternative that the dispersion effect is stronger in the low IO sample, which again supports Hypothesis 1 that difference-in-opinion is the more plausible driver of the dispersion effect.

4. The Dispersion Effect: The Russell Sample

4.1 Bias from Endogenous Institutional Ownership

Institutional ownership is likely endogenous in nature. This is because institutional investors often have an information advantage over ordinary investors, so stocks with low future returns (bad news) would attract fewer institutional investors hence have low institutional ownership. This effect would be more significant in stocks with greater information asymmetry or uncertainty, for which analyst forecast dispersion is a good proxy. Consequently, the endogeneity of institutional ownership and the fact that dispersion is a good proxy for information uncertainty can result in a large return spread between high dispersion stocks in the low institutional ownership quintile and high dispersion stocks in the high institutional ownership quintile. As clear from what we noted in an earlier section, this large return spread drives the result that the dispersion effect is stronger in the low institutional ownership quintile than in the high institutional ownership quintile. In other words, any results that we obtain by comparing dispersion effects among stocks with high and low institutional ownership may simply arise from comparing stocks with good and bad news, and/or from institutional investors' information advantage, neither of which has anything to do with the two explanations we are trying to disentangle.

To address this endogeneity in institutional ownership, we make use of the exogenous variation in institutional ownership resulting from the annual Russell 3000 index reconstitution. In what follows, we first introduce the Russell 3000 index and then describe how it generates the exogenous variation in institutional ownership. Finally, we test our hypotheses using the Russell sample.

4.2 Background on the Russell 3000 Index

Russell Investment constitutes the Russell 3000 index comprising the Russell 1000 and 2000 indexes each year starting from 1984. Russell Investment ranks stocks traded in the U.S. by their end-of-May market capitalization in descending order to determine the membership of each index. Stocks ranked between 1 and 1000 are assigned to the Russell 1000 index. Stocks ranked between 1001 and 3000 are assigned to the Russell 2000 index. After determining the membership, Russell Investment first derives float-adjusted capitalization by making free-float adjustment to its market capitalization, and then calculates the index weight of stocks by their end-of-June float-adjusted market capitalization within each index. Following Boone and White (2015), we use Russell-assigned index-weight ranking in this paper as it is most relevant for institutional ownership.[9,10] The Russell 1000 and 2000 indexes are value-weighted. Consequently, stocks in the bottom of the

---

[9] Appel, Gormley, and Keim (2016b) point out using Russell-assigned index-weight ranking can cause a jump in float-adjusted market capitalization at the threshold, which violates the underlying identification assumption of the regression discontinuity and overstates the actual impact of index assignment. But as we have pointed out in an earlier footnote, in this paper, we are not running a regression discontinuity study to measure the impact of index assignment on dispersion effect. Rather, we take advantage of index reconstitution to construct a sample of stocks with a large variation in institutional ownership around a cutoff that is largely exogenously induced, in order to conduct tests to distinguish among the explanations of dispersion effect that is free from the endogeneity concern of the institutional ownership.

[10] To ensure that our finding is not due to the free-float adjustment, we also perform a robustness check using CRSP end-of-May market capitalization (as a proxy for the Russell end-of-May ranking, which is proprietary) as follows. We require a firm not only is in the 200 bandwidth in terms of Russell index weight ranking around the cutoff but also is in the 200 bandwidth in terms of CRSP end-of-May market capitalization ranking around the cutoff to be included in the sample. (Note that we use 200 instead of 100 here because using 100 would eliminate many firms with a distance to the cutoff around 100). We then use 100 firms around the cutoff in terms of the June Russell ranking on each side chosen from stocks that satisfy the criteria we just mentioned as our sample. We perform the tests in Table 6 using this sample and present the results in Table A1. We find qualitatively similar results. However, these additional requirements reduce our already limited number of observations even further, which, in turn, results in an insignificant t-test. But we do observe strong dispersion effect only in the top of the Russell 2000 index, which supports the analysts self-censoring explanation.

Russell 1000 index (e.g. 1000) receive a low index weight because they are the smallest of the Russell 1000 index stocks. However, stocks ranked below 1000 (e.g., 1001) that are in the top of the Russell 2000 index receive a high index weight because they are the largest of the Russell 2000 stocks. These differences in index weighting for stocks on each side of the 1001st ranking generate exogenous variation in institutional ownership, either by passive institutional investors who directly track the index, or by active institutional investors whose performances are benchmarked against these indexes. Chang, Hong, and Liskovich (2015) show that as of 2006, there are 273 investment products provided by institutional investors that are benchmarked to the Russell 2000 index with a dollar amount of 221.1 billion while there are 52 products that are benchmarked to the Russell 1000 index with a dollar amount of 146.1 billion. The index weights of Russell index matters because these investment products presumably will adjust their holdings according to the index weight. Figure 1 plots average institutional ownership (the ratio of shares held by professional money managers to total shares outstanding) for stocks around the threshold ranking (1001st) that determines the index membership. The horizontal axis is the difference between the ranking of a stock and 1001. Stocks to the left (negative distance) are those in the bottom of the Russell 1000 index, while stocks to the right (non-negative distance) are those in the top of the Russell 2000 index. On each side, we group stocks into 20 non-overlapping bins (bin width is 10 distance units). The dots represent the average of institutional ownership of stocks within each bin, and the curve is a third-order polynomial fit of the dots on each side. Figure 1 shows a clear discontinuity in institutional ownership, which is substantially lower for stocks in [-100, 0), i.e., the smallest 100 stocks in the Russell 1000 index. We therefore use 100 stocks around each side of the threshold. Boone and White (2015) show Russell index reconstitution leads to a difference of 19.2% in overall institutional ownership when comparing 100 stocks around the 1000th cutoff, out of which, 14% is due to institutional ownership by passive type institutional investors while the rest 5.2% is due to others. Moreover, the Russell index reconstitutes every year, stocks in the bottom (top) of the Russell 1000 (2000) may remain there for consecutive years, leading to pre-assignment differences in institutional ownership. This may also relate to float adjustment that stocks with low free float are adjusted to the bottom of the Russell 1000 index. But these alone do not invalidate our study as emphasis earlier we only require the difference in institutional ownership not the active choice of institutional investors to shun away from stocks with bad news, especially those that are high in dispersion (uncertainty). In other words, the pre-assignment difference by itself doesn't affect our comparing the dispersion effect of stocks in the bottom of the Russell 1000 index with those in the top of the Russell 2000 index, whose institutional ownership differ mainly because of the index assignment.[11]

[Insert Figure 1 Here]

This exogenously induced difference in the institutional ownership between stocks on each side of the threshold provides us a good setting to test the difference-in-opinion explanation that is free from the endogeneity concern of the institutional ownership. However, one may have concern that this setting may not be appropriate for the test of the self-censoring explanation, because if the induced difference in the institutional ownership mainly comes from passive investors, it may not have impact on the demand for analyst service, violating the requirement of the self-censoring explanation. Fortunately, this concern is unwarranted. Bushee (1998) classifies institutional

[11] The jump in the float-adjustment market capitalization at the threshold, as discussed in footnote 8, reflects that some less liquid (lower float) firms have been float-adjusted to the bottom of the Russell 1000 index. These adjustments would bias us against finding the results supportive of the analyst self-censoring explanation. This is because those firms are less liquid and more difficult for short sellers to arbitrage, and thus should exhibit strong dispersion effect if short-sale constraint is the main explanation.

investors into three types: quasi-indexer (passive), transient (active), and dedicated. Boone and White (2015) and Crane, Michenaud, and Weston (2016) show that the institutional ownership variation around the Russell 1000/2000 threshold is due to transient institutional investors and quasi-indexer, with the latter account for the majority of the variation. Clearly, transient institutional investors, being active, would have a demand for analyst service. But Boone and White (2015) show that quasi-indexer also have a demand. Quasi-indexer include index funds and other well-diversified institutions that have long investment horizon and typically allocate large weights of their holdings to index funds. For example, fund like Teachers Insurance Annuity Association–College Retirement Equities Fund (TIAA-CREF) puts about 80% of their holdings in indexed funds (Carleton, Nelson, and Weisbach (1998)). With their long investor horizon and large weight in index funds, these quasi-indexers lost the ability to follow the Wall Street Rule (vote by exiting). Thus, they do care about the performance of the firms in the indices and would need information provided by analysts to monitor them.

We obtain index membership data from Russell Investment from 1984 to 2006. After 2006, Russell Investment changed its method of determining membership, rendering the data unsuitable for research purposes. In order to reduce turnover across indexes, Russell now keeps a past year Russell 1000 (2000) stock in the new Russell 1000 (2000) index if the market capitalization of the stock drops (increases) to within a small band of the new index cutoff. This practice makes it possible for a bottom-ranking Russell 1000 stock to have a lower market capitalization than a top-ranking Russell 2000 firm. In that case, the lower institutional ownership of the bottom Russell 1000 stocks could then be partly a reflection of their lower market capitalization.

Instead of using the general sample, we focus on the 100 stocks at the bottom of the Russell 1000 index and the 100 stocks at the top of the Russell 2000 index. Stocks that are further away from the 1001st threshold ranking may be substantially different from those around the 1001st threshold ranking in unobservable characteristics that may confound our identification strategy. The choice of 100 as the bandwidth is a compromise between sample size and difference in characteristics between the two groups of stocks on each side. Note that in Panel B of Table 1, the characteristics of the bottom 100 Russell 1000 stocks and those of the top 100 Russell 2000 stocks are very much alike except for institutional ownership. Comparing the magnitude of the dispersion effect on each side of the threshold gauges the impact of institutional ownership on the dispersion effect. The impact is plausibly causal as the differences in institutional ownership are induced by different index weights.

4.3 Portfolio Strategy

We perform the same portfolio analysis for the Russell sample that we performed for the general sample. At the end of each month after reconstitution in year t (July of year t to June of year t+1), we sort stocks (with a price of over five dollars) into five quintiles (D1 to D5) based on analysts' forecast dispersion, irrespective of their index assignment. D1 contains stocks with the lowest dispersion, while D5 consists of stocks with the highest dispersion. Instead of sorting independently on institutional ownership as in the general sample, we rely on index membership to identify low and high institutional ownership stocks in the Russell sample. We label the stocks in the bottom of the Russell 1000 index R1. These are the stocks with low institutional ownership. We label the stocks in the top of the Russell 2000 index R2. These are the stocks with high institutional ownership. We end up with 10 (2X5) portfolios each month. The portfolios are rebalanced each month.

Before presenting our results, we again verify that institutional ownership is indeed associated with less binding short-sale constraints in this Russell sample. We again measure short-sale constraint as the supply of shares available for shorting times 1000 scaled by total number of shares outstanding for each stock in each month. For each group (R1 and R2), we calculate the monthly median level of short-sale constraints and take the time-series average of the median. Table 5 reports the results. We see that short-sale constraint is more binding for stocks in R1 than for those in R2. The difference in short-sale constraint is statistically significant at 5% with a t-stat of 2.38. Institutional ownership is also statistically higher for stocks in the top of the Russell 2000 index (R2).

[Insert Table 5 Here]

We next present our 2X5 portfolio strategy results. Panel A of Table 6 reveals an insignificant dispersion effect among the low institutional ownership portfolio R1, but a strong and significant dispersion effect in the high institutional ownership portfolio R2 irrespective of the return measures. For example, the FF4 alpha of high dispersion firms (D5) is lower than that of low dispersion firms (D1) by 0.76% per month in R2. These results are remarkable as they are completely the opposite of what we found in Table 3 with the general sample, even though short-sale constraint is more binding for stocks in R1 than for those in R2.

[Insert Table 6 Here]

Next, we are interested in finding out after we address the endogeneity issue via Russell sample if self-censoring would become more plausible explanation for the dispersion effect than difference in opinion (i.e., whether Hypothesis 2 is supported by the data instead). To this end, we set up a test with prediction of Hypothesis 2 being the alternative H_A: R2D1-R2D5 > R1D1-R1D5 and with H_0: R2D1-R2D5 $\leq$ R1D1-R1D5 being the null. Again, because the alternative is specified as a strict inequality, this would be a right-tailed test. Panel B of Table 6 shows that although the test is handicapped by the relatively low power resulting from the small number of stocks in each dispersion quintile (as reported in Panel C of Table 6), the null can be rejected at the 10% level based on raw returns. The results are similar if we risk-adjust the return by CAPM, or the FF4 model. When adjusted by FF3 model, the test results are close to be significant, with a p-value at 0.11. Rejection of the null constitutes support for Hypothesis 2 and hence the rejection of Hypothesis 1. This suggests that the oft-neglected analyst self-censoring story, as opposed to the difference-in-opinion story, is the more plausible explanation of the dispersion effect in a sample that is largely free from the bias caused by the endogeneity of institutional ownership.

4.4 Regression Approach

We repeat the regression tests on the Russell sample in Table 4. Stocks with low institutional ownership (R1) are those in the bottom of the Russell 1000 index, and stocks with high institutional ownership (R2) are those in the top of the Russell 2000 index. The results are reported in Panel A of Table 7.

[Insert Table 7 Here]

In column (1), the coefficient estimate on *DISP* is not statistically significant, suggesting that there is no dispersion effect in R1. In contrast, the negative and significant coefficient on *DISP* in column (2) suggests a significant dispersion effect in stocks with higher institutional ownership in R2. As a formal test, we test the null H_0: Low IO $\beta_1 \leq$ High IO β_1 against the prediction of Hypothesis 2 (i.e., H_A: Low IO $\beta_1 >$ High IO β_1). The test results in Panel B reveal that in the

Russell sample we reject the null and accept the alternative (Hypothesis 2) at the 5% significance level. This is consistent with the earlier finding from the portfolio tests reported in Table 6.

Overall, our results in this section indicate that both portfolio tests and regression tests reject difference-in-opinion in favor of self-censoring as a more plausible explanation of the dispersion effect in a sample that is not contaminated by the endogeneity of institutional ownership. The fact that we draw a completely opposite conclusion in such a sample suggests that we are justified in our concern about the endogeneity of institutional ownership and the bias it causes in the general sample.

5. The Dispersion-Bias Relationship

So far, we have found the dispersion effect to be stronger for stocks with high institutional ownership in a sample where the variation of institutional ownership is exogenously induced, which is consistent with the view that self-censoring is a more plausible story than difference-in-opinion as an explanation of the dispersion effect. For further validation, we subject the self-censoring story to an additional test on the dispersion-bias relationship. As stated in Hypothesis 2, self-censoring implies a positive dispersion-bias relationship that is stronger among stocks with higher institutional ownership.

We use the following panel regression to test the dispersion-bias relationship:

$$BIAS_{i,t} = \beta_0 + \beta_1 DISP_{i,t} + \beta_2 LOGMV_{i,t} + \beta_3 LOGBM_{i,t} + \beta_4 MOM_{i,t} + \beta_5 MOM_{i,t} + \beta_6 ROE_{i,t} + \beta_7 INV_{i,t} + \beta_8 COV_{i,t} + \mu_t + \varepsilon_{i,t}. \qquad (2)$$

$BIAS_{i,t}$ is the bias in the consensus forecast for stock *i* in month *t*. $DISP_{i,t}$ is the dispersion in analysts' forecasts at month *t*. Control variables are defined in Appendix A. We include month fixed effects in the model as well. Standard errors are double clustered at both the stock and month level.

Table 8 reports the regression results. In the Russell sample (Panel A), we see positive dispersion-bias relationships in both the low and high institutional ownership subsamples, consistent with the prediction of self-censoring story. More importantly the magnitude of the coefficient on *DISP* is larger in the subsample with high institutional ownership (0.012 vs. 0.003). The difference of the coefficients is statistically significant with a p-value of <0.01, indicating that the dispersion-bias relationship is significantly stronger in the Russell sample among stocks with higher institutional ownership. This substantiates the stronger positive dispersion-bias relationship in the high institutional ownership stocks predicted by the analyst self-censoring explanation.

For the general sample (Panel B), we again observe precisely the opposite results: the predictive power of *DISP* is stronger in the sample of stocks with *low* institutional ownership. The difference is also significant. This is again likely due to the endogeneity issue of institutional ownership in the general sample, institutional investors have an information advantage and thus are less likely to hold stocks with bad news. This, combined with analysts' incentive to self-censor only unfavorable forecasts, makes self-censoring more likely among low institutional ownership stocks with high dispersion. In other words, unlike in the Russell sample, in the general sample, the endogenous nature of institutional ownership should manifest itself in a stronger dispersion-bias relationship among stocks with lower institutional ownership, which would have led us to wrongly reject the self-censoring story.

[Insert Table 8 Here]

To sum up, we find that a positive dispersion-bias relationship is present in both the general sample and the Russell sample. By itself, this result supports the self-censoring explanation over the difference-in-opinion explanation, as the latter is silent on the direction of the relationship. Furthermore, we find a stronger (more positive) dispersion-bias relationship among stocks with high institutional ownership in the Russell sample and the exact opposite in the general sample. These results are consistent with both the self-censoring hypothesis and the existence of a significant endogeneity bias of institutional ownership in the general sample. These results are also parallel to what we find with the dispersion effect (the negative dispersion-return relationship). Both sets of results demonstrate the importance of controlling for the endogeneity of institutional ownership in the investigation of the dispersion effect, and they suggest that self-censoring is the more plausible explanation for the dispersion effect, at least in a sample in which the endogeneity of institutional ownership is not a concern.

6. The Role of Forecast Bias

Our results indicate that dispersion effect is plausibly due to the self-censoring of analysts in which higher dispersion in analyst forecast translates into higher bias in the forecast. An implication of the result is that once we control for the bias of analyst forecast, the predictive power of dispersion should become weaker. To check if it is the case, we add forecast bias to the return-dispersion regression (equation (1)) and run similar regressions for the general and the Russell low/high institutional ownership samples as we previously did.

The results are reported in Table 9 for both the general and the Russell sample. It can be seen in all four columns, irrespective of institutional ownership, the coefficients on *BIAS* are negative and significant while those on *DISP* no longer significantly predict future stock returns. This further substantiates our argument that analyst self-censoring explains the dispersion effect. However, we caution that the evidence is only suggestive because *BIAS* is calculated using look-ahead information at the time of portfolio formation while *DISP* only utilizes current information.

[Insert Table 9 Here]

7. Conclusion

This paper attempts to disentangle two plausible explanations of the dispersion effect (analyst self-censoring vs. difference-in-opinion) by constructing two mutually exclusive hypotheses regarding the effect of institutional ownership on the strength of the dispersion effect. The difference-in-opinion explanation predicts a stronger dispersion effect among stocks with lower institutional ownership. The self-censoring explanation predicts the exact opposite. Although our initial results confirm the findings in the literature that the dispersion effect is stronger among stocks with lower institutional ownership and seem to support the difference-in-opinion explanation, we raise the concern that this conclusion can be tainted by the possible endogeneity of institutional ownership. Using a sample from Russell index reconstitution, in which the variation of institutional ownership is exogenously induced, we find the opposite results. We find parallel results when we focus on the relationship between dispersion and bias. These findings suggest that the self-censoring story may be more plausible than the popular difference-in-opinion story as an explanation of the dispersion effect, and it deserves more attention in the literature.

Our paper highlights the importance of controlling for the endogeneity of institutional ownership when institutional ownership is part of the hypothesis for the explanation of characteristics-sorted return effects, if the characteristics also proxy for the degree of the institutional investors' information advantage.

References

Abarbanell, Jeffery S., William N. Lanen, and Robert E. Verrecchia, 1995, Analysts' forecasts as proxies for investor beliefs in empirical research, *Journal of Accounting and Economics* 20, 31-60.

Appel, Ian, Todd A. Gormley, and Donald B. Keim, 2016a, Passive investors, not passive owners, *Journal of Financial Economics* 121(1), 111-141.

Appel, Ian, Todd A. Gormley, and Donald B. Keim, 2016b, Identification using Russell 1000/2000 index assignments: A discussion of methodologies, working paper.

Asquith, Paul, Parag A. Pathak, and Jay R. Ritter, 2005, Short interest, institutional ownership, and stock returns, *Journal of Financial Economics* 78, 243-276.

Avramov, Doron, Tarun Chordia, Gergana Jostova, and Alexander Philipov, 2009, Dispersion in analysts' earnings forecasts and credit rating, *Journal of Financial Economics* 91, 83-101.

Baik, Bok, Jun-Koo Kang, and Jin-Mo Kim, 2010, Local institutional investors, information asymmetries, and equity returns, *Journal of Financial Economics* 97(1), 81-106.

Barniv Ran, Mark J. Myring, and Wayne B. Thomas, 2005, The association between the legal and financial reporting environments and forecast performance of individual analysts, *Contemporary Accounting Research* 22, 727-758.

Barron, Orie E., Oliver Kim, Steve C. Lim, and Douglas E. Stevens, 1998, Using analysts' forecasts to measure properties of analysts' information environment, *Accounting Review* 421-433.

Boehme, Rodney D., Bartley R. Danielsen, and Sorin M. Sorescu, 2006, Short-sale constraints, differences of opinion, and overvaluation, *Journal of Financial and Quantitative Analysis* 41, 455-487.

Boone, Audra L., and Joshua T. White, 2015, The effect of institutional ownership on firm transparency and information production, *Journal of Financial Economics* 117, 508-533.

Bushee, Brian J., 1998, The influence of institutional investors on myopic R&D investment behavior, *Accounting Review*, 305-333.

Bushee, Brian J., and Theodore H. Goodman, 2007, Which institutional investors trade based on private information about earnings and returns?, *Journal of Accounting Research* 45(2), 289-321.

Carleton, Willard T., James M. Nelson, and Michael S. Weisbach, 1998, The influence of institutions on corporate governance through private negotiations: Evidence from TIAA-CREF, *The Journal of Finance* 53(4), 1335-1362.

Chang, Eric C., Tse-Chun Lin, and Xiaorong Ma, 2015, Does short-selling threat discipline managers in mergers and acquisitions decisions?, working paper.

Chang, Yen-Cheng, Harrison Hong, and Inessa Liskovich, 2014, Regression discontinuity and the price effects of stock market indexing, *Review of Financial Studies* 28(1), 212-246.

Chen, Tao, Hui Dong, and Chen Lin, 2019, Institutional shareholders and corporate social responsibility: Evidence from two quasi-natural experiments, *Journal of Financial Economics*, forthcoming.

Crane, Alan D., Sébastien Michenaud, and James P. Weston, 2016, The effect of institutional ownership on payout policy: Evidence from index thresholds, *Review of Financial Studies* 29(6), 1377-1408.

Diether, Karl B., Christopher J. Malloy, and Anna Scherbina, 2002, Differences of opinion and the cross section of stock returns, *Journal of Finance* 57, 2113-2141.

Frankel, Richard, S. P. Kothari, and Joseph Weber, 2006, Determinants of the informativeness of analyst research, *Journal of Accounting and Economics* 41, 29-54.

Harvey, Campbell R., Yan Liu, and Heqing Zhu, 2016, … and the cross-section of expected returns, *The Review of Financial Studies* 29(1), 5-68.

Hwang, Chuan-Yang, and Yuan Li, 2018, Analysts' reputation concerns, self-censoring and the international dispersion effect, *Management Science*, Vol. 64, 2289-2307.

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *The Review of Financial Studies* 28(3), 650-705.

Johnson, Timothy C., 2004, Forecast dispersion and the cross section of expected returns, *Journal of Finance* 59, 1957-1978.

Ljungqvist, Alexander, Felicia Marston, Laura T. Starks, Kelsey D. Wei, and Hong Yan, 2007, Conflicts of interest in sell-side research and the moderating role of institutional investors, *Journal of Financial Economics* 85, 420-456.

Maffett, Mark, 2012, Financial reporting opacity and informed trading by international institutional investors, *Journal of Accounting and Economics* 54(2-3), 201-220.

Mann, Prem S., 2010, *Introductory Statistics*, 7th Edition, Wiley.

McNichols, Maureen F., and Patricia C. O'Brien, 1997, Self-selection and analyst coverage, *Journal of Accounting Research* 35, 167-199.

Miller, Edward M., 1977, Risk, uncertainty, and divergence of opinion, *Journal of Finance* 32, 1151-1168.

Nagel, Stefan, 2005, Short sales, institutional investors and the cross-section of stock returns, *Journal of Financial Economics* 78, 277-309.

O'Brien, Patricia C., Maureen F. McNichols, and Hsiou-wei Lin, 2005, Analyst impartiality and investment banking relationships, *Journal of Accounting Research* 43, 623-650.

Petersen, Mitchell A., 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *The Review of Financial Studies* 22(1), 435-480.

Richardson, Scott, Irem Tuna, and Peter Wysocki, 2010. Accounting anomalies and fundamental analysis: A review of recent research advances. *Journal of Accounting and Economics* 50, 410-454

Scherbina, Anna, 2005, Analyst disagreement, forecast bias and stock returns, working paper, University of California, Davis.

Scherbina, Anna, 2008, Suppressed negative information and future underperformance, *Review of Finance* 12, 533-565.

Figure 1 **Average Institutional Ownership around the Russell 1000/2000 Index Threshold**

This figure displays the average institutional ownership of stocks around the Russell 1000/2000 index threshold in the sample period. The vertical axis is the total institutional ownership. The horizontal axis is the distance from the threshold (1001st ranking) market capitalization ranking that separates the Russell 1000 index from the Russell 2000 index. Stocks with positive distance are in the Russell 2000 index, while stocks with negative distance are in the Russell 1000 index. The sample spans from 1984 to 2006. Russell index data are obtained from Russell Investment. For each stock each year, total institutional ownership is calculated as the ratio of total shares held by institutional money managers to total shares outstanding. The information regarding shares held by institutional money managers is obtained from quarterly 13-F filings data from Thompson Reuters (from September year t to June year t+1 for the Russell index in year t). Stocks are divided into 20 non-overlapping bins on each side of the threshold. The dots in the figure represent the average institutional ownership of stocks within each bin. The curves represent a third-order polynomial fit of the dots.

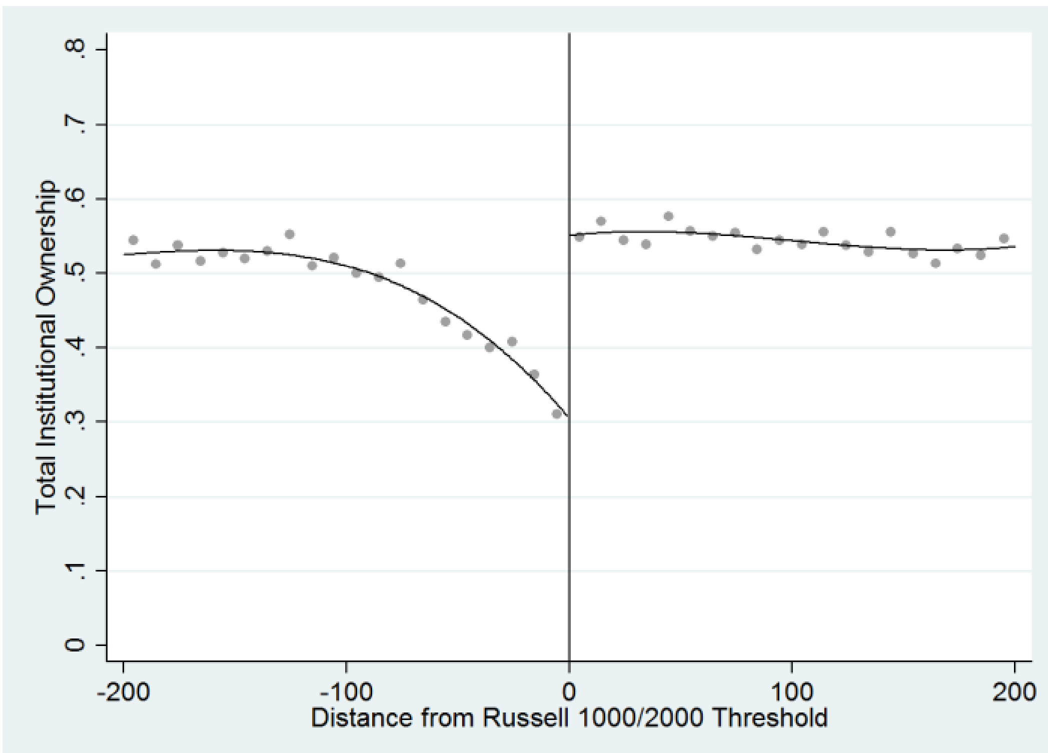

Table 1 **Summary Statistics**

This table presents summary statistics for the main variables used in this study. The general sample covers all NYSE, AMEX, and NASDAQ stocks from 1984 to 2006. At the end of each month, we divide all stocks in the general sample into quintiles based on percentage institutional ownership, which is calculated as the total number of shares, obtained from their latest quarterly 13-F filings, divided by the number of shares outstanding. We report the summary statistics for the bottom quintile (Low IO) and top quintile (High IO) in Panel A. The Russell sample, consists of 100 stocks in the bottom of the Russell 1000 index and another 100 stocks in the top of the Russell 2000 index each year from 1984 to 2006. We report summary statistics for the Russell sample in Panel B. Dispersion (*DISP*) in analysts' earnings forecasts is calculated each month as the ratio of the standard deviation of analysts' current fiscal-year annual earnings-per-share forecasts to the absolute value of the mean forecast. Forecast bias (*BIAS*) is defined as the analysts' consensus earnings-per-share forecast in the current month minus the corresponding actual earnings-per-share announced in the future, scaled by the current price of the stock. *COV* is the log of one plus the number of analysts who have issued fiscal year one earnings forecasts for the stock in each month. *IO* for the general sample is the percentage institutional ownership as explained above. Since Russell reconstitution takes effect each June, IO for the Russell sample in a given year is the average of four institutional ownership, each calculated respectively based on data from 13-F filings in September, December and in the following March and June. *DIFF(MEAN)* is the difference in the mean in characteristics between the two groups of stocks. t-statistics are reported. ***, **, and * correspond to significance at 1%, 5%, and 10%, respectively.

Panel A. The General Sample

Low IO				**High IO**					
	N	*MEAN*	*STD*		*N*	*MEAN*	*STD*	*DIFF (MEAN)*	*t-stat*
DISP	106798	0.18	0.90	*DISP*	129433	0.13	1.03	0.05***	12.77
BIAS	101098	1.48%	8.66%	*BIAS*	125932	1.07%	5.05%	0.41%***	14.35
COV	106798	1.56	0.49	*COV*	129433	2.40	0.60	-0.84***	-369.08
IO	106798	0.17	0.10	*IO*	129433	0.77	0.17	-0.60***	-1057.17

Panel B. The Russell Sample

Low IO				**High IO**					
	N	*MEAN*	*STD*		*N*	*MEAN*	*STD*	*DIFF (MEAN)*	*t-stat*
DISP	17764	0.16	1.11	*DISP*	20115	0.15	0.83	0.01	1.61
BIAS	17189	1.36%	13.17%	*BIAS*	19447	1.47%	9.30%	-0.11%	-0.87
COV	17764	2.15	0.48	*COV*	20115	2.14	0.46	0.01	-1.40
IO	17757	0.45	0.26	*IO*	20104	0.57	0.26	-0.12***	-43.86

Table 2 **Institutional Ownership and Short-Sale Constraint: The General Sample**

This table presents short-sale constraint and institutional ownership across institutional ownership groups for the general sample using data from 1984 to 2006 (The short-sale data is only available from 2002). At the end of each month, we divide all stocks in the general sample into quintiles based on percentage institutional ownership, which is calculated as the total number of shares, obtained from their latest quarterly 13-F filings, divided by the number of shares outstanding (labelled I1 to I5). I1 (I5) includes stocks with the lowest (highest) institutional ownership. Monthly short-sale constraint is defined as the supply of shares available for short sale times 1000 scaled by total number of shares outstanding for each stock in each month. A higher value means less binding short-sale constraints. For each group, we first calculate the monthly median of short-sale constraint and institutional ownership and then calculate the time-series average of the median. The difference (*DIFF*) is the difference in value between group I1 and group I5, and the associated t-statistics is given. ***, **, and * correspond to significance at 1%, 5%, and 10%, respectively.

	Low IO				High IO	
	I1	**I2**	**I3**	**I4**	**I5**	**DIFF**
Short-Sale Constraint (‰)	30.60	62.27	78.27	89.88	103.37	72.76***
t-stat						4.36
Institutional Ownership (%)	16.92	33.60	47.20	59.92	74.47	57.55***
t-stat						66.67

Table 3 **Institutional Ownership and the Dispersion Effect: The General Sample**

Panel A of this table presents the dispersion effect for all stocks in the general sample. Panel B of this table presents the dispersion effect across institutional ownership groups for the general sample. Panel C presents the one-tailed t-test of whether the difference-in-opinion explanation is more plausible than the self-censoring explanation. Panel D reports the average number of stocks in each portfolio across the sample period. The sample period is from 1984 to 2006. At the end of each month, we first divide all stocks in the general sample into five groups based on dispersion in analysts' earnings forecasts (labelled D1 to D5. D1 (D5) includes stocks with the lowest (highest) dispersion. Dispersion in analysts' earnings forecasts is calculated each month as the ratio of the standard deviation of analysts' current fiscal-year annual earnings-per-share forecasts to the absolute value of the mean forecast. We divide all stocks in the general sample into quintiles based on percentage institutional ownership, which is calculated as the total number of shares, obtained from their latest quarterly 13-F filings, divided by the number of shares outstanding (labelled I1 to I5). I1 (I5) includes stocks with the lowest (highest) institutional ownership. We end up with 25 (5X5) portfolios. IaDb represents a stock selected from groups Ia and Db. After being assigned to portfolios, stocks are held for one month. D1-D5 is the return of the hedge portfolio that holds a long position in stocks of D1 and a short position in stocks of D5. The monthly return (in percentage) of each portfolio is the equal-weighted average of the returns of all the stocks in the portfolio. We report returns in **bold**, including raw monthly returns (RAW), the CAPM alpha, alpha from the Fama-French three-factor model (FF3), and alpha from the Fama-French three-factor model augmented by the Carhart momentum factor (FF4). Newy-West t-statistics are reported. ***, **, and * correspond to significance at 1%, 5%, and 10%, respectively.

Panel A. Dispersion Effect

All Stocks	*Dispersion*					
	D1	**D2**	**D3**	**D4**	**D5**	**D1-D5**
RAW	**1.38**	**1.29**	**1.15**	**0.95**	**0.44**	**0.95*******
t-stat	6.35	5.95	5.14	3.79	1.54	4.51
CAPM	**0.99**	**0.89**	**0.74**	**0.53**	**0.02**	**0.97*******
t-stat	4.28	3.89	3.22	2.10	0.07	4.61
FF3	**0.96**	**0.86**	**0.72**	**0.50**	**-0.01**	**0.98*******
t-stat	4.20	3.73	3.09	1.92	-0.04	4.61
FF4	**1.04**	**0.92**	**0.78**	**0.54**	**0.03**	**1.01*******
t-stat	4.21	3.76	3.18	1.93	0.09	4.33

Panel B. Institutional Ownership and the Dispersion Effect

Raw Return	*Dispersion*					
Institutional Ownership	**D1**	**D2**	**D3**	**D4**	**D5**	**D1-D5**
I1	**1.27**	**1.27**	**0.95**	**0.58**	**-0.14**	**1.41*****
t-stat	5.08	4.95	3.53	2.03	-0.43	5.05
I2	**1.41**	**1.16**	**1.01**	**0.88**	**0.42**	**0.99*****
t-stat	5.64	5.29	4.26	3.35	1.33	3.61
I3	**1.40**	**1.29**	**1.21**	**1.06**	**0.56**	**0.84*****
t-stat	6.67	5.57	5.25	4.22	1.84	3.70
I4	**1.39**	**1.35**	**1.25**	**1.18**	**0.74**	**0.65*****
t-stat	5.88	6.23	5.63	4.67	2.79	3.21
I5	**1.44**	**1.37**	**1.28**	**1.08**	**0.92**	**0.53****
t-stat	6.84	6.14	5.64	4.42	3.33	2.51

CAPM	*Dispersion*					
Institutional Ownership	**D1**	**D2**	**D3**	**D4**	**D5**	**D1-D5**
I1	**0.86**	**0.85**	**0.53**	**0.15**	**-0.57**	**1.42*****
t-stat	3.25	3.11	1.90	0.50	-1.64	5.01
I2	**1.00**	**0.76**	**0.59**	**0.47**	**-0.01**	**1.01*****
t-stat	3.92	3.33	2.45	1.73	-0.03	3.69
I3	**1.00**	**0.89**	**0.80**	**0.66**	**0.15**	**0.85*****
t-stat	4.51	3.68	3.36	2.56	0.50	3.86
I4	**1.00**	**0.96**	**0.83**	**0.77**	**0.31**	**0.68*****
t-stat	4.06	4.24	3.71	3.09	1.16	3.34
I5	**1.06**	**0.97**	**0.88**	**0.68**	**0.54**	**0.52****
t-stat	4.72	4.13	3.85	2.73	2.00	2.51

FF3	*Dispersion*					
Institutional Ownership	**D1**	**D2**	**D3**	**D4**	**D5**	**D1-D5**
I1	**0.80**	**0.81**	**0.51**	**0.06**	**-0.59**	**1.39*****
t-stat	3.11	2.97	1.85	0.21	-1.76	5.04
I2	**0.97**	**0.72**	**0.55**	**0.44**	**-0.04**	**1.01*****
t-stat	3.84	3.23	2.31	1.61	-0.13	3.72
I3	**0.99**	**0.84**	**0.79**	**0.62**	**0.12**	**0.86*****
t-stat	4.40	3.46	3.24	2.38	0.39	3.81
I4	**0.98**	**0.95**	**0.82**	**0.76**	**0.28**	**0.70*****
t-stat	3.97	4.00	3.59	2.87	0.99	3.46
I5	**1.05**	**0.96**	**0.87**	**0.67**	**0.50**	**0.56****
t-stat	4.63	3.96	3.67	2.48	1.72	2.40

FF4	*Dispersion*					
Institutional Ownership	**D1**	**D2**	**D3**	**D4**	**D5**	**D1-D5**
I1	**0.85**	**0.83**	**0.56**	**0.09**	**-0.51**	**1.36*** **
t-stat	3.15	2.89	1.96	0.28	-1.39	4.69
I2	**1.02**	**0.76**	**0.63**	**0.47**	**-0.02**	**1.04*** **
t-stat	3.84	3.19	2.50	1.61	-0.05	3.61
I3	**1.06**	**0.90**	**0.83**	**0.66**	**0.16**	**0.91*** **
t-stat	4.33	3.49	3.25	2.34	0.45	3.62
I4	**1.07**	**1.02**	**0.91**	**0.81**	**0.33**	**0.75*** **
t-stat	4.05	4.15	3.73	2.88	1.11	3.38
I5	**1.15**	**1.06**	**0.93**	**0.74**	**0.48**	**0.68*** **
t-stat	4.69	4.08	3.74	2.59	1.63	2.86

Panel C. Test for Hypothesis 1 (*Difference-in-Opinion*)

H_0 **:** **(**I1D1-I1D5 $\leq$ I5D1-I5D5**)**

H_A **:** *Hypothesis 1* **(**I1D1-I1D5> I5D1-I5D5**)**

	RAW	**CAPM**	**FF3**	**FF4**
(I1D1-I1D5)-(I5D1-I5D5)	**0.88*** **	**0.90*** **	**0.83*** **	**0.68*** **
p-value	0.00	0.00	0.00	0.00

Panel D. Average Number of Stocks in Each Portfolio

Observations	*Dispersion*				
Institutional Ownership	**D1**	**D2**	**D3**	**D4**	**D5**
I1	100	83	89	100	119
I2	95	96	96	98	107
I3	94	99	100	99	101
I4	101	103	101	101	87
I5	104	112	107	94	76

Table 4 **Institutional Ownership and the Dispersion Effect: The General Sample**

Panel A of this table presents the regression results from the following model:

$$RET_{i,t+1} = \beta_0 + \beta_1 DISP_{i,t} + \beta_2 LOGMV_{i,t} + \beta_3 LOGBM_{i,t} + \beta_4 MOM_{i,t} + \beta_5 ROE_{i,t} + \beta_6 INV_{i,t} + \beta_7 COV_{i,t} + \mu_{t+1} + \varepsilon_{i,t+1}$$

for the general sample using data from 1984 to 2006. The dependent variable *RET* is monthly stock return. *DISP*, *LOGMV*, *LOGBM*, *MOM*, *ROE*, *INV*, *COV* are dispersion, log of market capitalization, log of book-to-market ratio, momentum, investment, and analyst coverage as defined in Appendix A. At the end of each month, we divide the sample into quintiles based on institutional ownership, calculated as the total number of shares, obtained from their latest quarterly 13-F filings, divided by the total number of shares outstanding. We run the regression separately for stocks in bottom quintile (Low IO) and for stocks in the top quintile (High IO). We include month fixed effects in each regression. Standard errors are clustered at both the month and the stock level. t-statistics are in parentheses. Panel B reports the right-tailed t-test implied by the difference-in-opinion hypothesis, with standard errors clustered at month level. ***, **, and * correspond to significance at 1%, 5%, and 10%, respectively

Panel A. Panel Regression Results

	The General Sample	
	(1)	(2)
	Low IO	High IO
DISP	**-0.163*** **	**-0.032**
	(-2.75)	**(-0.86)**
LOGBM	0.517***	0.081
	(4.57)	(0.69)
LOGMV	0.018	0.027
	(0.27)	(0.38)
MOM	0.579**	0.155
	(2.19)	(0.34)
ROE	-0.000	-0.008
	(-0.53)	(-1.29)
INV	0.000	0.000
	(0.65)	(0.06)
COV	-0.152	0.116
	(-1.35)	(1.14)
Intercept	14.949***	13.318***
	(17.13)	(14.94)
Time Fixed Effects	Yes	Yes
Adjusted R-squared	0.128	0.195
Observations	99,665	125,426

Panel B. Test for Hypothesis 1 (difference-in-opinion)

H_0 **:** Low IO $\beta_1 \geq$ High IO β_1

H_A **:** *Hypothesis 1* **(**Low IO $\beta_1 <$ High IO β_1**)**

Low IO β_1 - High IO β_1	**-0.131****
p-value	0.021

Table 5 **Institutional Ownership and Short-Sale Constraint: The Russell Sample**

This table presents short-sale constraint and institutional ownership across institutional ownership groups for the Russell sample using data from 1984 to 2006 (The short-sale data is only available from 2002). Stocks in the bottom of the Russell 1000 are labeled group R1 and have low institutional ownership (Low IO). Stocks in the top of the Russell 2000 index are labeled group R2 and have high institutional ownership (High IO). For each stock, institutional ownership is calculated as total shares held by institutional money managers who file 13-F filings divided by the total shares outstanding. Since Russell reconstitution takes effect each June, IO for the Russell sample in a given year is the average of four institutional ownership, each calculated respectively based on data from 13-F filings in September, December and in the following March and June. Monthly short-sale constraint is defined as the supply of shares available for short sale times 1000 scaled by total number of shares outstanding for each stock in each month. A higher value means less binding short-sale constraints. For each group, we first calculate the monthly median of short-sale constraint and institutional ownership and then calculate the time-series average of the median. The difference (*DIFF*) is the difference in value between group R1 and group R2, and the associated t-stat is given. ***, **, and * correspond to significance at 1%, 5%, and 10%, respectively.

	Low IO	High IO	
	R1	**R2**	**DIFF**
Short-Sale Constraint (‰)	56.07	98.58	42.51**
t-stat			2.38
Institutional Ownership (%)	44.05	59.54	15.49***
t-stat			11.41

Table 6 **Institutional Ownership and the Dispersion Effect: The Russell Sample**

Panel A of this table presents the dispersion effect across institutional ownership groups for the Russell sample. Panel B presents the test of whether the self-censoring explanation is more plausible than the difference-in-opinion explanation. Panel C reports the average number of stocks in each portfolio across the sample period. The sample period is from 1984 to 2006. At the end of each month, we first divide all stocks in the Russell sample into five groups based on dispersion in analysts' earnings forecasts (labelled D1 to D5. D1 (D5) includes stocks with the lowest (highest) dispersion). Dispersion in analysts' earnings forecasts is calculated each month as the ratio of the standard deviation of analysts' current fiscal-year annual earnings-per-share forecasts to the absolute value of the mean forecast. Stocks in the bottom of the Russell 1000 are labeled group R1 and have low institutional ownership (Low IO). Stocks in the top of the Russell 2000 index are labeled group R2 and have high institutional ownership (High IO). We have 10 (2x5) portfolios. After being assigned to portfolios, stocks are held for one month. RaDb represents a stock selected from group Ra and Db. D1-D5 is the return of the hedge portfolio that holds a long position in stocks of D1 and a short position in stocks of D5. The monthly return (in percentage) of each portfolio is the equal-weighted average of the returns of all the stocks in the portfolio. We report returns in **bold**, including raw monthly returns (RAW), the CAPM alpha, alpha from the Fama-French three-factor model (FF3), and alpha from the Fama-French three-factor model augmented by the Carhart momentum factor (FF4). Newey-West t-statistics are reported. ***, **, and * correspond to significance at 1%, 5%, and 10%, respectively.

Panel A. Institutional Ownership and the Dispersion Effect

Raw Return	*Dispersion*					
Institutional Ownership	**D1**	**D2**	**D3**	**D4**	**D5**	**D1-D5**
R1	**1.20**	**1.05**	**1.43**	**0.86**	**0.87**	**0.33**
t-stat	4.16	3.31	5.19	2.93	2.69	0.99
R2	**1.44**	**1.04**	**1.12**	**0.91**	**0.68**	**0.77*** **
t-stat	6.36	3.87	4.16	2.75	2.13	2.82

CAPM	*Dispersion*					
Institutional Ownership	**D1**	**D2**	**D3**	**D4**	**D5**	**D1-D5**
R1	**0.78**	**0.67**	**1.00**	**0.44**	**0.47**	**0.31**
t-stat	2.65	2.06	3.51	1.53	1.43	0.91
R2	**1.07**	**0.66**	**0.74**	**0.52**	**0.33**	**0.73*** **
t-stat	4.32	2.35	2.59	1.51	1.02	2.64

FF3	*Dispersion*					
Institutional Ownership	**D1**	**D2**	**D3**	**D4**	**D5**	**D1-D5**
R1	**0.74**	**0.71**	**1.00**	**0.38**	**0.46**	**0.28**
t-stat	2.55	2.09	3.67	1.22	1.37	0.81
R2	**1.02**	**0.66**	**0.70**	**0.47**	**0.35**	**0.68** **
t-stat	4.47	2.29	2.50	1.34	1.01	2.44

FF4	*Dispersion*					
Institutional Ownership	**D1**	**D2**	**D3**	**D4**	**D5**	**D1-D5**
R1	**0.83**	**0.79**	**1.16**	**0.42**	**0.59**	**0.24**
t-stat	2.75	2.22	4.39	1.28	1.68	0.66
R2	**1.10**	**0.76**	**0.74**	**0.53**	**0.34**	**0.76*** **
t-stat	4.31	2.51	2.46	1.45	0.94	2.61

Panel B. Test of Hypothesis 2 (*Self-Censoring*)

H_0 **: (**R2D1-R2D5 $\leq$ R1D1-R1D5**)**

H_A **:** *Hypothesis 2* **(**R2D1-R2D5>R1D1-R1D5**)**

	RAW	**CAPM**	**FF3**	**FF4**
(R2D1-R2D5)-(R1D1-R1D5**)**	**0.43***	**0.43***	**0.40**	**0.52***
p-value	0.08	0.09	0.11	0.06

Panel C. Average Number of Stocks in Each Portfolio

Observations	*Dispersion*				
Institutional Ownership	**D1**	**D2**	**D3**	**D4**	**D5**
R1	13	13	14	14	17
R2	15	16	15	15	15

Table 7 **Institutional Ownership and the Dispersion Effect: The Russell Sample**

Panel A of this table presents the regression results from the following model:

$$RET_{i,t+1} = \beta_0 + \beta_1 DISP_{i,t} + \beta_2 LOGMV_{i,t} + \beta_3 LOGBM_{i,t} + \beta_4 MOM_{i,t} + \beta_5 ROE_{i,t} + \beta_6 INV_{i,t} + \beta_7 COV_{i,t} + \mu_{t+1} + \varepsilon_{i,t+1}$$

for the Russell sample using data from 1984 to 2006. The dependent variable *RET* is monthly stock return. *DISP*, *LOGMV*, *LOGBM*, *MOM*, *ROE*, *INV*, *COV* are dispersion, log of market capitalization, log of book-to-market ratio, momentum, investment, and analyst coverage as defined in Appendix A. Stocks in the bottom of the Russell 1000 are labeled group R1 and have low institutional ownership (Low IO). Stocks in the top of the Russell 2000 index are labeled group R2 and have high institutional ownership (High IO). We run the regression for stocks in R1 and R2 separately. We include month fixed effects in each regression. Standard errors are clustered at both the month and the stock level. t-statistics are in parentheses. Panel B reports the left-tailed test implied by the self-censoring hypothesis, with standard errors clustered at month level. ***, **, and * correspond to significance at 1%, 5%, and 10%, respectively.

Panel A. Panel Regression Results

	The Russell Sample	
	(1)	(2)
	Low IO (R1)	High IO (R2)
DISP	**0.012**	**-0.213*** **
	(0.11)	**(-2.72)**
LOGBM	0.367***	0.168
	(2.77)	(1.15)
LOGMV	0.254	-0.227
	(1.23)	(-0.64)
MOM	0.674**	0.491**
	(2.15)	(2.51)
ROE	-0.092***	0.188
	(-3.42)	(1.38)
INV	-0.000	0.001***
	(-0.48)	(4.63)
COV	0.076	-0.023
	(0.46)	(-0.15)
Intercept	10.898***	13.925***
	(4.93)	(3.51)
Time Fixed Effects	Yes	Yes
Adjusted R-squared	0.178	0.178
Observations	16,985	19,390

Panel B. Test of Hypothesis 2 (*Self-Censoring*)

H_0 **:** Low IO $\beta_1 \leq$ High IO β_1

H_A **:** *Hypothesis 2* **(**Low IO $\beta_1 >$ High IO β_1**)**

Low IO β_1 - High IO β_1	**0.225** **
p-value	0.043

Table 8 **Institutional Ownership, Forecast Dispersion and Forecast Bias**

This table reports the results of the following regressions:

$$BIAS_{i,t} = \beta_0 + \beta_1 DISP_{i,t} + \beta_2 LOGMV_{i,t} + \beta_3 LOGBM_{i,t} + \beta_4 MOM_{i,t} + \beta_5 ROE_{i,t} + \beta_6 INV_{i,t} + \beta_7 COV_{i,t} + \mu_t + \varepsilon_{i,t}$$

for the Russell sample (Panel A) and the general sample (Panel B) using data from 1984 to 2006. The dependent variable *BIAS* is forecast bias. It is defined as the difference between analysts' consensus earnings-per-share forecast in the current month minus the corresponding actual earnings-per-share announced in the future, scaled by the current price of the stock. *DISP*, *LOGMV*, *LOGBM*, *MOM, ROE, INV, COV* are dispersion, log of market capitalization, log of book-to-market ratio, momentum, investment, and analyst coverage as defined in Appendix A. For the Russell sample, stocks in the bottom of the Russell 1000 are labeled group R1 and have low institutional ownership (Low IO). Stocks in the top of the Russell 2000 index are labeled group R2 and have high institutional ownership (High IO). For the general sample, at the end of each month stocks are sorted into quintiles based on institutional ownership, calculated as the total number of shares held by institutional money managers, obtained from the latest quarterly 13-F filings, divided by the total number of shares outstanding. Stocks in the top (bottom) quintile is labelled High IO (Low IO). For both the Russell sample and the general sample, we do the following. We first run the regression separately for High IO and Low IO stocks. We then perform a one-sided t-test for the difference of the coefficients of *DISP* from the two regressions. We include month fixed effects in each regression. Standard errors are clustered at both the month and stock level. t-statistics are in parentheses. ***, **, and * correspond to significance at 1%, 5%, and 10%, respectively.

Panel A. Panel Regression Results: The Russell Sample

	The Russell Sample	
	(1)	(2)
	Low IO	High IO
DISP	**0.003****	**0.012*****
	(2.44)	**(3.11)**
LOGBM	0.001	0.010***
	(0.56)	(3.35)
LOGMV	-0.010	-0.013*
	(-1.29)	(-1.87)
MOM	-0.018***	-0.004*
	(-4.23)	(-1.68)
ROE	-0.001***	-0.002
	(-3.58)	(-1.31)
INV	0.000**	0.000
	(2.12)	(0.37)
COV	0.005**	0.005
	(2.02)	(1.51)
Intercept	0.141*	0.240***
	(1.79)	(2.72)
H_0 **:**Low IO $\beta_1 \geq$ High IO β_1		
H_A **:**Low IO $\beta_1 <$ High IO β_1	-0.009***	
p-value	0.000	
Time Fixed Effects	Yes	Yes
Adjusted R-squared	0.021	0.059
Observations	16,985	19,390

Panel B. Panel Regression Results: The General Sample

	The General Sample	
	(1)	(2)
	Low IO	High IO
DISP	**0.008*****	**0.003*****
	(8.21)	**(3.17)**
LOGBM	0.004***	0.005***
	(4.11)	(7.51)
LOGMV	-0.005***	-0.003***
	(-7.80)	(-5.64)
MOM	-0.015***	-0.017***
	(-6.63)	(-9.08)
ROE	-0.000**	-0.000
	(-2.05)	(-1.64)
INV	0.000	-0.000***
	(0.65)	(-5.37)
COV	0.005***	0.000
	(3.04)	(0.32)
Intercept	0.101***	0.089***
	(10.68)	(12.23)
H_0 **:**Low IO $\beta_1 \geq$ High IO β_1		
H_A **:**Low IO $\beta_1 <$ High IO β_1	0.005	
p-value	1.000	
Time Fixed Effects	Yes	Yes
Adjusted R-squared	0.033	0.056
Observations	99,665	125,426

Table 9 **Institutional Ownership and the Dispersion Effect: The Role of Bias**

Panel A of this table presents the regression results from the following model:

$RET_{i,t+1} = \beta_0 + \beta_1 DISP_{i,t} + \beta_2 BIAS_{i,t} + \beta_3 LOGMV_{i,t} + \beta_4 LOGBM_{i,t} + \beta_5 MOM_{i,t} + \beta_6 ROE_{i,t} + \beta_7 INV_{i,t} + \beta_8 COV_{i,t} + \mu_{t+1} + \varepsilon_{i,t+1}$

for the general and the Russell sample using data from 1984 to 2006. The dependent variable *RET* is monthly stock return. *DISP*, *BIAS*, *LOGMV*, *LOGBM*, *MOM*, *ROE*, *INV*, *COV* are dispersion, bias, log of market capitalization, log of book-to-market ratio, momentum, investment, and analyst coverage as defined in Appendix A. For the Russell sample, stocks in the bottom of the Russell 1000 are labeled group R1 and have low institutional ownership (Low IO). Stocks in the top of the Russell 2000 index are labeled group R2 and have high institutional ownership (High IO). For the general sample, at the end of each month, stocks are sorted into quintiles based on institutional ownership, calculated as the total number of shares held by institutional money managers, obtained from the latest quarterly 13-F filings, divided by the total number of shares outstanding. Stocks in top (bottom) quintile is labelled High IO (Low IO). For both the Russell sample and the general sample, we do the following. We first run the regression separately for High IO and Low IO stocks. Standard errors are clustered at both the month and the stock level. t-statistics are in parentheses. ***, **, and * correspond to significance at 1%, 5%, and 10%, respectively.

	The General Sample		The Russell Sample	
	(1)	(2)	(3)	(4)
	Low IO	High IO	Low IO	High IO
DISP	**-0.093**	**0.019**	**0.024**	**-0.087**
	(-1.53)	**(0.64)**	**(0.23)**	**(-1.15)**
BIAS	**-8.987*** **	**-15.681*** **	**-4.030** **	**-10.212*** **
	(-3.40)	**(-7.83)**	**(-2.00)**	**(-8.26)**
LOGBM	0.552***	0.156	0.373***	0.272*
	(4.85)	(1.33)	(2.84)	(1.86)
LOGMV	-0.031	-0.013	0.213	-0.364
	(-0.46)	(-0.18)	(1.06)	(-1.06)
MOM	0.444*	-0.116	0.601*	0.448**
	(1.72)	(-0.25)	(1.93)	(2.37)
ROE	-0.000	-0.010*	-0.095***	0.170
	(-0.76)	(-1.87)	(-3.56)	(1.25)
INV	0.000	-0.000	-0.000	0.001***
	(0.80)	(-1.11)	(-0.40)	(4.59)
COV	-0.108	0.121	0.095	0.026
	(-0.97)	(1.21)	(0.57)	(0.17)
Intercept	15.857***	14.720***	11.466***	16.372***
	(17.72)	(16.37)	(5.24)	(4.27)
Time Fixed Effects	Yes	Yes	Yes	Yes
Adjusted R-squared	0.133	0.200	0.181	0.185
Observations	99,665	125,426	16,985	19,390

Appendix A. Definition of Variables

DISP: Dispersion in analysts' earnings forecasts is calculated each month as the ratio of the standard deviation of analysts' current fiscal-year annual earnings-per-share forecasts to the absolute value of the mean forecast.

BIAS: Forecast bias is defined as the analysts' consensus earnings-per-share forecast in the current month minus the corresponding actual earnings-per-share announced in the future, scaled by the current price of the stock.

IO: institutional ownership is calculated by dividing total shares held by institutional money managers who file 13-F filings to total shares outstanding. We use latest quarterly filing for the general sample while the average of intuitional ownership using September to June next year's filing information is used for the Russell sample.

LOGMV: the natural log of market capitalization. Market capitalization is calculated as the product of share prices and total shares outstanding.

LOGBM: is the natural log of the book value of equity to the market value of equity (market capitalization). Book value is calculated as book value of stockholders' equity, plus balance sheet deferred taxes and investment tax credit (if available), minus the book value of preferred stock. Depending on availability, we use the redemption, liquidation, or par value (in that order) to estimate the book value of preferred stock. Stockholders' equity is the value reported by COMPUSTAT, if it is available. If not, we measure stockholders' equity as the book value of common equity plus the par value of preferred stock, or the book value of assets minus total liabilities (in that order). Fiscal year t-1 book value is paired with market capitalization in the current calendar year t if the portfolio is formed on or after June; otherwise we use the book value of fiscal year t-2.

COV: the log of the number of analysts who have issued fiscal year one earnings forecasts for the stock in each month plus one.

ROE: return on equity calculated as the ratio of income before extraordinary items to total common equity.

INV: total investment as measured by the growth rate of total assets.

Table A1 **Institutional Ownership and the Dispersion Effect: The Russell Sample (Robustness)**

This table provides robustness test in which we require a firm not only is in the 200 bandwidth in terms of Russell index ranking around the cutoff but also is in the 200 bandwidth in terms of CRSP end-of-May market capitalization ranking around the cutoff to be included in the sample. We then use 100 firms around the cutoff in terms of the Russell ranking on each side as our sample. Panel A of this table presents the dispersion effect across institutional ownership groups for the sample. Panel B presents the test of whether the self-censoring explanation is more plausible than the difference-in-opinion explanation. Panel C reports the average number of stocks in each portfolio across the sample period. The sample period is from 1984 to 2006. At the end of each month, we first divide all stocks in the general sample into five groups based on dispersion in analysts' earnings forecasts (labelled D1 to D5. D1 (D5) includes stocks with the lowest (highest) dispersion. Dispersion in analysts' earnings forecasts is calculated each month as the ratio of the standard deviation of analysts' current fiscal-year annual earnings-per-share forecasts to the absolute value of the mean forecast. Institutional ownership is determined by index assignment, with stocks in the bottom of the Russell 1000 index labeled R1 and those in the top of the Russell 2000 index labeled R2. Stocks in R1 have lower institutional ownership than those in R2. We end up with 10 (2X5) portfolios. After being assigned to portfolios, stocks are held for one month. RaDb represents a stock selected from group Ra and Db. D1-D5 is the return of the hedge portfolio that holds a long position in stocks of D1 and a short position in stocks of D5. The monthly return (in percentage) of each portfolio is the equal-weighted average of the returns of all the stocks in the portfolio. We report returns in **bold**, including raw monthly returns (RAW), the CAPM alpha, alpha from the Fama-French three-factor model (FF3), and alpha from the Fama-French three-factor model augmented by the Carhart momentum factor (FF4). Newy-West t-statistics are reported. ***, **, and * correspond to significance at 1%, 5%, and 10%, respectively.

Panel A. Institutional Ownership and the Dispersion Effect

Raw Return	*Dispersion*					
Institutional Ownership	**D1**	**D2**	**D3**	**D4**	**D5**	**D1-D5**
R1	**1.27**	**0.98**	**1.52**	**0.90**	**0.82**	**0.45**
t-stat	3.79	3.10	5.45	2.97	2.55	1.21
R2	**1.48**	**1.06**	**1.16**	**0.96**	**0.74**	**0.74*****
t-stat	5.99	4.02	4.37	2.68	2.50	2.69

CAPM	*Dispersion*					
Institutional Ownership	**D1**	**D2**	**D3**	**D4**	**D5**	**D1-D5**
R1	**0.85**	**0.61**	**1.14**	**0.42**	**0.41**	**0.44**
t-stat	2.55	1.92	3.92	1.41	1.23	1.17
R2	**1.12**	**0.67**	**0.76**	**0.57**	**0.40**	**0.71*****
t-stat	4.14	2.44	2.77	1.54	1.32	2.61

FF3	*Dispersion*					
Institutional Ownership	**D1**	**D2**	**D3**	**D4**	**D5**	**D1-D5**
R1	**0.79**	**0.62**	**1.15**	**0.32**	**0.43**	**0.36**
t-stat	2.39	1.87	3.85	1.01	1.30	0.94
R2	**1.06**	**0.69**	**0.74**	**0.53**	**0.41**	**0.65****
t-stat	4.19	2.48	2.73	1.41	1.29	2.39
FF4	*Dispersion*					
Institutional Ownership	**D1**	**D2**	**D3**	**D4**	**D5**	**D1-D5**

R1	**0.86**	**0.64**	**1.25**	**0.44**	**0.57**	**0.29**
t-stat	2.51	1.81	4.10	1.27	1.63	0.73
R2	**1.13**	**0.80**	**0.81**	**0.58**	**0.38**	**0.75****
t-stat	4.05	2.76	2.68	1.52	1.15	2.58

Panel B. Test of Hypothesis 2 (*Self-Censoring*)

H_0 **:** **(**R2D1-R2D5 $\leq$ R1D1-R1D5**)**

H_A **:** *Hypothesis 2* **(**R2D1-R2D5>R1D1-R1D5**)**

	RAW	**CAPM**	**FF3**	**FF4**
(R2D1-R2D5)-(R1D1-R1D5**)**	**0.29**	**0.27**	**0.29**	**0.46**
p-value	0.24	0.25	0.24	0.14

Panel C. Average Number of Stocks in Each Portfolio

Observations	*Dispersion*				
Institutional Ownership	**D1**	**D2**	**D3**	**D4**	**D5**
R1	10	10	11	10	12
R2	14	14	14	14	13