Ambiguity, Investor Disagreement, and Expected Stock Returns

Lawrence Hsiao*

Job Market Paper

click here for the latest version

Abstract

Motivated by an implication of the model of market-trading with disagreement by Kandel and Pearson (1995), I measure investor disagreement (ID) as the correlation coefficient between trading volume and absolute price change, multiplied by -1. This leads to my main novel empirical findings: (1) stocks in the highest ID decile outperform stocks in the lowest ID decile by 9.24 percent annually, adjusted for exposures to the market return as well as size, value, momentum, and liquidity factors; (2) stocks in the highest ID decile prior to earnings announcements earn significantly higher earnings announcement returns. Furthermore, I develop a theory consistent with these findings by extending Kandel and Pearson (1995) to incorporate traders' ambiguity aversion and their ambiguity about other traders' interpretations. Specifically, my model generates a positive relation between ID and expected stock return.

*Lawrence Hsiao is at Kellogg School of Management, Northwestern University at 2211 Campus Dr, Evanston, IL 60208 (e-mail: lawrence.hsiao@kellogg.northwestern.edu; phone: 773-681-2335). I am extremely grateful to my dissertation committee, Robert Korajczyk (co-chair), Charles Nathanson, Pietro Ortoleva (co-chair), and Viktor Todorov, for their continuous support, mentoring, and encouragement. The paper has also benefited from Fotis Grigoris (discussant), Peter Klibanoff, Anna Scherbina (discussant), Laura Starks, Wei Wu, and seminar participants at American Finance Association (scheduled), European Finance Association Doctoral Tutorial, Yale SOM Lynne & Andrew Redleaf Foundation Graduate Student Conference, Dauphine Finance Ph.D. Workshop, University of Texas at Austin PhD Student Symposium on Financial Market Policy Development and Research, Financial Markets and Corporate Governance Conference, Wharton Inter-Finance PhD seminar, and Kellogg Brown Bag Seminar. I am grateful to research grants provided by 2021 Ministry of Science and Technology Taiwanese Overseas Pioneers Grants (TOP Grants) for PhD Candidates. All errors are my own. First draft: December 15, 2019.

1 Introduction

People disagree. In financial markets, investors have different interpretations of public information. Many theoretical models in the economics and finance literature also assume that investors can differ in how they interpret fundamentals.[1] In this paper, I focus on two questions: how do we measure investor disagreement and how is investor disagreement related to the cross-section of expected stock returns?

Empirically, it is difficult to measure investor disagreement. Past studies typically measure investor disagreement by analyst forecast dispersion, trading volume or volatility.[2] There are, however, some concerns with theses approaches. For instance, analysts receive similar training and interact with other analysts frequently, so the views of analyst forecasts may under-represent the views among investors, as proposed by Daniel et al. (2002), Anderson et al. (2005), and Erturk (2006). Furthermore, analysts are biased in their forecasts due to agency problems, and thus tend to make overly optimistic forecasts, incorporate negative news into their forecasts sluggishly, and follow trends.[3] On the other hand, standard turnover or volatility variation can be more consistent with classical asset pricing stories that do not feature investor disagreement at all, making it difficult to attribute those variations to dispersion in beliefs.

In this paper, I use the correlation coefficient between trading volume and absolute price change, multiplied by -1, to measure investor disagreement. To motivate this measure, I use simulation to show that the correlation coefficient between volume and absolute price change is smaller when investor disagreement is higher based on an implication of Kandel and Pearson (1995)'s market-trading with disagreement model. The intuition is that, if investors actively trade in the exact opposite directions (high investor disagreement), large trading volume can be accompanied by a small price change (correlation between volume and absolute price change is low).[4] In other words, the correlation coefficient of daily trading volume and absolute price change serves as a negative indicator for investor disagreement

[1] See, for example, Harrison and Kreps (1978), Harris and Raviv (1993), Scheinkman and Xiong (2003), Cao and Ou-Yang (2008), and Banerjee and Kremer (2010).

[2] For analyst forecast dispersion, see, e.g., Diether et al. (2002), Doukas et al. (2006), Sadka and Scherbina (2007), and Barinov (2013). For trading volume, see, e.g., Garfinkel and Sokobin (2006), Garfinkel (2009), and Berkman et al. (2009). For volatility, see, e.g., Boehme et al. (2006) and Chatterjee et al. (2012).

[3] For analyst optimism, see, e.g., De Bondt and Thaler (1985), La Porta (1996), Dechow and Sloan (1997), and Brown (2001). For slow incorporation of negative information by analysts, see, e.g., Chan et al. (1996), Easterwood and Nutt (1999), Lim (2001), and Conrad et al. (2006). For analyst herding, see, e.g., Graham (1999), Welch (2000), Lamont (2002), and Hong and Kubik (2003).

[4] Kim and Verrecchia (1991) and Harris and Raviv (1993) indicate that when there's no investor disagreement, volume should be perfectly proportional to absolute price change.

(ID). Empirically, ID for a stock at the end of a given month is defined as the contemporaneous correlation coefficient of daily trading volume and absolute price change over the past two months, multiplied by -1.[5]

Now that we have an investor disagreement measure, ID, I proceed to examine the relation between ID and expected stock returns in the cross section. At the end of each month, I sort stocks into ten decile portfolios based on ID. Then, I examine the returns for each ID decile in the subsequent month. I find that when moving from the lowest ID decile to the highest ID decile, mean returns increase almost monotonically. In particular, stocks in the highest ID decile significantly outperform stocks in the lowest ID decile by 65 basis points per month with a Newey and West (1987) t-statistic of 3.91 from January 1983 to December 2019. The corresponding monthly differences in CAPM, three-, four-, and five-factor alphas are 0.87% (t-statistic = 5.64), 0.71% (t-statistic = 5.47), 0.75% (t-statistic = 5.94), and 0.77% (t-statistic = 6.28), respectively.[6] The univariate portfolio results indicate a strongly positive relation between ID and expected stock returns.

In addition to univariate portfolio analysis, I perform bivariate portfolio analysis to ensure that the significantly positive return differences between high and low ID decile are not driven by well-known stock characteristics or risk factors. In particular, I control one at a time for 12 return predictors, including firm size (SIZE), book-to-market (BM) ratio, the cumulative return over the 11 months prior to the portfolio formation month (MOM), short-term reversal (REV), average turnover ratio (TURN), idiosyncratic volatility (IVOL) as defined in Ang et al. (2006), Amihud (2002) illiquidity ratio (ILLIQ), demand for lottery stocks with extreme positive returns (MAX) as defined in Bali et al. (2011), institutional ownership ratio (IOR), the stock beta (BETA), co-skewness (COSKEW) as defined in Harvey and Siddique (2000), and analyst forecast dispersion (DISP) as defined in Diether et al. (2002). After controlling for each of the above variables, the return differences between high ID and low ID decile portfolios are in the range of 0.34% and 0.68% per month with Newey and West (1987) t-statistics ranging from 3.41 to 6.78. The corresponding 5-factor alpha differences are in the range of 0.45% to 0.72% and are all highly significant.

It is, however, possible that some information is lost via portfolio aggregation. Hence, in order to control for multiple variables simultaneously, I implement Fama and MacBeth

[5]The number of trading days is around 44 in two months. Using only one trading month to compute a correlation coefficient may be subject to lack of statistical power. The asset pricing implications of ID are the same if I instead use turnover ratio to substitute for trading volume, or use squared return to substitute for absolute price change.

[6]The returns reported here are equal-weighted as in Table 1. I also present value-weighted returns in Table 2. For robustness checks, I also use DGTW-adjusted returns following Daniel et al. (1997) and Wermers (2003), and the results remain the same. The results are available upon request.

(1973) regressions to examine the cross-sectional relation at the stock level. The results suggest that the relation between ID and future stock returns remains positive and highly statistically significant when a large set of control variables is included. I also perform a battery of robustness checks. I find that the significantly positive relation between ID and future stock returns persists in high and low sentiment periods (Baker and Wurgler (2006)), NBER recessions and expansions, and high and low economic uncertainty periods (Jurado et al. (2015), Ludvigson et al. (2015), and Baker et al. (2016)). All these results provide strong evidence for a positive relation between ID and expected stock returns.

Next, I provide solid evidence that investor disagreement (ID) is persistent. First, I run ID on lagged ID with a large set of control variables. The coefficient on lagged ID is 0.468 with an adjusted R-squared of 34.54%, which implies that ID is highly persistent. In addition, I examine the average month-to-month ID transition matrix and find that all diagonal probabilities exceed 10%. In particular, the diagonal probabilities are 43.03% and 38.06% for the lowest and highest ID decile, respectively. I further vary both the number of months in the formation of ID and the portfolio holding periods and find that the long-short ID strategy is robust to those variations.

I also examine whether the positive relation between ID and future stock returns holds in the earnings announcement setting. As firms typically use earnings announcements to communicate relevant information to the market, there exists a sudden increase of information prior to the earnings announcement for investors to disagree on.[7] Using portfolio sorts and stock level cross-sectional regressions, I find that stocks with high ID prior to the earnings announcement experience significantly higher cumulative abnormal returns around the earnings announcement period compared to stocks with low ID. In particular, stocks in the highest ID decile prior to earnings announcements outperform stocks in the lowest ID decile by 65 basis points in the 3-day window around earnings announcements with a Newey and West (1987) t-statistic of 4.77. The positive relation is also robust to variations in different windows to compute ID prior to earnings announcements or different earnings announcement windows.

The above empirical findings indicate a positive relation between investor disagreement and expected stock returns. A natural question is, what is the underlying mechanism? The literature still disagrees on how investor disagreement should be related to expected stock returns. The two competing views are represented by Miller (1977) and Merton (1987). First, Miller (1977) posits that in the presence of short-sales constraints, stock prices are biased

[7] Ball and Brown (1968), Krinsky and Lee (1996), Back et al. (2018), and Yang et al. (2020) argue that the leaking of information is pervasive prior to earnings announcements.

upward (lower future returns) when disagreement among investors is high. This is because when pessimists can't freely trade on the negative information as a result of short-sales constraints, asset prices are mainly set by optimists. Morris (1996), Viswanathan (2001), and Chen et al. (2002) also suggest that prices typically reflect a more optimistic valuation due to high short-sale costs.

In a traditional CAPM world, idiosyncratic risk is not priced since investors can hold efficiently diversified portfolios. Merton (1987), however, argues that investors tend to hold stocks they are familiar with and thus hold under-diversified portfolios. Naturally, they demand compensation to hold low visibility stocks with idiosyncratic risk. Since high disagreement indicates higher variation in earning streams, stocks with high divergence of opinion should earn higher future returns. However, as Merton (1987)'s hypothesis only applies to stocks with low visibility, it is difficult to explain the relation between disagreement and expected stock returns for high visibility stocks.[8]

On the other hand, some models consider disagreement as a source of "speculation risk" and also predict a positive relation between disagreement and future stock returns. For instance, David (2008) constructs a general equilibrium model in which two types of agents have heterogeneous beliefs about future fundamental growth. Agents face the risk that market prices move more in line with the trading models of competing agents than with their own, and thus speculate with each other. Gao et al. (2019) also argue that when investors agree to disagree, they both expect to profit at the expense of their trading counterparties. One assumption of David (2008) is that each trader is absolutely convinced that his own belief is correct.[9] In other words, traders agree to disagree all the time and never use others' beliefs to update their priors. This assumption seems too strong as in reality traders may adjust their beliefs after observing others' beliefs.

To motivate my empirical findings of the positive relation between investor disagreement and expected stock returns, I extend Kandel and Pearson (1995) model by incorporating traders' ambiguity aversion and their ambiguity to others' beliefs. To begin with, there is a public signal which is equal to the sum of the true value of a stock and some noise. Two types of traders know the true variance of the noise but have different prior beliefs on the mean of the noise. For simplicity, I use the word "interpretation" to describe traders' prior beliefs on the mean of the noise, and investor disagreement is defined as the absolute value

[8] Boehme et al. (2009) provide a detailed summary and empirically show that idiosyncratic risk is positively related to future stock returns only within stocks with low visibility.

[9] This assumption is similar to that of the overconfidence model of Daniel et al. (2005), in which each investor assigns an excessively large weight on his own model.

of the difference in two types of interpretations. By setup, higher interpretation of the signal corresponds to a more pessimistic valuation.

I relax the strong assumption in David (2008) that traders agree to disagree all the time. In particular, I assume that when traders observe the other type's interpretation, they believe that it can range from being less accurate to being more accurate than their own interpretation. This assumption stems from the fact that most traders in the market lack the ability, time, and information to accurately interpret others' interpretations. Formally, since traders are ambiguous about the information quality of the other type's interpretation, they assign a range of information precision to it. In addition, in most ambiguity models, it is the variance of the signal that is difficult to judge.[10] In this model, however, the public signal is not ambiguous as its variance is known. In other words, traders are ambiguous about the other type's interpretation of the signal rather than the signal itself.

The essential behavioral assumption in this model is that traders are ambiguity-averse.[11] Several papers have provided evidence that many people exhibit ambiguity aversion. For example, in an experiment involving 104 individuals who are asked to choose between an ambiguous urn and a risky urn, Halevy (2007) finds that 61% are ambiguity-averse, 22% are ambiguity neutral, and 17% prefer the ambiguous urn. Using a Unicredit sample of 1,686 Italian retail investors, Butler et al. (2014) find that 52% are ambiguity-averse and 25% are ambiguity-neutral. Dimmock et al. (2016) find that out of $3,258$ respondents in the American Life Panel (ALP), 52% are ambiguity averse, 10% are ambiguity-neutral, and 38% are ambiguity-seeking. In order to model ambiguity aversion, traders' preferences will be represented using the maxmin expected utility model of Gilboa and Schmeidler (1989).[12] Under the maxmin expected utility, agents have a set of probability measures and evaluate any action using the probability that minimizes the expected utility of that action.

I show that the expected stock return is increasing in investor disagreement, holding fixed the average interpretation in the market. The intuition is that ambiguity aversion induces traders to take into account the other type's interpretation asymmetrically, i.e., traders give more (less) weight to the other type's interpretation if it is higher (lower) than their own

[10] See, for example, Epstein and Schneider (2007), Epstein and Schneider (2008), Easley and O'Hara (2010), and Illeditsch (2011).

[11] The concept of ambiguity aversion in economics can be traced back to at least the Ellsberg Paradox (Ellsberg (1961)), which suggests that individuals are averse to vague probabilities and may not act as if they have a single prior.

[12] There are different forms of preferences in the literature that reflect ambiguity aversion, including the smooth ambiguity model by Klibanoff et al. (2005) and Klibanoff et al. (2009), "α-maxmin" model of Ghirardato et al. (2004), and "robust control" by Hansen and Sargent (2007) and Hansen and Sargent (2011). Though this paper adopts the maxmin expected utility formulation, a brief discussion of the extent to which alternative models would push on the results can be found in Section 4.2.

interpretation. In other words, when updating their beliefs, the optimistic trader assigns relatively more weight to the pessimistic view while the pessimistic trader assigns relatively less weight to the optimistic view. Hence, when investor disagreement is higher, stock price reflects a more pessimistic valuation (higher expected return).

In addition, I examine whether earnings announcements resolve disagreement among investors. I find that ID on average increases after earnings announcements. In particular, compared to good earnings news, bad earnings news trigger a larger increase in ID. I also obtain firm-specific public news stories from RavenPack and classify them into six different news categories (Financial, Legal, M&A, Operational, Ratings, and Others). It turns out that ID also increases after these firm-specific news stories. In contrast, ID before and after macroeconomic announcements like FOMC meetings remain virtually the same.

This paper mainly contributes to the disagreement literature, both theoretically and empirically. First, the correlation coefficient between ID and analyst forecast dispersion, turnover ratio, and idiosyncratic volatility are 0.001, -0.201, and -0.199, respectively, suggesting that ID is not picking up these existing disagreement or uncertainty measures.[13] In addition, the positive relation between ID and future stock returns remains highly significant after controlling for these three proxies. Second, to the best of my knowledge, this is the first paper to incorporate ambiguity aversion into a disagreement model to study the relation between investor disagreement and expected stock returns. To the extent that investors experience ambiguity, it seems especially plausible that they do so when thinking about the interpretations belonging to other investors. This observation plus the experimental evidence on ambiguity aversion motivates the way it is used in our model. In addition, the model's prediction applies to all stocks, while Merton (1987)'s hypothesis applies only to stock with low visibility.

The paper also sheds light on the earnings announcement and disclosure literature. As Berkman et al. (2009) point out, it has been difficult to isolate the effect of disagreement from other effects such as momentum or post-earnings announcement drift in the traditional monthly returns setting. Hence, examining the relation between disagreement and expected stock returns in the earnings announcement setting is important. In addition, how uncertainty evolves following disclosure events has long received considerable attention.[14] This paper

[13]As many firms are covered by only 2 to 3 analysts, it is unsurprising that the correlation coefficient between ID and analyst forecast dispersion is close to 0. When I restrict my observations to the ones that are covered by more than 20 analysts, the correlation coefficient between ID and analyst forecast dispersion is 0.020, suggesting a weakly positive relation between the two.

[14]For instance, Patell and Wolfson (1979) have documented immediate decline in volatility after earnings announcements, which reflects the resolution of uncertainty. In addition, Billings et al. (2015) find that implied volatility decreases after guidance announcements, while Rogers et al. (2009) find that earnings

complements the literature by providing the evolution of investor disagreement following both firm-specific and macroeconomic news.

The paper is organized as follows. Section 2 introduces Kandel and Pearson (1995)'s model to motivate the investor disagreement (ID) measure. Section 3 describes the data, variables, and the main empirical tests. Section 4 incorporates ambiguity aversion into the model. Section 5 examines the evolution of ID from before to after earnings announcements, firm-specific news stories, and FOMC announcements. Section 6 provides a summary and concludes.

2 Measuring investor disagreement

In this section I introduce a simplified version of the Kandel and Pearson (1995) model. There are three time periods ($t = 0$, $t = 1$, and $t = 2$). There are two assets in a competitive market: a risk-free asset with a zero rate of return and a stock with an uncertain payoff X that is realized at $t = 2$. The stock is assumed to be in zero net supply. There is a continuum of type 1 traders and a continuum of type 2 traders in the market, with each type constituting half of the total traders.

At $t = 0$, traders can have different prior beliefs on X. In particular, type i traders' prior beliefs of X are given by normal distributions of mean X_i and precision Z_i, where $i \in \{1, 2\}$. In addition, traders don't know others' beliefs and likelihood functions. From an outsider's point of view, X_1 and X_2 are unbiased estimates of X, which means that $E[X_1] = E[X_2] = X$.

At $t = 1$, a public signal S arrives and traders observe S. The informative while noisy signal is given by $S = X + \eta$, where η is independent of X, $\eta \sim N(\mu_\eta, \sigma_\eta^2)$, and $0 < \sigma_\eta^2 < \infty$. Everything about S is common knowledge except for the mean μ_η. In particular, type i traders believe that

$$\mu_\eta \sim N(\mu_i, \sigma^2), \tag{1}$$

where μ_i denotes type i traders' interpretation of S and $0 < \sigma^2 < \infty$. Hence, from type i traders' point of view,

$$S \sim N(X + \mu_i, \sigma^2 + \sigma_\eta^2). \tag{2}$$

In other words, type i traders think that S is higher than X if $\mu_i > 0$, and higher μ_i implies a more negative view of the same signal. In particular, following the setup in Kandel and Pearson (1995), traders at $t = 0$ don't take into account the fact that prices will be "wrong"

announcements increase short-term volatility.

at $t = 1$ because others are potentially using different likelihood functions to update their beliefs and trade. Traders strive to maximize their final wealth W at $t = 2$, and are endowed with negative exponential utility functions $U(W) = -e^{-\lambda W}$, where λ is the coefficient of absolute risk aversion.

Since there are only two types of traders in the market, the equilibrium trading volume from $t = 0$ to $t = 1$, $V^*_{0,1}$, is the absolute change in traders' respective equilibrium holdings from $t = 0$ to $t = 1$. Let $m_{i,t}$ denote the position held by each of type i traders at time t. In particular,

$$V^*_{0,1} = |\frac{1}{2}m_{1,1}(P^*_1) - \frac{1}{2}m_{1,0}(P^*_0)| = |\frac{1}{2}m_{2,1}(P^*_1) - \frac{1}{2}m_{2,0}(P^*_0)|. \tag{3}$$

The following proposition describes the joint behavior of volume and price change.

Proposition 1. *Let $b = (\sigma^2 + \sigma^2_\eta)^{-1}$. It can be shown that*

$$V^*_{0,1} = |A + B\Delta P^*_{0,1}|, \tag{4}$$

where $\Delta P^*_{0,1} = (P^*_1 - P^*_0)$,

$$A = \frac{b(\mu_1 - \mu_2)}{4\lambda}, \tag{5}$$

and

$$B = \frac{(Z_1 - Z_2)}{4\lambda}. \tag{6}$$

Proof. See page 847-850 in Kandel and Pearson (1995) . □

Note that when there is no disagreement in the market, i.e., $\mu_1 = \mu_2$, equilibrium trading volume is perfectly proportional to absolute price change (Kim and Verrecchia (1991) and Harris and Raviv (1993)), and there exists no trading volume given zero price change. However, when disagreement exists so that $\mu_1 \neq \mu_2$, there can exist trading volume given zero price change. Kandel and Pearson (1995) uses (4) to provide explanation for the existence of large trading volume around some earnings announcements with small price changes.

One implication of the model is that, when investor disagreement, $|\mu_1 - \mu_2|$, is higher, the relation between the equilibrium trading volume ($V^*_{0,1}$) and absolute price change ($|\Delta P^*_{0,1}|$) is weaker. To illustrate this idea, Figure 1 plots the correlation between equilibrium trading volume and absolute price change over different values of $(\mu_1 - \mu_2)$. Without loss of generality, μ_1 is fixed to 0, so $(\mu_1 - \mu_2)$ varies under different values of μ_2. For a given value of $(\mu_1 - \mu_2)$, I draw 100,000 observations from the distribution of $\eta \sim N(\mu_\eta, \sigma_\eta^2)$ and thus acquire 100,000 observations of S since $S = X + \eta$. The equilibrium trading volume, absolute price change, and the correlation between the two can be computed accordingly.

[Insert Figure 1 about here]

Figure 1 suggests that, the correlation coefficient between equilibrium trading volume and absolute price change is decreasing in investor disagreement, $|\mu_1 - \mu_2|$. When there is no disagreement, equilibrium trading volume and absolute price change are perfectly correlated. In addition, as long as investor disagreement is not too high, the correlation coefficient of trading volume and absolute price change is positive, which is consistent with the findings in the past literature.[15]

Hence, the contemporaneous correlation coefficient of daily trading volume and absolute price change serves as a negative indicator for investor disagreement (ID).[16] When the correlation coefficient between the two is smaller, it is more likely that disagreement among investors is higher.

3 Data and variable definitions

I proceed to examine the relation between ID and expected stock returns in the cross-section. This section contains detail of the empirical analysis. First, I describe the data and sample selection. Second, I define the ID measure. Third, I perform univariate portfolio-level analysis. Fourth, I discuss average stock characteristics in each ID decile portfolio. Fifth, I perform bivariate portfolio-level analysis to examine the return-predicting power of ID after controlling for commonly known stock characteristics and risk factors. Sixth, I implement Fama and MacBeth (1973) regressions to examine the cross-sectional relation at the stock level while controlling for multiple variables simultaneously. Finally, I perform robustness checks.

3.1 Data

The stock sample includes all common stocks (share code 10 or 11) traded on NYSE, AMEX, and Nasdaq from the Center for Research in Security Prices (CRSP) for the period from

[15] Past literature has documented a positive contemporaneous relation of volume and volatility. See for example, Clark (1973), Tauchen and Pitts (1983), Karpoff (1987), Gallant et al. (1992), and Andersen (1996).

[16] In cross-section asset pricing, the implicit assumption is that firms are ex-ante identical in a way that b and $|Z_1 - Z_2|$ are virtually the same across firms.

January 1983 to December 2019. The second data set is Compustat, which is used to obtain the equity book values for computing the book-to-market ratios of individual firms. Stocks are required to have non-missing firm size (SIZE), book-to-market (BM) ratio, and momentum (MOM), which are defined in detail in Appendix I.

3.2 Measuring investor disagreement

The first step involves measuring investor disagreement (ID) for each stock-month. I define ID at the end of a given month as the contemporaneous correlation coefficient of daily trading volume and absolute price change over the past two months (around 44 trading days), multiplied by -1. For example, investor disagreement of a stock at the end of October is defined as the correlation coefficient between its daily trading volume and absolute price change over September and October, multiplied by -1.

A stock trading day t is eligible if the price per share on $t-31$ is at least 5 dollars and has non-missing return and volume. This is to ensure the results are not driven by small, illiquid stocks or by bid-ask bounce. All returns are delisting-adjusted. Stocks are required to have at least 30 eligible trading days to compute ID. Figure 2 plots the time-series distribution of the number of all CRSP common stocks and eligible stocks to compute ID.

[Insert Figure 2 about here]

3.3 Univariate portfolio-level analysis

I first perform univariate portfolio-level analysis to examine the relation between ID and expected stock returns in the cross section. At the end of each month, I sort stocks into ten decile portfolios based on ID. Decile 1 (low ID) is the portfolio of stocks with the lowest investor disagreement, and decile 10 (high ID) is the portfolio of stocks with the highest investor disagreement. Stocks are held for one month after being assigned into ID decile portfolios.

Table 1 presents the equal-weighted monthly average returns of ID-sorted decile portfolios. When moving from the lowest to highest ID decile, the next-month average excess return increases almost monotonically from 0.18% to 0.82%. The average excess return difference between decile 10 (high ID) and decile (low ID) is 0.65% with a corresponding Newey and West (1987) t-statistic of 3.91.

[Insert Table 1 about here]

In addition to the average excess returns, Table 1 also presents the risk-adjusted returns (alphas) from regressing monthly excess return on contemporaneous risk factors. CAPM alpha is the intercept from the regression of excess portfolio returns on a constant and excess market return (MKT). Three-factor alpha is the intercept from the regression of excess portfolio returns on a constant, the excess market return (MKT), a size factor (SMB), and a book-to-market factor (HML). Four-factor alpha is the intercept from the regression of excess portfolio returns on a constant, the excess market return (MKT), a size factor (SMB), a book-to-market factor (HML), and a momentum factor (UMD) of Carhart (1997). Five-factor alpha is the intercept from the regression of excess portfolio returns on a constant, the excess market return (MKT), a size factor (SMB), a book-to-market factor (HML), a momentum factor (UMD), and a liquidity factor (LIQ) of Pástor and Stambaugh (2003). If the factor model can capture the cross-sectional variation in stock returns, then the corresponding alpha should be statistically indistinguishable from zero.

As shown in the third column in Table 1, CAPM alpha increases from −0.63% to 0.23% per month when moving from the lowest to highest ID decile. The difference in CAPM alphas between the high and low ID portfolios is 0.87% per month with a Newey and West (1987) t-statistic of 5.64. The next three columns present similar alpha results from the three-factor, four-factor, and five-factor models. When moving from the lowest to the highest ID decile, the three-factor alpha increases from −0.57% to 0.14%, the four-factor alpha increases from −0.50% to 0.25%, and the five-factor alpha increases from −0.51% to 0.26%. The difference in alphas between the high ID and low ID portfolios is 0.71% (t-statistic=5.47), 0.75% (t-statistic=5.94), and 0.77% (t-statistic=6.28) per month for the three-factor, four-factor, and five-factor model, respectively.

In addition, I examine the source of the risk-adjusted return difference between high ID and low ID portfolios. Is it generated by outperformance of high ID stocks or underperformance of low ID stocks? The last column of Table 1 indicates the strongly significant five-factor alpha spread (t-statistic=6.28) is driven by both the outperformance of high ID stocks (significantly positive with a t-statistic of 2.69) and the underperformance of low ID stocks (significantly negative with a t-statistic of −6.43).

Table 2 presents evidence from the value-weighted decile portfolios of ID. The results are slightly weaker but in general consistent with the equal-weighted portfolio results.

[Insert Table 2 about here]

Stocks in decile 1 (low ID) generate a value-weighted average excess return of 0.50% per month, while stocks in decile 10 (high ID) generate higher value-weighted average excess return of 0.91% per month. The average return differential is 0.50% per month with a Newey and West (1987) t-statistic of 2.78. The difference in alphas between the high ID and low ID portfolios is 0.52% (t-statistic=3.56), 0.40% (t-statistic=3.20), 0.37% (t-statistic=2.95), and 0.36% (t-statistic=2.84) per month for the CAPM, three-factor, four-factor, and five-factor model, respectively.

In Table 1 and Table 2, I also report betas with respect to MKT, SMB, HML, UMD, and LIQ risk factors. In both cases, MKT betas and HML betas are significantly negative and significantly positive, respectively, suggesting that compared to stocks in the lowest ID decile, stocks in the highest ID decile are less exposed to market risk and have a tilt towards value stocks. To test the hypothesis that all 10 alphas are jointly equal to zero, I implement GRS test of Gibbons et al. (1989). For both equal-weighted and value-weighted and for all regression models, the GRS test rejects at 1% level.[17]

Overall, the univariate portfolio analysis suggests a positive relation between ID and expected stock returns.

3.4 Average stock characteristics

Next, I examine the stock composition of investor disagreement (ID) decile portfolios. In particular, Table 3 presents for each ID decile, the time-series average of mean values of stock characteristics, including firm size (SIZE), book-to-market (BM) ratio, the cumulative return (in percent) over the 11 months prior to the portfolio formation month (MOM), the return (in percent) in the portfolio formation month (REV), average turnover ratio (TURN), idiosyncratic volatility (IVOL) as defined in Ang et al. (2006), Amihud (2002) illiquidity ratio (ILLIQ), lottery demand (MAX) as defined in Bali et al. (2011), institutional owernship ratio (IOR) defined the ratio of shares owned by institutions as reported in 13F filings in the last quarter, the stock beta (BETA), and co-skewness (COSKEW) as defined in Harvey and Siddique (2000). Definitions of these variables are given in Appendix I. The weights are based on the number of observations in each portfolio in each month and there is an average of 306 stocks per decile portfolio.

The first row of Table 3 reports that the average investor disagreement (ID) increases from -0.77 to 0.07 when moving from the lowest to highest ID decile. The average ID in the subsequent month increases monotonically from -0.58 to -0.12 from the lowest to highest ID decile, which sheds light on the persistence of investor disagreement.

[Insert Table 3 about here]

[17] For brevity I didn't report the results here. However, all the test statistics are available upon request.

Fama and French (1992) and Fama and French (1993) report that on average, small stocks earn higher future returns than large stocks. The third row of Table 3 indicates the average market capitalization (SIZE) slightly increases and then decreases when moving from the low ID decile to high ID decile. In fact, SIZE is relatively large around middle ID deciles. This is perhaps because large firms benefit more from their disclosure policy compared to small firms (Diamond and Verrecchia (1991)) due to lower information and proprietary costs. As large firms on average tend to be more transparent, investors disagree less.

This result provides further support for the return differences between high and low ID decile in Table 2 and Table 3 since if small stocks do earn higher subsequent returns, then low ID decile should earn higher returns than middle ID decile.

The average book-to-market (BM) ratio for each investor disagreement (ID) decile is reported in the fourth row. As ID increases across the deciles, BM increases monotonically. The concentration of high book-to-market stocks in the high ID deciles casts doubt on the positive relation between ID and expected stock returns, as Fama and French (1992) and Fama and French (1993) document that stocks with high BM ratio stocks (value stocks) earn higher subsequent returns than stocks with low BM ratio (growth stock).

Looking at the fifth and sixth row of Table 3, one observes that as ID increases across the deciles, both momentum (MOM) and short-term reversal (REV) decrease. The decrease in MOM is good news as Jegadeesh and Titman (1993) shows that stocks that perform the best (worst) over intermediate horizons tend to do well (poorly) in the future. If past losers do continue to perform badly in the future, high ID stocks should experience low instead of high returns. However, the decrease in REV across ID deciles casts doubt on the significance of the long-short ID strategy, as stocks tend to exhibit return reversal due to initial price overreaction to good news and bid-ask bounce (Jegadeesh (1990) and Lehmann (1990)).

Gervais et al. (2001) find that stocks withe higher volume earn higher returns, which is known as the high volume return premium. Looking at the seventh row of Table 3, stock turnover ratio (TURN) decreases monotonically when ID increases. The pattern is good news for the positive relation between ID and expected stock returns, as the concentration of high trading volume stocks in low ID deciles would suggest these portfolios earn higher instead of lower returns observed in the data.

Next, the eight row of Table 3 indicates that as ID increases across deciles, average idiosyncratic volatility (IVOL) decreases. As Ang et al. (2006) present evidence that stocks with high idiosyncratic volatility generate lower future returns, the negative relation between ID and idiosyncratic volatility raises concern on the positive relation between ID and future stock returns. On the other hand, Amihud (2002) suggests that expected stock returns increase in illiquidity. Looking at the ninth row of Table 3, there exists no striking pattern of illiquidity across ID deciles.

As shown in the tenth row of Table 3, the average demand for lottery stocks with extreme positive returns (MAX) is lower for stocks in high ID deciles. Since Bali et al. (2011) and Bali et al. (2017) document that low MAX stocks earn higher expected returns than high MAX stocks, the negative relation between ID and MAX casts doubt on the positive relation between ID and future stock returns.

Looking at the eleventh row of Table 3, institutional ownership ratio (IOR) decreases as ID increases. This negative relation between ID and IOR provides support of the positive relation between ID and expected stock returns, as Asquith et al. (2005) find that short-sale constrained stocks with low institutional ownership significantly underperform than high institutional ownership stocks.

Next, the twelfth row of Table 3 indicates that when ID increases across deciles, average stock beta (BETA) decreases monotonically. This pattern suggests that high ID stocks are less exposed to market risk. If stocks are compensated more for bearing more exposure to market risk, stocks with higher ID should instead earn lower future returns. Hence, the negative relation between ID and BETA is good news for the return differences between the high and low ID decile as reported in Table 1 and Table 2.

On the other hand, in the thirteenth row of Table 3, average co-skewness (COSKEW) first increases then decreases when moving from the lowest to the highest ID decile. Compared to low ID deciles, high ID deciles on average have lower co-skewness, which further provides support for the positive relation between ID and future stock returns, since Harvey and Siddique (2000) report that stocks with high co-skewness generate lower one-month-ahead returns. In the fourteenth row of Table 3, average analyst forecast dispersion (DISP) appears to be stable across ID deciles, which indicates that ID is not picking up or related to dispersion in beliefs among analysts.

In sum, Table 3 indicates that compared to low ID stocks, high ID stocks on average have high book-to-market (BM) ratio, low intermediate-horizon momentum (MOM), low short-term reversal (REV), low turnover ratio (TURN), low idiosyncratic volatility (IVOL), low demand for lottery stocks (MAX), and low exposure to market risk (BETA). In particular, the fact that high ID stocks having high BM, low REV, low IVOL, and low MAX seems to dampen the validity of the positive relation between ID and expected stock returns. In the next section, I use bivariate portfolio sorts to show that the positive relation between ID and expected stock returns is not driven by the above return predictors.

3.5 Bivariate portfolio-level analysis

The section studies whether the relation between investor disagreement (ID) and expected stocks returns still holds after controlling for the well-known cross-sectional return predictors: market capitalization (SIZE), book-to-market (BM) ratio, momentum (MOM), short-term reversal (REV), turnover ratio (TURN), idiosyncratic volatility (IVOL), illiquidity (ILILQ), demand for lottery stocks with extreme positive returns (MAX), institutional ownership ratio (IOR), the stock beta (BETA), co-skewness (COSKEW), and analyst forecast dispersion (DISP).

I first examine whether the results in Table 1 and Table 2 are simply capturing a size effect. Each month, I assign stocks to one of five quintiles based on firm size (SIZE).[18] Within each size quintile, stocks are further sorted into deciles based on ID in the previous month. I then examine the next month returns in each portfolio. Table 4 shows that the return differential is positive and highly significant in all size quintiles. In addition, the average equal-weighted monthly return differential between high ID and low ID stocks decreases when moving from the smallest to the largest size quintile (except when going from the second to the third size quintile).

[Insert Table 4 about here]

In particular, the long-short ID strategy for the smallest and the largest size quintile on average generates a return of 1.07% and 0.28% per month, with a Newey and West (1987) t-statistic of 3.75 and 2.14, respectively. In addition, the corresponding CAPM, three-factor, four-factor, and five-factor alphas are all significantly positive. Specifically, the five-factor alpha differences are in the range of 0.29% to 1.32% per month with t-statistics ranging from 2.29 to 4.57. The above results indicate that the strongly positive relation between ID and expected stock returns is not driven by size effect.

Table 5 presents the results of two-way cuts on book-to-market (BM) ratio and ID. The return differential and corresponding CAPM, three-factor, four-factor, and five-factor alphas between low and high ID stocks are highly significant in all book-to-market quintiles, indicating that the positive relation between ID and expected stock returns is not simply capturing a book-to-market effect. In addition, compared to other BM quintiles, the long-short

[18] As a robustness check, I also form portfolios using NYSE-based market capitalization. The results are similar and are available upon request.

ID strategy in the lowest BM quintile generates the highest return of 0.97% per month with a Newey and West (1987) t-statistic of 5.33.

[Insert Table 5 about here]

Table 6 presents the double sorts results on momentum (MOM) and ID. Again, the return differential and CAPM, three-factor, four-factor, and five-factor alphas between high and low ID stocks remain highly significant across all momentum quintiles. In particular, the return differential between high and low ID stocks is the highest in the stocks that are past losers. In particular, the long-short ID strategy generates a five-factor alpha of 1.54% with a Newey and West (1987) t-statistic of 6.57 in the lowest momentum quintile.

[Insert Table 6 about here]

Overall, Table 5, Table 6, and Table 7 indicate that the significantly positive relation between ID and future stock returns cannot be explained by the well-known size, value, or momentum effect. In addition, the return differential between high and low ID stocks is most pronounced in small stocks, growth stocks, and stocks that perform poorly over the past year.

I proceed to control for other commonly used return-predicting stock characteristics. In each month, stocks are first sorted into deciles based on a control variable and then, within each decile I sort stocks into deciles based on ID. Stocks are held for month and portfolio returns are equal-weighted. For brevity, I do not report returns for all 100 (10×10) portfolios. Instead, the ten investor disagreement decile portfolios are averaged over each of the ten control variable decile portfolios. Table 8 reports for each control variable the time-series average of excess returns, high-minus-low excess returns, and corresponding five-factor alphas, together with Newey and West (1987) t-statistics to examine their statistical significance.

[Insert Table 8 about here]

Table 8 shows that after controlling for many cross-sectional return predictors, the return differences between high ID and low ID decile portfolios are in the range of 0.34% and 0.68% per month with Newey and West (1987) t-statistics ranging from 3.41 to 6.78. The corresponding five-factor alpha differences are in the range of 0.45% to 0.72% and are all highly significant. In particular, when controlling for TURN, IVOL, and DISP, the five-factor alpha difference is 0.65% (t-statistic=6.50), 0.45% (t-statistic=5.29), and 0.50% (t-statistic=5.03), respectively, which provides evidence that the positive relation between ID and expected stock returns is not simply picking up existing disagreement measures. Overall, the results in this section indicate that well-known firm characteristics or risk factors cannot explain the positive relation between ID and expected stock returns.

3.6 Firm-level cross-sectional regressions

So far, the significance of investor disagreement (ID) as a determinant of the cross-section of expected returns has been examined at the portfolio level (both univariate and bivariate). The portfolio-level analysis is non-parametric since no functional form on the relation between the ID and the future returns is imposed. In addition, it is possible that some information is lost via portfolio aggregation and it is difficult to control for multiple variables simultaneously via portfolio analysis. Moreover, the Gibbons et al. (1989) tests seldom come close to rejecting the hypothesis that the three-factor, four-factor, or five-factor model explains average returns. Hence, I examine the cross-sectional relation between ID and expected returns at the stock level using Fama and MacBeth (1973) regressions. The incremental predictive power of ID can be examined relative to other control variables known to explain the cross-section of returns.

Table 9 reports the time-series averages of the slope coefficients from the regressions of one-month-ahead stock returns on ID with and without control variables. The average slopes provide standard Fama and MacBeth (1973) tests for determining which explanatory variables on average have nonzero premiums. Specifically, I run the following monthly cross-sectional regressions at a monthly frequency from January 1983 to December 2019:

$$R_{i,m+1} = \alpha_m + \beta_m ID_{i,m} + \lambda_m X_{i,m} + \epsilon_{i,m+1}, \tag{7}$$

where $R_{i,m+1}$ is the realized excess return on stock i in month $m+1$, ID is the investor disagreement of stock i at the end of month m, and $X_{i,m}$ is the same set of stock-specific control variables at time m for stock i, including firm size (SIZE), book-to-market (BM) ratio, momentum (MOM), short-term reversal (REV), turnover ratio (TURN), idiosyncratic volatility (IVOL), illiquidity (ILILQ), demand for lottery stocks with extreme positive returns (MAX), institutional ownership ratio (IOR), the stock beta (BETA), co-skewness (COSKEW), and analyst forecast dispersion (DISP).

[Insert Table 9 about here]

Table 9 reports the time-series averages of the slope coefficients with corresponding Newey and West (1987) t-statistics in parentheses. In the first column, the average slope coefficient from regressing realized returns on ID alone is 0.780 and highly significant (t-statistic = 3.85), indicating a strongly positive relation between ID and expected stock returns.

Column 2 of Table 9 controls for firm size (SIZE), book-to-market (BM) ratio, and momentum (MOM), and the coefficient on ID remains economically and statistically significant. Column 3 further controls for the short-term reversal (REV) and turnover ratio (TURN). Still, the average slope on ID is positive and highly significant. Column 4 of Table 9 shows that after including idiosyncratic volatility (IVOL), illiquidity (ILLIQ), demand for lottery sock with extreme positive returns (MAX), and institutional ownership ratio, the average slope on ID becomes 0.367 with a highly significant Newey and West (1987) t-statistic of 3.73. Column 5 further includes market beta (BETA) and co-skewness (COSKEW), and the coefficient on ID is still significantly positive. Finally, Column 6 incorporates analyst forecast dispersion and the coefficient on ID shrinks to 0.251 with a Newey and West (1987) t-statistic of 2.86.

The coefficients on most control variables are consistent with evidence in the literature. Stocks exhibit strong intermediate-horizon momentum and short-term reversals. The average slopes are significantly negative for idiosyncratic volatility, institutional ownership ratio, and analyst forecast dispersion, which is consistent with the evidence in Ang et al. (2006), Asquith et al. (2005), and Diether et al. (2002).

Overall, the multivariate Fama-MacBeth regression results in Table 9 indicate that when simultaneously controlling for various stock characteristics and risk factors, the average slopes on ID remain positive and highly significant, indicating a strongly positive relation between ID and the cross-section of expected stock returns.

3.7 Robustness checks

In this section, I provide a variety of robustness checks to examine whether the positive relation between investor disagreement (ID) and future stocks returns is nonlinear and thus changes over time. I also examine the persistence of ID.

3.7.1 Business cycles, investor sentiment, and economic uncertainty

I first examine whether the long-short ID strategy is sensitive to business cycles and investor sentiment in Table 10. In the second and the third column, the five-factor alphas and corresponding Newey and West (1987) t-statistics of each ID decile and the long-short ID strategy are reported under economic expansions and recessions. The expansions and recessions months are issued by the National Bureau of Economic Research's (NBER) Business

Cycle Dating Committee.[19] Specifically, a recession is the period between a peak of economic activity and its subsequent trough. Between trough and peak, the economy is in an expansion. There are 410 expansions and 34 recessions from January 1983 to December 2019.

[Insert Table 10 about here]

The equal-weighted five-factor alpha increases from −0.49% to 0.22% and from −0.55% to 0.69% per month for expansions and recessions, respectively. In particular, the difference in alphas is 0.71% (t-statistic=5.38) for expansions and 1.23% (t-statistic=2.04) for recessions. The results provide strong evidence that the significantly positive relation between ID and future stock returns is robust to different business cycles.

In addition, it is possible that the positive relation between ID and future stock returns is concentrated in certain investor sentiment periods. To mitigate this concern, I first classify each month as following either a high-sentiment month or a low-sentiment month. A high-sentiment (low-sentiment) month is one in which the value of the BW (Baker and Wurgler (2006)) sentiment index in the previous month is above (below) the median value for the sample period.[20] The fourth and the fifth column show that long-short ID strategy generates a five-factor alpha of 0.65% (t-statistic=5.25) and 0.92% (t-statistic=4.53) per month for low sentiment and high sentiment periods, respectively. The results indicate that the significantly positive relation between ID and expected stock returns is robust to investor sentiment.

Another robustness check is to examine whether macroeconomic uncertainty affects the positive relation between ID and expected stock returns. I use four economic uncertainty measures (macro, real, financial, and policy-related economic uncertainty) in the literature to classify each month as either a high-uncertainty month or a low-sentiment month. A high-sentiment month is one in which the value of the economic uncertainty index is above the median value for the sample period, and the low-sentiment months are those with below-median values.

Jurado et al. (2015) and Ludvigson et al. (2015) introduce time series measures of macroeconomic, real, and financial uncertainty.[21] In the two papers, real activity shocks are originated from technology, monetary policy, preferences, or government expenditure innovations, financial uncertainty arises because of expected volatility in financial markets,

[19] https://www.nber.org/research/business-cycle-dating

[20] The latest investor sentiment data is available till year 2018 and can be obtained from Professor Jeffrey Wurgler's website.

[21] The data is obtained from Professor Ludvigson's website.

and macro uncertainty arises because of expected volatility in the macro economy, such as an expectation of greater difficulty in predicting future productivity, future monetary policy or future fiscal policy. Baker et al. (2016) constructs policy-related economic uncertainty index by combining newspaper coverage of policy-related economic uncertainty, the number of federal tax code provisions set to expire in future years, and disagreement among economic forecasters.[22]

[Insert Table 11 about here]

Table 11 reports the five-factor alphas and corresponding Newey and West (1987) t-statistics of each ID decile portfolio and the long-short ID strategy. In all columns, the five-factor alphas increase when moving from the lowest ID to the highest ID decile. In addition, the five-factor alphas of the long-short ID strategy are in the range of 0.72% to 0.92% per month, with t-statistics between 3.95 and 6.68. The results indicate that the long-short ID strategy prevails in either high- or low- macro, financial, real, and policy-related economic uncertainty periods.

3.7.2 Persistence of ID and the long-short ID strategy performance

First, I examine whether investor disagreement (ID) is persistent. To address this question, I examine the persistence of ID by running firm-level cross-sectional regressions of ID on lagged ID and 12 lagged cross-sectional predictors including firm size (SIZE), book-to-market (BM) ratio, momentum (MOM), short-term reversal (REV), turnover ratio (TURN), idiosyncratic volatility (IVOL), illiquidity (ILILQ), demand for lottery stocks with extreme positive returns (MAX), institutional ownership ratio (IOR), the stock beta (BETA), co-skewness (COSKEW), and analyst forecast dispersion (DISP).

Panel A in Table 11 reports the average cross-sectional coefficients on ID from the univariate and multivariate cross-sectional regressions. The coefficients on ID are 0.552 and 0.468 for univariate and multivariate cross-sectional regressions, respectively, and are both extremely significant. The adjusted R-squared in both regressions are above 30%, indicating substantial cross-sectional explanatory power. The regression results suggest that stocks with high ID in one month on average tend to be of high ID in the subsequent month.

[Insert Table 11 about here]

[22] The data is obtained from https://www.policyuncertainty.com/index.html.

Another way to examine the persistence of ID is to compute the average month-to-month decile portfolio transition matrix. Panel B in Table 11 reports the results, where column (i, j) is the average probability that a stock in ID decile i in month will be in ID decile j in the following month. If ID is completely random, then all the diagonal probabilities should be approximately 10%. First, all the diagonal elements of the transition matrix exceeds 10%, indicating that ID is indeed persistent. In particular, the persistence is especially strong within the extreme deciles. Stocks in decile 10 (high ID) have a 38.06% chance of remaining in the same decile in the subsequent month, and stocks in decile 1 (low ID) have a 42.87% chance of appearing in the same decile in the following month.

In addition, I vary the number of months in the formation of ID and examine the significance and magnitude and the corresponding long-short ID strategy. In particular, ID at the end of a given month is computed as the contemporaneous correlation coefficient of daily trading volume and absolute price change over the past T months, multiplied by -1. For different formation periods ranging from 3 to 12 months, Table 12 reports the next-month equal-weighted excess returns, CAPM alpha, three-factor alpha, four-factor alpha, and five-factor alpha between the highest and the lowest ID decile.

[Insert Table 12 about here]

The excess return of the long-ID strategy ranges from 0.46% to 0.58% per month, with Newey and West (1987) t-statistics between 2.20 and 3.44. The corresponding risk-adjusted returns are all positive and highly significant, indicating that the positive relation between ID and expected stock returns is robust to different formation months of ID.

Next, I examine the long-short ID strategy under different holding periods to ensure that the high returns generated by the long-short ID strategy are not caused by a statistical fluke. In particular, I vary the number of months one holds each ID portfolio after it has been formed following Jegadeesh and Titman (1993). For example, when the holding period equals to 3 months, the portfolio return in month t is the average return of the decile portfolios formed in $t - 1$, $t - 2$, and $t - 3$. Hence, each decile portfolio changes one-third of its composition each month.

[Insert Table 13 about here]

Table 13 reports the equal-weighted excess returns, CAPM alpha, three-factor alpha, four-factor alpha, and five-factor alpha between the highest and the lowest ID decile for different holding months. Both excess returns and risk-adjusted returns remain significantly positive under all holding periods up to 12 months. In addition, the five-factor alpha decreases from 0.68% (t-statistic=6.57) to 0.37% (t-statistic=3.68) per month as the number of holding month increases. The results suggest that the positive relation between ID and future stock returns is most significant for short to intermediate horizons.

3.8 ID and earnings announcements

So far I've only examined the significance of investor disagreement (ID) in the cross-sectional monthly pricing of stocks. Some investors, however, tend to trade stocks around earnings announcements (Kaniel et al. (2012) and Yang et al. (2020)). If there does exists a positive relation between ID and expected stocks returns, then it should also be the case that stocks with high ID prior to the earnings announcement significantly outperform those with low ID around the earnings announcement.

3.8.1 Data and variable definitions

To test this hypothesis, I first identify earnings announcement dates of firms with common stocks traded on NYSE, NASDAQ, and AMEX from Compustat, which according to WRDS are more reliable compared to announcement dates from IBES.[23] Next, I define reference and earnings announcement period as the 44-day window $[-45, -2]$ and 3-day window $[-1, 1]$, respectively, where $t = 0$ is the earnings announcement date. As a robustness check, I also use four other variations of reference and earnings announcement period in all my following tests.[24]

ID is defined as the contemporaneous correlation coefficient of daily trading volume and absolute price change in the reference period, multiplied by -1. Again, a stock trading day t is eligible if the price per share on $t - 31$ is at least 5 dollars and has non-missing return and volume. Stocks are required to have at least 30 eligible trading days in the reference period to compute ID. Figure 3 plots the number of eligible stocks issuing earnings announcement in each calendar quarter from the first quarter of 1983 to the fourth quarter of 2019.

[Insert Figure 3 about here]

Following most literature studying earnings announcements, stock performance around an earnings announcement is defined as the stock's cumulative abnormal return (CAR), which

[23] https://wrds-www.wharton.upenn.edu/pages/support/support-articles/ibes/

[24] REF period $[-48, -5]$ with EAR period $[-4, 4]$, REF period $[-47, -4]$ with EAR period $[-3, 3]$, REF period $[-46, -3]$ with EAR period $[-2, 2]$, and REF period $[-45, -2]$ with EAR period $[-1, 1]$.

is the difference between the compounded stock return and value-weighted market return (in percent) over the earnings announcement period.

Other control variables are defined similarly as in Appendix I. SIZE is the log of market capitalization in millions of dollars and BM is book-to-market ratio. RET is the return (in percent) compounded over the reference period. TURN and IVOL are the average turnover ratio and idiosyncratic volatility in the reference period. IOR is the ratio of shares owned by institutions as reported in 13F filings in the last quarter.[25] NUMEST is the number of unique analysts that have eligible fiscal year one earnings estimates on IBES in the reference period.[26]

3.8.2 Portfolio analysis

I start my analysis by examining the relation between investor disagreement (ID) in the reference period and cumulative abnormal returns (CAR) around earnings announcements. First, every calendar quarters are classified into deciles based on their ID in the reference period. Then, I compute the cross-sectional mean CAR around earnings announcements for each ID decile. Then, I compute the time-series (weighted) averages of these cross-sectional means across all quarters. The weights are based on the number of observations in each ID decile each quarter .

Table 14 presents time-series average of quarterly mean values of CAR around earnings announcement period within ID deciles. Looking at the second column, the average $CAR_{-1,1}$ increases from 0.12% to 0.76% when moving from the lowest to the highest ID decile. The difference in CARs is 0.65% with a significant Newey and West (1987) t-statistic of 4.77. As a robustness check, in the third to sixth column, I also examine other variations of reference and earnings announcement periods, and the results are similar. For example, in the last column the average $CAR_{-5,5}$ increases from -0.51% to 0.22% when moving from the lowest to the highest ID decile, and the return differential is 0.73% with a Newey and West (1987) t-statistic of 7.10.

[Insert Table 14 about here]

Overall, the results in Table 14 suggest that stocks with high ID prior to the earnings announcement experience significantly higher cumulative abnormal returns in the earnings announcement period.

[25] Nagel (2005) emphasizes the relation between short-sales constraints and divergence of opinion when examining stock returns.

[26] See Israelsen (2016), Lee and So (2017), and Ali and Hirshleifer (2020) for evidence of the relation between analyst coverage and stock returns.

3.8.3 Regression analysis

Next, I perform a cross-sectional regression analysis that controls for various stock characteristics that may potentially affect the relation between investor disagreement (ID) in the reference period and cumulative abnormal returns (CAR) in the earnings announcement period. I implement Fama and MacBeth (1973) regressions in which the dependent variable is CAR in the earnings announcement period. In particular, I run the following cross-sectional regression every quarter:

$$CAR_{i,q} = \alpha_q + \beta_q ID_{i,q} + \lambda_q X_{i,q} + \epsilon_{i,q}, \tag{8}$$

where i refers to the stock, q refers to the calendar quarter, $\mathrm{ID}_{i,q}$ is investor disagreement in the reference period with respect to quarter q for stock i, and $X_{i,q}$ is the set of stock-specific control variables for stock i in quarter q, and $\mathrm{CAR}_{i,q}$ is cumulative abnormal return in the earnings announcement period for firm i in quarter q. Then, I average (weighted) the cross-sectional coefficients across all quarters, where the weights correspond to the number of observations in each quarterly cross-sectional regression. The choice of quarterly frequency is consistent with other papers in the earnings announcement literature (e.g., Garfinkel and Sokobin (2006), Johnson and So (2012), and Akbas (2016)). The coefficient of interest is ID in the reference period. If there indeed exists a positive relation between ID in the reference period and earnings announcement premium, β_q should be significantly positive.

Table 15 presents the results. The coefficients on ID are positive and highly significant across five different reference and earnings announcement periods. For example, looking at the second column, when the reference period is $[-45, -2]$ and the earnings announcement period is $[-1, 1]$, the coefficient on ID is 0.337 with a Newey and West (1987) t-statistic of 3.96. In other words, stocks with high ID prior to earnings announcements on average experience significantly higher cumulative abnormal returns around earnings announcements.

[Insert Table 15 about here]

The coefficients on control variables are mostly consistent with the literature. The coefficients on RET are significantly negative, which is consistent with the well-known reversal effect. The coefficients on BM are significantly positive, which implies that value stocks in the reference periods tend to perform better around earnings announcements periods (Porta et al. (1997)). The coefficients on IVOL are significantly negative, which is consistent with Ang et al. (2006). The coefficients of SIZE, however, are positive, while Chari et al. (1988) and

Ball and Kothari (1991) suggest that earnings announcement returns are larger for smaller firms.

Overall, the results in this section provide evidence that the positive relation between ID and expected stock returns exists not only on a monthly basis, but also in the earnings announcement setting.

4 A disagreement model with ambiguity aversion

In this section I extend Kandel and Pearson (1995) to motivate the empirical results. In particular, I introduce one additional time period ($t = 1'$) and Figure 4 presents the model timeline. In particular, at $t = 1'$ traders observe the other type's interpretation and trade. The basic setup and other assumptions are the same as those in Section 2.

At $t = 1'$ (shortly after $t = 1$), type 1 traders observe μ_2, and type 2 traders observe μ_1. The fact that traders do not observe the other type's interpretation until $t = 1'$ aims to capture the existence of information acquisition costs and traders' limited attention, which delays the information transmission process.[27] To simplify notations, let μ_{-i} denote the other type's interpretation from type i traders' point of view. In other words, type i traders observe μ_{-i} at $t = 1'$.

First, as most traders in reality don't have enough information or skills to correctly evaluate the information quality of the other type's interpretation, I make the following assumption:

Assumption 1. *When traders observe the other type's interpretation at $t = 1'$, they believe that it can range from being less precise to being more precise than their own interpretation.*

Assumption 1 indicates that traders take into account the other type's interpretation in the loosest possible sense: they think that it can range from being less accurate to being more accurate than their own interpretation and update their beliefs accordingly. Formally, type i traders think that

$$\mu_{-i} = \mu_\eta + \epsilon, \quad \mu_\eta \perp \epsilon, \quad \epsilon \sim N(0, \sigma_\epsilon^2), \quad \sigma_\epsilon^2 \in [\underline{\sigma_\epsilon^2}, \overline{\sigma_\epsilon^2}], \tag{9}$$

where $0 < \underline{\sigma_\epsilon^2} < \sigma^2 < \overline{\sigma_\epsilon^2} < \infty$ and $i \in \{1, 2\}$. When σ_ϵ^2 is higher (lower) than σ^2, type i traders believe that compared to their own interpretation μ_i, the other type's interpretation

[27] See, for example, Hirshleifer and Teoh (2003), Peng and Xiong (2006), and Corwin and Coughenour (2008).

μ_{-i} is more imprecise (precise).[28] The information quality of μ_{-i} is thus captured by the range of precisions $[1/\overline{\sigma_\epsilon^2}, 1/\underline{\sigma_\epsilon^2}]$.

In order to update their priors on μ_η, type i traders apply Bayes's rule to obtain a family of posteriors:

$$\mu_\eta \sim N(\mu_i + \frac{\sigma^2}{\sigma^2+\sigma_\epsilon^2}(\mu_{-i}-\mu_i), \frac{\sigma^2\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2}), \quad \sigma_\epsilon^2 \in [\underline{\sigma_\epsilon^2}, \overline{\sigma_\epsilon^2}]. \tag{10}$$

For tractability, let $\underline{\sigma_\epsilon^2} = \sigma^2(1-\alpha)$, $\overline{\sigma_\epsilon^2} = \sigma^2(1+\beta)$, $0 < \alpha < 1$, and $0 < \beta < \infty$. Next, motivated by many experimental studies in the literature, I assume that traders exhibit ambiguity aversion.[29]

Assumption 2. *Traders are ambiguity-averse.*

Traders' preferences of ambiguity aversion will be represented by the maxmin expected utility of Gilboa and Schmeidler (1989). Under ambiguity aversion, traders have a set of probability measures and evaluate any action using the probability that maximizes the expected utility of that action. Next, I introduce traders' problem at $t = 1'$.

4.1 $t = 1'$

At $t = 1'$, each type i trader solves the following problem:

$$\max_{m_{i,1'}} \min_{\sigma_\epsilon^2 \in [\underline{\sigma_\epsilon^2}, \overline{\sigma_\epsilon^2}]} E_{i,1'} \ -e^{-\lambda m_{i,1'}(X-P_{1'})}, \tag{11}$$

where $E_{i,1'}$ denotes expectation with respect to X of type i traders at $t = 1'$ and $m_{i,1'}$ denotes the position held by each of them at $t = 1'$. Note that traders' posterior mean on μ_η is negatively related to the expected utility before maximization. Hence, ambiguity-averse traders select an information quality $\sigma_\epsilon^2 \in [\underline{\sigma_\epsilon^2}, \overline{\sigma_\epsilon^2}]$ that generates the highest posterior mean on μ_η. That is, if the other type's interpretation is higher ($\mu_{-i} - \mu_i > 0$), type i traders act as if μ_{-i} is precise ($\sigma_\epsilon^2 = \underline{\sigma_\epsilon^2}$). In contrast, if the other type's interpretation is lower ($\mu_{-i} - \mu_i < 0$), type i traders act as if μ_{-i} is imprecise ($\sigma_\epsilon^2 = \overline{\sigma_\epsilon^2}$). Formally, at $t = 1'$, type i traders' posterior belief on μ_η is given by

$$\begin{cases} \mu_\eta \sim N(\mu_i + \frac{\sigma^2}{\sigma^2+\underline{\sigma_\epsilon^2}}(\mu_{-i}-\mu_i), \frac{\sigma^2\underline{\sigma_\epsilon^2}}{\sigma^2+\underline{\sigma_\epsilon^2}}) & \text{, if } \mu_{-i}-\mu_i > 0 \\ \\ \mu_\eta \sim N(\mu_i + \frac{\sigma^2}{\sigma^2+\overline{\sigma_\epsilon^2}}(\mu_{-i}-\mu_i), \frac{\sigma^2\overline{\sigma_\epsilon^2}}{\sigma^2+\overline{\sigma_\epsilon^2}}) & \text{, if } \mu_{-i}-\mu_i < 0. \end{cases} \tag{12}$$

[28]Since $\mu_{-i} = \mu_\eta + \epsilon$, $\mu_\eta = \mu_{-i} - \epsilon$. Hence, $\mu_\eta \sim N(\mu_{-i}, \sigma_\epsilon^2)$. From (1) we know that $\mu_\eta \sim N(\mu_i, \sigma^2)$.
[29]See, for example, Halevy (2007), Butler et al. (2014), and Dimmock et al. (2016).

Next, plug in $\underline{\sigma_\epsilon^2} = \sigma^2(1-\alpha)$ and $\overline{\sigma_\epsilon^2} = \sigma^2(1+\beta)$. Hence, type 1 traders' posterior belief on μ_η is given by

$$\begin{cases} \mu_\eta \sim N(\frac{(1-\alpha)\mu_1+\mu_2}{2-\alpha}, \frac{1-\alpha}{2-\alpha}\sigma^2) & \text{, if } \mu_2 - \mu_1 > 0 \\ \\ \mu_\eta \sim N(\frac{(1+\beta)\mu_1+\mu_2}{2+\beta}, \frac{1+\beta}{2+\beta}\sigma^2) & \text{, if } \mu_2 - \mu_1 < 0. \end{cases} \tag{13}$$

Similarly, type 2 traders' posterior belief on μ_η is given by

$$\begin{cases} \mu_\eta \sim N(\frac{(1-\alpha)\mu_2+\mu_1}{2-\alpha}, \frac{1-\alpha}{2-\alpha}\sigma^2) & \text{, if } \mu_1 - \mu_2 > 0 \\ \\ \mu_\eta \sim N(\frac{(1+\beta)\mu_2+\mu_1}{2+\beta}, \frac{1+\beta}{2+\beta}\sigma^2) & \text{, if } \mu_1 - \mu_2 < 0. \end{cases} \tag{14}$$

The market-clearing condition is $\frac{1}{2}m_{1,1'} + \frac{1}{2}m_{2,1'} = 0$.

4.2 Model prediction

The following proposition presents the main results of the model.

Proposition 2. *Suppose $\sigma^2 << \sigma_\eta^2$. Let $R = P_2^* - P_{1'}^*$. Then, the expected stock return*

$$E[R] = \frac{\sigma_\eta^{-2}[-2\mu_\eta + (\mu_1+\mu_2) + \frac{(\alpha+\beta)}{(2-\alpha)(2+\beta)}|\mu_1-\mu_2|]}{Z_1 + Z_2 + 2\sigma_\eta^{-2}} \tag{15}$$

and is increasing in investor disagreement, $|\mu_1 - \mu_2|$, holding the average interpretation in the market, $(\mu_1 + \mu_2)$, fixed.

Proof. See Appendix II. □

The intuition is as follows. At $t = 1'$, traders are uncertain about the other type's interpretation and since they are ambiguity-averse, they choose an information quality $\sigma_\epsilon^2 \in [\underline{\sigma_\epsilon^2}, \overline{\sigma_\epsilon^2}]$ that generates the highest posterior mean on μ_η before maximization. That is, traders place more emphasis on the larger of the two interpretations. In other words, ambiguity aversion motivates traders to form a more negative view on S after observing the other type's interpretation.

Hence, when the average interpretation of S is fixed in the market, the stock price reflects the pessimistic view more as the two types of interpretations are more polarized (high investor disagreement). Furthermore, since X is realized at $t = 2$ and is not affected by traders' interpretations, there is no speculation risk involved.

The assumption of $\sigma^2 << \sigma_\eta^2$ requires that the public signal S is extremely imprecise. This is consistent with that fact that in reality, firms do not communicate with investors on a daily basis, so publicly available information reveals very less about the true value of the stock.

One thing worth mentioning is that I use the maxmin expected utility to model traders' ambiguity aversion. However, the only use of ambiguity aversion is to induce traders to give more (less) weight to the other type's interpretation when it is higher (low) than their own. Other ambiguity aversion preferences would also generate traders' asymmetric response and thus lead to the same conclusion. The only difference is the extent of this asymmetry. As the maxmin expected utility is the most strict version of ambiguity aversion, it leads to the most asymmetric traders' response.

5 Investor disagreement and news

In this section, I examine investor disagreement (ID) before and after earnings announcements, firm-specific news stories, and FOMC meetings, respectively.

5.1 Evolution of ID: earnings announcements

I first compute ID before and after the earnings announcement. ID before the earnings announcement date (day 0) is defined as the correlation coefficient between daily trading volume and absolute price change over the 44-day window $[-45, -2]$, multiplied by -1. ID after the earnings announcement is defined similarly over the 44-day window $[2, 45]$. Then, ΔID is defined as $\Delta\text{ID} = \text{ID}_{\text{after}} - \text{ID}_{\text{before}}$.

Figure 5 plots the time-series average of mean values of ID before and after the earnings announcement. First, ID after earnings announcements seems to be higher than ID before earnings announcements, although the difference in magnitude is small. This is consistent with findings in Table 11 that ID is highly persistent.

[Insert Figure 5 about here]

In particular, the mean ΔID is 0.0084 with a significant Newey and West (1987) t-statistic of 4.39. In addition, I compute the mean ΔID for the sample period before and after the implementation of Regulation Fair Disclosure (Reg FD), which prevents firms from doing selective disclosure. Specifically, the pre-Reg-FD and post-Reg-FD mean ΔID are 0.0089 (t-statistic=5.67) and 0.0062 (t-statistic=2.16), respectively. The slight decrease in ID following the implementation of Reg FD could be a result of more transparent and valid firm disclosures. Figure 5 suggests that on average, ID increases after earnings announcements.

5.2 Evolution of ID: firm-specific public news stories

I proceed to examine whether ID also increases following firm-specific public news stories. I first obtain public news data from RavenPack News Analytics on WRDS. I select news events with a relevance score equal to 100, which according to RavenPack, means that the entity plays a key role in the news story and is considered highly relevant.

I further classify news events into six categories; Financial, Legal, M&A, Operational, Ratings, and Others. The news date is defined as the date when the first news story reporting an event about one or more entities is announced. To avoid double counting issue, subsequent news stories reporting the the same news events are not included. ID before and after news events are computed in the same fashion treating the news date as day 0.

[Insert Figure 6 about here]

Figure 6 plots the time-series average of mean values of ID before and after six types of news stories. Again, ID before and after news stories behave very similarly, reassuring the persistence of ID. In particular, the mean ΔID is 0.0177 (t-statistic=5.66) for financial news, 0.0058 (t-statistic=1.78) for legal news, 0.0089 (t-statistic=2.66) for M&A news, 0.0066 (t-statistic=2.09) for operational news, 0.0074 (t-statistic=2.21) for ratings news, and 0.0156 (t-statistic=2.66) for other news. The results indicate that other than earnings announcements, ID also increases after different types of firm-specific public news events.

5.3 Evolution of ID: FOMC meetings

Next, since investor disagreement proxies for security-level behavioral bias (Harvey et al. (2016)), it should not respond to macroeconomic events. If ID significantly increases after macroeconomic events, then it is possible that ID measure in this paper simply captures economic uncertainty instead of investor disagreement.

To mitigate this concern, I examine ID before and after Federal Open Market Committee (FOMC) meetings. There are eight regularly scheduled FOMC meetings each year and meeting minutes are made public following the meetings. Prior work (See, for example, Cieslak et al. (2019), Lucca and Moench (2015), and Bernanke and Kuttner (2005), etc.) study stock market's reaction in the form of realized stock returns to FOMC announcements.

In our context, however, the hypothesis is that mean ΔID should be insignificantly different from 0.

I obtain FOMC scheduled meetings from 1994 to 2019, as in the first meeting (February 3-4, 1994) a reasonable portion of the discussion centered on the need to make the committee's intentions clear to the public. I then examine ID before and after FOMC meetings following the same approach.

[Insert Figure 7 about here]

Figure 7 plots the time-series average of mean values of ID before and after FOMC meetings. ID before and after behave virtually the same. In particular, the mean ΔID is 0.0007 with a significant Newey and West (1987) insignificant t-statistic of 0.33, which is consistent with the conjecture that ID proxies for firm-specific investor disagreement.

Together, Figure 5, Figure 6, and Figure 7 show that ID is sensitive to firm-specific information disclosure events but indifferent to macroeconomic news. When firm-specific news bring in a sudden influx of information, investors on average tend to slightly disagree more.

5.4 Evolution of ID: good and bad earnings news

Next, I examine whether the "sentiment" of earnings announcement affects the evolution of investor disagreement. Rogers et al. (2009), for example, find that short-term increase in uncertainty is attributable to forecasts that convey bad news. In a similar fashion, I test if the evolution of investor disagreement (ID) from before to after earnings announcements is asymmetric to good and bad earnings news.

I first use three-day cumulative abnormal returns centered at the earnings announcement date, $CAR_{-1,1}$, to determine whether an earnings announcement convey good news or bad news. In particular, $CAR_{-1,1}$ reveals investors' expectations regarding the firm's future cash flow prospects. $CAR_{-1,1}$ is defined as the compounded return over the $[-1, 1]$ window around the earnings announcement date ($t = 0$) in excess of the compounded value-weighted market return.

Table 16 presents the results. The mean ΔID for good earnings news ($CAR_{-1,1} > 0$) is 0.0032 with a Newey and West (1987) t-statistic of 2.08, while the mean ΔID for bad earnings news ($CAR_{-1,1} \leq 0$) is 0.0107 with a Newey and West (1987) t-statistic of 6.59. In particular, the difference in mean ΔID between bad news and good news is 0.0075 (t-statistic=7.42).

The results suggest that ID increases more following bad earnings news than good earnings news.

[Insert Table 16 about here]

In addition, I examine the difference in mean ΔID between bad news and good news when controlling for firm size (SIZE). In each calendar quarter, I first sort stocks with earnings announcements into quintiles based on SIZE. Next, within each SIZE quintile, I examine the difference in mean ΔID between bad earnings news ($CAR_{-1,1} \leq 0$) and good earnings news ($CAR_{-1,1} > 0$). In particular, the mean ΔID differences are significantly positive in three out of the five SIZE quintiles, providing further support that the increase in ID is larger following bad earnings news than good earnings news. The SUE and SUEAF sample are smaller and largely comprise bigger firms.

As a robustness check, I use two other measures, standardized unexpected earnings (SUE) and standardized unexpected earnings using analysts' forecasts (SUEAF), to capture earnings surprises at earnings announcements. Following Livnat and Mendenhall (2006), Garfinkel and Sokobin (2006), Johnson and So (2012), and Akbas (2016), SUE is defined as the difference in EPS before extraordinary items in quarters q and $q-4$ divided by quarter $q-4$ end price. SUEAF is defined as the difference between the median analyst forecast over the 90-day period before the announcement and actual earnings divided by quarter $q-4$ end price.

Panel A of Table 17 presents ID before and after earnings announcements for good earnings news (SUE $>$ 0) and bad earnings news (SUE $\leq$ 0) in each SIZE quintile. The differences in mean ΔID between bad news and good news are in the range of 0.0101 and 0.0171, with t-statistics between 4.47 and 8.83. Similarly, Panel B of Table 17 presents ID before and after earnings announcements for good earnings news for good earnings news (SUEAF $>$ 0) and bad earnings news (SUEAF $\leq$ 0) across SIZE quintiles. The differences in mean ΔID are positively significant in all but the largest SIZE quintile. Together, Table 16 and Table 17 provide strong evidence that investor disagreement increases more following bad earnings news than good earnings news.

In addition, I run stock-level cross-sectional regression of ΔID on bad news indicator variables ($1_{CAR_{-1,1}\leq 0}$, $1_{SUE\leq 0}$, and $1_{SUEAF\leq 0}$) in Table 18 to control for multiple variables simultaneously. The coefficients on the bad news indicator variables are all positive and highly significant, which indicates that compared to good earnings news, bad earnings news triggers a larger increase in investor disagreement.

6 Conclusion

This paper studies the role of investor disagreement (ID) in the cross-section of expected stock returns. Motivated by an implication of Kandel and Pearson (1995), I compute a stock's ID at the end of a given month as the contemporaneous correlation coefficient of daily trading volume and absolute price change over the past two months, multiplied by -1.

I find that stocks in the highest ID decile significantly outperform stocks in the lowest ID decile by an annualized risk-adjusted return of 9.24%. Bivariate portfolio-level analyses and stock-level cross-sectional regressions that control for firm size, book-to-market ratio, momentum, short-term reversal, turnover ratio, illiquidity, market beta, co-skewness, demand for lottery stocks with extreme positive returns, idiosyncratic volatility, and analyst forecast dispersion provide further support for the positive relation. In addition, the positive relation between ID and future stock returns persists in high and low sentiment periods, recessions and expansions, and high and low macro, financial, real, and policy-related economic uncertainty periods.

Next, I use AR(1) regression and ID transition matrix to show that ID is persistent. In addition, the positive relation between ID and expected stock returns is also robust to different numbers of months in the formation of ID and portfolio holding periods. Besides using monthly returns to examine the asset pricing implications of ID, I also examine the relation between ID and expected stock returns in the earnings announcement setting. Using portfolio analyses and stock-level cross-sectional regressions, I find that compared to stocks with low ID, stocks with high ID prior to earnings announcements earn significantly higher cumulative abnormal returns around earnings announcements.

To explain these novel empirical findings, I develop a theory by extending Kandel and Pearson (1995) to incorporate traders' ambiguity aversion and their ambiguity about other traders' interpretations. In particular, traders believe that others' interpretations can range from being less accurate to being more accurate than their own interpretations. Since traders are ambiguity-averse, the optimistic trader places relatively more weight on the pessimistic view while the pessimistic trader places relatively less weight on the optimistic view. As a result, the stock price reflects pessimism (higher expected return) more as investor disagreement becomes higher.

In addition, I find that ID increases following earnings announcements, with the size of the increase being bigger following bad news. Moreover, ID increases after firm-specific news stories but remains virtually the same following FOMC scheduled announcements. A potentially fruitful area of inquiry would be whether investor disagreement among economically related stocks cross-predicts each other. A preliminary glance provides a positive answer, and this is the next step on my research agenda.

References

Akbas, F. (2016). The calm before the storm. *The Journal of Finance 71*(1), 225–266.

Ali, U. and D. Hirshleifer (2020). Shared analyst coverage: Unifying momentum spillover effects. *Journal of Financial Economics 136*(3), 649–675.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of financial markets 5*(1), 31–56.

Andersen, T. G. (1996). Return volatility and trading volume: An information flow interpretation of stochastic volatility. *The Journal of Finance 51*(1), 169–204.

Anderson, E. W., E. Ghysels, and J. L. Juergens (2005). Do heterogeneous beliefs matter for asset pricing? *The Review of Financial Studies 18*(3), 875–924.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang (2006). The cross-section of volatility and expected returns. *The Journal of Finance 61*(1), 259–299.

Asquith, P., P. A. Pathak, and J. R. Ritter (2005). Short interest, institutional ownership, and stock returns. *Journal of Financial Economics 78*(2), 243–276.

Back, K., K. Crotty, and T. Li (2018). Identifying information asymmetry in securities markets. *The Review of Financial Studies 31*(6), 2277–2325.

Baker, M. and J. Wurgler (2006). Investor sentiment and the cross-section of stock returns. *The journal of Finance 61*(4), 1645–1680.

Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring economic policy uncertainty. *The quarterly journal of economics 131*(4), 1593–1636.

Bali, T. G., S. J. Brown, and Y. Tang (2017). Is economic uncertainty priced in the cross-section of stock returns? *Journal of Financial Economics 126*(3), 471–489.

Bali, T. G., N. Cakici, and R. F. Whitelaw (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics 99*(2), 427–446.

Ball, R. and P. Brown (1968). An empirical evaluation of accounting income numbers. *Journal of accounting research*, 159–178.

Ball, R. and S. P. Kothari (1991). Security returns around earnings announcements. *Accounting Review*, 718–738.

Banerjee, S. and I. Kremer (2010). Disagreement and learning: Dynamic patterns of trade. *The Journal of Finance 65*(4), 1269–1302.

Barinov, A. (2013). Analyst disagreement and aggregate volatility risk. *Journal of Financial and Quantitative Analysis*, 1877–1900.

Berkman, H., V. Dimitrov, P. C. Jain, P. D. Koch, and S. Tice (2009). Sell on the news: Differences of opinion, short-sales constraints, and returns around earnings announcements. *Journal of Financial Economics 92*(3), 376–399.

Bernanke, B. S. and K. N. Kuttner (2005). What explains the stock market's reaction to federal reserve policy? *The Journal of finance 60*(3), 1221–1257.

Billings, M. B., R. Jennings, and B. Lev (2015). On guidance and volatility. *Journal of Accounting and Economics 60*(2-3), 161–180.

Boehme, R. D., B. R. Danielsen, P. Kumar, and S. M. Sorescu (2009). Idiosyncratic risk and the cross-section of stock returns: Merton (1987) meets miller (1977). *Journal of Financial Markets 12*(3), 438–468.

Boehme, R. D., B. R. Danielsen, and S. M. Sorescu (2006). Short-sale constraints, differences of opinion, and overvaluation. *Journal of Financial and Quantitative Analysis 41*(2), 455–487.

Brown, L. D. (2001). A temporal analysis of earnings surprises: Profits versus losses. *Journal of accounting research 39*(2), 221–241.

Butler, J. V., L. Guiso, and T. Jappelli (2014). The role of intuition and reasoning in driving aversion to risk and ambiguity. *Theory and decision 77*(4), 455–484.

Cao, H. H. and H. Ou-Yang (2008). Differences of opinion of public information and speculative trading in stocks and options. *The Review of Financial Studies 22*(1), 299–335.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of finance 52*(1), 57–82.

Chan, L. K., N. Jegadeesh, and J. Lakonishok (1996). Momentum strategies. *The Journal of Finance 51*(5), 1681–1713.

Chari, V. V., R. Jagannathan, and A. R. Ofer (1988). Seasonalities in security returns: The case of earnings announcements. *Journal of Financial Economics 21*(1), 101–121.

Chatterjee, S., K. John, and A. Yan (2012). Takeovers and divergence of investor opinion. *The Review of Financial Studies 25*(1), 227–277.

Chen, J., H. Hong, and J. C. Stein (2002). Breadth of ownership and stock returns. *Journal of financial Economics 66*(2-3), 171–205.

Cieslak, A., A. Morse, and A. Vissing-Jorgensen (2019). Stock returns over the fomc cycle. *The Journal of Finance 74*(5), 2201–2248.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica: journal of the Econometric Society*, 135–155.

Conrad, J., B. Cornell, W. R. Landsman, and B. R. Rountree (2006). How do analyst recommendations respond to major news? *Journal of Financial and Quantitative Analysis 41*(1), 25–49.

Corwin, S. A. and J. F. Coughenour (2008). Limited attention and the allocation of effort in securities trading. *The Journal of Finance 63*(6), 3031–3067.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers (1997). Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of finance 52*(3), 1035–1058.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam (2005). Investor psychology and security market under-and overreaction. *Advances in Behavioral Finance, Volume II*, 460–501.

Daniel, K., D. Hirshleifer, and S. H. Teoh (2002). Investor psychology in capital markets: Evidence and policy implications. *Journal of monetary economics 49*(1), 139–209.

Daniel, K. and S. Titman (2006). Market reactions to tangible and intangible information. *The Journal of Finance 61*(4), 1605–1643.

David, A. (2008). Heterogeneous beliefs, speculation, and the equity premium. *The Journal of Finance 63*(1), 41–83.

Davis, J. L., E. F. Fama, and K. R. French (2000). Characteristics, covariances, and average returns: 1929 to 1997. *The Journal of Finance 55*(1), 389–406.

De Bondt, W. F. and R. Thaler (1985). Does the stock market overreact? *The Journal of finance 40*(3), 793–805.

Dechow, P. M. and R. G. Sloan (1997). Returns to contrarian investment strategies: Tests of naive expectations hypotheses. *Journal of financial economics 43*(1), 3–27.

Diamond, D. W. and R. E. Verrecchia (1991). Disclosure, liquidity, and the cost of capital. *The journal of Finance 46*(4), 1325–1359.

Diether, K. B., C. J. Malloy, and A. Scherbina (2002). Differences of opinion and the cross section of stock returns. *The Journal of Finance 57*(5), 2113–2141.

Dimmock, S. G., R. Kouwenberg, O. S. Mitchell, and K. Peijnenburg (2016). Ambiguity aversion and household portfolio choice puzzles: Empirical evidence. *Journal of Financial Economics 119*(3), 559–577.

Doukas, J. A., C. F. Kim, and C. Pantzalis (2006). Divergence of opinion and equity returns. *Journal of Financial and Quantitative Analysis 41*(3), 573–606.

Easley, D. and M. O'Hara (2010). Microstructure and ambiguity. *The Journal of Finance 65*(5), 1817–1846.

Easterwood, J. C. and S. R. Nutt (1999). Inefficiency in analysts' earnings forecasts: Systematic misreaction or systematic optimism? *The Journal of Finance 54*(5), 1777–1797.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The quarterly journal of economics*, 643–669.

Epstein, L. G. and M. Schneider (2007). Learning under ambiguity. *The Review of Economic Studies 74*(4), 1275–1303.

Epstein, L. G. and M. Schneider (2008). Ambiguity, information quality, and asset pricing. *The Journal of Finance 63*(1), 197–228.

Erturk, B. (2006). Dispersion of forecasts and stock returns. In *AFA 2007 Chicago Meetings Paper*.

Fama, E. F. and K. R. French (1992). The cross-section of expected stock returns. *the Journal of Finance 47*(2), 427–465.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of*.

Fama, E. F. and K. R. French (1996). Multifactor explanations of asset pricing anomalies. *The journal of finance 51*(1), 55–84.

Fama, E. F. and J. D. MacBeth (1973). Risk, return, and equilibrium: Empirical tests. *Journal of political economy 81*(3), 607–636.

Gallant, A. R., P. E. Rossi, and G. Tauchen (1992). Stock prices and volume. *The Review of Financial Studies 5*(2), 199–242.

Gao, G. P., X. Lu, Z. Song, and H. Yan (2019). Disagreement beta. *Journal of Monetary Economics 107*, 96–113.

Garfinkel, J. A. (2009). Measuring investors' opinion divergence. *Journal of Accounting Research 47*(5), 1317–1348.

Garfinkel, J. A. and J. Sokobin (2006). Volume, opinion divergence, and returns: A study of post–earnings announcement drift. *Journal of Accounting Research 44*(1), 85–112.

Gervais, S., R. Kaniel, and D. H. Mingelgrin (2001). The high-volume return premium. *The Journal of Finance 56*(3), 877–919.

Ghirardato, P., F. Maccheroni, and M. Marinacci (2004). Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory 118*(2), 133–173.

Gibbons, M. R., S. A. Ross, and J. Shanken (1989). A test of the efficiency of a given portfolio. *Econometrica: Journal of the Econometric Society*, 1121–1152.

Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of mathematical economics 18*(2), 141–153.

Graham, J. R. (1999). Herding among investment newsletters: Theory and evidence. *The Journal of Finance 54*(1), 237–268.

Halevy, Y. (2007). Ellsberg revisited: An experimental study. *Econometrica 75*(2), 503–536.

Hansen, L. P. and T. J. Sargent (2007). Recursive robust estimation and control without commitment. *Journal of Economic Theory 136*(1), 1–27.

Hansen, L. P. and T. J. Sargent (2011). Robustness and ambiguity in continuous time. *Journal of Economic Theory 146*(3), 1195–1223.

Harris, M. and A. Raviv (1993). Differences of opinion make a horse race. *The Review of Financial Studies 6*(3), 473–506.

Harrison, J. M. and D. M. Kreps (1978). Speculative investor behavior in a stock market with heterogeneous expectations. *The Quarterly Journal of Economics 92*(2), 323–336.

Harvey, C. R., Y. Liu, and H. Zhu (2016). and the cross-section of expected returns. *The Review of Financial Studies 29*(1), 5–68.

Harvey, C. R. and A. Siddique (2000). Conditional skewness in asset pricing tests. *The Journal of Finance 55*(3), 1263–1295.

Hirshleifer, D. and S. H. Teoh (2003). Limited attention, information disclosure, and financial reporting. *Journal of accounting and economics 36*(1-3), 337–386.

Hong, H. and J. D. Kubik (2003). Analyzing the analysts: Career concerns and biased earnings forecasts. *The Journal of Finance 58*(1), 313–351.

Illeditsch, P. K. (2011). Ambiguous information, portfolio inertia, and excess volatility. *The Journal of Finance 66*(6), 2213–2247.

Israelsen, R. D. (2016). Does common analyst coverage explain excess comovement? *Journal of Financial and Quantitative Analysis 51*(4), 1193–1229.

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *The Journal of finance 45*(3), 881–898.

Jegadeesh, N. and S. Titman (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance 48*(1), 65–91.

Johnson, T. L. and E. C. So (2012). The option to stock volume ratio and future returns. *Journal of Financial Economics 106*(2), 262–286.

Jurado, K., S. C. Ludvigson, and S. Ng (2015). Measuring uncertainty. *American Economic Review 105*(3), 1177–1216.

Kandel, E. and N. D. Pearson (1995). Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy 103*(4), 831–872.

Kaniel, R., S. Liu, G. Saar, and S. Titman (2012). Individual investor trading and return patterns around earnings announcements. *The Journal of Finance 67*(2), 639–680.

Karpoff, J. M. (1987). The relation between price changes and trading volume: A survey. *Journal of Financial and quantitative Analysis 22*(1), 109–126.

Kim, O. and R. E. Verrecchia (1991). Trading volume and price reactions to public announcements. *Journal of accounting research 29*(2), 302–321.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005). A smooth model of decision making under ambiguity. *Econometrica 73*(6), 1849–1892.

Klibanoff, P., M. Marinacci, and S. Mukerji (2009). Recursive smooth ambiguity preferences. *Journal of Economic Theory 144*(3), 930–976.

Krinsky, I. and J. Lee (1996). Earnings announcements and the components of the bid-ask spread. *The Journal of Finance 51*(4), 1523–1535.

La Porta, R. (1996). Expectations and the cross-section of stock returns. *The Journal of Finance 51*(5), 1715–1742.

Lamont, O. A. (2002). Macroeconomic forecasts and microeconomic forecasters. *Journal of economic behavior & organization 48*(3), 265–280.

Lee, C. M. and E. C. So (2017). Uncovering expected returns: Information in analyst coverage proxies. *Journal of Financial Economics 124*(2), 331–348.

Lehmann, B. N. (1990). Fads, martingales, and market efficiency. *The Quarterly Journal of Economics 105*(1), 1–28.

Lim, T. (2001). Rationality and analysts' forecast bias. *The journal of Finance 56*(1), 369–385.

Livnat, J. and R. R. Mendenhall (2006). Comparing the post–earnings announcement drift for surprises calculated from analyst and time series forecasts. *Journal of accounting research 44*(1), 177–205.

Lucca, D. O. and E. Moench (2015). The pre-fomc announcement drift. *The Journal of finance 70*(1), 329–371.

Ludvigson, S. C., S. Ma, and S. Ng (2015). Uncertainty and business cycles: exogenous impulse or endogenous response? Technical report, National Bureau of Economic Research.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The Journal of Finance 42*(3), 483–510.

Miller, E. M. (1977). Risk, uncertainty, and divergence of opinion. *The Journal of finance 32*(4), 1151–1168.

Morris, S. (1996). Speculative investor behavior and learning. *The Quarterly Journal of Economics 111*(4), 1111–1133.

Nagel, S. (2005). Short sales, institutional investors and the cross-section of stock returns. *Journal of financial economics 78*(2), 277–309.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation. *Econometrica 55*(3), 703–708.

Pástor, L. and R. F. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political economy 111*(3), 642–685.

Patell, J. M. and M. A. Wolfson (1979). Anticipated information releases reflected in call option prices. *Journal of accounting and economics 1*(2), 117–140.

Peng, L. and W. Xiong (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics 80*(3), 563–602.

Porta, R. L., J. Lakonishok, A. Shleifer, and R. Vishny (1997). Good news for value stocks: Further evidence on market efficiency. *The Journal of Finance 52*(2), 859–874.

Rogers, J. L., D. J. Skinner, and A. Van Buskirk (2009). Earnings guidance and market uncertainty. *Journal of Accounting and Economics 48*(1), 90–109.

Sadka, R. and A. Scherbina (2007). Analyst disagreement, mispricing, and liquidity. *The Journal of Finance 62*(5), 2367–2403.

Scheinkman, J. A. and W. Xiong (2003). Overconfidence and speculative bubbles. *Journal of political Economy 111*(6), 1183–1220.

Tauchen, G. E. and M. Pitts (1983). The price variability-volume relationship on speculative markets. *Econometrica: Journal of the Econometric Society*, 485–505.

Viswanathan, S. (2001). Strategic trading, heterogeneous beliefs/information and short constraints. *Unpublished working paper, Duke University*.

Welch, I. (2000). Herding among security analysts. *Journal of Financial economics 58*(3), 369–396.

Wermers, R. (2003). Is money really'smart'? new evidence on the relation between mutual fund flows, manager behavior, and performance persistence. *New Evidence on the Relation Between Mutual Fund Flows, Manager Behavior, and Performance Persistence (May 2003)*.

Yang, Y. C., B. Zhang, and C. Zhang (2020). Is information risk priced? evidence from abnormal idiosyncratic volatility. *Journal of Financial Economics 135*(2), 528–554.

Appendix I: Variable definitions

In this section, I define various variables used in the paper. Following Fama and French (1992), Fama and French (1993), and Davis et al. (2000), firm size (SIZE) for July of year t to June of t+1 is defined as the natural logarithm of market value of equity at the end of December of year $t-1$, and the book-to-market (BM) ratio from July of year t through June of year $t+1$, is computed as the shareholders' book value of equity plus deferred taxes and investment tax credit (if available) minus the book value of preferred stock at the end of the last fiscal year, $t-1$, divided by the market value of equity at the end of December of year $t-1$. Depending on availability, the redemption, liquidation, or par value is used to estimate the book value of preferred stock. Following Daniel and Titman (2006), the minimum 6-month lag is to ensure the firm's annual report is publicly available information.

Following Jegadeesh and Titman (1993), momentum (MOM) is computed as the cumulative return of a stock of 11 months ending one month prior to the given month. Following Jegadeesh (1990), short-term reversal (REV) is defined as the stock return over the portfolio formation month. Turnover ratio (TURN) is computed as the percentage of trading volume divided by the total number of shares outstanding shares over the portfolio formation month. A minimum of 15 daily observations in the given month is required to calculate TURN.

Following Amihud (2002), stock illiquidity for each stock in month m as the ratio of the absolute monthly stock return to its dollar trading volume, multiplied by 10^6:

$$ILLIQ_{i,m} = 10^6 \times Avg\left[\frac{|R_{i,d}|}{DTV_{i,d}}\right], \tag{16}$$

where $R_{i,d}$ and $DTV_{i,d}$ are the daily return and dollar trading volume for stock i on day d, respectively. A minimum of 15 daily observations in the given month is required to calculate ILLIQ.

Stock beta (BETA), is computed by regressing the stock's monthly excess return on monthly market excess return and lagged market excess return to accommodate non-synchronous trading effects:

$$R_{i,m} = \alpha_i + \beta_{i,1}R_{M,m} + \beta_{i,2}R_{M,m-1} + \epsilon_{i,m}, \tag{17}$$

where $R_{i,m}$ and $R_{M,m}$ are the monthly excess returns on stock i and the CRSP value-weighted market index, respectively. Following Fama and French (1992), I run the regression each month over a moving window covering the most recent 60 months, requiring at least 36 months of non-missing data. The stock's monthly beta is defined as $\widehat{\beta_{i,1}} + \widehat{\beta_{i,2}}$.

Following Bali et al. (2011) and Bali et al. (2017), demand for lottery-like stocks (MAX) is defined as the average of the five highest daily daily returns of the stock during the portfolio formation month. A minimum of 15 daily observations in the given month is required to calculate MAX.

Following Harvey and Siddique (2000), the co-skewness (COSKEW) of stock i in month m is defined as the estimated slope $\widehat{\gamma_{i,m}}$ in the following regression:

$$R_{i,m} = \alpha_i + \beta_i R_{M,m} + \gamma_i R_{M,m}^2 + \epsilon_{i,m}. \tag{18}$$

Similar to stock beta, regression are performed over a moving window covering the most recent 60 months, requiring at least 36 months of non-missing data.

Following Ang et al. (2006), the monthly idiosyncratic volatility of stock i (IVOL) is computed as the standard deviation of the daily residuals estimated from the following regression:

$$R_{i,d} = \alpha_i + \beta_i MKT_{M,d} + \gamma_i SMB_d + \phi_i HML_d + \gamma_i UMD_d \epsilon_{i,d}, \tag{19}$$

where $R_{i,d}$ and $MKT_{M,d}$ are the daily excess returns on stock i and the CRSP value-weighted market index, respectively. SMB_d and HML_d are the daily size and book-to-market factors of Fama and French (1996), respectively. UMD_d is the momentum factor.

Following Diether et al. (2002), Analyst forecast dispersion (DISP) is defined as the standard deviation of fiscal year one earnings forecasts scaled by the absolute value of the mean earnings forecast in a given month. To compute analyst forecast dispersion, each stock must be covered by two or more analysts during that month.

Appendix II: Proof of Proposition 1

The posterior beliefs of type i traders on X at $t = 1'$ are represented by

$$\begin{cases} X \sim N(\frac{Z_i X_i + (\frac{1-\alpha}{2-\alpha}\sigma^2+\sigma_\eta^2)^{-1}(S-\frac{(1-\alpha)\mu_i+\mu_{-i}}{2-\alpha})}{Z_i+(\frac{1-\alpha}{2-\alpha}\sigma^2+\sigma_\eta^2)^{-1}}, (Z_i + (\frac{1-\alpha}{2-\alpha}\sigma^2+\sigma_\eta^2)^{-1})^{-1}) & \text{if } \mu_{-i}-\mu_i > 0 \\ \\ X \sim N(\frac{Z_i X_i + (\frac{1+\beta}{2+\beta}\sigma^2+\sigma_\eta^2)^{-1}(S-\frac{(1+\beta)\mu_i+\mu_{-i}}{2+\beta})}{Z_i+(\frac{1+\beta}{2+\beta}\sigma^2+\sigma_\eta^2)^{-1}}, (Z_i + (\frac{1+\beta}{2+\beta}\sigma^2+\sigma_\eta^2)^{-1})^{-1}) & \text{if } \mu_{-i}-\mu_i < 0 \end{cases} \tag{20}$$

Since $\sigma^2 << \sigma_\eta^2$, we have $\frac{1-\alpha}{2-\alpha}\sigma^2 < \frac{1+\beta}{2+\beta}\sigma^2 << \sigma_\eta^2$. Therefore, $(\frac{1+\beta}{2+\beta}\sigma^2+\sigma_\eta^2) \approx \sigma_\eta^2$ and $(\frac{1-\alpha}{2-\alpha}\sigma^2+\sigma_\eta^2) \approx \sigma_\eta^2$. Using the above properties, (20) is given by

$$\begin{cases} X \sim N(\frac{Z_i X_i + \sigma_\eta^{-2}(S-\frac{(1-\alpha)\mu_i+\mu_{-i}}{2-\alpha})}{Z_i+\sigma_\eta^{-2}}, (Z_i+\sigma_\eta^{-2})^{-1}) & \text{if } \mu_{-i}-\mu_i > 0 \\ \\ X \sim N(\frac{Z_i X_i + \sigma_\eta^{-2}(S-\frac{(1+\beta)\mu_i+\mu_{-i}}{2+\beta})}{Z_i+\sigma_\eta^{-2}}, (Z_i+\sigma_\eta^{-2})^{-1}) & \text{if } \mu_{-i}-\mu_i < 0. \end{cases} \tag{21}$$

The resulting demand for type i traders is given by

$$\begin{cases} m_{i,1'} = (\frac{Z_i X_i + \sigma_\eta^{-2}(S-\frac{(1-\alpha)\mu_i+\mu_{-i}}{2-\alpha})}{Z_i+\sigma_\eta^{-2}} - P_{1'})\frac{Z_i+\sigma_\eta^{-2}}{\lambda} & \text{if } \mu_{-i}-\mu_i > 0 \\ \\ m_{i,1'} = (\frac{Z_i X_i + \sigma_\eta^{-2}(S-\frac{(1+\beta)\mu_i+\mu_{-i}}{2+\beta})}{Z_i+\sigma_\eta^{-2}} - P_{1'})\frac{Z_i+\sigma_\eta^{-2}}{\lambda} & \text{if } \mu_{-i}-\mu_i < 0. \end{cases} \tag{22}$$

Using the market-clearing condition ($0.5m_{1,1'} + 0.5m_{2,1'} = 0$), the market-clearing price at $t = 1'$ is given by

$$P_{1'}^* = \frac{Z_1X_1 + Z_2X_2 + \sigma_\eta^{-2}\{(S-\mu_1)+(S-\mu_2) - \frac{(\alpha+\beta)}{(2-\alpha)(2+\beta)}|\mu_1-\mu_2|\}}{Z_1+Z_2+2\sigma_\eta^{-2}}. \tag{23}$$

Next, $P_2^* = X$ since X is revealed at $t = 2$. Denote $R = P_2^* - P_{1'}^*$. Using $E[X_1] = E[X_2] = X$ and $E[S] = E[X+\eta] = X + \mu_\eta$, we have

$$\begin{aligned} E[R] &= E[P_2^* - P_{1'}^*] && (24) \\ &= E[X - \frac{Z_1X_1 + Z_2X_2 + \sigma_\eta^{-2}\{(S-\mu_1)+(S-\mu_2) - \frac{(\alpha+\beta)}{(2-\alpha)(2+\beta)}|\mu_1-\mu_2|\}}{Z_1+Z_2+2\sigma_\eta^{-2}}] && (25) \\ &= \frac{\sigma_\eta^{-2}\{-2\mu_\eta + (\mu_1+\mu_2) + \frac{(\alpha+\beta)}{(2-\alpha)(2+\beta)}|\mu_1-\mu_2|\}}{Z_1+Z_2+2\sigma_\eta^{-2}}. && (26) \end{aligned}$$

Since $0 < \alpha < 1$ and $0 < \beta < \infty$, $E[R]$ is increasing in investor disagreement, $|\mu_1 - \mu_2|$, holding the average interpretation in the market $(\mu_1 + \mu_2)$ fixed.

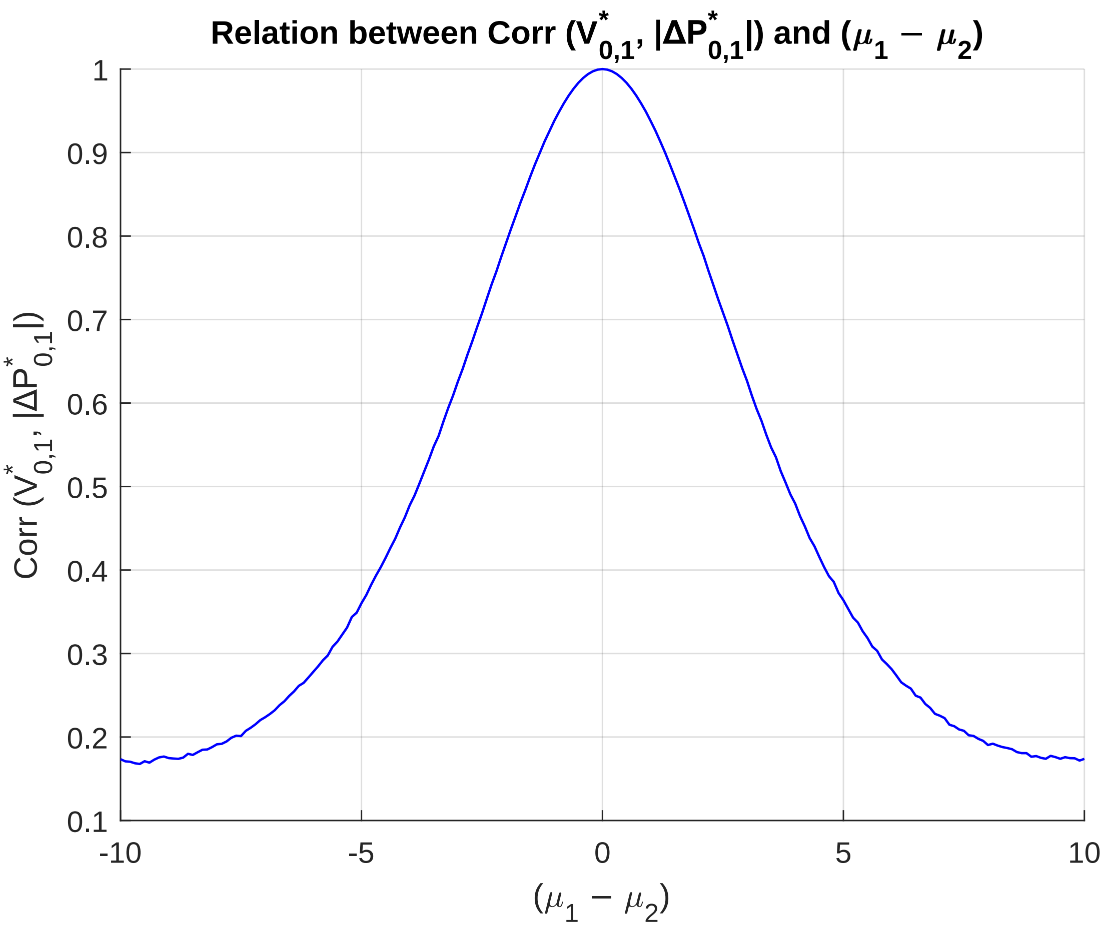


Figure 1: Relation between $Corr(V^*_{0,1}, |\Delta P^*_{0,1}|)$ and $(\mu_1 - \mu_2)$. The figure plots the correlation coefficient between equilibrium trading volume and absolute price change as a function of $(\mu_1 - \mu_2)$. Without loss of generality, μ_1 is fixed to 0, so $(\mu_1 - \mu_2)$ varies under different values of μ_2. For a given value of $(\mu_1 - \mu_2)$, I draw 100,000 observations from the distribution of $\eta \sim N(\mu_\eta = 0, \sigma^2_\eta = 2,000)$ and thus acquire 100,000 observations of S since $S = X + \eta$. The equilibrium trading volume, absolute price change, and the correlation between the two can be computed accordingly. Other model parameters are as follows: $X = 50$, $X_1 = 49$, $X_2 = 51$, $Z_1 = 1.05$, $Z_2 = 0.95$, $\lambda = 0.5$, and $\sigma^2 = 0.01$.

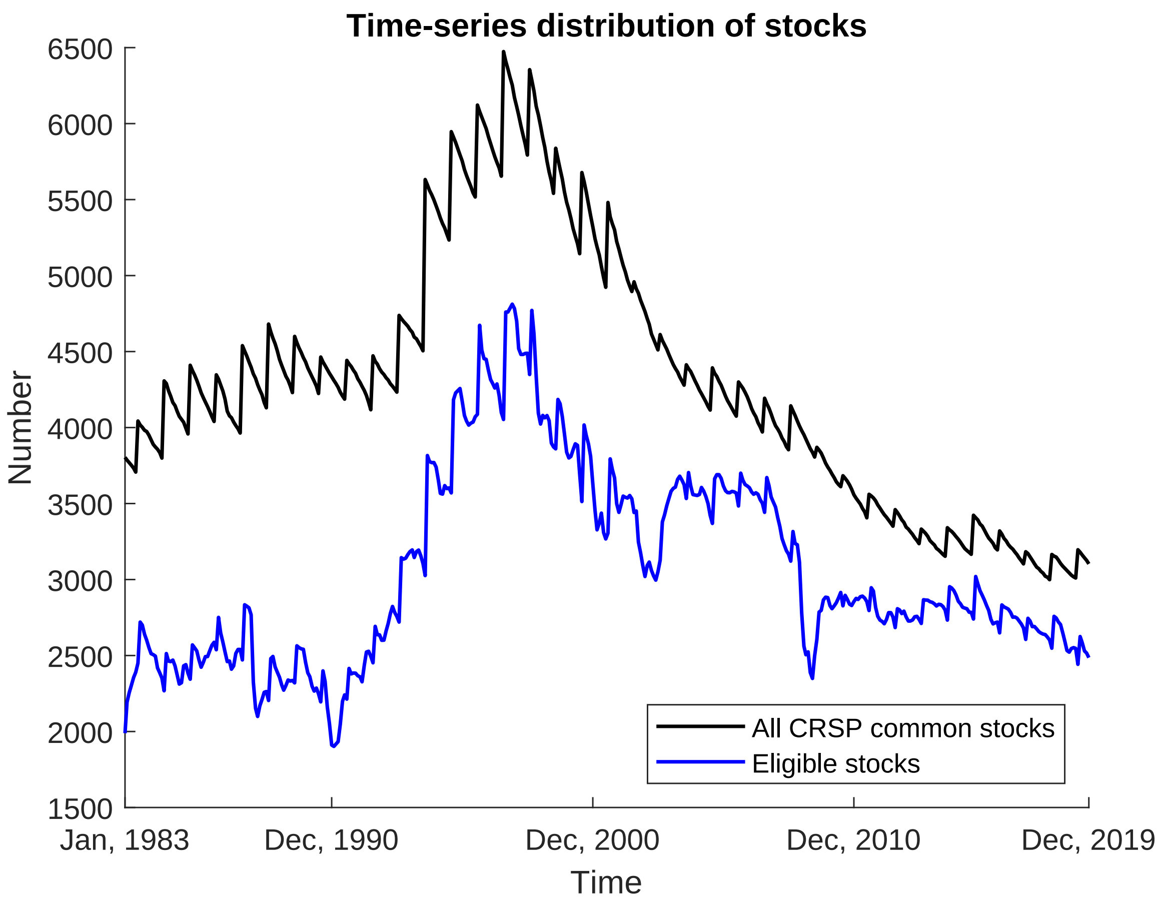


Figure 2: Time-series distribution of stocks. The figure plots the time-series distribution of all CRSP common stocks and eligible stocks. Eligible stocks are stocks with non-missing investor disagreement (ID) at the end of each month. ID at the end of a given month is defined as the contemporaneous correlation coefficient of daily trading volume and absolute price change over the past two months (around 44 trading days), multiplied by −1. 3. The sample period is from January 1983 to December 2019 (444 months).

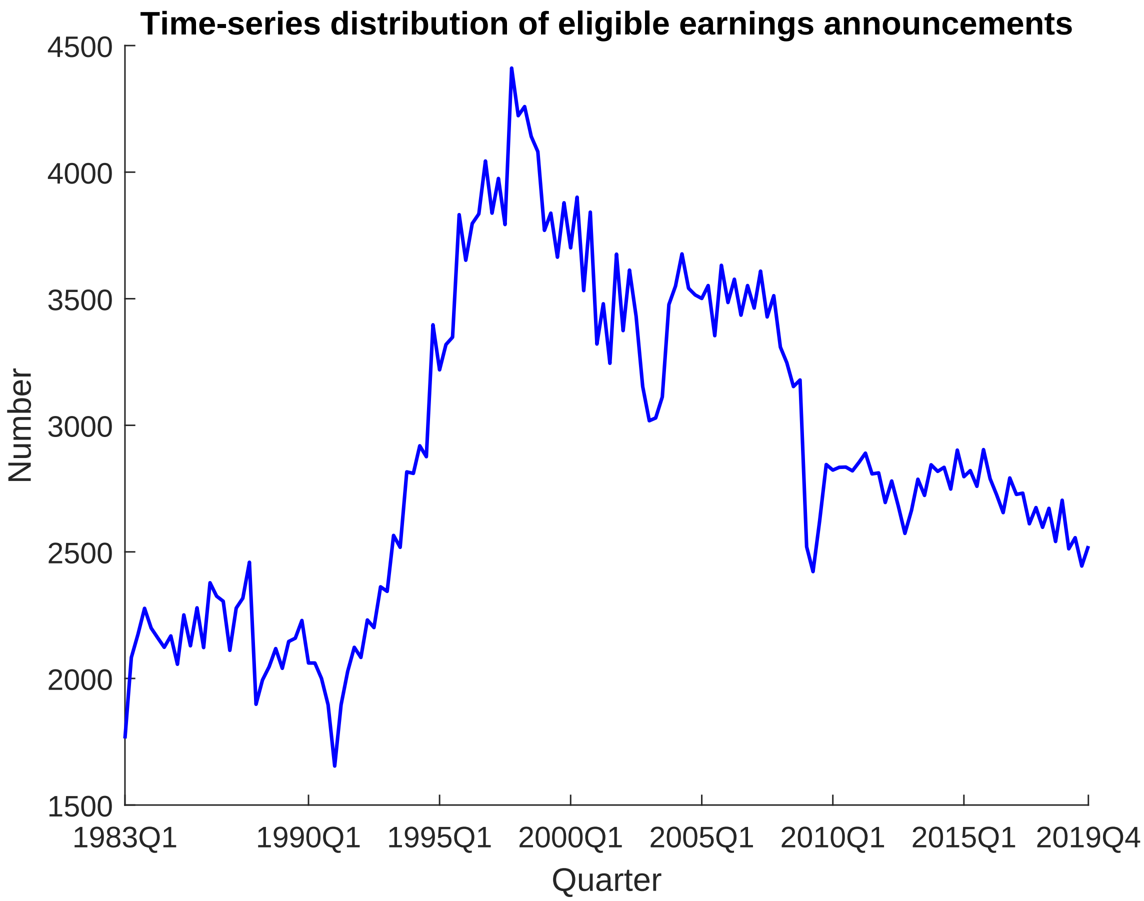


Figure 3: Time-series distribution of eligible earnings announcements. This figure plots the total number of eligible quarterly earnings announcements over time. It covers all NYSE, NASDAQ, and Amex firms available from the Compustat quarterly file with non-missing earnings. In addition, investor disagreement before the earnings announcement ($\mathrm{ID}_{\mathrm{before}}$) is required to be non-missing. $\mathrm{ID}_{\mathrm{before}}$ is defined as the contemporaneous correlation coefficient of daily trading volume and absolute price change over the window $[-45, -2]$, multiplied by -1, where $t = 0$ is the earnings announcement date. The sample period is from the first quarter of 1983 to the fourth quarter of 2019 (148 quarters).

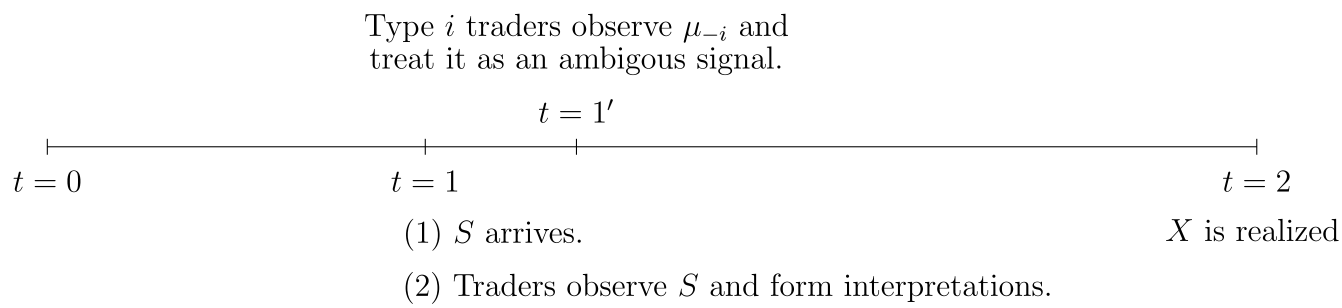


Figure 4: Model Timeline. There are two types of traders in the market with equal mass indexed by $i = 1, 2$. X is the true value of the stock and is realized at $t = 2$. At $t = 0$, type i traders' prior on $X \sim N(X_i, Z_i^{-1})$. At $t = 1$, the public signal $S = X + \eta$ arrives, where $\eta \perp X$ and $\eta \sim N(\mu_\eta, \sigma_\eta^2)$. Everything about S is common knowledge except for the mean μ_η. Type i traders believe that $\mu_\eta \sim N(\mu_i, \sigma^2)$, where μ_i denotes type i traders' interpretation of S. Let μ_{-i} denote the other type's interpretation of S from type i traders' perspective. At $t = 1'$, type i traders observe μ_{-i} and believe that $\mu_{-i} = \mu_\eta + \epsilon$, $\mu_\eta \perp \epsilon$, $\epsilon \sim N(0, \sigma_\epsilon^2)$, and $\sigma_\epsilon^2 \in [\underline{\sigma_\epsilon^2}, \overline{\sigma_\epsilon^2}]$, where $0 < \underline{\sigma_\epsilon^2} < \sigma^2 < \overline{\sigma_\epsilon^2} < \infty$.

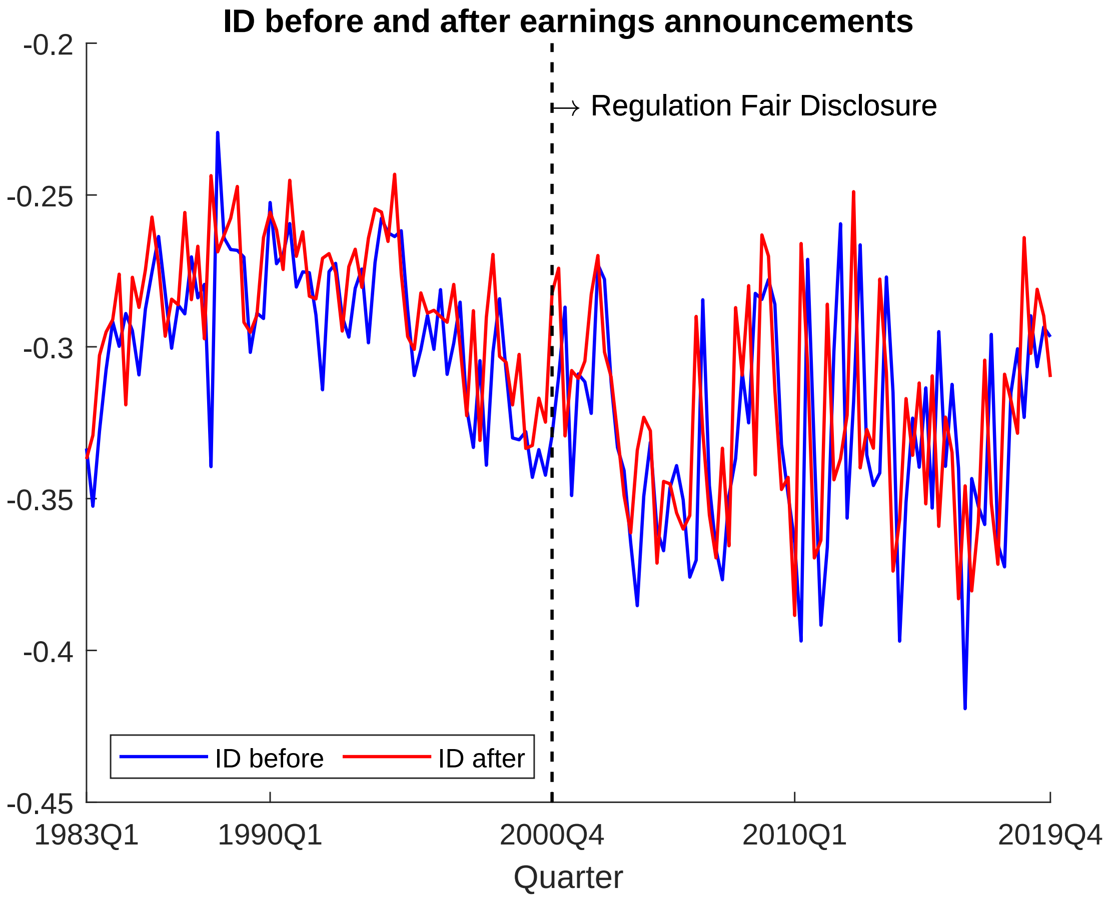
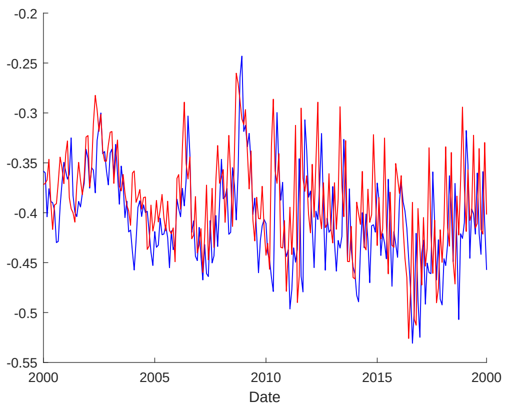
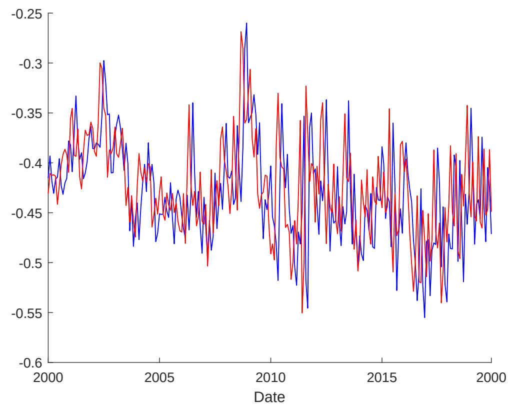


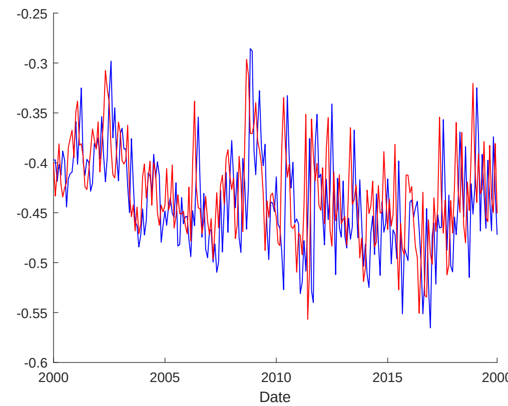
Figure 5: Investor disagreement (ID) before and after earnings announcement. The figure presents time series of cross-sectional average investor disagreement (ID) before and after the earnings announcement. $\mathrm{ID}_{\mathrm{before}}$ and $\mathrm{ID}_{\mathrm{after}}$ are defined as the contemporaneous correlation coefficient of daily trading volume and absolute price change over the window $[-45, -2]$ and $[2, 45]$, multiplied by -1, where $t = 0$ is the earnings announcement date. The sample period is from the first quarter of 1983 to the fourth quarter of 2019 (148 quarters). There are 413,454 observations.



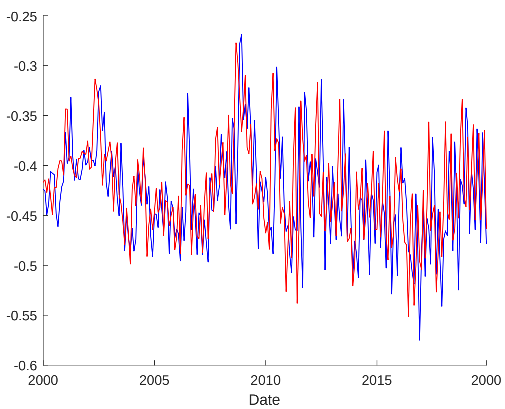
(a) ID before and after financial news



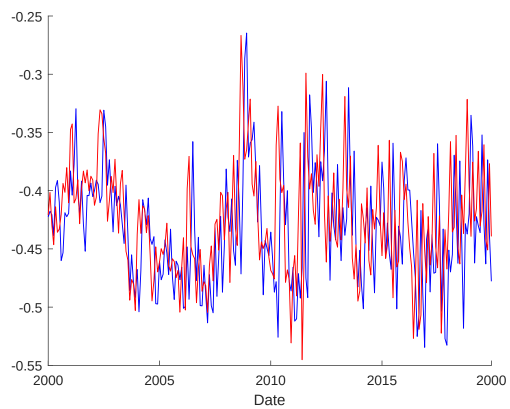
(b) ID before and after legal news



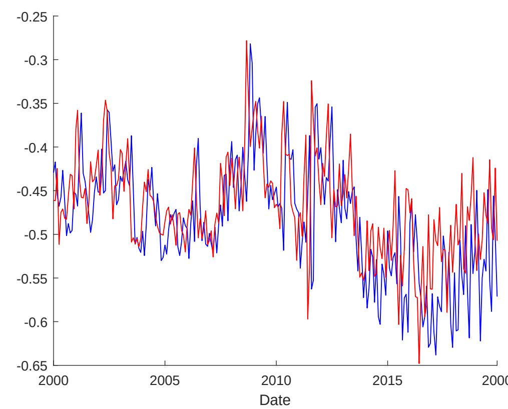
(c) ID before and after M&A news



(d) ID before and after operational news



(e) ID before and after ratings news



(f) ID before and after other news

Figure 6: Investor disagreement (ID) before and after news stories. The figure presents time series of cross-sectional average investor disagreement (ID) before (blue) and after (red) 6 types (financial, legal, M&A, operational, ratings, and others) of news stories. ID before the news (day 0) is defined as the correlation coefficient between daily trading volume and absolute price change over the 44-day window $[-45, -2]$, multiplied by -1. ID after the earnings announcement is defined similarly over the 44-day window $[2, 45]$. News data is obtained from RavenPack News Analytics on WRDS. The sample period is from January 2000 to December 2019.

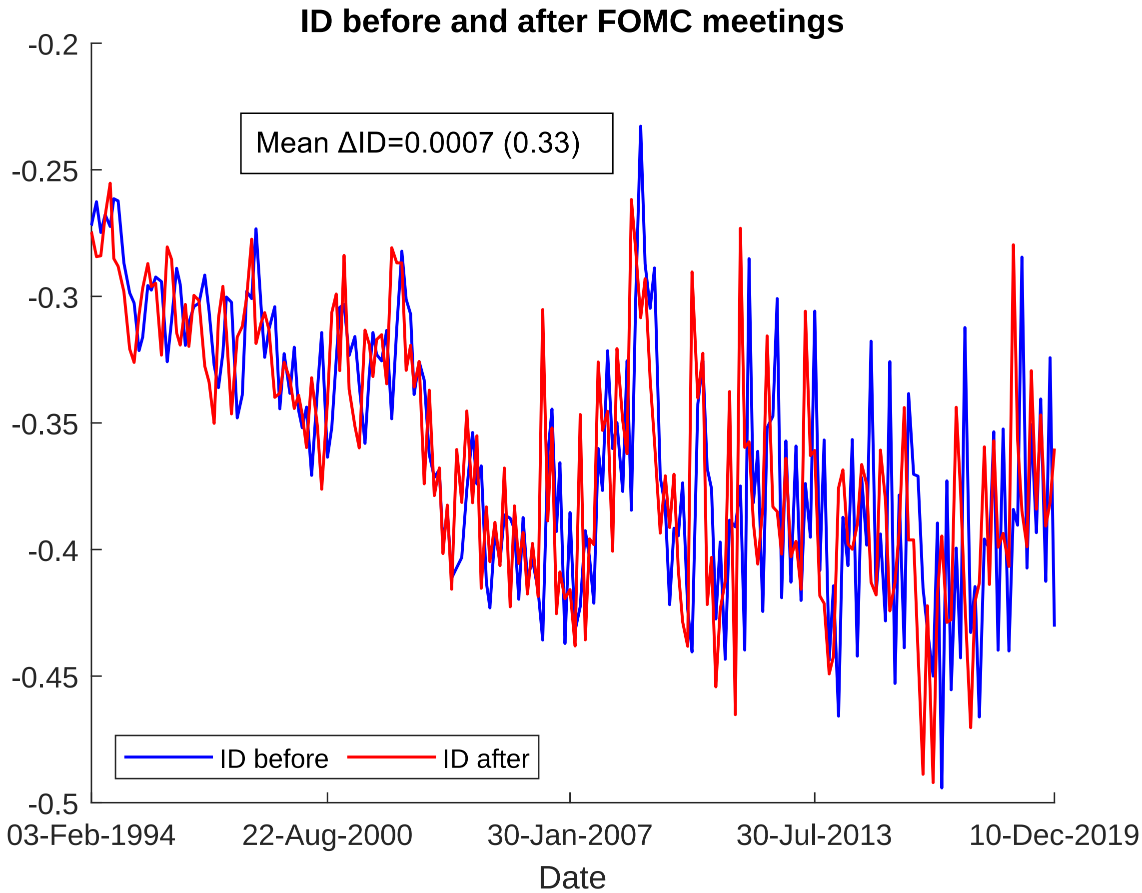


Figure 7: Investor disagreement (ID) before and after FOMC meetings. The figure presents time series of cross-sectional average investor disagreement before ($\text{ID}_{\text{before}}$) and after ($\text{ID}_{\text{after}}$) the earnings announcement. $\text{ID}_{\text{before}}$ and ID_{after} are defined as the contemporaneous correlation coefficient of daily trading volume and absolute price change over the window $[-45, -2]$ and $[2, 45]$, multiplied by -1, where $t = 0$ is the FOMC date. The sample period is from 1994 to 2019 (208 FOMC announcements). There are 665,013 observations.

Table 1: Returns of equal-weighted portfolios sorted on investor disagreement. For each month, decile portfolios are formed by sorting individual stocks based on their investor disagreement (ID) at the end of previous month. Portfolio 1 (10) is the portfolio of stocks with the lowest (highest) ID at the end of last month. Stocks are held for one month and portfolio returns are equal-weighted. ID at the end of a given month is defined as the correlation coefficient of daily trading volume and absolute price change in the past 2 months, multiplied by -1. The second column reports the time series average of monthly excess returns. The third to fifth column report corresponding alphas with respect to the CAPM, Fama-French three-factor model, and Fama-French-Carhart four-factor model. The sixth column reports the alpha of the five-factor model that in addition includes the liquidity factor of Pástor and Stambaugh (2003). The row labeled "$10-1$" presents the the differences in monthly excess returns and alphas between decile 10 (High ID) and decile 1 (Low ID). Newey and West (1987) t-statistics are reported in parentheses. The sample period is from January 1983 to December 2019.

ID deciles	Excess return	CAPM alpha	3-factor alpha	4-factor alpha	5-factor alpha
1 (Low)	0.18	-0.63	-0.57	-0.50	-0.51
	(0.60)	(-4.07)	(-6.93)	(-6.23)	(-6.43)
2	0.32	-0.49	-0.44	-0.39	-0.40
	(1.13)	(-3.7)	(-7.06)	(-6.47)	(-6.79)
3	0.39	-0.40	-0.39	-0.31	-0.31
	(1.43)	(-3.69)	(-7.26)	(-5.73)	(-5.91)
4	0.46	-0.33	-0.34	-0.25	-0.25
	(1.67)	(-2.88)	(-5.69)	(-4.78)	(-4.80)
5	0.56	-0.19	-0.21	-0.14	-0.14
	(2.20)	(-1.76)	(-4.10)	(-2.62)	(-2.66)
6	0.57	-0.17	-0.21	-0.10	-0.10
	(2.24)	(-1.49)	(-3.48)	(-1.51)	(-1.55)
7	0.69	-0.01	-0.07	0.04	0.05
	(2.79)	(-0.11)	(-1.07)	(0.56)	(0.67)
8	0.73	0.04	-0.02	0.10	0.10
	(3.04)	(0.36)	(-0.27)	(1.22)	(1.37)
9	0.65	0.00	-0.08	0.03	0.03
	(2.72)	(0.02)	(-0.89)	(0.31)	(0.37)
10 (High)	0.82	0.23	0.14	0.25	0.26
	(3.63)	(1.61)	(1.42)	(2.46)	(2.69)
$10-1$	0.65	0.87	0.71	0.75	0.77
	(3.91)	(5.64)	(5.47)	(5.94)	(6.28)
MKT BETA		-0.31	-0.20	-0.21	-0.21
		(-8.25)	(-5.66)	(-6.05)	(-5.80)
SMB BETA			-0.37	-0.37	-0.37
			(-3.04)	(-3.19)	(-3.18)
HML BETA			0.38	0.36	0.36
			(6.17)	(6.81)	(6.73)
UMD BETA				-0.06	-0.06
				(-1.02)	(-1.00)
LIQ BETA					-0.06
					(-1.69)
Adj. R^2		18.82%	48.30%	48.78%	49.08%

Table 2: Returns of value-weighted portfolios sorted on investor disagreement. For each month, decile portfolios are formed by sorting individual stocks based on their investor disagreement (ID) at the end of previous month. Portfolio 1 (10) is the portfolio of stocks with the lowest (highest) ID at the end of last month. Stocks are held for one month and portfolio returns are value-weighted. ID at the end of a given month is defined as the correlation coefficient of daily trading volume and absolute price change in the past 2 months, multiplied by -1. The second column reports the time series average of monthly excess returns. The third to fifth column report corresponding alphas with respect to the CAPM, Fama-French three-factor model, and Fama-French-Carhart four-factor model. The sixth column reports the alpha of the five-factor model that in addition includes the liquidity factor of Pástor and Stambaugh (2003). The row labeled "$10-1$" presents the the differences in monthly excess returns and alphas between decile 10 (High ID) and decile 1 (Low ID). Newey and West (1987) t-statistics are reported in parentheses. The sample period is from January 1983 to December 2019.

ID deciles	Excess return	CAPM alpha	3-factor alpha	4-factor alpha	5-factor alpha
1 (Low)	0.50	-0.23	-0.21	-0.07	-0.07
	(2.06)	(-2.55)	(-2.30)	(-0.71)	(-0.74)
2	0.62	-0.08	-0.08	-0.01	-0.00
	(2.91)	(-1.05)	(-1.11)	(-0.08)	(-0.01)
3	0.54	-0.17	-0.19	-0.10	-0.10
	(2.38)	(-2.37)	(-2.76)	(-1.36)	(-1.41)
4	0.67	-0.04	-0.07	0.00	0.01
	(3.03)	(-0.59)	(-1.02)	(0.07)	(0.14)
5	0.67	-0.05	-0.06	0.03	0.04
	(2.92)	(-0.62)	(-0.98)	(0.50)	(0.54)
6	0.61	-0.09	-0.12	-0.04	-0.05
	(2.88)	(-1.31)	(-1.70)	(-0.56)	(-0.67)
7	0.67	-0.04	-0.09	-0.01	-0.01
	(3)	(-0.56)	(-1.27)	(-0.09)	(-0.08)
8	0.76	0.09	0.04	0.10	0.10
	(3.45)	(1.03)	(0.51)	(1.25)	(1.20)
9	0.71	0.05	-0.03	0.09	0.06
	(3.28)	(0.53)	(-0.29)	(0.88)	(0.70)
10 (High)	0.91	0.29	0.18	0.29	0.29
	(4.51)	(2.31)	(1.72)	(2.65)	(2.71)
$10-1$	0.41	0.52	0.40	0.37	0.36
	(2.78)	(3.56)	(3.12)	(2.67)	(2.68)
MKT BETA		-0.15	-0.09	-0.08	-0.08
		(-4.02)	(-2.76)	(-2.26)	(-2.26)
SMB BETA			-0.02	-0.02	-0.02
			(-0.36)	(-0.36)	(-0.36)
HML BETA			0.36	0.37	0.37
			(6.59)	(6.57)	(6.54)
UMD BETA				0.04	0.04
				(0.83)	(0.83)
LIQ BETA					0.01
					(0.23)
Adj. R^2		5.33%	18.45%	18.59%	18.41%

Table 3: Investor disagreement decile: average stock characteristics. Stocks are sorted into decile portfolios based on investor disagreement (ID) at the end of each month. Decile 1 (10) is the portfolio of stocks with the lowest (highest) investor disagreement at the end of each month. Investor disagreement (ID) at the end of a given month is defined as the correlation coefficient of daily trading volume and absolute price change in the past 2 months, multiplied by -1. The table presents for each ID decile, the time-series average of mean values of stock characteristics, including firm size (SIZE), book-to-market (BM) ratio, the cumulative return (in percent) over the 11 months prior to the portfolio formation month (MOM), the return (in percent) in the portfolio formation month (REV), average turnover ratio (TURN), idiosyncratic volatility (IVOL) as defined in Ang et al. (2006), Amihud (2002) illiquidity ratio (ILLIQ), lottery demand (MAX) as defined in Bali et al. (2011), institutional owernship ratio (IOR) defined the ratio of shares owned by institutions as reported in 13F filings in the last quarter, the stock beta (BETA), co-skewness (COSKEW) as defined in Harvey and Siddique (2000), and analyst forecast dispersion (DISP) as defined in Diether et al. (2002). The weights are based on the number of observations in each portfolio in each month, and the variables are defined in detail in Appendix I. Newey and West (1987) t-statistics are reported in parentheses. The sample period is from January 1983 to December 2019 and there is an average of 306 stocks per decile portfolio.

	Investor disagreement (ID) decile portfolio									
	1 (Low)	2	3	4	5	6	7	8	9	10 (High)
ID	-0.77	-0.61	-0.52	-0.44	-0.37	-0.3	-0.23	-0.16	-0.07	0.07
ID (next month)	-0.58	-0.49	-0.44	-0.39	-0.35	-0.32	-0.28	-0.24	-0.19	-0.12
SIZE	5.90	6.10	6.16	6.16	6.11	6.04	5.94	5.82	5.66	5.37
BM	0.62	0.63	0.64	0.65	0.67	0.68	0.70	0.72	0.74	0.77
MOM	24.91	26.48	24.7	22.65	20.46	18.83	17.21	15.93	14.62	12.82
REV	2.57	1.75	1.23	0.89	0.69	0.45	0.26	0.15	0.03	-0.19
TURN	1.00	0.70	0.60	0.53	0.47	0.43	0.39	0.35	0.31	0.26
IVOL	3.30	2.56	2.33	2.18	2.09	2.03	1.98	1.95	1.93	1.96
ILLIQ	0.08	0.12	0.16	0.20	0.27	0.24	0.09	0.35	0.43	0.30
MAX	9.99	7.45	6.65	6.12	5.82	5.56	5.38	5.21	5.09	4.99
IOR	0.49	0.48	0.48	0.47	0.47	0.46	0.45	0.44	0.42	0.39
BETA	1.47	1.42	1.37	1.32	1.27	1.22	1.18	1.13	1.09	1.01
COSKEW	0.10	0.14	0.15	0.07	0.02	-0.05	-0.16	-0.20	-0.37	-0.59
DISP	0.02	0.01	0.02	0.03	0.01	0.02	0.04	0.01	0.02	0.01

Table 4: Mean portfolio returns by firm size (SIZE) and investor disagreement (ID). Each month, individual stocks are first sorted into quintiles based on firm size (SIZE) in the previous month. Next, within each SIZE decile, stocks are further sorted into deciles based on investor disagreement (ID) in the previous month. Stocks are held for one month, and portfolio returns are equal-weighted. The table reports time series averages of monthly excess returns. SIZE is the log of market capitalization in millions of dollars. Investor disagreement (ID) at the end of a given month is defined as the correlation coefficient of daily trading volume and absolute price change in the past 2 months, multiplied by -1. "$10-1$", "CAPM alpha" , "3-factor alpha", "4-factor alpha", and "5-factor alpha" report the difference in excess returns, CAPM alpha, three-factor alpha (MKT, SMB, and HML), four-factor alpha (MKT, SMB, HML, and UMD), and five-factor alpha (MKT, SMB, HML, UMD, and LIQ) between high ID and low ID decile in each SIZE quintile, respectively. The corresponding Newey and West (1987) t-statistics are reported in parentheses. The sample period is from January 1983 to December 2019.

	Size quintiles				
ID deciles	Small caps	2	3	4	Large caps
1 (Low)	-0.39	0.16	0.07	0.37	0.56
2	-0.14	0.16	0.40	0.42	0.68
3	0.06	0.30	0.33	0.56	0.71
4	0.16	0.33	0.57	0.71	0.6
5	0.25	0.49	0.55	0.75	0.80
6	0.37	0.53	0.73	0.63	0.77
7	0.47	0.34	0.67	0.77	0.64
8	0.49	0.56	0.88	0.85	0.77
9	0.63	0.67	0.76	0.69	0.81
High	0.68	0.77	0.83	0.89	0.84
$10-1$	1.07	0.62	0.77	0.52	0.28
	(3.75)	(2.86)	(3.80)	(2.94)	(2.14)
CAPM alpha	1.32	0.81	0.96	0.69	0.41
	(4.68)	(3.92)	(4.81)	(3.97)	(3.37)
3-factor alpha	1.17	0.65	0.77	0.53	0.30
	(4.19)	(3.18)	(4.55)	(3.81)	(2.67)
4-factor alpha	1.27	0.75	0.79	0.50	0.30
	(4.33)	(3.63)	(4.73)	(3.33)	(2.26)
5-factor alpha	1.32	0.78	0.80	0.49	0.29
	(4.57)	(3.82)	(4.82)	(3.28)	(2.29)

Table 5: Mean portfolio returns by book-to-market (BM) ratio and investor disagreement (ID). Each month, individual stocks are first sorted into quintiles based on book-to-market (BM) ratio in the previous month. Next, within each BM decile, stocks are further sorted into deciles based on investor disagreement (ID) in the previous month. Stocks are held for one month, and portfolio returns are equal-weighted. The table reports time series averages of monthly excess returns. Investor disagreement (ID) at the end of a given month is defined as the correlation coefficient of daily trading volume and absolute price change in the past 2 months, multiplied by -1. "10$-$1", "CAPM alpha" , "3-factor alpha", "4-factor alpha", and "5-factor alpha" report the difference in excess returns, CAPM alpha, three-factor alpha (MKT, SMB, and HML), four-factor alpha (MKT, SMB, HML, and UMD), and five-factor alpha (MKT, SMB, HML, UMD, and LIQ) between high ID and low ID decile in each BM quintile, respectively. The corresponding Newey and West (1987) t-statistics are reported in parentheses. The sample period is from January 1983 to December 2019.

	Book-to-market quintiles				
ID deciles	Low BM	2	3	4	High BM
Low	-0.40	0.31	0.37	0.37	0.38
2	-0.28	0.49	0.50	0.53	0.59
3	-0.11	0.42	0.57	0.61	0.61
4	0.01	0.32	0.67	0.73	0.65
5	0.02	0.61	0.69	0.69	0.80
6	0.17	0.45	0.69	0.84	0.78
7	0.20	0.58	0.69	0.84	0.84
8	0.30	0.58	0.85	0.82	0.82
9	0.38	0.51	0.81	0.79	0.81
High	0.56	0.72	0.76	0.89	0.99
$10-1$	0.97	0.41	0.39	0.52	0.61
	(5.33)	(2.07)	(2.47)	(3.08)	(3.26)
CAPM alpha	1.07	0.56	0.56	0.72	0.82
	(5.85)	(2.91)	(3.78)	(4.68)	(4.66)
3-factor alpha	0.95	0.43	0.44	0.63	0.76
	(5.41)	(2.23)	(3.24)	(4.20)	(4.26)
4-factor alpha	1.01	0.53	0.47	0.66	0.81
	(5.13)	(2.92)	(3.33)	(4.57)	(4.79)
5-factor alpha	1.00	0.56	0.51	0.66	0.84
	(5.17)	(3.15)	(3.65)	(4.71)	(5.01)

Table 6: Mean portfolio returns by momentum (MOM) and investor disagreement (ID). Each month, individual stocks are first sorted into quintiles based on momentum (MOM) in the previous month. Next, within each MOM decile, stocks are further sorted into deciles based on investor disagreement (ID) in the previous month. Stocks are held for one month, and portfolio returns are equal-weighted. The table reports time series averages of monthly excess returns. MOM is computed as the cumulative return of a stock of 11 months ending one month prior to the portfolio formation month. Investor disagreement (ID) at the end of a given month is defined as the correlation coefficient of daily trading volume and absolute price change in the past 2 months, multiplied by -1. "$10-1$", "CAPM alpha" , "3-factor alpha", "4-factor alpha", and "5-factor alpha" report the difference in excess returns, CAPM alpha, three-factor alpha (MKT, SMB, and HML), four-factor alpha (MKT, SMB, HML, and UMD), and five-factor alpha (MKT, SMB, HML, UMD, and LIQ) between high ID and low ID decile in each MOM quintile, respectively. The corresponding Newey and West (1987) t-statistics are reported in parentheses. The sample period is from January 1983 to December 2019.

	Momentum quintiles				
ID deciles	Losers	2	3	4	Winners
Low	-0.69	0.35	0.55	0.50	0.44
2	-0.66	0.40	0.68	0.78	0.56
3	-0.49	0.52	0.70	0.65	0.68
4	-0.40	0.39	0.58	0.82	0.68
5	-0.29	0.53	0.69	0.81	0.83
6	-0.40	0.63	0.76	0.84	0.91
7	0.12	0.75	0.75	0.86	0.77
8	0.25	0.66	0.87	0.83	0.86
9	0.12	0.77	0.78	0.91	0.73
High	0.68	0.74	0.86	0.83	0.94
$10-1$	1.37	0.39	0.32	0.35	0.50
	(5.74)	(2.69)	(1.89)	(2.22)	(3.00)
CAPM alpha	1.54	0.53	0.48	0.54	0.64
	(6.58)	(3.73)	(3.14)	(3.68)	(3.68)
3-factor alpha	1.40	0.44	0.36	0.40	0.48
	(6.44)	(3.44)	(2.54)	(2.68)	(3.33)
4-factor alpha	1.50	0.46	0.36	0.40	0.48
	(6.37)	(3.68)	(2.73)	(2.78)	(3.02)
5-factor alpha	1.54	0.49	0.38	0.42	0.47
	(6.57)	(3.87)	(2.95)	(2.93)	(2.98)

Table 7: Bivariate portfolio sorts on investor disagreement and control variables. Double-sorted, equally-weighted decile portfolios are formed every month based on investor disagreement (ID) after controlling for market capitalization (SIZE), book-to-market ratio (BM), momentum (MOM), short-term reversals (REV), turnover (TURN), idiosyncratic volatility (IVOL), illiquidity (ILLIQ), demand for lottery stocks (MAX), institutional ownership ratio (IOR), stock market beta (BETA), co-skewness (COSKEW), and analyst forecast dispersion (DISP). ID at the end of a given month is computed as the contemporaneous correlation coefficient of daily trading volume and absolute price change over the past 2 months, multiplied by -1. The other control variables are defined in Appendix I. In each case, I first sort stocks into deciles using the control variable, then within each decile I sort stocks into decile portfolios based on ID. The ten ID portfolios are then averaged over each of the ten control deciles to compute excess returns. "$10-1$" and "5-factor alpha" report the differences in average monthly excess returns and alphas with respect to the five-factor model (MKT, SMB, HML, UMD, and LIQ) between the High ID and Low ID decile portfolios, respectively. The sample period is from January 1983 to December 2019. Newey and West (1987) t-statistics are reported in parentheses.

	Investor disagreement (ID) decile											
	1 (Low)	2	3	4	5	6	7	8	9	10 (High)	$10-1$	5-factor alpha
SIZE	0.16	0.32	0.38	0.47	0.57	0.60	0.64	0.75	0.68	0.80	0.64	0.72
	(0.56)	(1.14)	(1.35)	(1.76)	(2.20)	(2.35)	(2.57)	(3.14)	(2.88)	(3.57)	(4.14)	(6.50)
BM	0.22	0.36	0.45	0.46	0.53	0.61	0.60	0.68	0.65	0.79	0.57	0.71
	(0.77)	(1.30)	(1.68)	(1.76)	(2.04)	(2.42)	(2.42)	(2.86)	(2.68)	(3.45)	(4.19)	(6.32)
MOM	0.26	0.34	0.42	0.41	0.55	0.58	0.69	0.66	0.65	0.81	0.55	0.62
	(0.94)	(1.24)	(1.55)	(1.55)	(2.18)	(2.26)	(2.75)	(2.71)	(2.66)	(3.50)	(4.21)	(6.05)
REV	0.22	0.37	0.48	0.53	0.57	0.58	0.60	0.68	0.62	0.73	0.52	0.60
	(0.78)	(1.40)	(1.77)	(2.00)	(2.16)	(2.25)	(2.38)	(2.79)	(2.54)	(3.16)	(4.51)	(6.57)
TURN	0.20	0.32	0.40	0.50	0.55	0.56	0.62	0.70	0.67	0.84	0.64	0.65
	(0.75)	(1.27)	(1.50)	(1.93)	(2.12)	(2.18)	(2.46)	(2.81)	(2.72)	(3.43)	(6.78)	(6.50)
IVOL	0.40	0.41	0.50	0.53	0.52	0.41	0.60	0.63	0.61	0.75	0.34	0.45
	(1.62)	(1.57)	(1.91)	(1.99)	(1.96)	(1.58)	(2.31)	(2.47)	(2.45)	(3.13)	(3.67)	(5.29)
ILLIQ	0.13	0.31	0.34	0.52	0.56	0.64	0.67	0.67	0.70	0.81	0.68	0.72
	(0.46)	(1.10)	(1.27)	(1.95)	(2.20)	(2.56)	(2.73)	(2.81)	(2.98)	(3.73)	(4.58)	(6.90)
MAX	0.38	0.43	0.49	0.51	0.55	0.51	0.62	0.60	0.57	0.72	0.34	0.45
	(1.45)	(1.62)	(1.89)	(1.96)	(2.12)	(1.98)	(2.38)	(2.45)	(2.27)	(2.96)	(3.41)	(4.65)
IOR	0.24	0.32	0.43	0.51	0.57	0.56	0.66	0.70	0.68	0.79	0.55	0.63
	(0.83)	(1.17)	(1.57)	(1.95)	(2.23)	(2.26)	(2.74)	(2.92)	(2.82)	(3.39)	(3.58)	(5.68)
BETA	0.20	0.30	0.42	0.50	0.52	0.58	0.64	0.72	0.66	0.81	0.61	0.66
	(0.74)	(1.13)	(1.59)	(1.91)	(1.99)	(2.32)	(2.54)	(2.95)	(2.67)	(3.40)	(5.41)	(7.09)
COSKEW	0.23	0.28	0.43	0.47	0.51	0.62	0.61	0.73	0.66	0.83	0.60	0.69
	(0.80)	(1.02)	(1.63)	(1.74)	(1.96)	(2.45)	(2.47)	(3.00)	(2.70)	(3.59)	(4.19)	(6.26)
DISP	0.27	0.34	0.43	0.46	0.59	0.58	0.57	0.62	0.63	0.73	0.45	0.50
	(0.97)	(1.22)	(1.61)	(1.69)	(2.29)	(2.29)	(2.30)	(2.48)	(2.59)	(3.10)	(3.42)	(5.03)

Table 8: Fama-Macbeth Cross-Sectional Regressions. The table reports the time-series averages of the slope coefficients obtained from regression monthly excess returns on investor disagreement (ID) in the previous month and a set of lagged predictive variables using the Fama-Macbeth (1973) approach. The control variables are the log market capitalization in millions of dollars (SIZE), book-to market (BM) ratio, momentum (MOM), short-term reversal (REV), turnover ratio (TURN), idiosyncratic volatility (IVOL), illiquidity (ILLIQ), demand for lottery stocks (MAX), institutional ownership ratio (IOR), stock beta (BETA), co-skewness (COSKEW), and analyst forecast dispersion (DISP). ID at the end of a given month is computed as the contemporaneous correlation coefficient of daily trading volume and absolute price change over the past 2 months, multiplied by -1. The other control variables are defined in Appendix I. The sample period is from January 1983 to December 2019. Newey and West (1987) t-statistics are reported in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)
ID	0.780	0.824	0.599	0.367	0.300	0.251
	(3.85)	(4.51)	(4.67)	(3.73)	(3.44)	(2.86)
SIZE		0.095	0.104	0.037	0.032	0.023
		(2.83)	(2.97)	(1.21)	(1.11)	(0.76)
BM		0.208	0.207	0.166	0.111	0.058
		(2.12)	(2.18)	(1.83)	(1.29)	(0.57)
MOM		0.003	0.004	0.003	0.003	0.005
		(1.93)	(2.26)	(2.00)	(2.61)	(3.45)
REV			-0.025	-0.031	-0.035	-0.038
			(-6.13)	(-7.03)	(-8.00)	(7.32)
TURN			-0.241	0.091	0.041	0.075
			(-1.91)	(0.91)	(0.47)	(0.77)
IVOL				-0.361	-0.350	-0.376
				(-6.81)	(-6.88)	(-6.72)
ILLIQ				0.017	0.012	-0.041
				(1.61)	(1.12)	(-0.29)
MAX				0.040	0.041	0.042
				(3.81)	(3.90)	(3.36)
IOR				-0.722	-0.808	-1.297
				(-4.54)	(-5.85)	(-8.79)
BETA					0.050	0.070
					(0.51)	(0.63)
COSKEW					-0.002	0.001
					(-0.44)	(0.16)
DISP						-0.058
						(-3.04)
Intercept	1.084	0.243	0.213	1.237	1.345	1.574
	(4.68)	(0.69)	(0.59)	(4.27)	(5.45)	(4.96)
Adj. R^2	0.33%	2.37%	3.73%	5.09%	6.09%	8.07%
observations	1,355,834	1,355,834	1,355,831	1,293,882	1,178,701	707,856

Table 9. Investor disagreement premium: business cycles and investor sentiment. Stocks are sorted into decile portfolios based on investor disagreement (ID) at the end of each month. Decile 1 (10) is the portfolio of stocks with the lowest (highest) investor disagreement at the end of each month. The table reports alphas with respect to the five-factor model (MKT, SMB, HML, UMD, and LIQ) in different sample periods. NBER expansion and recession periods are set by the NBER's Business Cycle Dating Committee. A high-sentiment month (low-sentiment) month is one in which the value of the BW (Baker and Wurgler (2006)) sentiment index in the previous month is above (below) the median value for the sample period. The row labeled "10 − 1" reports the five-factor alphas of the long-short ID strategy. The sample period for business cycles is January 1983 to December 2019, and the sampe period for investor sentiment is from January 1983 to December 2018. Newey and West (1987) t-statistics are reported in parentheses.

ID decile	NBER Expansions	NBER Recessions	Low Sentiment	High Sentiment
1 (Low)	-0.49	-0.55	-0.39	-0.65
	(-6.00)	(-1.52)	(-3.78)	(-5.38)
2	-0.39	-0.48	-0.33	-0.46
	(-6.27)	(-2.19)	(-4.27)	(-5.22)
3	-0.32	-0.17	-0.19	-0.42
	(-5.91)	(-0.60)	(-2.56)	(-5.23)
4	-0.24	-0.03	-0.10	-0.35
	(-4.54)	(-0.14)	(-1.58)	(-4.10)
5	-0.14	-0.12	-0.07	-0.18
	(-2.51)	(-0.77)	(-1.01)	(-2.15)
6	-0.09	-0.05	-0.05	-0.10
	(-1.45)	(-0.18)	(-0.17)	(-0.92)
7	0.05	0.14	0.02	0.17
	(0.67)	(0.53)	(0.22)	(1.56)
8	0.09	0.50	0.07	0.18
	(1.17)	(1.73)	(1.00)	(1.41)
9	0.06	0.04	-0.03	0.16
	(0.62)	(0.10)	(-0.39)	(1.08)
10 (High)	0.22	0.69	0.27	0.27
	(2.09)	(1.37)	(3.33)	(1.54)
10 − 1	0.71	1.23	0.65	0.92
	(5.38)	(2.04)	(5.25)	(4.53)
# of months	410	34	215	217

Table 10. Investor disagreement premium: economic uncertainty. Stocks are sorted into decile portfolios based on investor disagreement (ID) at the end of each month. Decile 1 (10) is the portfolio of stocks with the lowest (highest) investor disagreement at the end of each month. The table reports alphas with respect to the five-factor model (MKT, SMB, HML, UMD, and LIQ) over high and low economic uncertainty periods. The row labeled "$10-1$" reports the five-factor alphas of the long-short ID strategy. A high-sentiment month is one in which the value of the economic uncertainty index is above the median value for the sample period, and the low-sentiment months are those with below-median values. Macro, financial, real economic uncertainty measures are defined in Jurado et al. (2015) and Ludvigson et al. (2015). Policy-related economic uncertainty is defined in Baker et al. (2016). The sample period for macro, financial, and real economic uncertainty is from January 1983 to December 2019, and the sample period for policy-related economic uncertainty index is from January 1985 to December 2019. Newey and West (1987) t-statistics are reported in parentheses.

ID decile	Low Macro_UNC	High Macro_UNC	Low Fin_UNC	High Fin_UNC	Low Real_UNC	High Real_UNC	Low Policy_UNC	High Policy_UNC
1 (Low)	-0.58	-0.44	-0.59	-0.46	-0.50	-0.52	-0.59	-0.47
	(-5.64)	(-3.91)	(-6.71)	(-3.74)	(-4.71)	(-4.44)	(-4.80)	(-4.61)
2	-0.47	-0.34	-0.43	-0.36	-0.47	-0.33	-0.35	-0.40
	(-6.20)	(-4.28)	(-5.43)	(-4.16)	(-6.48)	(-4.05)	(-3.67)	(-5.28)
3	-0.34	-0.27	-0.37	-0.27	-0.38	-0.23	-0.34	-0.23
	(-4.48)	(-3.75)	(-5.35)	(-3.33)	(-5.18)	(-3.21)	(-4.2)	(-2.92)
4	-0.30	-0.15	-0.23	-0.22	-0.21	-0.21	-0.21	-0.18
	(-4.15)	(-1.96)	(-3.29)	(-2.84)	(-2.88)	(-2.95)	(-2.83)	(-2.39)
5	-0.14	-0.10	-0.15	-0.14	-0.19	-0.05	-0.07	-0.11
	(-2.36)	(-1.10)	(-2.07)	(-1.58)	(-2.81)	(-0.59)	(-0.96)	(-1.76)
6	-0.11	-0.07	-0.15	-0.08	-0.17	0.02	-0.08	-0.03
	(-1.66)	(-0.69)	(-2.19)	(-0.81)	(-2.76)	(0.20)	(-0.91)	(-0.42)
7	-0.02	0.09	0.05	0.04	0.05	0.08	0.12	0.05
	(-0.24)	(0.92)	(0.85)	(0.34)	(0.70)	(0.84)	(1.52)	(0.52)
8	0.13	0.06	0.03	0.12	0.14	0.09	0.16	0.11
	(-2.33)	(0.55)	(0.47)	(1.03)	(2.36)	(0.83)	(2.00)	(1.04)
9	0.07	0.02	0.06	0.06	0.02	0.10	0.15	0.04
	(1.08)	(0.11)	(0.87)	(0.42)	(0.22)	(0.82)	(1.73)	(0.32)
10 (High)	0.23	0.27	0.15	0.37	0.22	0.31	0.33	0.32
	(3.19)	(1.85)	(2.17)	(2.40)	(2.75)	(2.23)	(3.20)	(2.26)
$10-1$	0.81	0.72	0.74	0.83	0.72	0.83	0.92	0.79
	(6.65)	(3.95)	(6.68)	(4.21)	(5.53)	(4.84)	(5.30)	(5.43)
# of months	222	222	222	222	222	222	210	211

Table 11: Persistence of investor disagreement. The table examines the persistence of investor disagreement (ID). Panel A reports coefficients of regressing firm-level ID on lagged ID and lagged cross-sectional variables, including firm size (SIZE), book-to market (BM) ratio, momentum (MOM), short-term reversal (REV), turnover ratio (TURN), idiosyncratic volatility (IVOL), illiquidity (ILLIQ), demand for lottery stocks (MAX), institutional ownership ratio (IOR), stock beta (BETA), and co-skewness (COSKEW). Panel B presents the average month-to-month investor disagreement (ID) decile transition matrix. Column (i, j) is the average probability (in percentage) that a stock in ID decile i in one month will be in decile j in the subsequent month. Newey and West (1987) t-statistics are reported in parentheses. The sample period is from January 1983 to December 2019.

Panel A: Predictive regression	
Univariate predictive regression	0.552
	(185.20)
Adj. R^2	30.58%
Controlling for lagged variables	0.468
	(104.08)
Adj. R^2	34.54%

Panel B: Transition matrix (in %)

	ID deciles in next month									
ID deciles	Low	2	3	4	5	6	7	8	9	High
Low	42.87	18.26	9.99	7.14	5.64	4.48	3.80	3.23	2.57	2.02
2	17.15	24.20	17.38	11.72	8.47	6.47	5.07	4.14	3.10	2.31
3	9.76	16.82	18.61	15.30	11.65	8.86	6.81	5.28	4.10	2.80
4	7.10	11.16	15.07	16.24	14.01	11.48	9.07	6.94	5.28	3.67
5	5.68	8.09	11.39	11.93	14.94	13.83	11.27	9.06	7.09	4.71
6	4.77	6.40	8.62	11.16	13.49	14.76	13.85	11.69	9.13	6.12
7	4.09	5.19	6.71	8.66	11.35	13.66	15.27	14.45	12.13	8.50
8	3.50	4.22	5.21	6.91	8.87	11.41	14.50	16.85	16.27	12.26
9	2.81	3.29	4.12	5.30	6.91	8.99	11.97	16.18	20.96	19.46
High	2.17	2.30	2.89	3.59	4.64	6.17	8.44	12.25	19.48	38.06

Table 12: Investor disagreement premium: formation periods. At the end of each month, stocks are sorted into deciles based on investor disagreement (ID) and assigned into portfolios. Stocks are then held in the portfolio for the subsequent month. ID at the end of a given month is computed as the contemporaneous correlation coefficient of daily trading volume and absolute price change over the past T months, multiplied by -1. Portfolio returns are equal-weighted. The table presents the difference in excess returns, CAPM alpha, three-factor alpha (MKT, SMB, and HML), four-factor alpha (MKT, SMB, HML, and UMD), and five-factor alpha (MKT, SMB, HML, UMD, and LIQ) between the highest and the lowest ID decile. The sample period is from January 1983 to December 2019. Newey and West (1987) t-statistics are reported in parentheses.

Formation period (in months)	Excess return	CAPM alpha	3-factor alpha	4-factor alpha	5-factor alpha
3	0.58	0.82	0.66	0.70	0.72
	(3.44)	(5.29)	(5.10)	(5.57)	(5.85)
4	0.50	0.77	0.59	0.59	0.61
	(2.91)	(4.67)	(4.45)	(4.81)	(5.01)
5	0.56	0.84	0.66	0.65	0.67
	(3.17)	(5.15)	(5.07)	(5.35)	(5.58)
6	0.50	0.80	0.61	0.58	0.60
	(2.73)	(4.60)	(4.42)	(4.52)	(4.73)
7	0.47	0.78	0.57	0.54	0.57
	(2.42)	(4.19)	(3.90)	(4.00)	(4.28)
8	0.50	0.82	0.61	0.58	0.61
	(2.56)	(4.41)	(4.30)	(4.37)	(4.70)
9	0.52	0.85	0.62	0.60	0.63
	(2.57)	(4.42)	(4.35)	(4.46)	(4.81)
10	0.50	0.83	0.60	0.54	0.57
	(2.50)	(4.41)	(4.26)	(3.81)	(4.21)
11	0.49	0.84	0.60	0.53	0.56
	(2.41)	(4.40)	(4.11)	(3.43)	(3.89)
12	0.46	0.81	0.57	0.48	0.52
	(2.20)	(4.21)	(3.86)	(3.19)	(3.59)

Table 13: Investor disagreement premium: holding periods. At the end of each month, stocks are sorted into deciles based on investor disagreement (ID) and assigned into portfolios. Stocks are then held in the portfolio for T months, with $\frac{1}{T}$ of each portfolio reinvested monthly. ID at the end of a given month is computed as the contemporaneous correlation coefficient of daily trading volume and absolute price change over the past 2 months, multiplied by -1. Portfolio returns are equal-weighted. The table presents the difference in excess returns, CAPM alpha, three-factor alpha (MKT, SMB, and HML), four-factor alpha (MKT, SMB, HML, and UMD), and five-factor alpha (MKT, SMB, HML, UMD, and LIQ) between the highest and the lowest ID decile. The sample period is from January 1983 to December 2019. Newey and West (1987) t-statistics are reported in parentheses.

Holding period (in months)	Excess return	CAPM alpha	3-factor alpha	4-factor alpha	5-factor alpha
2	0.57	0.80	0.65	0.66	0.68
	(3.85)	(5.82)	(5.80)	(6.20)	(6.57)
3	0.48	0.71	0.57	0.55	0.58
	(3.42)	(5.49)	(5.64)	(5.62)	(6.05)
4	0.45	0.69	0.55	0.52	0.55
	(3.33)	(5.57)	(5.76)	(5.4)	(5.85)
5	0.44	0.68	0.54	0.51	0.53
	(3.27)	(5.57)	(5.69)	(5.13)	(5.57)
6	0.49	0.65	0.51	0.47	0.49
	(5.21)	(5.34)	(5.4)	(4.80)	(5.21)
7	0.38	0.63	0.48	0.44	0.46
	(2.83)	(5.07)	(5.11)	(4.51)	(4.91)
8	0.37	0.61	0.47	0.42	0.43
	(2.70)	(4.92)	(4.96)	(4.25)	(4.64)
9	0.34	0.59	0.44	0.38	0.40
	(2.50)	(4.66)	(4.65)	(3.80)	(4.18)
10	0.34	0.58	0.43	0.37	0.39
	(2.46)	(4.60)	(4.59)	(3.61)	(3.99)
11	0.33	0.58	0.44	0.36	0.38
	(2.42)	(4.57)	(4.55)	(3.47)	(3.84)
12	0.33	0.58	0.43	0.35	0.37
	(2.4)	(4.55)	(4.52)	(3.32)	(3.68)

Table 14: Average cumulative abnormal return (CAR) around earnings announcement by investor disagreement (ID). The table presents time-series average of quarterly mean values of cumulative market-adjusted returns (CAR) within investor disagreement (ID) deciles. The weights are based on the number of observations in each portfolio in each calendar quarter. CAR_{-t_1,t_1} is defined as the compounded return over the $[-t_1, t_1]$ window around the earnings announcement date ($t = 0$) in excess of the compounded value-weighted market return (in percent). The corresponding reference period is defined as the 44-day $[t_1 - 44, t_1 - 1]$ window prior to the earnings announcement date. In each calendar quarter, stocks are sorted into deciles by ID, which is defined as the contemporaneous correlation coefficient of volume and absolute price change in the reference period, multiplied by -1. The row labeled "$10 - 1$" reports the difference in CAR between decile 10 (High ID) and decile 1 (Low ID). The sample period is from the first quarter of 1983 to the fourth quarter of 2019 (444 quarters). Newey and West (1987) t-statistics are reported in parentheses.

ID deciles	$CAR_{-1,1}$ (1)	$CAR_{-2,2}$ (2)	$CAR_{-3,3}$ (3)	$CAR_{-4,4}$ (4)	$CAR_{-5,5}$ (5)
1 (Low)	0.12	-0.54	-0.57	-0.51	-0.51
	(0.45)	(-3.99)	(-8.02)	(-7.12)	(-7.07)
2	0.27	-0.39	-0.44	-0.39	-0.40
	(1.06)	(-3.36)	(-8.24)	(-7.49)	(-7.62)
3	0.37	-0.30	-0.37	-0.29	-0.29
	(1.43)	(-2.87)	(-7.59)	(-6.03)	(-6.13)
4	0.42	-0.23	-0.31	-0.23	-0.22
	(1.65)	(-2.16)	(-6.05)	(-4.87)	(-4.72)
5	0.53	-0.10	-0.19	-0.13	-0.12
	(2.22)	(-1.03)	(-4.06)	(-2.58)	(-2.46)
6	0.52	-0.10	-0.21	-0.11	-0.11
	(2.19)	(-1.01)	(-4.09)	(-2.07)	(-2.01)
7	0.66	0.06	-0.07	0.04	0.05
	(2.84)	(0.62)	(-1.20)	(0.66)	(0.78)
8	0.66	0.08	-0.06	0.07	0.07
	(2.95)	(0.82)	(-0.93)	(1.04)	(1.13)
9	0.69	0.13	-0.02	0.10	0.10
	(3.09)	(1.20)	(-0.25)	(1.29)	(1.41)
10 (High)	0.76	0.25	0.09	0.21	0.22
	(3.56)	(2.15)	(1.05)	(2.53)	(2.72)
$10 - 1$	0.65	0.79	0.65	0.72	0.73
	(4.77)	(6.18)	(5.90)	(6.88)	(7.10)

Table 15. Investor disagreement and earnings announcement returns. The table presents results of quarterly weighted Fama and MacBeth (1973) regressions using cumulative abnormal returns around the earnings announcement date, CAR_{-t_1,t_1}, as the dependent variable. The weights correspond to the number of observations used in each quarterly cross-sectional regression. CAR_{-t_1,t_1} is defined as the compounded return over the $[-t_1, t_1]$ window around the earnings announcement date ($t = 0$) in excess of the compounded value-weighted market return (in percent). The corresponding reference period defined as the 44-day $[t_1-44, t_1-1]$ window prior to the earnings announcement date. ID (investor disagreement) is defined as the contemporaneous correlation coefficient of volume and absolute price change in the reference period, multiplied by -1. SIZE is the log of market capitalization in millions of dollars and BM is book-to-market ratio. RET is the return (in percent) compounded over the reference period. TURN and IVOL are the average turnover ratio and idiosyncratic volatility in the reference period, respectively. IOR is the ratio of shares owned by institutions as reported in 13F filings in the last quarter. NUMEST is the number of unique analysts that have eligible fiscal year one earnings estimates on IBES in the reference period. The sample period is from the first quarter of 1983 to the fourth quarter of 2019. Newey and West (1987) t-statistics are reported in parentheses.

	$CAR_{-1,1}$ (1)	$CAR_{-2,2}$ (2)	$CAR_{-3,3}$ (3)	$CAR_{-4,4}$ (4)	$CAR_{-5,5}$ (5)
ID	0.337	0.379	0.431	0.493	0.539
	(3.96)	(4.33)	(4.18)	(4.35)	(4.58)
SIZE	0.031	0.04	0.058	0.061	0.064
	(1.52)	(1.83)	(2.35)	(2.28)	(2.18)
BM	0.189	0.198	0.220	0.198	0.189
	(4.46)	(3.78)	(3.78)	(2.80)	(2.13)
RET	-0.013	-0.015	-0.017	-0.019	-0.022
	(-7.84)	(-7.21)	(-6.68)	(-6.64)	(-6.66)
TURN	-9.405	-7.964	-6.324	-5.409	1.008
	(-1.75)	(-1.20)	(-0.78)	(-0.56)	(0.09)
IVOL	-0.121	-0.158	-0.175	-0.186	-0.212
	(-4.44)	(-5.16)	(-4.77)	(-4.43)	(-4.79)
IOR	-0.650	-0.710	-0.779	-0.827	-0.904
	(-6.57)	(-6.48)	(-5.99)	(-5.79)	(-5.84)
NUMEST	0.011	0.012	0.012	0.015	0.016
	(2.68)	(2.55)	(2.29)	(2.6)	(2.59)
Intercept	0.331	0.352	0.274	0.279	0.323
	(2.52)	(2.55)	(1.54)	(1.31)	(1.39)
Adj. R^2	0.87%	0.98%	1.34%	1.64%	1.93%
# of observations	428,194	428,003	427,892	427,764	427,648

Table 16. Investor disagreement: good and bad earnings. In each calendar quarter, stocks are first sorted into quintiles based on firm size (SIZE). Next, within each SIZE decile, stocks are further sorted into good earnings ($CAR_{-1,1} > 0$) and bad earnings ($CAR_{-1,1} \leq 0$) portfolios. The table presents the time-seires of quarterly mean values of ΔID in each portfolio. ΔID is defined as $\Delta\text{ID} = \text{ID}_{\text{after}} - \text{ID}_{\text{before}}$. ID_{after} and $\text{ID}_{\text{before}}$ are defined as the contemporaneous correlation coefficient of daily trading volume and absolute price change over the window $[2, 45]$ and $[-45, -2]$, multiplied by -1, where $t = 0$ is the earnings announcement date. $CAR_{-1,1}$ is defined as the compounded return over the $[-1, 1]$ window around the earnings announcement date ($t = 0$) in excess of the compounded value-weighted market return (in percent).There are 413,454 stock-quarter observations. The sample period is from the first quarter of 1983 to the fourth quarter of 2019 (444 quarters). Newey and West (1987) t-statistics are reported in parentheses. All coefficients are multiplied by 100.

	Good news ($CAR > 0$)	Bad news ($CAR \leq 0$)	Bad-Good
Small	0.02	1.58	1.56
	(0.07)	(8.55)	(6.76)
2	0.18	1.53	1.35
	(0.82)	(6.44)	(6.03)
3	0.34	1.03	0.69
	(1.69)	(5.09)	(3.42)
4	0.51	0.78	0.27
	(2.36)	(3.27)	(1.17)
Large	0.57	0.30	-0.27
	(2.36)	(1.27)	(-1.46)
All	0.32	1.07	0.75
	(2.08)	(6.59)	(7.42)

Table 17. Investor disagreement: good and bad earnings announcement news. In each calendar quarter, stocks are first sorted into quintiles based on firm size (SIZE). Next, within each SIZE decile, stocks are further sorted into good earnings portfolio($SUE > 0$) and bad earnings portfolio ($SUE \leq 0$) in Panel A and into good earnings portfolio ($SUEAF > 0$) and bad earnings portfolio ($SUEAF \leq 0$) in Panel B. The table presents the time-series of quarterly mean values of ΔID in each portfolio. ID_{after} and $\text{ID}_{\text{before}}$ are defined as the contemporaneous correlation coefficient of daily trading volume and absolute price change over the window $[2, 45]$ and $[-45, -2]$, multiplied by -1, where $t = 0$ is the earnings announcement date. *SUE* is standardized unexpected earnings defined as the difference in EPS before extraordinary items between quarters q and $q-4$ divided by the $q-4$ quarter-end price. *SUEAF* is the difference between the median analyst forecast over the 90-day period before the announcement and actual earnings divided by the $q-4$ quarter-end price. There are 382,724 and 258,544 stock-quarter observations for *SUE* and *SUEAF* sample, respectively. The sample period is from the first quarter of 1983 to the fourth quarter of 2019 (444 quarters). Newey and West (1987) t-statistics are reported in parentheses. All coefficients are multiplied by 100.

	Panel A: sorted on SIZE and *SUE*		
SIZE quintile	Good news ($SUE > 0$)	Bad news ($SUE \leq 0$)	Bad-Good
Small	0.34	1.35	1.01
	(1.52)	(6.41)	(4.49)
2	0.23	1.58	1.35
	(1.04)	(6.56)	(6.63)
3	-0.01	1.70	1.71
	(-0.05)	(7.85)	(8.83)
4	-0.05	1.61	1.66
	(-0.25)	(5.89)	(6.96)
Large	0.04	1.12	1.08
	(0.17)	(4.29)	(4.47)
All	0.10	1.47	1.38
	(0.61)	(8.54)	(11.30)

	Panel B: sorted on SIZE and *SUEAF*		
SIZE quintile	Good news ($SUEAF > 0$)	Bad news ($SUEAF \leq 0$)	Bad-Good
Small	-0.12	1.44	1.56
	(-0.21)	(4.07)	(3.02)
2	0.38	1.52	1.14
	(1.36)	(4.88)	(2.90)
3	0.64	1.57	0.93
	(2.26)	(7.04)	(3.16)
4	0.32	0.89	0.57
	(1.42)	(2.61)	(2.11)
Large	0.31	0.67	0.36
	(1.17)	(2.46)	(1.38)
All	0.41	1.12	0.71
	(2.27)	(5.47)	(4.74)

Table 18. Change in investor disagreement: bad earnings announcement. The table reports stock-level cross-sectional regression of of ΔID on 3 bad earnings announcement indicator variable ($1_{CAR_{-1,1}\leq 0}$, $1_{SUE\leq 0}$, and $1_{SUEAF\leq 0}$) with a variety of control variables. ΔID is defined as $\Delta\text{ID} = \text{ID}_{\text{after}} - \text{ID}_{\text{before}}$. ID_{after} (ID_{before}) is defined as the contemporaneous correlation coefficient of daily trading volume and absolute price change over the window $[2, 45]$ and $[-45, -2]$, multiplied by -1, where $t = 0$ is the earnings announcement date. $1_{CAR_{-1,1}\leq 0}$ is an indicator variable that equals to one if $CAR_{-1,1} \leq 0$. $1_{SUE\leq 0}$ is an indicator variable that equals to one if $SUE \leq 0$. $1_{SUEAF\leq 0}$ is an indicator variable that equals to one if $SUEAF \leq 0$. $CAR_{-1,1}$ is defined as the compounded return over the $[-1, 1]$ window around the earnings announcement date ($t = 0$) in excess of the compounded value-weighted market return (in percent). SUE is standardized unexpected earnings defined as the difference in EPS before extraordinary items between quarters q and $q - 4$ divided by the $q - 4$ quarter-end price. $SUEAF$ is the difference between the median analyst forecast over the 90-day period before the announcement and actual earnings divided by the $q - 4$ quarter-end price. SIZE is the log of market capitalization in millions of dollars and BM is book-to-market ratio. RET is the return (in percent) compounded over the period $[-45, -2]$. TURN and IVOL are the average turnover ratio and idiosyncratic volatility in the period $[-45, -2]$, respectively. IOR is the ratio of shares owned by institutions as reported in 13F filings in the last quarter. NUMEST is the number of unique analysts that have eligible fiscal year one earnings estimates on IBES in the period $[-45, -2]$. The sample period is from the first quarter of 1983 to the fourth quarter of 2019. Newey and West (1987) t-statistics are reported in parentheses.

	(1)	(2)	(3)
$1_{CAR_{-1,1}\leq 0}$	0.007		
	(6.51)		
$1_{SUE\leq 0}$		0.011	
		(9.28)	
$1_{SUEAF\leq 0}$			0.006
			(4.73)
SIZE	0.018	0.018	0.017
	(9.85)	(9.55)	(6.56)
BEME	-0.003	-0.003	-0.003
	(-1.94)	(-1.87)	(-1.49)
RET	0.001	0.001	0.0004
	(7.38)	(7.45)	(5.16)
TURN	-0.335	-0.465	-1.518
	(-2.10)	(-2.63)	(-6.36)
IVOL	0.047	0.05	0.073
	(20.84)	(19.99)	(20.29)
IOR	-0.011	-0.014	-0.008
	(-2.65)	(-3.31)	(-1.70)
NUMEST	0.001	0.001	0.001
	(3.11)	(2.85)	(3.88)
Intercept	-0.203	-0.214	-0.251
	(-15.65)	(-15.15)	(-12.10)
Firm FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Adj. R^2	3.50%	3.69%	5.04%
# of observations	412,620	382,091	257,699