Option-Implied Firm-Level Sentiment and Stock Returns[1]

Gady Jacoby
School of Economics
College of Management Academic Studies
Email: gady.jacoby@colman.ac.il

Fuwei Jiang
School of Finance
Central University of Finance and Economics
Email: jfuwei@gmail.com

Lei Lu
Asper School of Business
University of Manitoba
Email: lu.lei@umanitoba.ca

Qi Zhang
Asper School of Business
University of Manitoba
Email: zhangq22@myumanitoba.ca

Abstract

We propose a novel firm-specific investor sentiment measure—the daily change in the open interest weighted implied volatility ratio of out-of-the-money (OTM) calls over that of OTM puts. To validate this ratio as a sentiment measure, we show that, at the aggregate, our firm-level sentiment measure is highly correlated with existing market-level sentiment indices. We find that an increase in our aggregate market sentiment measure predicts a short-term stock return increase with a subsequent reversal. We then proceed to examine how our firm-level sentiment measure affects returns in the cross section. A long-short portfolio strategy, based on a long position in the high-sentiment portfolio and a short position in the low-sentiment portfolio, generates a significant abnormal return of 70 bps per month (8.73% annualized). This effect of sentiment on stock returns is more pronounced for hard-to-value stocks, which are small, young, high-volatility, and less-liquid stocks. Finally, we apply a Fama-MacBeth regression at the stock level to show that a higher sentiment in the current period predicts a higher return next period.

***Keywords*:** Firm-specific Investor Sentiment, Options, Return Predictability, Asset Pricing

***JEL Classification*:** G10; G12; G13

---

[1] We are grateful for valuable comments from comments from Shiu-Yik Au, Frank Chen, Zhenzhen Fan, Alex Paseka, Tianyang Wang, Tong Yu, S M Zahid, Hong Zou, and seminar participants from FMA 2021 Annual Meeting, University of Manitoba, Central University of Finance and Economics, MRS PhD student seminar.

1. Introduction

Retail investors are increasingly trading options for speculation (Choy and Wei, 2012; Lakonishok, Lee, Pearson, and Poteshman, 2007), resulting in a platform where investor sentiment plays a significant role. In this paper, we contribute to the literature on market-based sentiment measures by proposing a novel, option-based firm-specific investor sentiment measure, and by examining its effect on cross-sectional stock returns. Current market-based sentiment measures include the Baker and Wurgler aggregate-level sentiment index and overnight return (Baker and Wurgler, 2006; Aboody, Even-Tov, Lehavy, and Trueman, 2018). Growing studies use Baker and Wurgler index to examine firm-level issues, such as stock price reactions to earnings and analyst forecast accuracy (Livnat and Petrovits, 2019; Hribar and Mclnnis, 2012; Mian and Sankaraguruswamy, 2012). However, an aggregate-level sentiment measure may ignore important cross-sectional variations, causing problems in examining a firm's announcement events and investment decision making regarding a specific firm (Baker and Wurgler, 2006; Brown and Cliff, 2004; Aboody et al., 2018).

Considering the above limitations of aggregate level investor sentiment, Aboody et al. (2018) show that a stock's overnight return is a qualified firm-level sentiment measure. However, the suitability of overnight return as a firm-level sentiment proxy is challenged by the mixed effect in international equity markets (Xiong, Meng, Li, and Shen, 2020). Considering the above limitations of overnight return as a proxy for firm-level investor sentiment, we propose a novel, option-based firm-specific investor sentiment measure: the change in the open interest-weighted implied volatility ratio of out-of-the-money (OTM) calls over puts.

This differs from the Baker and Wurgler sentiment index and overnight return proxy that is based on historical equity trading data. Our market-based firm-level sentiment is a forward-looking measure using the options market. Implied volatility is a direct and forward-looking measure of what the market believes about the underlying securities price movement. Current literature on option price discovery supports a sentiment measure that is a forward-looking measure (Chakravarty, Gulen, and Mayhew, 2004; Collin-Dufrsne, Fos, and Muravyev, 2021). Another advantage of our option-based sentiment measure is its high frequency, which has significant practical implications for fund managers' market timing. This sentiment proxy can be constructed at the hourly level using tick-by-tick data, compared to the monthly Baker and Wurgler sentiment index and weekly overnight return proxy.

Baker and Wurgler (2006) suggest that investor sentiment can be defined as "optimism or pessimism about stocks in general." Optimistic investors, who expect a dramatic increase in the stock price, choose to long in out-of-the-money (OTM) call options to maximize their profit potential, while pessimistic investors long in OTM put options (Buraschi and Jiltsov, 2006; Xing, Zhang, and Zhao, 2010). The net buying pressure of OTM options causes an increase in the option's implied volatility (Bollen and Whaley, 2004; Garleanu, Pedersen, and Poteshman, 2008). Therefore, optimistic (pessimistic) opinion is given by the increase of OTM calls (puts) implied volatility. As a result, an equity option's relative implied volatility ratio of OTM calls over puts captures a representative investor's optimistic sentiment relative to pessimistic sentiment. The implied volatility ratio is highly persistent, so we measure investor sentiment $Sent^{IVR}$ as the change in the open interest-weighted implied volatility ratio of OTM calls over puts.

Our intuition is consistent with Buraschi and Jiltsov (2006), who suggest that OTM calls (puts) are traded for bullish (bearish) expectations. Our idea is also supported by the demand-based option pricing model (Garleanu et al., 2008), which shows that an option's implied volatility is affected by its demand pressure. In addition, the plausibility of our sentiment measure is built on the fact that unsophisticated investors actively participate in equity options trading for speculative purposes (Han, 2008; Choy and Wei, 2012; Lakonishok et al., 2007). Instead of trading in the stock market, optimists prefer to invest in call options because of their limited downside loss and the benefits of leveraged gains.

Using option data from January 1996 to June 2019, we construct a firm-specific investor sentiment measure. We first find that our sentiment measure, consistent with our intuition, reflects investors' bullish beliefs about the underlying stock's future movement. Our aggregated sentiment measure has a significant correlation with conventional market-wide sentiment indexes while aggregating the firm-specific implied volatility ratio at a monthly frequency. We also aggregate firm-level investor sentiment $Sent^{IVR}$ at a daily frequency and empirically show that an increase in our aggregate sentiment measure predicts a short-term stock return increase, followed by a reversal. This reflects that fact that an increase of our daily aggregated sentiment measure causes the mispricing of S&P 500 index returns with a subsequent correction.

We then proceed to examine how our sentiment measure affects returns in the cross section. Options markets are forward-looking, so our firm-specific investor sentiment could positively predict future stock returns at the monthly frequency. In addition, the abnormal return of being long in the optimistic portfolio and short in the pessimistic portfolio is 70 bps per month (8.73% annualized). This effect is robust when we control for firm-level and option-level variables.

We also empirically support the predictions from Baker and Wurgler (2006) that investor sentiment has a more pronounced effect on subjectively hard-to-value stocks. Sentiment-driven investors prefer to trade hard-to-value stocks to suit their sentiment (optimistic or pessimistic) because these stocks are apparently speculative and have virtually unlimited growth opportunities. Using size, age, idiosyncratic volatility, and Amihud's illiquidity to proxy for hard-to-valuate, we find that the predictability effect of investor sentiment on stock return is more significant and pronounced for small, young, volatile, and illiquid stocks. Lastly, we find that a higher sentiment leads to higher volatility, which is consistent with the intuition that sentiment pushes stock price away from its fundamentals.

This paper contributes to the sentiment literature by proposing a novel, market-based, firm-specific sentiment measure directly from the option market. The advantage of our firm-specific investor sentiment measure is that it is a direct measure of the market's beliefs about the underlying stock price movement. Another advantage is that we can aggregate at any frequency or in any industry to represent the industry-specific sentiment measure. Aboody et al. (2018) empirically examine the qualification of a stock's overnight return as a firm-specific investor sentiment measure. Coqueret (2020) examines the predictability of Bloomberg news-based sentiment on a stock's daily return.

This paper also contributes to the literature that examines sentiment in options markets. Han (2008) examines the relationship between the slope of the index option's implied volatility function and the aggregate sentiment. Choy and Wei (2012) empirically demonstrate that the main driver of options trading is investor opinion. This paper extends their findings and explores the measure of stock-specific sentiment from equity options markets.

The remainder of this paper is organized as follows: Section 2 describes the measure of our firm-specific investor sentiment while Section 3 describes the data used in this study. Section 4 reports the cross-sectional level and aggregate level relationship between our sentiment and returns. Section 5 provides additional empirical analysis and Section 6 presents some brief conclusions.

2. Measuring Investor Sentiment

In this section, we describe current investor sentiment proxies before intuitively introducing our option-based, firm-specific investor sentiment metric. We then describe a theoretical framework to claim that this option-based investor sentiment measure captures investor sentiment variation in a formal approach.

2.1 Conventional Investor Sentiment Measures

Investor sentiment can be defined as "optimism or pessimism about stocks" in general (Baker and Wurgler, 2006). When investors hold bullish (bearish) sentiments, it may be a rational reflection of the upcoming prosperity (recession), or an irrational hope (fear) for the future, or some combination of both (Brown and Cliff, 2005). There are three approaches to measuring investor sentiment. The first approach is very intuitive and relies on surveys conducted by institutions or universities. For example, well-designed survey-based sentiment indexes include the American Association of Individual Investors survey (S^{AAII}), the Investors' Intelligence index (S^{II}), and the University of Michigan Consumer Sentiment Index (S^{MCSI}). While directly measuring subjects' bearish or bullish expectations, survey-based sentiment indexes are at a disadvantage because of their low scope and frequency, which results from the high cost and time-consuming nature of data collection. Survey-based sentiment measures are, at best, available only weekly, and they represent beliefs from only a segment of market participants.

The second approach measures investor sentiment by directly analyzing the views of investors from markets, including the equity market, the mutual fund market, and the derivatives market. Examples include the PCA-based Baker and Wurgler investor sentiment index (Baker and Wurgler, 2006), the PLS-based sentiment index (Huang, Jiang, Tu, and Zhou, 2015), the closed-end mutual fund discount (Lee, Shleifer, and Thaler, 1991), mutual fund flow (Ben-Rephael, Kandel, and Wohl, 2012), and the put-call volume ratio.

The third approach utilizes “big data” techniques to measure sentiment by analyzing the context of news, searching websites, and analyzing social media content (see Tetlock, 2007; Garcia, 2013; Engelberg and Parsons, 2011; Das and Chen, 2007; Da, Engelberg, and Gao, 2015, among others). For example, Tetlock (2007) measures sentiment by analyzing the context of the “Abreast of the Market” column reports in the *Wall Street Journal* and Da, Engelberg, and Gao (2015) construct a daily FEARS index using the Google search volume index.

2.2 Measuring Firm-level Investor Sentiment

Earlier researchers and institutions mainly constructed investor sentiment measures from the aggregate level, and paid less attention to firm-level measures (which should be preferred to examine firm-level issues and investment decisions regarding a specific firm). Using the aggregate-level investor sentiment measure for investment and fund management may cause problems because there are cross-sectional variations among stock-specific investor sentiment. Current firm-specific sentiment metrics include overnight return (Aboody et al., 2018), Bloomberg news sentiment (Coqueret, 2020), and Twitter sentiment.

Our paper contributes to this field of research by proposing a forward-looking, market-based, stock-level sentiment measure: the change in open interest-weighted implied volatility ratio of OTM calls over puts. Compared with overnight return and text-based firm-specific sentiment measures, our measure shows advantages because the implied volatility is a direct and forward-looking measure of what the market believes regarding the underlying equity returns.

The options market provides a suitable context to investigate both optimistic and pessimistic beliefs about stocks. Optimists—who expect a quick and dramatic rise in the underlying stock beyond a certain price—could maximize profit potential by buying *call* options at a strike price that is likely to be below the future price of the stock. By contrast, pessimists tend to choose out-of-the-money (OTM) *put* options when they expect a decline in the price of a stock (Buraschi and Jiltsov, 2006; Xing et al., 2010). Net buying pressure of OTM calls (puts) causes an increase in the option's implied volatility (Bollen and Whaley, 2004; Garleanu et al., 2008). Therefore, optimistic (pessimistic) opinion is given by the implied volatility of OTM calls (puts). As a result, an equity option's relative implied volatility of OTM calls over puts captures representative investor optimistic sentiment relative to their pessimistic sentiment.

There are three advantages of using implied volatility to measure investor sentiment. First, implied volatility is a direct measure of the market's beliefs about the movement of underlying stock prices. An increase in the implied volatility of calls (puts) directly demonstrates that the market holds an optimistic (pessimistic) belief about the underlying stock. Second, implied volatility is a forward-looking variable that indicates investors' real-time assessments regarding how likely the price of underlying securities is expected to move. For example, a 30-day out-of-the-money (OTM) option's implied volatility right now measures the market's current positive expectations about the underlying stock's price movement over the next 30 days. Third, there are benefits from implied volatility's forward-looking characteristics and real-time updates because the investor sentiment measure can be constructed at a very high frequency (even hourly) to express the market's beliefs over the next several months. This has significant implications for fund managers and investors for timing their investments.

There are two reasons why we use only information from out-of-the-money (OTM) options to construct our investor sentiment measure. First, OTM calls are traded for bullish expectations while OTM puts are traded for bearish expectations (Buraschi and Jiltsov, 2006). Optimistic (pessimistic) investors who are long on OTM call (put) options and expect their positions to be in-the-money (ITM) eventually, are inherently more optimistic (pessimistic) than those who are long on ITM or at-the-money (ATM) calls (puts).

Second, a narrow definition of investor sentiment suggests that investor sentiment is caused by noise and the trading of unsophisticated investors, which pushes the price away from fundamentals. While there is disagreement on the existence of informed trading in the options market (Choy and Wei, 2012), the authors supporting informed trading clearly show that ITM options (Johnson and So, 2012; Chung, Ryu et al., 2016) and near-the-money options (Chakravarty, Gulen, and Mayhew, 2004; Ge, Lin, and Pearson, 2016) are favorable for informed investors. Hence, informed investors try to avoid OTM options in general. The role of informed investors in OTM options is less significant than that of uninformed investors (who are more likely to trade based on sentiment). Thus, OTM call (put) options are used in this paper to capture the severity of optimistic (pessimistic) beliefs and eliminate the influence of informed investors.

We first calculate the open interest-weighted sum of implied volatility for both OTM calls and puts separately across moneyness and maturities at each day. Then we calculate the implied volatility ratio *IVR* as the fraction of the weighted average of the implied volatility of calls to puts, where an option's open interest is the weight. The daily implied volatility ratio *IVR* is thus calculated as:

$$IVR_{i,t} = \frac{\sum_{K,\tau} w_{i,t,K,\tau} * IV(OTMC)_{i,t,K,\tau}}{\sum_{K,\tau} w_{i,t,K,\tau} * IV(OTMP)_{i,t,K,\tau}} \quad (1)$$

where $w_{i,t,K,\tau} = \frac{open\ interest_{i,t,K,\tau}}{\sum_{K,T} open\ interest_{i,t,K,\tau}}$ is the weight, *i* represents the specific firm, *K* represents the various strike price, *τ* refers to maturity, *t* is the date of option trading, *IV* is the implied volatility, *OTMC (OTMP)* refers to OTM call (put) options. When we choose the weight, volume and open interest are the two candidates. When observing the data, we find that some option contracts only have the open interest, but no volume traded. If we use volume to calculate weights, the measure would eliminate the information contained in these options. Thus, in this paper, we mainly focus on the open interest-weighted call/put ratio, since it captures all available information.

Empirical analyses have shown that the implied volatility ratio is persistent (with a daily autocorrelation coefficient of 0.64). We therefore measure monthly investor sentiment as the changes in the implied volatility ratio:

$$Sent_{i,t}^{IVR} = IVR_{i,t} - IVR_{i,t-1} \tag{2}$$

The reason we use the first difference is that it is the simplest approach to cope with the persistence issue. In the robustness check, we find robust results when we calculate investor sentiment as the percentage change, AR (1) innovation, and percentage AR (1) innovation.

The foundation of constructing investor sentiment measures from options markets is that the main purpose of options trading is speculation, rather than hedging or arbitrage. This has been supported by recent literature. Lakonishok et al. (2007) suggest that betting on volatility accounts for a small percentage of options trading and speculation is the main driver of trading. Choy and Wei (2012) also support this argument by finding speculation during the earnings announcement periods.

A subsequent question raised is this: whose sentiment does this measure capture? Recent literature suggests that unsophisticated investors actively participate in equity options trading for speculative purposes (Han, 2008; Choy and Wei, 2012). Han (2008) suggests that equity options are actively utilized by individual investors for speculation purposes. Lemmon and Ni (2014) state that unsophisticated individual investors account for more than 30% of equity options, while around 40% of stock options trading is initiated by full-service (sophisticated) traders. However, Houlihan and Creamer (2019) show that from 2005 to 2012, around 70% of options trading was done by customer traders, who are unsophisticated investors generally. While more recent studies support the idea that unsophisticated investors are the main market participants of the options market, we cannot be certain that our sentiment measure demonstrates retail investors' behaviors. Thus, our sentiment measure captures sentiment from both retail and institutional investors, who have also shown sentiment trading (Chen, Han, and Pan, 2021; DeVault, Sias, and Starks, 2019).

2.3 A Theoretical Framework

2.3.1 The market

We consider a one-period discrete time equilibrium economy in which two investors are assumed in the market: an optimistic investor and a pessimistic investor. The two investors contribute to market-based sentiment concerning individual risky assets. The two investors have identical endowments but differ in their beliefs about the stock return. Both investors make their investment selections between a riskless asset, a stock, and options to maximize their expected utility of terminal wealth. There is a risk-free asset whose gross return, represented as R_f, equals one, there is one stock (indexed as *s*), one call option (indexed as *c*), and one put option (indexed as *p*) in the economy.

We assume that optimistic and pessimistic investors have a perceived return misperception error relative to the stock's fundamental return R_s that is sentiment-free and follows a normal distribution of $R_s \sim N(\mu_s, \sigma_s^2)$. The optimistic and pessimistic investor's sentiment-related misperception error on the stock return are represented by ρ_i and ρ_j, where *i* and *j* index the optimistic and pessimistic investor, respectively. ρ_i and ρ_j are positive and normally distributed with $\rho_i \sim N(\mu_{\rho_i}, \sigma_{\rho_i}^2)$ and $\rho_j \sim N(\mu_{\rho_j}, \sigma_{\rho_j}^2)$. Thus, the optimistic (pessimistic) investor's perceived stock return is given by $R_{s,i} = R_s + \rho_i$ ($R_{s,j} = R_s + \rho_j$).

The optimistic (pessimistic) investor would include one European call (put) option that expires at t_1, in its portfolio at t_0. The strike price of the European call (put) option is denoted as *K*, whose present value equals the stock price $S_0 = Ke^{-r_f}$ by assumption for simplicity. Thus, the call (put) option return is shown by: $R_c = (S_1 - K)^+/C$ ($R_p = (S_1 - K)^+/P$) where *C (P)* denotes the equilibrium price of the call (put) option and S_1 represents the equilibrium stock price at t_1. A risk-neutral and sentiment-free market maker trade with optimistic and pessimistic investors in the options market. To be specific, the market maker shorts (longs) the European call (put) while the optimistic (pessimistic) investor purchases it.

Let W^0 represent the initial wealth of the optimistic investor and the pessimistic investor; W_i^1 and W_j^1 are the optimistic and pessimistic investors' terminal random wealth; $D_{i,x}$ and $D_{j,x}$ are optimistic and pessimistic investors' dollar investments in the asset x, which is indexed by f (the risk-free asset), s (the stock), c (the call), and p (the put). We also assume that both sentiment-affected investors are risk-averse and maximize the expected utility of their terminal wealth with a utility function U, such that U'>0 and U''<0. Stocks are infinitely divisible with fixed quantities, while call and put options are not. The equilibrium of both optimistic and pessimistic investors generates the price for the risky assets.

2.3.2 The investor's problem and equilibrium

Based on above assumptions, the optimistic investor *i* solves the problem as follows:

$$\max_{x_{i,f},\ x_{i,s},\ x_c} \mathrm{E}[\,U(W_i^1)] \quad (3)$$

$$\text{where } W_i^1 = W^0(x_{i,f}R_f + x_{i,s}R_s + x_cR_c) \quad (4)$$

$$\text{subject to } x_{i,f} + x_{i,s} + x_c = 1 \quad (5)$$

where $x_{i,f} = {}^{D_{i,f}}/_{W^0}$, $x_{i,s} = {}^{D_{i,s}}/_{W^0}$, and $x_{i,c} = {}^{D_{i,c}}/_{W^0}$ represents the proportion of wealth that is invested in the risk-free asset, the stock, and the European call option, respectively.

For the pessimistic investor (noted by *j*), its utility function, the wealth accumulation constraint, and the budget constraint can be represented below:

$$\max_{x_{j,f},\ x_{j,s},\ x_p} \mathrm{E}[\mathrm{U}(\,W_j^1)] \quad (6)$$

$$\text{where } W_j^1 = W^0(x_{j,f} * R_f + x_{j,s} * R_s + x_p * R_p) \quad (7)$$

$$\text{subject to } x_{j,f} + x_{j,s} + x_p = 1 \quad (8)$$

where $x_{j,f} = {}^{D_{j,f}}/_{W^0}$, $x_{j,s} = {}^{D_{j,s}}/_{W^0}$, and $x_{j,p} = {}^{D_{j,p}}/_{W^0}$ demonstrates the proportion of wealth that is invested in the risk-free asset, the stock, and the European put option, respectively.

In equilibrium, the risk-free asset market, the stock market, and the option market clear. We can obtain four market clearing conditions: $D_{i,f} + D_{j,f} = 0$, $D_{i,s} + D_{j,s} = S_0$, $D_{i,c} = C$, and $D_{j,p} = P$.

2.3.3 Analysis

When put-call parity holds and stock price equals the present value of the exercise price ($S_0 = Ke^{-r_f}$), the call price should be equal to the put price (C=P) under the assumption of no investor sentiment. However, we expect that the optimistic (pessimistic) investor increases its demand for the call (put) option when investor sentiment increases (decreases). Thus, we show that the ratio of the call price to the put price will be greater (less) than 1 when investors are more optimistic (pessimistic).

Corrado and Miller (1996) derive an approximation of the implied volatility from the Black-Scholes model shows that:

$$\sigma_C\sqrt{T} \approx \frac{\sqrt{2\pi}}{S+Ke^{-r\tau}}\left(C - \frac{S-Ke^{-r\tau}}{2} + \sqrt{\left(C - \frac{1}{2}(S - Ke^{-r\tau})\right)^2 - \frac{(S-Ke^{-r\tau})^2}{\pi}}\right) \tag{9}$$

where σ_c represents the implied volatility of call options, *T* refers to the maturity, *S* is the stock price, *K* refers to the strike price, *r* measures the risk-free rate, and *C* denotes the price of the call option. When the stock price equals the present value of the strike price, the above equation can be simplified to $\sigma_C\sqrt{T} = \sqrt{\frac{\pi}{2}}\frac{C}{S}$. For a put option with the same strike price, we obtain $\sigma_P\sqrt{T} = \sqrt{\frac{\pi}{2}}\frac{P}{S}$.

The approximation of implied volatility suggests that the option price (*C* or *P*) is a function of implied volatility. The ratio of *C* over *P* equals the ratio of the implied volatility of calls to the implied volatility of puts.

$$\frac{C}{P} = \frac{\sigma_C}{\sigma_P} = \frac{IV(C)}{IV(P)} \tag{10}$$

From our theoretical framework, we derive the ratio of call price to put price as the ratio of call demand to put demand.

$$\frac{C}{P} = \frac{D_{i,c}}{D_{j,p}} \tag{11}$$

Equations (10) and (11) suggest that the implied volatility of the call over the put should equal the demand of the call over the put.

$$\frac{IV(C)}{IV(P)} = \frac{D_{i,c}}{D_{j,p}} \tag{12}$$

Thus, the implied volatility of the call over the put captures the demand of call options over the demand of put options from investors. From our model setup, we observe that when the relative sentiment is higher $(\rho_i > \rho_j)$, optimistic investors will increase their positions on the call option. When sentiment declines, pessimistic investors are more willing to buy the put option, causing the demand for put options to increase. Therefore, the demand for call options over put options captures investor sentiment in the market. In this case, the implied volatility of the call over the put also captures the changes of sentiment in the market.

With this theoretical setup, the put is in-the-money, which means that its price will be higher than the OTM put on the same stock. Therefore, ceteris paribus, the ratio of implied volatility (i.e., our sentiment measure) should be lower for the ITM put relative to the OTM put. Thus, when we use OTM puts, which are less expensive than ITM puts, our sentiment measure will be higher. This is the limitation of our theoretical framework, which considers limited variation of option moneyness. However, this limitation does not actually have much practical significance since earlier studies show that OTM calls are traded for bullish expectations while OTM puts are traded for bearish expectations (Buraschi and Jiltsov, 2006). Using OTM calls and puts is more consistent with practical implications.

3. Data

In this section, we describe the options and other databases we used for empirical analysis, as well as the summary statistics.

3.1 Options Data

The major database in this paper is IvyDB US from OptionMetrics; we use historical options trading for the period January 1996 to June 2019. IvyDB US provides trading price and volume, open interest, implied volatility, and the so-called option greeks (delta, gamma, vega, and theta) for all U.S. exchange-listed options, with various underlying assets, including common stocks, market index, mutual funds, ADR/ADS, and so on. We obtain historical open interest, volume, delta, and implied volatility from the Option Price File. OptionMetrics also provides information about Volatility Surface and Standardized Options, but we use historical trading data from the Option Price File since the historical data provide detailed information about market trading, such as trading volume, open interest, and implied volatility for each option quote. Open interest and the trading volume contain important messages about the distribution of options trading across maturity, moneyness, and option price. Since the open interest data are lagged for one day after November 28, 2011, in the OptionMetrics database, we adjusted for this in our empirical analysis.

For our analysis, we exclude options whose underlying assets are index funds, mutual funds, or ADR/ADS and only use common stock options. In addition, referring to Xing et al., (2010), Han (2008), and Seo and Kim (2015), we require the equity options contracts to satisfy the following criteria to avoid illiquid options: (1) positive volume or open interest; (2) implied volatility must be greater than 0.03 and smaller than 2; (3) the mean of the best bid and best ask must be greater than \$0.125; (4) maturity must be within 30–365 days; and 5) the bid price is non-zero. In addition, we set OTM calls (puts) delta changes from 0.125 to 0.375 (-0.375 to -0.125) (Bollen and Whaley, 2004).

In the empirical analysis, we control for option-level predictors that affect stock returns. Bali and Hovakimian (2009) suggest a negative relationship between the realized-implied volatility spread (RV-IV) and stock returns. Following Bali and Hovakimian (2009), we measure RV-IV as the difference between the monthly realized volatility and the average of the implied volatility of ATM calls and puts with 30-day maturity. Bali and Hovakimian (2009) also find a positive effect of the call/put implied volatility spread (IVC-IVP) on stock returns, and we measure it as the difference between the implied volatility of 30-day matured ATM calls and puts (with a delta of 0.5). Referring to An et al. (2014), we also compute ΔIVC and ΔIVP as changes in the ATM calls and puts (30-day maturity and 0.5 delta) implied volatility, respectively. Put/Call Parity Deviation (PC Dev) is measured as the open interest weighted average difference between pairwise (strike and maturity) call and put implied volatility based on Cremers and Weinbaum (2010). We also measure skewness (Skew) as the difference between the OTM put implied volatility and the average of the ATM call and put implied volatility following Xing, Zhang, and Zhao (2010). Finally, the options to stock volume ratio (O/S) is computed as the ratio of options volume to stock volume (Johnson and So, 2012).

3.2 Additional Data

3.2.1 Firm-level characteristics

We obtain equity returns from CRSP and firm-level accounting data from COMPUSTAT. We merge IvyDB and CRSP using CUSIPs at a monthly frequency. Then, we allocate COMPUSTAT firm-level data of June at fiscal year-end of t+1 to IvyDB data of year t. We next construct variables that have shown an effect on stock returns. Monthly market beta β_{MKT} is calculated as the market beta via the Carhart four-factor model using daily stock return over the month as follows:

$$Ret_{i,d} - Rf_d = \alpha + \beta_{MKT}MKT_d + \beta_{SMB}SMB_d + \beta_{HML}HML_d + \beta_{UMD}UMD_d + \varepsilon_{i,d} \qquad (13)$$

where $Ret_{i,d}$ and Rf_d refer to the daily stock return and risk-free return, and MKT_d, SMB_d, HML_d, and UMD_d are daily market, size, value, and a momentum factor, downloaded from Kenneth French’s website. *Size* is computed as the natural logarithm of market equity, which is the product of stock price and the number of shares outstanding at the end of each month. The book-to-market ratio (*B/M*) is measured as the ratio of a firm’s book value at the end of June to its current month’s market value.

We measure momentum (*Mom*) as the cumulative return over the previous 11 months, following (Jegadeesh and Titman, 1993). Following (Amihud, 2002), illiquidity (*Illiq*) is computed as the ratio of the absolute monthly stock return to dollar volume. We measure volatility (*Vol*) as the standard deviation of monthly returns over the past 12 months and idiosyncratic Volatility (*Idvol*) as the standard deviation of daily residuals ($\varepsilon_{i,d}$) from Carhart's four-factor model. Turnover is computed as the ratio of the number of shares traded in a month to shares outstanding. Profitability (*E/BE*) is calculated as the earnings to book value ratio, and age refers to the years since the equity was first shown in CRSP.

3.2.2 Market-wide sentiment measures

To examine the characteristics of our sentiment measure, we aggregate it at the market level and check its correlation with conventional aggregate investor sentiment indexes. The sentiment index used in this paper includes the Baker and Wurgler investor sentiment index, obtained from Jeffrey Wurgler's website. Baker and Wurgler's investor sentiment index is calculated as the principal components of five indexes: value-weighted dividend premium, first-day returns on IPOs, IPO volume, closed-end fund discount, and equity share in new issues with both lagged variables and current variables.[2] We also download the monthly University of Michigan Consumer Sentiment and the weekly bull-bear spread of the American Association of Individual Investors from the Federal Reserve Bank of St. Louis and the American Association of Individual Investors' official website. We convert the weekly AAII sentiment measure into the monthly index by keeping the end-of-month observations. We obtain the PLS-based investor sentiment index from Dashan Huang's personal website.

3.3 Summary Statistics

[2] This index dropped the NYSE turnover because the authors argue that turnover does not mean what it once did, given the explosion of institutional high-frequency trading and the migration of trading to a variety of venues.

In this subsection, we describe the monthly summary statistics for investor sentiment and stock- and option-level variables in our sample, followed by correlation analysis. We then examine how our aggregated sentiment measure is correlated with conventional market-wide investor sentiment indexes.

Panel A of Table 1 reports that the mean of monthly $Sent^{IVR}$ is around 0 and the median is -0.001, suggesting that the distribution of firm-level monthly investor sentiment measure is quite symmetrical. Panel A also summarizes the statistics for both stock- and option-level characteristics. It suggests that option covered stocks are relatively large and liquid stocks with low book-to-market ratio compared with the whole sample from CRSP in the earlier literature. The option covered stocks show an average size of 7.434, a mean book-to-market ratio of 0.526, and Amihud's illiquidity measure is 0.009. This is consistent with earlier studies regarding the equity options market.

[INSERT TABLE 1 ABOUT HERE]

Panel B reports the summary statistics for portfolios sorted by sentiment $Sent^{IVR}$. Quintile 1 refers to the portfolio with the lowest sentiment measure, and quintile 5 refers to that of the highest sentiment measure. We find that, consistent with Baker and Wurgler (2006), small, illiquid, less profitable, and extreme growth stocks are more subject to sentiment. The results show that stocks in quintile 1 (most pessimistic) and quintile 5 (most optimistic) represent smaller size, less liquidity, higher volatility, lower profitability, and higher book-to-market ratios.

The results in Panel B also suggest a positive relationship between our sentiment measure and the change of the call's implied volatility and call minus put implied volatility spread. This is reasonable because our measures inherently are correlated with these variables. Panel C provides a more formal test of correlation, and it shows that our sentiment measure is positively correlated with the change of the call's implied volatility and call minus put implied volatility at 0.147 and 0.128, respectively. Even though they are significantly and positively correlated, the small correlation coefficients suggest that our measure is different from the call minus put implied volatility spread.

Aggregated firm-level sentiment measures should be highly correlated with conventional market-wide sentiment indexes. The intuition is simple: if these measures both contain the same common factor-investor sentiment, they should move in the same direction. The idiosyncratic components of firm-specific investor sentiment cancel out through aggregating, and we obtain a time-series of market-wide investor sentiment proxy.

To validate our sentiment measure, we calculate the pairwise correlation between our aggregated sentiment measure and conventional market-level investor sentiment measures. We choose four conventional market-wide investor sentiment measures to calculate the following pairwise correlations: Baker and Wurgler's Sentiment index (S^{BW}), the University of Michigan Consumer Sentiment (S^{MCSI}), the bull-bear spread of the American Association of Individual Investors (S^{AAII}), and the PLS-based investor sentiment measure (S^{PLS}) from Huang et al. (2015).

Our aggregate level sentiment is calculated as the cross-sectional average of the firm-level implied volatility ratio at a monthly frequency. There are two reasons for doing this. First, S^{BW}, the most common-used sentiment, is constructed using the principal components of five sentiment measures with both lag and contemporaneous observations; it is therefore not appropriate to calculate a difference measure of sentiment. To be consistent with S^{BW}, we also use level, which is the implied volatility ratio to calculate the correlation. Second, most conventional sentiment indexes are low-frequency indexes, such as S^{BW}, S^{MCSI}, and S^{PLS}. Thus, to be consistent, we aggregate our firm-level investor sentiment at a monthly frequency.

Panel D of Table 1 shows that our aggregated sentiment index is significantly correlated with PCA-based S^{BW} and PLS-based S^{PLS}, with a correlation coefficient of 0.548 and 0.376, respectively. It is also highly correlated with the survey data of S^{MCSI} and S^{AAII}, with a coefficient of 0.730 and 0.404, respectively. Brown and Cliff (2004) show that MCSI and AAII surveys individual investors, thus representing individual investor sentiment. The significantly high correlation supports the argument that our sentiment measure captures retail investor sentiment. Additionally, Da et al.(2015) suggest a search-based investor sentiment, which is a difference measure of investor sentiment; thus, we aggregate $Sent_{i,t}^{IVR}$ cross-sectionally and check the correlation with FEARS. The contemporaneous correlation coefficient is -0.066, which is reasonable because FEARS captures the negative sentiment and should be negatively correlated with our sentiment measure.

4. Return Predictability

In this section, we empirically examine the relationship between our sentiment measure $Sent^{IVR}$ and cross-sectional stock returns using the portfolio and Fama-MacBeth regression approach. We also examine the aggregate level predictability of return reversal.

4.1 Univariate portfolio analysis

Our firm-level investor sentiment measure is constructed as the difference in the implied volatility ratio of OTM calls over OTM puts. A higher value reflects investors' optimistic beliefs about underlying stock price in the next 30-365 days. When investors have a more optimistic view about the movement of the stock price, then they will trade and hold more of this stock, resulting in an increase of the stock price and its realized return. Thus, in the short run, the sentiment measure should have a positive effect on stock return.

At the end of each month, we sort all option covered stocks into quintile portfolios based on their monthly sentiment $Sent^{IVR}$. We then hold each portfolio for one month and compute both equal-weighted and value-weighted portfolio returns for the holding period. The optimistic portfolio (portfolio 5) contains stocks with the highest increase in the implied volatility ratio of OTM calls over puts, while the pessimistic portfolio (portfolio 1) includes stocks with the largest decline in the implied volatility ratio.

Panel A of Table 2 reports the equal-weighted portfolio returns and risk-adjusted abnormal returns of quintile portfolios. By sorting optionable stocks based on their changes in the implied volatility ratio, we find that all five portfolios show statistically significant and positive returns. They monotonically increase from the pessimistic portfolios (with an average return of 56 bps per month) to the optimistic portfolios (126 bps per month), demonstrating that optimistic sentiment pushes the stock price higher and causes higher portfolio returns. The zero-cost portfolio that is long in the optimistic portfolio and short in the pessimistic portfolio obtains 70 bps for an average month (8.73% annualized) with a t-stat of 7.20. Since small size stock return are more influenced by investor sentiment, we also report value-weighted portfolio returns in Panel B of Table 2 to mitigate the effect of size on portfolio returns. The results show a robust and monotonic increase in value-weighted portfolio returns from the pessimistic portfolio (0.56% per month) to the optimistic portfolio (1.21% per month), while the return spread of the long-short portfolio (0.65% per month) shrinks compared to the equal-weighted portfolio (0.70%), which is caused by the size effect. These results support the conclusion that the return pattern of portfolios sorted by our firm-specific sentiment $Sent^{IVR}$ is consistent across different weighting schemes.

[INSERT TABLE 2 ABOUT HERE]

In Table 2, we also report the risk-adjusted returns for each portfolio to examine whether the return pattern can be explained by conventional asset pricing models. In columns (2) to (8), we estimate the risk-adjusted returns while considering the simplest CAPM model to Fama French 5 factors augmented by liquidity and momentum factors (FF5+LIQ+UMD). Considering these conventional linear risk factor models, the zero-cost strategy of being long in the optimistic portfolio and short in the pessimistic portfolio still earns a statistically significant and positive abnormal return, ranging from 62 bps per month to 71 bps per month. This supports the conclusion that the effect of our firm-level sentiment on portfolio return is robust while we control for other conventional linear risk factor models.

4.2 Bivariate Portfolio Analysis

In this subsection, we examine whether our results are robust while controlling for both stock-level and option-level variables; we do this by conducting the double-sorted portfolio analysis. The stock-level characteristics (Panel A of Table 3) include market beta from the Carhart four-factor model (β_{MKT}), *size*, book-to-market ratio (*B/M*), Amihud's illiquidity (*illiquidity*), momentum (*Mom*), and profitability (*E/BE*). The option-level variables (Panel B of Table 3) consist of the options-to-stock volume ratio (*O/S*), changes in implied volatility of ATM calls (ΔIVC) and puts (ΔIVP), the ATM call minus put implied volatility spread (*IVC-IVP*), the realized minus implied volatility spread (*RV-IV*), put-call parity deviation (*CP dev*), and skewness (*Skew*).

Option covered stocks are first sorted by their target characteristics into five portfolios and subsequently ranked by their sentiment measures $Sent^{IVR}$ into quintiles within each characteristic portfolio. Thus, we form 25 equal-weighted portfolios and then equally aggregate the five target characteristics-sorted portfolios within each sentiment-sorted quintile portfolios. This leaves five sentiment-sorted portfolios with similar exposure to the target characteristics, and eliminates the effect of these characteristics on the relationship between our stock-specific sentiment measure $Sent^{IVR}$ and stock returns.

The bivariate sorted portfolio returns and alphas by market beta β_{MKT} and sentiment $Sent^{IVR}$ are reported in Row (β_{MKT}) of Panel A in Table 3. It shows that the equal-weighted portfolio return is monotonically increasing from the pessimistic portfolio with a monthly return of 0.56% to the optimistic portfolio with a 1.20% monthly return. The return spread between the optimistic and pessimistic portfolio is 7.96% annualized, and is significant at the 1% level with a t-stat of 7.13. This demonstrates that optimistic stocks earn higher stock returns while controlling for the effect of market beta. We also report the risk-adjusted alphas considering conventional linear risk factor models, including the Carhart four-factor model (C4α) and the Carhart 4 augmented by a liquidity factor (C4PSα). The alphas of the net-cost portfolio that are long in the optimistic portfolios and short in the pessimistic portfolios are significant and positive, with an annualized alpha of 7.57%.

[INSERT TABLE 3 ABOUT HERE]

We next control for firm size and report the time-series average of double-sorted portfolio monthly equal-weighted returns and risk-adjusted alphas in Row (Size) of Panel A in Table 3. The results suggest that the equal-weighted optimistic portfolio earns a higher monthly return of 1.23% relative to the pessimistic portfolio of 0.58% while controlling for the size effect. The returns and Carhart four-factor adjusted abnormal returns for the zero-cost long-short portfolio are 0.65% and 0.63% per month, respectively, and are significant at the 1% level. This indicates that the positive relationship between our sentiment and stock returns is robust when controlling for the size effect.

In Panel A of Table 3, we also report the portfolio analysis results sorted by the book-to-market ratio (*B/M*), Amihud's illiquidity (*Illiquidity*), momentum (*Mom*), and profitability (*E/BE*). The results again suggest that the positive effect of sentiment $Sent^{IVR}$ on stock returns is robust while controlling for the effect of stock-level characteristics. Panel B of Table 3 reports the portfolio returns and risk-adjusted alphas double-sorted by option-level variables and sentiment $Sent^{IVR}$. The Row (*O/S*) reports the equal-weighted portfolio return sorted by the option-to-stock volume ratio (*O/S*) and sentiment $Sent^{IVR}$, showing that the portfolio return increases monotonically from the pessimistic portfolio (0.50% per month) to the optimistic portfolio (1.18% per month). The return difference between the optimistic and pessimistic portfolio is 0.68% per month and is significant at the 1% level with a t-stat of 7.20. Controlling for the Carhart four-factor model, the abnormal return spread is also positive (0.65% per month) and significant at the 1% level. This supports the conclusion that optimistic stocks earn higher stock returns when we control for the option-to-stock volume ratio, which shows predictive power for the underlying stock returns.

We also control for changes in the implied volatility of ATM calls (ΔIVC) and report the bivariate sorting portfolio returns in Row (ΔIVC) of Panel B in Table 3. The results suggest that the equal-weighted optimistic portfolio earns a higher monthly return of 1.15% relative to the pessimistic portfolio of 0.62%, indicating that the return of the long-short portfolio is 0.52% per month. Even though the return spread (0.52% per month) shrinks slightly compared to univariate-sorted portfolio results (0.70%), the value is still positive and significant at the 1% level. This indicates that the positive relationship between our sentiment and stock returns is robust while controlling for changes in the implied volatility of ATM calls. In Panel B of Table

3, we also report the results after controlling for changes in the implied volatility of ATM puts (Δ*IVP*), the ATM call minus put implied volatility spread (*IVC-IVP*), realized minus implied volatility spread (*RV-IV*), the put-call parity deviation (*CP dev*), and skewness (*Skew*). The positive and significant return for the long-short portfolios demonstrates that more optimistic sentiment indicates that a higher return is robust considering option-level predictors.

4.3 Fama-MacBeth Regression

While the foregoing results suggest that the positive relationship between sentiment $Sent^{IVR}$ and portfolio returns is robust while controlling for most stock-specific and option-level predictors, portfolio analysis is not firm-specific and cannot control for multiple variables simultaneously. Thus, we employ the standard Fama-MacBeth regression to examine how our monthly firm-specific investor sentiment predicts stock returns at the firm-level. In the first stage, for each month *t*, we run the following regression model at a cross-section level and generate the estimated coefficient for each variable. Then, we report the time-series average of the coefficients and their t statistics.

$$Ret_{i,t+1} = \alpha_i + \beta_1 Sent_{i,t}^{IVR} + Xb' + \varepsilon_{i,t} \quad (14)$$

where $Ret_{i,t+1}$ is stock *i*'s realized return in month *t+1*, $Sent_{i,t}^{IVR}$ is stock *i*'s monthly investor sentiment in month *t*, and X refers to a vector of lagged control variables, including stock-level (market beta, size, book-to-market ratio, Amihud's illiquidity, momentum, idiosyncratic volatility, profitability, and stock turnover) and option-level variables (the options to stock volume ratio, changes in the implied volatility of ATM calls and puts, ATM call minus put implied volatility spread, realized minus implied volatility spread, put-call parity deviation, and skewness).

Table 4 reports the FM regression results. Column (1) in Table 4 reports the coefficients while considering only our sentiment measure $Sent^{IVR}$. It shows that the time-series average of coefficients of $Sent^{IVR}$ on $Ret_{i,t+1}$ is 0.018 which is significant at the 1% level (t-stat = 7.00), indicating a positive predictor of our sentiment measure.

[INSERT TABLE 4 ABOUT HERE]

This is consistent with the intuition that optimistic sentiment pushes prices higher and generates a higher return. Controlling for stock-level and option-level predictors; the coefficients of the sentiment measure $Sent^{IVR}$ remains significantly positive while they decline from 0.018 in Column (1) to 0.013 in Column (4) after including control variables in the model. This demonstrates that the positive relationship between $Sent^{IVR}$ and $Ret_{i,t+1}$ cannot be explained by these variables and supports the conclusion that our sentiment index is a proxy for a bullish belief about stock price movement.

To understand the economic significance of the slope coefficient of $Sent^{IVR}$, we report the cross-sectional distributions of $Sent^{IVR}$ in Table 1. The difference in $Sent^{IVR}$ values between the optimistic and pessimistic quintiles is 0.315. Thus, if a firm were to shift from the pessimistic quintile to the optimistic quintile while holding other characteristics constant, its next month realized return would increase 0.315*0.013 = 0.4095% per month. This number is smaller than the return spread in Tables 2 and 3 because we control for all firm-level and option-level characteristics simultaneously.

The slope coefficients of market beta (β_{MKT}), size, and book-to-market ratio are not significant in Column (2) of Table 4. We include additional stock-level characteristics such as momentum, illiquidity, realized volatility, idiosyncratic volatility, profitability, and turnover in Column (3). We find negative coefficients for size, volatility, and turnover, which are significant at the 10%, 10%, and 5% level, respectively. The weak effects of size, book-to-market ratio, momentum, and liquidity is consistent with An et al., (2014) because option-covered stocks are generally large and liquid stocks, leading to weak effects of size and liquidity.

Column (4) reports the regression results while controlling for both stock-level and option-level characteristics. The results show that changes in the implied volatility of the ATM calls (Δ*IVC*) and puts *(*Δ*IVP*) and the realized minus implied volatility spread (*RV-IV*) become insignificant after including our sentiment measure $Sent^{IVR}$. The ATM call minus put implied volatility spread (*IVC-IVP*) and the put-call parity deviation (*CP dev*) have significant and positive coefficients, which is consistent with Bali and Hovakimian (2009) and Cremers and Weinbaum (2010). The coefficients of option to stock volume ratio O/S and skewness (Skew) are also significantly negative, consistent with Seo and Kim (2015) and Xing et al. (2010).

4.4 Aggregate Predictability

In this subsection, we examine whether our aggregate market sentiment measure supports the theory of sentiment as mispricing. The theory predicts that when sentiment is high, the stock price is temporarily high because of retail investors' sentiment-driven trading, however, later it should correct to reflect the fundamentals. The regression model is:

$$Ret_{m,t+k} = \beta_0 + \beta_1 Sent_{m,t}^{IVR} + Xb' + \varepsilon_{m,t} \qquad (15)$$

where $Ret_{m,t+k}$ represents the stock market's return on day $t + k$ and $Sent_{m,t}^{IVR}$ represents the aggregated sentiment measure. X represents a vector of control variables, including lagged market return (up to 5 lags), changes in economic policy uncertainty (*EPU*), changes in the Aruoba-Diebold-Scotti (*ADS*) business conditions index, and the CBOE volatility index (*VIX*). These are consistent with Da et al. (2015). Since the options market closes at 4.02 PM for individual stock options, and the S&P 500 index is traded almost 24/7, there will be non-synchronous trading issues if we use the same-day prices for both equity and option. We therefore skip one day and examine the predictability model.

Table 5 reports the regression coefficients for the above model. When k = 1, the coefficient of $Sent_{m,t}^{IVR}$ on $Ret_{m,t+1}$ is 3.789 and is significant at the 1% level, demonstrating that an increase in the aggregated sentiment measure from the options market positively predicts the next day's market return. This suggests that when the current market is more optimistic in aggregate, the next day's market index price will increase. However, when k = 2, the coefficient of $Sent_{m,t}^{IVR}$ turns negative at -2.069 and is statistically significant at the 10% level, indicating that market price reversed on the second day. On day 3, the coefficient becomes insignificant and turns to around zero, suggesting that the effect of our aggregated sentiment measure on market return disappears on the third day. The R^2 in Table 5 is 0.015 on day 1, 0.009 on day 2, and 0.006 on day 3, suggesting decreasing explanatory power of our aggregated sentiment measure and control variables. Overall, our aggregated investor sentiment measure can predict the increase of S&P 500 index price on the first day and the reversal on the second day.

[INSERT TABLE 5 ABOUT HERE]

5. Hard-to-value Stocks and Stock Return Volatility

In this section, we empirically test two theoretical predictions of investor sentiment. First, investor sentiment should be more pronounced for hard-to-value stocks. Second, a higher absolute value of sentiment would lead to a higher level of volatility.

5.1 Hard-to-value Stocks

Baker and Wurgler (2006) propose that investor sentiment, defined as the propensity to speculate, has a more pronounced effect on subjectively hard-to-value stocks. Sentiment-driven investors prefer to trade hard-to-value stocks to suit their sentiment (optimistic or pessimistic) because these stocks are apparently speculative and have unlimited growth opportunities.

Another stream of hard-to-arbitrage analysis also draws the same conclusion. For opponents of the efficient market hypothesis, arbitrages are limited because they lack close substitutes and have short-horizon characteristics. In this section, we examine this theoretical implication using our firm-level sentiment measure. The theory predicts that if our measure captures firm-level investor sentiment, its forecasting power should be stronger for these hard-to-value, equivalently hard-to-arbitrage stocks.

We use size, age, idiosyncratic volatility, and Amihud's illiquidity as proxies for hard-to-value stocks, suggesting that small, young, highly volatile, and less liquid stocks are hard-to-value stocks. At the end of each month, we rank the stocks by the hard-to-value measures and form quintile portfolios. Then, in each of the quintiles, we sort stocks by $Sent^{IVR}$ and form quintile portfolios. We hold 25 portfolios for one month and calculate their equally weighted portfolio return, as well as the long-short portfolio return.

Table 6 shows the results for each of these four hard-to-value measures. The results are consistent with the theories, showing that firm-level sentiment has a stronger predictability effect for hard-to-value stocks. Panel A shows that the monthly average long-short portfolio return for the smallest firms is 1.43%, but is only 0.42% for the largest firms. The difference between these two portfolios is significantly different from zero. Panel B shows that the differences between the highest sentiment portfolio and the lowest sentiment portfolio return are 0.71% and 0.33% for the youngest and oldest firms, respectively. Panel C reports the results sorted by idiosyncratic volatility and shows that for the least and most volatile quintiles, the average long-short sentiment portfolio returns are 0.40% and 1.22%, respectively, with a significant difference of 0.82% per month. Illiquid stocks are also regarded as hard-to-value or hard-to-arbitrage stocks. Panel D reports the empirical analysis for this proxy. The results suggest that the difference between the highest and lowest sentiment is 1.20% for illiquid stocks and 0.35% for liquid stocks, with a significant difference of 0.86%. Overall, the empirical analysis suggests a stronger predictability effect for hard-to-value stocks.

[INSERT TABLE 6 ABOUT HERE]

5.2 Volatility

Earlier studies suggest that investor sentiment will cause stock prices to deviate from fundamentals and then reverse. To investigate this effect further, we study the relationship between stock volatility and contemporaneous absolute monthly investor sentiment. We use the absolute value of investor sentiment, because large changes in sentiment (either increases or decreases) should cause more volatility. We measure monthly stock volatility as the standard deviation of daily stock returns within a month.

To test the prediction, we estimate the following panel regression:

$$Vol_{i,t} = \alpha + \beta|Sent_{i,t}^{IVR}| + Xb + e_{i,t} \tag{16}$$

where $Vol_{i,t}$ represents the monthly volatility of stock i at month t, X represents a vector of firm-level characteristics, including age, size, B/M, profitability, liquidity, and so on. Since volatility is highly auto-correlated, we also include lagged volatility as a control variable. Table 7 reports the estimation results. We document a strong contemporaneous correlation between absolute investor sentiment and stock volatility. Column (1) shows that the regression coefficient of absolute value of sentiment is 0.006, which is statistically significant at the 1% level. This suggests that a one-standard-deviation increase in absolute investor sentiment is associated with a contemporaneous 0.6% increase in stock volatility using the Fama-MacBeth regression model. This effect is robust when we use a fixed effect model for panel data analysis in column (3). Our findings on stock volatility are consistent with the sentiment-induced stock price deviations from the fundamentals.

[INSERT TABLE 7 ABOUT HERE]

6. Conclusion

Using option data from January 1996 to June 2019, we construct a firm-specific investor sentiment measure $Sent^{IVR}$, which reflects the change in open interest weighted implied volatility ratio of OTM calls over puts. Our sentiment measure reflects investors' bullish beliefs about an underlying stock's future movement. At the firm level, our firm-specific investor sentiment positively predicts future stock returns at a monthly frequency. When investors long the optimistic portfolio and short the pessimistic portfolio, the abnormal return of this long-short strategy is 68 bps per month (8.16% annualized). Our empirical analysis supports the conclusion that the effect of sentiment on stock return is more pronounced for hard-to-value stocks, which are also limit-to-arbitrage stocks.

Our aggregated sentiment measure is significantly correlated with conventional market-wide sentiment measures: Baker and Wurgler's sentiment index, the PLS- based sentiment index, and survey-based investor sentiment measures like the University of Michigan's Consumer Sentiment and the American Association of Individual Investors bull-bear spread at monthly frequency. We find that an increase in our aggregated sentiment measure predicts an increase in the S&P 500 index price on day 1 and a return reversal on day 2. The limitation of this paper is that our firm-specific sentiment measure only covers optionable stocks.

References

Aboody, D., Even-Tov, O., Lehavy, R., & Trueman, B. (2018). Overnight returns and firm-specific investor sentiment. *Journal of Financial and Quantitative Analysis*, *53*(2), 485–505.

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

An, B. J., Ang, A., Bali, T. G., & Cakici, N. (2014). The joint cross section of stocks and options. *The Journal of Finance*, *69*(5), 2279–2337.

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

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives*, *21*(2), 129–152.

Bali, T. G., & Hovakimian, A. (2009). Volatility spreads and expected stock returns. *Management Science*, *55*(11), 1797–1812.

Ben-Rephael, A., Kandel, S., & Wohl, A. (2012). Measuring investor sentiment with mutual fund flows. *Journal of Financial Economics*, *104*(2), 363–382.

Bollen, N. P., & Whaley, R. E. (2004). Does net buying pressure affect the shape of implied volatility functions?. *The Journal of Finance*, *59*(2), 711–753.

Brown, G. W., & Cliff, M. T. (2004). Investor sentiment and the near-term stock market. *Journal of Empirical Finance*, *11*(1), 1–27.

Brown, G. W., & Cliff, M. T. (2005). Investor sentiment and asset valuation. *The Journal of Business*, *78*(2), 405–440.

Buraschi, A., & Jiltsov, A. (2006). Model uncertainty and option markets with heterogeneous beliefs. *The Journal of Finance*, *61*(6), 2841–2897.

Chakravarty, S., Gulen, H., & Mayhew, S. (2004). Informed trading in stock and option markets. *The Journal of Finance*, *59*(3), 1235–1257.

Chen, Y., Han, B., & Pan, J. (2021). Sentiment trading and hedge fund returns. *The Journal of Finance*, *76*(4), 2001–2033.

Choy, S. K., & Wei, J. (2012). Option trading: Information or differences of opinion?. *Journal of Banking & Finance*, *36*(8), 2299–2322.

Chung, K. H., & Ryu, D. (2016). Trade duration, informed trading, and option moneyness. *International Review of Economics & Finance*, *44*, 395–411.

Collin-Dufresne, P., Fos, V., & Muravyev, D. (2021). Informed Trading in the Stock Market and Option-Price Discovery. *Journal of Financial and Quantitative Analysis*, *56*(6), 1945–1984.

Coqueret, G. (2020). Stock-specific sentiment and return predictability. *Quantitative Finance*, *20*(9), 1531–1551.

Corrado, C. J., & Miller Jr, T. W. (1996). A note on a simple, accurate formula to compute implied standard deviations. *Journal of Banking & Finance*, *20*(3), 595–603.

Cremers, M., & Weinbaum, D. (2010). Deviations from put-call parity and stock return predictability. *Journal of Financial and Quantitative Analysis*, *45*(2), 335–367.

Da, Z., Engelberg, J., & Gao, P. (2015). The sum of all FEARS investor sentiment and asset prices. *The Review of Financial Studies*, *28*(1), 1–32.

Da, Z., Liu, Q., & Schaumburg, E. (2014). A closer look at the short-term return reversal. *Management Science*, *60*(3), 658–674.

Das, S. R., & Chen, M. Y. (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management Science*, *53*(9), 1375–1388.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance*, *45*(2), 379–395.

DeVault, L., Sias, R., & Starks, L. (2019). Sentiment metrics and investor demand. *The Journal of Finance*, *74*(2), 985–1024.

Engelberg, J. E., & Parsons, C. A. (2011). The causal impact of media in financial markets. *The Journal of Finance*, *66*(1), 67–97.

Garcia, D. (2013). Sentiment during recessions. *The Journal of Finance*, *68*(3), 1267–1300.

Garleanu, N., Pedersen, L. H., & Poteshman, A. M. (2008). Demand-based option pricing. *The Review of Financial Studies*, *22*(10), 4259–4299.

Ge, L., Lin, T. C., & Pearson, N. D. (2016). Why does the option to stock volume ratio predict stock returns?. *Journal of Financial Economics*, *120*(3), 601–622.

Han, B. (2008). Investor sentiment and option prices. *The Review of Financial Studies*, *21*(1), 387–414.

Houlihan, P., & Creamer, G. G. (2019). Leveraging a call-put ratio as a trading signal. *Quantitative Finance*, *19*(5), 763–777.

Hribar, P., & McInnis, J. (2012). Investor sentiment and analysts' earnings forecast errors. *Management Science*, *58*(2), 293–307.

Huang, D., Jiang, F., Tu, J., & Zhou, G. (2015). Investor sentiment aligned: A powerful predictor of stock returns. *The Review of Financial Studies*, *28*(3), 791–837.

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

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

Lakonishok, J., Lee, I., Pearson, N. D., & Poteshman, A. M. (2007). Option market activity. *The Review of Financial Studies*, *20*(3), 813–857.

Lee, C. M., Shleifer, A., & Thaler, R. H. (1991). Investor sentiment and the closed-end fund puzzle. *The Journal of Finance*, *46*(1), 75–109.

Lemmon, M., & Ni, S. X. (2014). Differences in trading and pricing between stock and index options. *Management Science*, *60*(8), 1985–2001.

Lemmon, M., & Portniaguina, E. (2006). Consumer confidence and asset prices: Some empirical evidence. *The Review of Financial Studies*, *19*(4), 1499–1529.

Livnat, J., & Petrovits, C. (2019). Investor sentiment, post-earnings announcement drift, and accruals. *The Journal of Applied Business and Economics*, *21*(8), 67–80.

Mian, G. M., & Sankaraguruswamy, S. (2012). Investor sentiment and stock market response to earnings news. *The Accounting Review*, *87*(4), 1357–1384.

Seo, S. W., & Kim, J. S. (2015). The information content of option-implied information for volatility forecasting with investor sentiment. *Journal of Banking & Finance*, *50*, 106–120.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, *62*(3), 1139–1168.

Xing, Y., Zhang, X., & Zhao, R. (2010). What does the individual option volatility smirk tell us about future equity returns?. *Journal of Financial and Quantitative Analysis*, *45*(3), 641–662.

Xiong, X., Meng, Y., Li, X., & Shen, D. (2020). Can overnight return really serve as a proxy for firm-specific investor sentiment? Cross-country evidence. *Journal of International Financial Markets, Institutions and Money*, *64*, 101173.

Table 1: Summary Statistics

Panel A: Descriptive Statistics

	$Sent^{IVR}$	β_{MKT}	Size	B/M	Mom	Illiq	E/BE	Vol	Idvol	ΔIVC	ΔIVP	IVC-IVP	CP dev	RV-IV	Skew	O/S
Mean	0.000	1.080	7.434	0.526	0.190	0.009	16.239	0.130	0.023	-0.001	-0.001	-0.008	-0.014	-0.031	0.066	0.026
Std.	0.152	1.310	1.579	0.504	0.639	0.034	139.188	0.084	0.016	0.127	0.127	0.099	0.048	0.206	0.063	0.064
P5	-0.177	-0.793	5.095	0.089	-0.459	0.000	0.000	0.049	0.008	-0.148	-0.151	-0.113	-0.075	-0.289	-0.003	0.000
P25	-0.054	0.429	6.305	0.237	-0.129	0.000	0.332	0.078	0.013	-0.042	-0.043	-0.025	-0.022	-0.108	0.031	0.002
Median	-0.001	1.034	7.289	0.413	0.088	0.002	9.167	0.111	0.019	-0.001	-0.001	-0.004	-0.008	-0.035	0.052	0.009
P75	0.053	1.694	8.412	0.675	0.350	0.006	16.709	0.160	0.029	0.039	0.040	0.014	0.001	0.036	0.087	0.028
P95	0.177	3.110	10.269	1.269	1.131	0.038	35.010	0.270	0.051	0.149	0.151	0.086	0.027	0.226	0.181	0.104
N	2075	1811	1817	1706	1767	1817	1706	1815	1811	2145	2145	2157	2163	2157	2157	2157

Panel B: Characteristics of Quintile Portfolios Sorted by Sentiment

	$Sent^{IVR}$	β_{MKT}	Size	B/M	Mom	Illiq	E/BE	Vol	Idvol	Age	ΔIVC	ΔIVP	IVC-IVP	CP dev	RV-IV	Skew	O/S
1	-0.157	1.038	7.087	0.568	0.138	0.011	15.55	0.131	0.024	19.21	-0.024	0.008	-0.025	-0.018	-0.042	0.082	0.019
2	-0.042	1.072	7.749	0.514	0.189	0.005	19.066	0.122	0.021	21.714	-0.009	0	-0.011	-0.014	-0.03	0.066	0.033
3	-0.001	1.09	7.923	0.499	0.205	0.005	19.599	0.121	0.021	22.257	-0.001	-0.001	-0.007	-0.013	-0.027	0.062	0.04
4	0.041	1.091	7.715	0.513	0.199	0.006	16.996	0.124	0.021	21.462	0.008	-0.002	-0.003	-0.013	-0.029	0.065	0.034
5	0.158	1.038	6.921	0.584	0.168	0.015	13.824	0.135	0.025	18.496	0.019	-0.007	0.005	-0.014	-0.04	0.077	0.016

Panel C: Correlation Matrix with Firm and Option Characteristics

	$Sent^{IVR}$	βMKT	Size	B/M	Mom	Illiq	E/BE	Vol	Idvol	Age	ΔIVC	ΔIVP	IVC-IVP	CP dev	RV-IV	Skew	O/S
$Sent^{IVR}$	1.000																
βMKT	-0.001	1.000															
Size	-0.005	-0.017	1.000														
B/M	-0.002	0.007	-0.123	1.000													
Mom	0.008	0.018	0.096	0.029	1.000												
Illiq	0.003	-0.015	-0.394	0.067	-0.107	1.000											
E/BE	0.000	-0.013	0.106	-0.129	-0.007	-0.051	1.000										
Vol	0.002	0.070	-0.445	0.001	0.151	0.179	-0.080	1.000									
Idvol	0.010	0.057	-0.434	-0.014	-0.014	0.220	-0.065	0.554	1.000								
Age	-0.001	-0.027	0.426	0.063	-0.029	-0.116	0.064	-0.314	-0.287	1.000							
ΔIVC	0.147	0.021	-0.001	-0.001	0.015	0.001	0.001	-0.023	0.001	0.006	1.000						
ΔIVP	-0.052	0.018	0.002	-0.001	0.006	0.000	0.001	-0.022	0.011	0.006	0.510	1.000					
IVC-IVP	0.128	0.006	0.037	-0.004	0.003	-0.024	0.003	-0.078	-0.069	0.030	0.319	-0.298	1.000				
CP dev	0.053	0.012	0.002	-0.028	-0.014	0.010	-0.003	-0.093	-0.056	-0.014	0.041	0.006	0.363	1.000			
RV-IV	0.002	0.095	0.109	-0.005	0.013	-0.080	0.014	0.004	0.499	0.042	-0.167	-0.166	0.023	0.044	1.000		
Skew	-0.021	-0.056	0.129	0.068	0.011	-0.103	0.014	-0.252	-0.242	0.140	-0.027	-0.023	-0.033	-0.126	0.037	1.000	
O/S	0.000	0.011	0.256	-0.119	0.096	-0.076	0.035	0.066	0.043	0.012	-0.010	-0.009	-0.034	-0.073	0.011	-0.088	1.000

Panel D: Correlation with Aggregate Sentiment

	S^{BW}	S^{PLS}	S^{MCSI}	S^{AAII}	$Sent^{IVR}$
S^{BW}	1				
S^{PLS}	0.676***	1			
S^{MCSI}	0.280***	0.161***	1		
S^{AAII}	0.175***	0.125**	0.449***	1	
$Sent^{IVR}$	0.548***	0.376***	0.730***	0.404***	1

This table reports summary statistics. Panel A describes the summary statistics for our sentiment measure, firm-level, and option-level characteristics for optionable stocks from January 1996 to June 2019. $\mathrm{Sent}^{\mathrm{IVR}}$ represents the firm-specific sentiment measure, which is the change in the open interest-weighted implied volatility ratio of OTM calls over puts. β_{MKT} is the market beta via the Carhart 4 factor model and Size is the natural logarithm of market equity. B/M represents the ratio of book value to the stock's market value and momentum (Mom) is the cumulative return over the previous 11 months. Illiq refers to the Amihud's illiquidity and is computed as the ratio of the absolute monthly stock return to dollar volume. Vol and Idvol represent the volatility, measured as the standard deviation of monthly returns over the past 12 months, and idiosyncratic volatility, which is the standard deviation of residuals from the Carhart 4 factor model. Turnover is the ratio of the number of shares traded in a month and the number of shares outstanding. E/BE refers to the earnings-to-book value ratio, while Age measures years since the equity first showed in the CRSP. RV-IV is calculated as the difference between the monthly realized volatility and the average of the ATM call and put implied volatility, while IVC-IVP is measured as the spread between the implied volatility of ATM calls and puts. ΔIVC and ΔIVP measure changes in ATM calls and puts implied volatility. CP dev is the put call parity deviation, which is measured as the open interest weighted average difference between pairwise (strike and maturity) call and put implied volatility. Skew is the difference between the OTM put implied volatility and the average of the ATM call and put implied volatility. O/S measures option to stock volume (O/S) and is computed as the ratio of options volume to stock volume. Panel B sorts the sample into 5 quintiles based on $\mathrm{Sent}^{\mathrm{IVR}}$ each month and the characteristic of each portfolio is described. Panel C reports the correlation between $\mathrm{Sent}^{\mathrm{IVR}}$ and other variables. Panel D describes the correlation coefficient between the aggregated sentiment measure and conventional investor sentiment indexes. S^{BW} refers to the sentiment index from Baker and Wurgler (2006), which is the first principal of five sentiment proxies, including closed-end fund discount, number of IPOs, the average first-day returns, equity share, and dividend premium. The University of Michigan Consumer Sentiment Index (S^{MCSI}) and the AAII bull-bear spread (S^{AAII}) are downloaded from the Federal Reserve Bank of St. Louis and the American Association of Individual Investors' official website. S^{PLS} is obtained from Dashan Huang's personal website.

Table 2: Quintile Portfolios of Stocks Sorted by sentiment

Panel A: Equal-weighted Portfolio Returns and Abnormal Returns

	(1) Ret	(2) CAPM	(3) FF3	(4) FF3+LIQ	(5) Carhart4	(6) Carhart4+LIQ	(7) FF5	(8) FF5+LIQ+UMD
1	0.56	-0.43	-0.50	-0.50	-0.35	-0.36	-0.49	-0.39
	(1.49)	(-2.60)	(-4.88)	(-5.05)	(-4.25)	(-4.44)	(-4.32)	(-4.91)
2	0.88	-0.12	-0.16	-0.17	-0.06	-0.07	-0.11	-0.06
	(2.36)	(-0.88)	(-1.77)	(-1.88)	(-0.74)	(-0.92)	(-1.32)	(-0.75)
3	0.93	-0.07	-0.10	-0.11	-0.01	-0.03	-0.02	0.02
	(2.52)	(-0.55)	(-1.30)	(-1.54)	(-0.16)	(-0.42)	(-0.30)	(0.30)
4	1.01	0.01	-0.03	-0.05	0.06	0.04	0.04	0.08
	(2.75)	(0.05)	(-0.39)	(-0.60)	(0.77)	(0.53)	(0.50)	(1.18)
5	1.26	0.25	0.18	0.17	0.32	0.31	0.22	0.31
	(3.40)	(1.44)	(1.68)	(1.62)	(3.13)	(3.03)	(1.95)	(3.19)
5-1	0.70***	0.68***	0.68***	0.68***	0.68***	0.67***	0.71***	0.71***
	(7.20)	(6.88)	(6.78)	(6.70)	(6.69)	(6.61)	(6.71)	(6.60)

Panel B: Value-weighted Portfolio Return and Abnormal Return

	(1) Ret	(2) CAPM	(3) FF3	(4) FF3+LIQ	(5) Carhart4	(6) Carhart4+LIQ	(7) FF5	(8) FF5+LIQ+UMD
1	0.56	-0.40	-0.46	-0.47	-0.34	-0.35	-0.48	-0.40
	(1.56)	(-2.66)	(-4.85)	(-5.06)	(-4.26)	(-4.51)	(-4.58)	(-5.21)
2	0.85	-0.12	-0.15	-0.16	-0.07	-0.09	-0.12	-0.08
	(2.40)	(-0.97)	(-1.87)	(-2.03)	(-0.97)	(-1.19)	(-1.63)	(-1.19)
3	0.94	-0.04	-0.07	-0.09	0.00	-0.02	-0.01	0.03
	(2.63)	(-0.39)	(-0.99)	(-1.27)	(0.03)	(-0.26)	(-0.08)	(0.41)
4	1.00	0.02	-0.01	-0.03	0.06	0.04	0.04	0.07
	(2.83)	(0.21)	(-0.14)	(-0.39)	(0.89)	(0.61)	(0.60)	(1.15)
5	1.21	0.23	0.16	0.15	0.28	0.27	0.19	0.26
	(3.43)	(1.48)	(1.69)	(1.61)	(3.04)	(2.91)	(1.87)	(2.95)
5-1	0.65***	0.63***	0.63***	0.63***	0.62***	0.62***	0.66***	0.66***
	(6.85)	(6.53)	(6.42)	(6.39)	(6.38)	(6.35)	(6.38)	(6.35)

This table reports equal-weighted (Panel A) and value-weighted (Panel B) return and abnormal returns for quintile portfolios, sorted on the previous month's sentiment measure $Sent^{IVR}$. Quintile 1 includes the stocks with the lowest sentiment measure, while stocks with the highest sentiment are shown in quintile 5. Column (1) reports the portfolio return; Column (2) shows the alpha obtained from CAPM; Column (3) shows the alpha obtained from the Fama and French 3 factor model; Column (4) reports the alpha from the FF3 and liquidity factor model; Column (5) shows the alpha from the Carhart 4 factor model; Column (6) is based on the Carhart 4 and liquidity factor model; Column (7) represents the alpha from the FF5 factor model; and Column (8) shows the alpha of FF5 and liquidity and momentum factor model. Our sample period is from January 1996 to June 2019. *, **, and *** show significance at the 10%, 5%, and 1% level, respectively.

Table 3: Bivariate Sorts: Controlling for Cross-sectional Characteristics

Panel A: Control for Stock-level Variables

	Low	2	3	4	High	H-L	t-stat	C4α	t-stat	C4PSα	t-stat
β_{MKT}	0.56	0.82	0.92	0.98	1.20	0.63***	(7.13)	0.61***	(6.85)	0.61***	(6.72)
Size	0.58	0.88	0.90	1.04	1.23	0.65***	(7.26)	0.63***	(6.84)	0.63***	(6.76)
B/M	0.57	0.89	0.96	1.01	1.25	0.68***	(6.85)	0.65***	(6.49)	0.65***	(6.39)
Illiquidity	0.60	0.88	0.89	1.04	1.24	0.65***	(7.16)	0.62***	(6.79)	0.63***	(6.74)
Mom	0.61	0.91	0.94	1.03	1.28	0.67***	(7.60)	0.66***	(7.08)	0.66***	(7.06)
E/BE	0.60	0.87	0.94	0.99	1.27	0.67***	(6.86)	0.65***	(6.42)	0.65***	(6.33)

Panel B: Control for Option Variables

	Low	2	3	4	High	H-L	t-stat	C4α	t-stat	C4PSα	t-stat
O/S	0.50	0.89	1.02	1.04	1.18	0.68***	(7.20)	0.65***	(6.59)	0.65***	(6.54)
$\Delta IV\ C$	0.62	0.90	0.90	1.06	1.15	0.52***	(6.51)	0.52***	(6.38)	0.53***	(6.45)
$\Delta IV\ P$	0.58	0.90	0.89	1.00	1.26	0.68***	(7.06)	0.66***	(6.54)	0.66***	(6.46)
IVC-IVP	0.71	0.85	0.97	0.95	1.15	0.44***	(5.32)	0.41***	(4.84)	0.42***	(4.92)
RV-IV	0.56	0.87	0.94	1.04	1.22	0.66***	(6.82)	0.64***	(6.45)	0.64***	(6.37)
CP dev	0.57	0.91	0.90	1.03	1.22	0.65***	(6.73)	0.62***	(6.55)	0.63***	(6.53)
Skew	0.54	0.93	0.90	1.04	1.22	0.68***	(7.19)	0.66***	(6.88)	0.66***	(6.71)

This table shows the average equal-weighted monthly returns of portfolios sorted on the following stock and option characteristics and the $\mathrm{Sent}^{\mathrm{IVR}}$. Each month, stocks are first sorted into 5 portfolios by the controlled characteristics and then within each characteristic portfolio, the stocks are sorted by $\mathrm{Sent}^{\mathrm{IVR}}$. Then the time-series average of returns sorted by $\mathrm{Sent}^{\mathrm{IVR}}$ across characteristic portfolios are calculated. The table also reports the H-L portfolio return, the Carhart 4 factor alpha, and the Carhart 4 factor augmented by the liquidity factor alpha. *, **, and *** show significance at the 10%, 5%, and 1% level, respectively. T-statistics are Newey-West adjusted t-statistics (with 3 lags).

Table 4: Fama-MacBeth Regression: 1-month

	(1) $Ret_{i,t+1}$	(2) $Ret_{i,t+1}$	(3) $Ret_{i,t+1}$	(4) $Ret_{i,t+1}$
$Sent_{i,t}^{IVR}$	0.018***	0.017***	0.017***	0.013***
	(7.00)	(6.99)	(7.20)	(6.04)
β_{MKT}		-0.001	-0.001	-0.001
		(-1.40)	(-0.88)	(-0.77)
Size		-0.000	-0.001*	-0.001
		(-0.45)	(-1.94)	(-1.49)
B/M		-0.000	-0.001	-0.001
		(-0.21)	(-0.64)	(-0.52)
Mom			0.001	0.001
			(0.46)	(0.25)
Illiq			0.022	0.035
			(0.89)	(1.52)
Vol			-0.024*	-0.011
			(-1.79)	(-0.91)
Idvol			-0.002	-0.030
			(-0.05)	(-0.34)
E/BE			0.000	0.000
			(1.48)	(1.48)
Turnover			-0.001**	-0.000*
			(-2.37)	(-1.88)
$\Delta IV\ C$				-0.000
				(-0.05)
$\Delta IV\ P$				0.000
				(0.03)
IVC-IVP				0.041***
				(5.13)
CP dev				0.043***
				(3.96)
RV-IV				0.002
				(0.40)
O/S				-0.029**
				(-2.25)
Skew				-0.024***
				(-3.71)
R^2	0.001	0.032	0.076	0.092

This table reports the standard Fama-MacBeth regression coefficients to examine the predictability of monthly firm-specific investor sentiment on stock returns. The dependent variable is stock return in the next month $Ret_{i,t+1}$ and the independent variable is the sentiment measure $Sent_{i,t}^{IVR}$. Column (2) controls for market β, size, and B/M ratio. Column (3) controls for all stock-level variables, and Column (4) adds option-level variables. *, **, and *** show significance at the 10%, 5%, and 1% level, respectively. T-statistics are reported in brackets.

Table 5: Aggregate Level Sentiment and Predictability

	(1) $Ret_{m.t+1}$	(2) $Ret_{m.t+2}$	(3) $Ret_{m.t+3}$
$Sent_{i,t}^{IVR}$	3.789***	-2.069*	-0.041
$Ret_{m,t}$	-0.039**	-0.060***	0.012
$Ret_{m,t-1}$	-0.068***	0.017	-0.012
$Ret_{m,t-2}$	0.005	-0.017	-0.047***
$Ret_{m,t-3}$	-0.016	-0.048***	-0.001
$Ret_{m,t-4}$	-0.048***	-0.001	-0.030**
$Ret_{m,t-5}$	-0.003	-0.033**	0.008
ΔADS	0.103***	0.088***	0.078***
ΔEPU	0.000	-0.000	-0.000
VIX	0.007***	0.006***	0.006***
R^2	0.015	0.009	0.006

This table reports the predictability of our aggregated firm-specific investor sentiment on S&P 500 index daily returns. The dependent variables are the daily return on days 1 to 3 (Columns 1-3). The independent variable includes our aggregated sentiment measure $Sent^{IVR}$ and lagged returns up to five lags, changes of news-based economic policy uncertainty (ΔEPU), changes in the Aruoba-Diebold-Scotti business conditions index (ΔADS), and the CBOE volatility index (VIX). *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 6: Hard-to-value Stocks

Panel A: Size

Sent	Small	2	3	4	Big	B-S
Low	0.43	0.45	0.62	0.81	0.64	0.22
2	1.10	0.93	0.80	0.86	0.82	-0.28
3	0.83	0.95	0.83	1.01	0.91	0.07
4	1.21	1.02	0.94	1.08	1.02	-0.20
High	1.85	1.20	1.10	1.08	1.06	-0.79*
H-L	1.43***	0.75***	0.48***	0.26**	0.42***	-1.01***

Panel B: Age

Sent	Young	2	3	4	Old	O-Y
Low	0.18	0.43	0.66	0.80	0.75	0.57*
2	0.70	0.87	1.14	0.90	0.81	0.11
3	0.70	0.86	0.88	1.08	1.05	0.35
4	0.86	0.87	1.15	1.21	1.03	0.17
High	0.89	1.40	1.52	1.35	1.08	0.20
H-L	0.71***	0.97***	0.87***	0.55***	0.33***	-0.38*

Panel C: Idiosyncratic Volatility

Sent	Low	2	3	4	High	H-L
Low	0.82	0.94	0.64	0.31	0.09	-0.73
2	0.91	0.93	0.78	0.92	0.76	-0.16
3	1.11	1.11	0.97	0.94	0.43	-0.68
4	1.06	1.13	1.14	0.80	0.65	-0.40
High	1.22	1.17	1.23	1.17	1.31	0.09
H-L	0.40***	0.23**	0.58***	0.86***	1.22***	0.82***

Panel D: Amihuid Illiquidity

Sent	Liq	2	3	4	Illiq	H-L
Low	0.71	0.79	0.61	0.44	0.44	-0.27
2	0.87	0.74	1.03	0.85	0.96	0.09
3	0.91	0.98	0.81	0.79	1.00	0.09
4	0.97	1.01	1.18	1.11	0.96	-0.00
High	1.06	1.04	1.30	1.29	1.64	0.59
H-L	0.35***	0.25*	0.69***	0.85***	1.20***	0.86***

This table reports the time-series average of equal weighted portfolio return sorted by size, age, idiosyncratic volatility and illiquidity. Stocks are first sorted into 5 quintiles based on the hard-to-value measures, and then within each bin, stocks are sorted by their sentiment. The portfolio is held for one month and rebalanced every month. *, **, and *** show significance at the 10%, 5%, and 1% level, respectively.

Table 7: Investor Sentiment and Volatility

	Vol FM	Vol FM	Vol Fixed Effect
$\| Sent^{IVR} \|$	0.006***	0.006***	0.003***
(t-stat)	(9.99)	(10.91)	(15.58)
Age	-0.000***	-0.000***	0.000***
(t-stat)	(-5.94)	(-6.41)	(3.14)
β_{MKT}	0.001***	0.001***	0.000***
(t-stat)	(3.63)	(3.51)	(10.71)
Size	-0.001***	-0.001***	-0.003***
(t-stat)	(-11.07)	(-11.35)	(-44.71)
Vol_{t-1}	0.941***	0.942***	0.931***
(t-stat)	(185.41)	(196.07)	(1875.98)
Mom		0.000	0.001***
(t-stat)		(0.35)	(20.5)
Illiquidity		-0.005*	-0.001
(t-stat)		(-1.86)	(-0.94)
B/M		-0.001***	-0.002***
(t-stat)		(-3.01)	(-19.47)
E/BE		-0.000***	0
(t-stat)		(-3.13)	(-0.07)
R^2	0.94	0.94	0.90

This table reports the effect of investor sentiment on stock volatility. *Vol* represents the volatility of the stock, $| Sent^{IVR} |$ refers to the absolute value of firm-level investor sentiment. T-statistics are reported in brackets. Columns (1) and (2) use Fama-MacBeth regression and Column (3) uses the panel fixed effect model. *, **, and *** show significance at the 10%, 5%, and 1% level, respectively.