Journal of Banking and Finance 73 (2016) 147–164

Contents lists available at ScienceDirect

Journal of Banking and Finance

journal homepage: www.elsevier.com/locate/jbf

Stock return predictability and investor sentiment: A high-frequency perspective☆

Licheng Sun[a,*], Mohammad Najand[a], Jiancheng Shen[b]

[a] Department of Finance, Strome College of Business, Old Dominion University, Norfolk, VA 23529, US
[b] Department of Business, Leadership, and Management College of Arts and Sciences, Regent University, Virginia Beach, VA 23464, US

ARTICLE INFO

Article history:
Received 19 November 2015
Accepted 25 September 2016
Available online 1 October 2016

JEL Classification:
G11
G14

Keywords:
Intraday
Investor sentiment
High frequency
Stock return predictability
Noise trading

ABSTRACT

We explore the predictive relation between high-frequency investor sentiment and stock market returns. Our results are based on a proprietary dataset of high-frequency investor sentiment, which is computed based on a comprehensive textual analysis of sources from news wires, internet news sources, and social media. We find substantial evidence that intraday S&P 500 index returns are predictable using lagged half-hour investor sentiment. The predictive power is also found in other stock and bond index ETFs. We document that this sentiment effect is independent of the intraday momentum effect, which is based on lagged half-hour returns. While the intraday momentum effect only exists in the last half hour, the sentiment effect persists in at least the last two hours of a trading day. From an investment perspective, high-frequency investor sentiment also appears to have significant economic value when evaluated with market timing trading strategies. We find evidence that the return predictability is most likely driven by the trading activities of noise traders.

1. Introduction

It has been well documented that investor sentiment plays an important role in financial markets. For instance, from a theoretical perspective, De Long et al. (1990) show that, with limits to arbitrage, changes in noise traders' sentiment will result in excess market volatility as well as deviation in stock prices away from their fundamental values. Barberis, Shleifer, and Vishny (1998) present a model of investor sentiment that can produce both underreaction and overreaction to news.

Empirically, investor sentiment has also been shown to impact asset prices as well as have explanatory power on some well-known asset pricing anomalies. For example, Hirshleifer and Shumway (2003) find that upbeat investor mood associated with morning sunshine in the city of a country's leading stock exchange is significantly correlated with daily market index returns across 26 countries. Lemmon and Portniaguina (2006) explore the time-series relationship between sentiment and the small-stock premium and find that consumer confidence can forecast small stock returns. Antoniou et al. (2013) document that momentum profits arise only under investor optimism. Baker and Wurgler (2006) examine the cross-sectional effect of investor sentiment. They show that when sentiment is low (high), subsequent returns are relatively high (low) for small stocks, young stocks, high volatility stocks, and distressed stocks.

Given the significant impact from investor sentiment on asset prices, it is imperative that researchers use high-quality measures of aggregate investor sentiment in their studies. In the extant literature, there are at least three approaches that attempt to measure investor sentiment with accuracy.

First, investor sentiment could be captured using certain market-based variables. Lee et al. (1991) document that fluctuations in discounts of closed-end funds are driven by changes in investor sentiment. Baker and Wurgler (2006); 2007) construct a measure of investor sentiment that is based on several market-based variables such as closed-end fund discount, IPO first-day returns, IPO volume, and trading volume. Other popular market-based sentiment measures include: option implied volatility index, and market state as defined by the sign of lagged three-year or one-year market returns.

However, as argued by Qiu and Welch (2006) as well as Da et al. (2015), market-based measures of sentiment have the drawback of being the equilibrium outcome of many economic forces other than investor sentiment. Thus to get a "cleaner"

☆ We are grateful to Richard Peterson, Managing Director at MarketPsych, and Elijah DePalma from Thomson Reuters for providing us with the proprietary Thomson Reuters MarketPsych Indices data. We also thank two anonymous referees as well as seminar participants at Old Dominion University for helpful comments.

* Corresponding author.
E-mail addresses: lsun@odu.edu (L. Sun), mnajand@odu.edu (M. Najand), jshen@regent.edu (J. Shen).

http://dx.doi.org/10.1016/j.jbankfin.2016.09.010
0378-4266/

measure of sentiment, one can use survey-based measures of investor sentiment. Examples of this approach include: the University of Michigan Consumer Sentiment Index, the AAII investor sentiment survey, and the UBS/GALLUP Index for Investor Sentiments. Survey-based sentiment measures are not without their own weakness. Da et al. (2015) note that they are not available in high frequency and become increasingly less reliable when non-response rates in surveys are high or the incentive for truth-telling is low.

More recently, sentiment metrics based on textual analysis of media contents such as newspaper columns, messages boards, blogs, and google search results have gained popularity. We call these media-based investor sentiment measure. Using daily content from a popular *Wall Street Journal* column, Tetlock (2007) find that high media pessimism predicts downward pressure on market prices followed by a reversion to fundamentals, and unusually high or low pessimism predicts high market trading volume. He concludes that these results are consistent with theoretical models of noise and liquidity traders, and are inconsistent with theories of media content as a proxy for new information about fundamental asset values or market volatility. Antweiler and Frank (2004) study the effect of more than 1.5 million messages posted on internet message boards and find significant evidence that the stock messages help predict market volatility. Da et al. (2015) construct a sentiment index based on google search results using key words such as “recession”, “unemployment”, and “bankruptcy”. They find that this index can predict both short-term return reversals and temporary increase in volatility.

An important advantage of using media-based sentiment measures is their availability in high-frequency. Da et al. (2015) note that: “to date, high frequency analysis of investor sentiment is found only in laboratory settings.” For example, The FEARS index constructed by Da et al. is available in daily frequency. In contrast, many survey-based measures are only available in monthly or quarterly frequency. The popular Baker–Wurgler index is available in monthly and annual frequency.

In this paper, we study the predictive relation between investor sentiment and stock market returns at the intraday level. To the best of our knowledge, this article is the first to study the relation between ultra-high frequency investor sentiment and the predictability in intraday stock returns at the market index level. In our view, there at least three reasons to study investor sentiment at the intraday level. First, from a big picture perspective, progress in scientific subjects are often made by studying behavior at infinitesimal time increments (high frequency). In finance, for example, we have gained new knowledge about the behavior of market price and liquidity, among many other things, by studying market microstructure. In the case of modeling the volatility of asset prices, the predictive power of GARCH models have benefited from using high frequency data. Given the importance of understanding the nature of investor sentiment, it is our belief that studying investor sentiment at intraday level will shed new light on this subject, which cannot be gained by studying sentiment at lower frequency (monthly or daily) levels. Second, from a modeling perspective, if we view investor sentiment as governed by a continuous time process, then studying investor sentiment at intraday level will give us a more precise estimate of the real changes and movements in investor sentiment. For instance, suppose that policy makers would like to gauge investors’ real time response to a recent change in monetary policy. In this case, the ability to estimate the change in sentiment at intraday level will be invaluable whereas the low frequency alternatives will be inadequate. Third, from a practical perspective, investors could potentially use intraday investor sentiment as their model inputs and improve their trading strategies.

Our intraday sentiment measure is obtained from a proprietary dataset from Thomson Reuters, which is based on a commercial-strength comprehensive textual analysis of sources from news wires, internet news sources, and social media. We note the prior studies that focus on textual analysis of media contents almost exclusively rely on a single source. For example, Tetlock (2007) and Garcia (2013) use columns from *Wall Street Journal* and *New York Times* respectively. Chen et al. (2014) conduct textual analysis of articles published on *seekingalpha.com*. Compared with these studies in the extant literature, our sentiment measure is constructed from a much broader and more comprehensive collection of both traditional and social media sources. For example its sources include financial news, social media, earnings conference call transcripts, and executive interviews. In our view, given the goal is to obtain an accurate measure of investor sentiment, the all-encompassing nature of our sentiment measure is an important advantage over other single-source sentiment measures.

To match with the frequency of our sentiment data, we naturally choose to study stock return predictability at the intraday level. Heston et al. (2010) provide a comprehensive study of the cross-section stock return patterns at the intraday level. They identify an interesting pattern of return continuation at half-hour intervals that are exact multiples of a trading day. More recently, Gao et al. (2015) (henceforth GHLZ) document an intriguing intraday momentum pattern for the S&P 500 index ETF. They show that the first half-hour return on the market predicts the last half-hour return on the market.

We find convincing evidence that intraday S&P 500 index returns are predictable using lagged half-hour changes in investor sentiment. We document that this sentiment effect is independent of the intraday momentum effect of GHLZ, which is based on lagged half-hour returns. While the intraday momentum effect only exists in the last half hour, the sentiment effect persists in at least the last two hours of a trading day. From an investment perspective, high-frequency investor sentiment also appears to have significant economic value when evaluated with market timing trading strategies. In addition, similar results are also found in other large-cap, small-cap, and international stock ETFs.

While the overall evidence are supportive of the hypothesis that the predictive power of our intraday sentiment measure is attributable to the activities of noise traders, we are open to the notion that our results could potentially be driven by rational factors such as some (unidentified) state variables and/or market frictions. To help differentiate the behavioral and rational explanations, we have conducted a battery of empirical tests. We find no evidence that our results are driven lagged macroeconomic variables, macroeconomic news announcement effects associated with the non-farm payroll unemployment report, or FOMC (Federal Open Market Committee) meetings. In contrast, we do find evidence that our intraday sentiment measure is significantly correlated with alternative sentiment measures such as the University of Michigan consumer sentiment index. More importantly, we find that the predictive value of investor sentiment measure mainly shows up in days with high trading volume, which, according to Odean (1998) and Barber and Odean (2000), is an indication of noise trading. Thus we believe that our results are consistent with the noise trading hypothesis and more aligned with a behavioral explanation.

The rest of this paper is organized as follows. The next section describes the data. Section 3 documents the empirical relation between high-frequency investor sentiment and intraday S&P 500 returns using predictive regressions. Section 4 provides a battery of robustness checks by examining monthly and weekday seasonality, the effects of macroeconomic variables and alternative measures of investor sentiment such as market states, and CBOE’s volatility index. Section 5 evaluates the economic significance of investor sentiment using market timing trading strategies. Section 6 evaluates alternative explanations by exploring the relation between investor