Stock Return Predictability and Investor Sentiment: A High-Frequency Perspective *

Licheng Sun, Mohammad Najand, and Jiancheng Shen

Department of Finance

Strome College of Business

Old Dominion University

Norfolk, VA 23529

First Draft: March 2015

This version: November 2015

*We are grateful to Richard Peterson, Managing Director at MarketPsych, for providing us with the proprietary Thomson Reuters MarketPsych Indices data. We also thank seminar participants at Old Dominion University for helpful comments. Please address comments to Licheng Sun (corresponding author, phone: 757-683-6552, e-mail: lsun@odu.edu), Mohammad Najand (phone: 757-683-3501, email: mnajand@odu.edu), or Jiancheng Shen (email: jshen@odu.edu).

Stock Return Predictability and Investor Sentiment: A High-Frequency Perspective

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 predictability is evident based on both in-sample and out-of-sample statistical metrics. 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.

JEL Classification: G11, G14

Keywords: Intraday, Investor Sentiment, High Frequency, Stock Return Predictability

I. Introduction

It has been well documented that investor sentiment plays an important role in financial markets. For instance, from a theoretical perspective, De Long, Shleifer, Summers, and Waldmann (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, Doukas, and Subrahmanyam (2013) find 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, Shleifer, and Thaler (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, Engelberg, and Gao (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" 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. The intraday frequency is a perfect setting for investigating the impact of noise traders because presumably fundamental asset values should not fluctuate much, if at all, at the intraday level.

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, De, Hu, and Hwang (2014) conduct textual analysis of articles published on *seekingalpha.com*. Compared to 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, Korajczyk, and Sadka (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, Han, Li, and Zhou (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. This predictability is evident based on both in-sample and out-of-sample statistical metrics. 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,

The rest of this paper is organized as follows. The next section describes the data. Section III documents the empirical relation between high-frequency investor sentiment and intraday S&P 500 returns using predictive regressions. Section IV 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 V evaluates the economic significance of investor sentiment using market timing trading strategies. Section VI provides some concluding remarks.

II. Data Description

Our intraday stock returns data are obtained from QuantQuote. The data vendor provides us with minute-by-minute price data starting from January 1998. Consistent with GHLZ (2015), we focus on the intraday returns of S&P 500 index ETF (ticker symbol SPY). According to Wikipedia, "for a long time, this fund was the largest ETF in the world." According to Yahoo!Finance, its average daily trading volume is more than 147 million shares. Due to arbitrage forces, its intraday price movements are almost identical to the underlying cash index as well as the S&P 500 index futures. We first filter out price data outside the regular trading hours. Then we convert 1-minute price data to half hour returns. Thus we obtain thirteen half hour returns for each trading day.

Following GHLZ, the half hour returns are calculated on the close-to-close basis. Therefore the first half hour return should contain information during the overnight hours after the closing of the previous day. For example many of the economic news are released at 8:30 am, which is probably why GHLZ find the first half hour return informative. GHLZ show that both the first and (to a lesser extent) the 12-th half hour returns can predict last half hour return. In this paper, our focus is broader than GHLZ. We are interested in knowing whether lagged changes in high frequency investor sentiment have predictive values for future market returns throughout the whole trading day, rather than just the last half hour.

Our intraday sentiment measure is based on the proprietary Thomson Reuters MarketPsych Indices (TRMI). Its website (**https://www.marketpsych.com/data/**) provides the following summary about the data.[1]

> "We have the world's most comprehensive finance-specific sentiment data, covering all major countries, currencies, commodities, equity sectors, and individual US and non-US equities. The data is produced by distilling a massive collection of news and social media content though through an extensively curated language framework, which not only measures different emotions (optimism, confusion, urgency etc.), but also financial language (price forecasts etc.) and specific topics (interest rate, mergers etc.). TRMI is produced from 1998 to present, on both a daily and minutely basis."

The construction of TRMI relies on the use of a patent-pending system to score sentiment-laden content in text. According to its documentation, TRMI's text analytic techniques are designed to score business-specific language for quantitative financial applications. This is important because, as shown by Loughran and McDonald (2011), word lists that is not unique to finance might not correctly reflect tone in the financial context.

TRMI's entire content set includes millions of articles and posts each day from various sources such as news wires, internet news sources, and social media. Due to the wide variety of content sources, cares have to be given when interpreting words or symbols that express emotions. For example, in social media someone may an acronym such as "LOL" that is unlikely to be used in a formal news press lease. To account for these differences, TRMI uses differentiated models for news, social media forums, tweets, SEC filings, as well as earnings conference call transcripts.

[1] The website also contains a more detailed description of the data in its user guide section. Please see **https://www.marketpsych.com/guide/**.

Although TRMI covers a wide range of asset markets including equities, currencies, and commodities, we only have access to sentiment data that is linked to the S&P 500 index (SPY) and for the sample period from 1998 to 2011. The data granularity is at minute level and covers twenty-four hours a day. More specifically, the TRMI sentiment measure provides the 24 hour rolling average score of total references in news and social media by counting overall positive references net of negative references. The scores are normalized so that its value ranges from -1 to 1. To match with our returns data, we convert the minute level data to half hour frequency and only focus on the regular trading hours. Perhaps due to the use of rolling average, we find that the levels of investor sentiment are highly persistent, with first order autocorrelation larger than 99%. It is well known that highly persistent variables can caused biased inferences in predictive regressions. Therefore we focus on the lagged *changes* in investor sentiment rather than its levels.

While the TRMI sentiment data is based on textual analysis, we also evaluate the use of alternative market-based sentiment proxies as control variables. We focus on two measures: the Chicago Board Options Exchange (CBOE) volatility index (VIX) and market state. The VIX is inferred from S&P 500 index option implied volatility data and is widely regarded by the market as the so-called "fear index". For example, Connolly, Stivers, and Sun (2005) show that heightened levels of the VIX is associated with episodes of high stock market uncertainty and related to the flight-to-quality effect between stocks and bonds. Since intraday VIX data are not available to us, we rely on lagged values of the VIX from the previous trading day. Cooper, Gutierrez, and Hameed (2004) find that momentum profits occur exclusively in the UP market state, defined as the time periods when lagged three-year or one-year aggregate market return is non-negative. They attribute this result to increased levels of investor overconfidence in the UP market states. Since lagged market states are associated with investor overconfidence (or lack thereof), it may be viewed as a measure of investor sentiment as well. In this article we define the current market state as UP (DOWN) if the lagged return of the S&P 500 index during the past 250 trading days (approximately one year) is non-negative (negative).

In addition to the sentiment measures, we also look at the role of lagged macroeconomic variables in terms of predicting intraday return of the stock market. We focus on the following set of macroeconomic variables: term premium (TERM) defined as the spread between yields on 10-year and 1-year Treasury notes, default spread (DEF) defined as the spread between Moody's BAA and AAA rated corporate bond yields, and short-term risk-free interest rate (Rate) using 3-month Treasury bills. All macroeconomic data are obtained from St. Louis Federal Reserve Bank. Since these macroeconomic variables are not available in intraday frequency, we use the daily values from the previous trading day.

In Figure 1 we plot the means and standard deviations of SPY intraday returns as well as changes in investor sentiment in each of the half hours. The sample period is from January 2, 1998 to December 31, 2011. We note that for the SPY intraday returns, the mean returns are very close to zero. The volatility on the other hand seems to exhibit an interesting U-Shape pattern. The market is very volatile during the first half hour, appears to calm down during the lunch hours, and picks up steam again near the end of a trading day. Change in sentiment and especially its volatility also appear to be large during the first half hour, but it flattens out during the rest of the day.

III. Main Empirical Evidence

Following GHLZ (2015) among many others in the literature, we focus on the predictive regression approach to assess the predictive value of lagged high-frequency investor sentiment for intraday stock market returns.

A. Predictive Regressions with High-Frequency Investor Sentiment

First, we evaluate if lagged changes in investor sentiment can predict the future half-hour SPY returns of a given trading day. More formally, we consider the following predictive regression model:

$$r_{i,t} = \beta_0 + \beta_1 \Delta s_{i-1,t} + \epsilon_t, \quad i = 2, \ldots, 13, \tag{1}$$

where $r_{i,t}$ is the i-th half-hour return on the S&P 500 index ETF on day t , and $\Delta s_{i-1,t}$ denotes the change in investor sentiment in the $(i-1)$-th half-hour. Note that we do not run this regression for the first half-hour market return because that will require the use of changes in investor sentiment from a previous trading day, but our focus in this paper is on predictive value of intraday investor sentiment.

Panels A Table I report results for the full sample period from 1998 to 2011. Newey and West (1987) robust t-statistics are shown in parentheses underneath the estimated coefficients.

In panel A, we find striking evidence that the change in investor sentiment variable Δs is significant in all half-hour periods except for the second half hour. In fact, other than the 3rd and 6th half hours where the significance are "only" at 10% level, it is significant at 1% levels in all other cases. Most notably, for the last four half-hour periods, the t-statistics are all larger than 5. More interestingly, we find that the signs of estimated coefficient for the Δs are positive in all predictive regressions. The closer we are to the end of a trading day, the higher the estimated values of the coefficient β_1 become. We notice that the R^2 values also increase with the passage of trading hours, and eventually exceed the 1% level for the last two hours of a trading day. The

R^2 values are 1.02%, 1.63%, 1.68%, and 1.43% for the 10-th, 11-th, 12-th, and last half hour, respectively.

We note that these R^2 values are comparable to those reported by GHLZ (2015) on intraday momentum. In their study, GHLZ report that their most powerful predictor r_1, first half-hour return, has a R^2 value of 1.6%. Their other predictor r_{12}, the 12-th half-hour return, has a R^2 value of 1.1%. GHLZ also note that these numbers match or even exceeds the typical predictive R^2 at the monthly frequency found in the literature.

In Panels B and C of Table I, we report the results for two sub-sample periods: the NBER-dated Recessions and Expansions. During our full sample period from 1998 to 2011, the U.S. economy has experienced two recessions: (1) from March 2001 to November 2001, and (2) from December 2007 to June 2009.[2]

Panel B shows that Δs is significant for the 4-th, 7-th, and the last four half hours of a trading day during the recessions. In terms of R^2 values, it exceeds the 1% level during 7-th and the last three half hours. Similar results from Panel C indicates that during expansions, the change in sentiment variable is significant for the 4-th, 5-th, as well as the last five half hours of a trading day. In terms of R^2 values, it exceeds the 1% level during the last four half hours.

Overall we find the signs for estimates of Δs stay positive in all cases. In terms of t-statistics, the sentiment effect appears stronger during economic expansions than recessions. Since recessions (expansions) are more likely to be associated with low (high) investor sentiment, we conclude that this business-cycle related phenomenon appears consistent with the noise trader model of Yu and Yuan (2011), which predicts that noise traders are more likely to participate in the market when sentiment is high.

The evidence presented in Table I suggests much stronger sentiment effect later

[2]NBER business cycle dates are obtained from http://www.nber.org/cycles.html.

in the day than during morning hours of a trading day. We conjecture that this is likely due to the fact that most noise traders from the west coast of the U.S. are not showing up during the morning hours, thanks to differences in time zones. Since the sentiment effect is strongest during the last four half-hour periods, our subsequent analysis will focus exclusively on them.

B. Predictive Regressions with High-Frequency Investor Sentiment and Lagged Returns

Next we run predictive regression where both lagged changes in sentiment and lagged returns are included as predictors. The inclusion of lagged returns is motivated by the study on intraday momentum by GHLZ (2015).

$$r_{i,t} = \beta_0 + \beta_1 \Delta s_{i-1,t} + \beta_2 r_{1,t} + \beta_3 r_{i-1,t} + \epsilon_t, \quad i = 10, \ldots, 13, \tag{2}$$

where $r_{i,t}$ is the i-th half-hour return on the S&P 500 index ETF on day t. Based on the results from Table I, we focus on the last four half-hour returns. $r_{1,t}$ is the first half-hour return, which GHLZ find to be a powerful predictor of the last half-hour return. Since the returns are calculated on close-to-close basis, r_1 actually contains useful information from overnight. $\Delta s_{i-1,t}$ and $r_{i-1,t}$ are the lagged change in investor sentiment and lagged half-hour return respectively.

Panel A of Table II reports that results for the full sample period. Panels B and C of the same table report results for the NBER-dated Recessions from March 2001 to November 2001 and from December 2007 to June 2009, and the NBER-dated Expansions, respectively.

We find strong evidence that lagged changes in investor sentiment is the most predominant predictor even in the present of the lagged return variables. For example, the estimated coefficients on Δs are significant in all cases. In fact they are highly significant at the 1% level both for the full sample and during the NBER-dated expansions. The results are slightly weaker during recession. But even in the weakest case, which is for the last half-hour during recessions, it is still significant at the 10% level.

In addition, we find that our results from the predictive regression for the last half-hour returns are remarkably consistent with the findings from GHLZ (2015). For example, based on a sample from 1993 to 2013, GHLZ report that in a predictive regression where the two lagged returns are the only two predictors, the estimated coefficient on r_1 is 0.068 with a Newey-West t-statistic of 4.14. Likewise we find that the estimated coefficient on r_1 is 0.066 with a Newey-West t-statistic of 3.67. Note that we include the change in sentiment variable in the predictive regression and our sample is from 1998 to 2011. The results on the other predictor r_{12} are also very similar. GHLZ report a coefficient of 0.114 with a t-statistic of 2.60. We find the estimated coefficient to be 0.101 with a t-statistic of 2.10. GHLZ also report an R^2 value of 2.6% with the two lagged return predictors. We report an adjusted R^2 of 3.66% with the addition of our sentiment variable.

However, we also find that the predictive power of the lagged return variables are only limited to the last-half hour. In fact, for the 10-th, 11-th, and 12-th half-hour predictive regressions, none of the lagged return variables are significant. The only exception is the r_9 variable in the 10-th half hour during NBER-dated expansions, which is significant at the 5% level. However, its estimated coefficient is negative, which is inconsistent with the interpretation of intraday momentum as argued by GHLZ.

Thus the overall evidence suggests that the sentiment effect is distinct from the intraday momentum effect. We find overwhelming evidence that impact from changes in investor sentiment on intraday market returns is much more pervasive than lagged return variables. The intraday momentum effect of GHLZ appears to exist only for the last half hour, whereas the sentiment effect persists in at least the last two hours of a trading day.

IV. Robustness Checks

In this section, we check the robustness of the sentiment effect by looking at its seasonality both monthly and by weekdays. We also check the roles of macroeconomic variables, alternative measures of investor sentiment, as well as the out-of-sample performance of predictive regressions.

A. Seasonality

Seasonality is often an important trait of an asset pricing anomaly. For example, Cooper, McConnell, and Ovtchinnikov (2006) document that there exists a so-called the Other January effect for the US market indexes to differentiate it from the well-known small-firm in January effect. Da et al. (2015) also find that their sentiment index derived from Google search volumes exhibits weekly seasonality. Thus to alleviate concerns about seasonality, we first run the following predictive regression by calendar months to see if our results are driven by any particular month.

$$r_{i,mt} = \beta_0 + \beta_1 \Delta s_{i-1,mt} + \beta_2 r_{1,mt} + \beta_3 r_{i-1,mt} + \epsilon_{mt}, \quad i = 10, \ldots, 13, \tag{3}$$

where $m = January, February, \ldots, December$.

Table III report the results based on calendar months. We find that the change in sentiment variable is significant for a majority number of months in a year. For example, it is significant from April to October for both the 10-th and 11-th half hours, from May to October as well as in January and December for the 12-th half hour, and from May to August as well as in January and October for the Last half hour. Interestingly we find that the sentiment variable is especially strong during the summer months when there is plenty of sunshine. This appears consistent with the hypothesis of Hirshleifer and Shumway (2003) that more sunshine leads to upbeat investor sentiment, and noise traders are more likely to participate in the market when their sentiment levels are elevated (Yu and Yuan (2011)).

In contrast, we find that the lagged return variables are mostly insignificant for the 10-th, 11-th, and 12-th half hours. For the last half hour, the r_1 variable is significant in January, March, May, and November. The r_{12} variable is significant in March, July, August, and November. Interestingly, we note that the intraday momentum effect is particular strong in November with an adjusted R^2 of 18.83%.

Next we run the following predictive regression to detect seasonality by weekdays.

$$r_{i,wt} = \beta_0 + \beta_1 \Delta s_{i-1,wt} + \beta_2 r_{1,wt} + \beta_3 r_{i-1,wt} + \epsilon_{wt}, \quad i = 10, \ldots, 13, \tag{4}$$

where $w = Monday, Tuesday, \ldots, Friday$.

We report the results in Table IV. Once again we find that the impact from investor sentiment is pervasive. It is significant in all weekday except on Thursday in the 10th half hour, significant from Tuesday to Friday in the 11th half hour, Monday to Friday for the 12th half hour, and from Tuesday to Thursday in the last half hour. In contrast, the lagged return variables are mostly insignificant (or significant but have a negative sign) for the 10th, 11th, and 12th half hours. For last half hour, the lagged returns appear to have gained significance. The r_1 variable is significant for Monday, Wednesday, and Friday. r_{12} however is only significant on Tuesday.

To sum up, we find that the sentiment effect is a widespread phenomenon. It shows up during majority of the months and almost every weekday. In contrast, we find that the intraday momentum effect seems to be less pervasive.

B. The Role of Macroeconomic Variables

We investigate the impact from a standard set of macroeconomic variables. Since intraday macroeconomic variables are typically not feasible, we rely on macroeconomic variables that are available at daily frequency. This leaves us with three commonly used macroeconomic variables: term premium (Term), default spread (Def), and short-term risk-free interest rate (Rate). For formal definitions of these variables, please see section II.

Table V reports results based on the following predictive regression model where lagged values (from the previous trading day) of the three macroeconomic variables are included:

$$r_{i,t} = \beta_0 + \beta_1 \Delta s_{i-1,t} + \beta_2 r_{1,t} + \beta_3 r_{i-1,t} + \beta_4 Term_{t-1} + \beta_5 Def_{t-1} + \beta_6 Rate_{t-1} + \epsilon_t, \quad (5)$$

where $i = 10th, \ldots, 13th$ half hour.

We find that none of these macroeconomic variables can explain intraday sentiment effect. The only case that a macroeconomic variable is significant is the default spread during the 11th half hour. However even in that case, the estimated coefficient is extremely small (0.001). In all other cases, the estimated coefficients on macroeconomic variables are indistinguishable from zero.

C. Alternative Measures of Investor Sentiment

Next we investigate whether our main results are robust to the inclusion of alternative measures of investor sentiment. We consider three alternative measures of sentiment: the CBOE volatility index (VIX), change in VIX, and market state as defined by the sign of lagged 250-day (one-year) market return. Similar to the case of macroeconomic variables, we include only lagged values from the previous trading day in the predictive regressions. For details about the construction of these variables, please refer to section II.

Specifically, we estimate the following predictive regression models:

$$r_{i,t} = \beta_0 + \beta_1 \Delta s_{i-1,t} + \beta_2 r_{1,t} + \beta_3 r_{i-1,t} + \beta_4 VIX_{t-1} + \beta_5 \Delta VIX_{t-1} + \beta_6 State_{t-1} + \epsilon_t, \quad (6)$$

where $i = 10th, \ldots, 13th$ half hour. VIX is the CBOE Volatility Index, ΔVIX refers to the change in VIX, and $State$ is the Market State variable.

Results from Table VI show that interestingly at least for the last hour of a trading day, these alternative sentiment measures do have some explanatory power. For example, in the 12th half hour, we find that lagged VIX and market state are significant at the 5% level. In the last half hour, the change in VIX appears highly significant at the 1% level.

In terms of adjusted R^2 values, the alternative sentiment measures also appear to be useful. For the 12th half hour, the adjusted R^2 increases from about 1.67% in Panel A of Table II to about 2.93%. For the last half hour, with ΔVIX, the adjusted R^2 also increases from about 3.66% in Panel A of Table II to about 4.76% in the current table.

However, in the other two half hours, these alternative measures do not appear to carry any significance. In contrast, our intraday sentiment variables remains highly significant in all cases.

D. Out-of-Sample Forecasting Performance

The previous tables present convincing evidence about the intraday sentiment variables' in-sample predictive power. However as emphasized by Welch and Goyal (2008) and, in the case of intraday market return, GHLZ (2015) among many others, it is also important to verify the out-of-sample predictability of the predictors. To this end, we evaluate the out-of-sample forecasting performance of our models with the following procedure.

To begin with, consider a sample of size T. We use the first M observations as our initial estimation window. In our empirical implementation, we choose $M = 250$ days. The predictive regressions are run using the first M observations and the estimated coefficients are retained. Starting from day $M + 1$, we calculate the forecasting errors by comparing the actual intraday return r_i at half hour i with the model predicted returns $\hat{r}_i$. The process is iterated until the end of sample is reached. The total of $T - M$ forecasting errors are then used to calculate various forecasting performance evaluation metrics. We consider the following four metrics: (1) Out-of-sample R^2 (OOS R^2); (2) Root Meant Squared Error (RMSE); (3) Mean Absolute Error (MAE); and (4) Theil U Statistic.

More formally, they are defined as follows.

$$OOS \quad R^2 = 1 - \frac{\sum_{t=M+1}^{T} (r_{i,t} - \hat{r}_{i,t})^2}{\sum_{t=M+1}^{T} (r_{i,t} - \bar{r}_{i,t})^2}, \tag{7}$$

where $\bar{r}_{i,t}$ is the sample mean of $r_{i,t}$, the $i - th$ half hour returns using the data up to $t - 1$.

$$RMSE = \sqrt{\frac{1}{T - M} \sum_{t=M+1}^{T} (r_{i,t} - \hat{r}_{i,t})^2}. \tag{8}$$

$$MAE = \frac{1}{T - M} \sum_{t=M+1}^{T} |r_{i,t} - \hat{r}_{i,t}|. \tag{9}$$

$$Theil \quad U = \sqrt{\frac{\sum_{t=M+1}^{T}(r_{i,t} - \hat{r}_{i,t})^2}{\sum_{t=M+1}^{T} r_{i,t}^2}}. \tag{10}$$

We choose OOS R^2 as this is also the preferred metric used by GHLZ. In addition, it is easy to compare its value with its in-sample counterpart. RMSE and MAE are chosen because they are probably the most commonly used metrics to evaluate out-of-sample forecasting accuracy. One drawback of RMSE and MAE is that they are not scale-free. Theil U statistic overcomes this drawback.

In Table VII we report the out-of-sample forecasting metrics using four models derived from equation (6) as follows:

- Model 1: we do not include any predictors other than the constant. Therefore this is a forecasting model based on the sample mean. We use this as a benchmark model.
- Model 2: we include only the predictor Δs.
- Model 3: we use the best in-sample model based on adjusted R^2 from Table VI.
- Model 4: we include all the predictors from equation (6).

We report the results for the 10th, 11th, 12th and last half hours. Note that for the 10th and 11th half hours, Model 2 coincides with Model 3 as Δs is the best in-sample model.

We find that the results based on the out-of-sample forecasting metrics are mostly in agreement with the in-sample results. For example, for the 10th and 11th half hours, all four metrics agree that Model 2 (in this case coincides with Model 3) is the best except for the MAE metric in the 10-th half hour, which finds that the sample mean model is the best. For the 12th and the last half hours, all four metrics agree that Model 3 is the best except for MAE in the 12th half hour, which again prefers the sample mean model.

Overall, we conclude that the results based on out-of-sample forecasting performance metrics are consistent with our in-sample results and suggest that the predictability by intraday sentiment continues to be valid when evaluated out of sample.

V. Market Timing Trading Strategies

In this section, we evaluate the performance of marketing timing trading strategies that are built upon the predictive regressions. We first describe these trading strategies and then evaluate their performance based on two popular metrics: Sharpe ratio and certainty-equivalent return.

A. Strategy Formulation

Our results so far have established the predictability of intraday returns by high-frequency investor sentiment from a statistical perspective. In addition to statistical significance, it is also important to evaluate the economic value of investor sentiment when implemented on some trading strategies. To this end, we propose the following market timing trading strategies.

To begin with, consider the trading strategies used by GHLZ. Since they focus on the return variables only, GHLZ's strategy takes a long position in the SPY if r_1, the first half hour return, is positive, and a short position if r_1 is negative. In the case of two predictors r_1 and r_{12}, GHLZ's strategy will take a long (short) position if and only if both predictors are positive (negative). While this is a simple strategy to implement, it cannot be easily extended to cases where there are more than two predictors or when non-return predictors are used. For example, GHLZ report that their trading strategy with two predictors r_1 and r_{12} has a success rate of 77% compared to the 54% by the strategy that uses only r_1. However, the mean return of the two-predictor strategy is only 4.49% compared to the 6.19% by the one-predictor strategy. Presumably this is because the strategy with two predictors is more likely to stay out of the market when the two predictors give conflicting signals. Now if we extend their approach to three or four predictors, then the strategies are likely to stay out of the market for longer periods of time and therefore make them ineffective.

Instead we propose a market timing strategy where the setup is similar to the out-of-sample predictive regressions described in the previous section. This allows us to implement the strategy where the number of predictors could be large and they are not necessarily return-based predictors. Specifically, given a sample of size T and a set of predictors that includes a constant $X \equiv [1, x_1, x_2, \ldots, x_n]$, we use an initial estimation window size of M, where $M < T$, and run the regression $r_i = X'\beta + \varepsilon$ to obtain the coefficient estimates $\widehat{\beta}$, where r_i is the market return at i-th half hour. Combined with the new observations at $t = M + 1$ for the predictors, we can proceed to calculate the model predicted value for r_i at $t = M + 1$, namely $\hat{r}_{i,M+1} = \widehat{\beta}' X_{M+1}$. This process is iterated by adding one more observation each time until we reach the end of the sample. Our trading signals are generated as follows: take a long position in the S&P 500 ETF if the predicted return for the i-th half hour is positive and otherwise take a short position. More formally, the strategy's return on day t based on the vector of predictors X is calculated as follows:

$$\eta(X) = \begin{cases} r_i & \text{if } \hat{r}_i > 0, \\ -r_i & \text{if } \hat{r}_i \leq 0. \end{cases} \tag{11}$$

These returns are then used to calculate the performance evaluation metrics for the sample period from $t = M + 1$ to $t = T$.

Similar to the out-of-sample predictive regressions, we consider the following trading strategies based on four predictive models: a benchmark model based on the sample mean, a model with Δs as the only predictor, the best in-sample predictive model, and the model with all predictors from Table VII. Note that the benchmark strategy is equivalent to the case where only the constant term is included in the predictive regression. Thus from a Bayesian perspective, investors who have a dogmatic prior belief that none of the predictors are useful should implement this benchmark strategy.

B. Performance Evaluation Metrics

We consider two performance evaluation metrics. The first is the Sharpe ratio defined as follows:

$$\widehat{SR}_j = \frac{\hat{\mu}_j - r_f}{\hat{\sigma}_j}, \tag{12}$$

where $\hat{\mu}_j$ and $\hat{\sigma}_j$ are the sample mean and standard deviation of returns for strategy j and r_f is the risk-free rate. Following GHLZ, we set daily risk-free rate to zero.

To test whether the Sharpe ratios of different strategies are statistically distinguishable, we follow DeMiguel, Garlappi, and Uppal (2009) (DGU) and report the p-value of the difference of the Sharpe ratios, using the approach suggested by Jobson and Korkie (1981) with adjustments proposed by Memmel (2003). Specifically, given two trading strategies j and k, denote the estimated means, standard deviations, and covariance of portfolio returns over a sample of size T as follows: $\hat{\mu}_j$, $\hat{\mu}_k$, $\hat{\sigma}_j$, $\hat{\sigma}_k$, and $\hat{\sigma}_{j,k}$ respectively. For the null hypothesis $H_0 : \hat{\mu}_j/\hat{\sigma}_j - \hat{\mu}_k/\hat{\sigma}_k = 0$, one can use the test statistic:

$$\hat{z}_{j,k} = \frac{\hat{\sigma}_k\hat{\mu}_j - \sigma_j\hat{\mu}_k}{\sqrt{\zeta}}, \tag{13}$$

where

$$\zeta = \frac{1}{T}\left(2\hat{\sigma}_j^2\hat{\sigma}_k^2 - 2\hat{\sigma}_j\hat{\sigma}_k\hat{\sigma}_{j,k} + \frac{1}{2}\hat{\sigma}_k^2\hat{\mu}_j^2 + \frac{1}{2}\hat{\sigma}_j^2\hat{\mu}_k^2 - \frac{\hat{\mu}_j\hat{\mu}_k}{\hat{\sigma}_j\hat{\sigma}_k}\hat{\sigma}_{j,k}^2\right) \tag{14}$$

which is asymptotically distributed as a standard normal.

The second performance evaluation metric is the certainty-equivalent return (CEQ) defined as the risk-free rate that an investor is willing to accept rather than adopting a particular risky strategy. Following DGU, it is computed as follows:

$$\widehat{CEQ}_j = \hat{\mu}_j - \frac{\gamma}{2}\hat{\sigma}_j^2, \tag{15}$$

where γ is the risk-aversion parameter. In our empirical results, we set $\gamma = 3$. This choice is based on a recent study by Guiso, Sapienza, and Zingales (2014), who find that empirically the average risk aversion increases from 2.87 before the 2008 financial crisis to 3.28 after the crisis.

The CEQ metric defined in this manner actually refers to the level of expected utility of a mean-variance investor. DGU note that this is approximately the CEQ of an investor with quadratic utility. Therefore it measures the economic significance of a given strategy. To test whether the CEQ returns from two strategies are statistically different, we can compute the p-value of the difference following the procedure described in DGU. Specifically, given two trading strategies j and k, let v denote the vector of moments $v = (\mu_j, \mu_k, \sigma_j^2, \sigma_k^2)$, $\hat{v}$ its empirical counterpart obtained from a sample of size T. Further denote $f(v) = (\mu_j - \frac{\gamma}{2}\sigma_j^2) - (\mu_k - \frac{\gamma}{2}\sigma_k^2)$ the difference in CEQ of the two strategies. DGU note that the asymptotic distribution of $f(v)$ is $\sqrt{T}(f(\hat{v}) - f(v)) \to N\left(0, \frac{\partial f}{\partial v}^T \Psi \frac{\partial f}{\partial v}\right)$, where

$$\Psi = \begin{pmatrix} \sigma_j^2 & \sigma_{j,k} & 0 & 0 \\ \sigma_{j,k} & \sigma_k^2 & 0 & 0 \\ 0 & 0 & 2\sigma_j^4 & 2\sigma_{j,k}^2 \\ 0 & 0 & 2\sigma_{j,k}^2 & 2\sigma_k^4 \end{pmatrix}$$

Table VIII report the annualized mean, standard deviation, Sharpe ratio, and

CEQ for each strategy under consideration. Our initial estimation window size is $M = 250$ days. Panels A, B, C, and D report the results for the 10th, 11th, 12th, and the last half hour returns respectively.

Interestingly we find that the strategies have remarkably similar volatilities. For example, for the 10th half-hour returns, the volatility of all strategies are around 4.66%. For the last half hour, they are around 6.5%. On the other hand, the performance of each strategy shows significant disparity in terms of their mean returns. For example, in the 11th half hour, the strategy that uses Δs as the only predictor earns about 9.43%. In contrast, the benchmark strategy based on sample mean only has a mean return of approximately 1%.

Consistent with the in-sample and out-of-sample predictive regression results, we find that the best performer for the 10th and 11th half hours is the strategy that relies only on the predictive power of Δs. This conclusion is based on mean return, Sharpe ratio, as well as CEQ. Surprisingly, we find that this strategy is also the best performer in the 12th half hour based on the rankings of mean return, Sharpe ratio, and CEQ. Recall from Tables VI and VII that the best predictive regression model in this case is the one with Δs, lagged VIX, and market state as predictors. However, it turns out that its performance lags behind the strategy where only Δs is used.

Results for the last half hour, however, is consistent with the in-sample and out-of-sample predictive regressions. We find that the best strategy is the one that uses changes in investor sentiment, lagged returns, as well as changes in VIX as predictors. The difference in performance however is not big. For example, when only Δs is used, its mean return is 8.34%, its Sharpe ratio is 1.28, and its CEQ 7.71%. This is comparable to the best strategy where the numbers are 9.19% (mean return), 1.42 (Sharpe ratio), and 8.56% (CEQ) respectively. Across the board, the worse strategy is the benchmark sample mean model.

Overall, the evidence from Table VIII confirms the economic value of utilizing the information from Δs to predict intraday market returns.

VI. Conclusion

In this article, we examine the predictability of intraday market return with changes in high-frequency investor sentiment. Our intraday investor sentiment measure is based on the proprietary Thomson Reuters MarketPsych index, which provides a commercial strength comprehensive textual analysis of various traditional as well social media sources. Our empirical findings can be summarized as follows.

First, we find strong evidence that changes in investor sentiment have predictive values for the intraday market returns. The predictability permeates throughout the whole trading day, but is particularly strong during the last two hours. In contrast, we find that the intraday momentum effect based on lagged returns are only significant for the last half hour. Thus the sentiment effect appears more pervasive than the intraday momentum effect. Moreover, the predictability is robust based on both in- and out-of-sample statistical metrics. We also find that intraday investor sentiment carries significant economic value as measured by both Sharpe ratio and CEQ of returns from market timing trading strategies.

Second, predictive value of high-frequency investor sentiment cannot be explained by macroeconomic variables or alternative sentiment measures.

Third, we also find interesting evidence that the sentiment effect appears stronger during the summer months as well as during economic expansions. We conjecture that this might be due to the fact that investors' mood is more upbeat during the summer months when there is ample sunshine (Hirshleifer and Shumway (2003)) and noise traders' participation in the market increases when sentiment level is more elevated (Yu and Yuan (2011)).

Our findings have important asset pricing implications. While prior studies have documented the impact from investor sentiment on asset prices, especially for small stocks, it has not been a consensus that sentiment can predict aggregate market index returns. We, on the other hand, provide strong evidence that indeed market index returns are predictable with investor sentiment at least at the intraday level. We attribute the predictability to the fact that, unlike prior studies that focus on a specific source, our sentiment measure is much more broadly based and constructed from all-encompassing traditional and social media sources. Thus it is able to pick up changes in investor sentiment with accuracy. With this improved measure of investor sentiment, our future research will continue to focus on the impact of investor sentiment on other important aspects of asset pricing, such as predictability in the medium and long horizons.

REFERENCES

Antoniou, Constantinos, John A. Doukas, and Avanidhar Subrahmanyam, 2013, Cognitive dissonance, sentiment, and momentum, *Journal of Financial and Quantitative Analysis* 48, 245–275.

Antweiler, Werner, and Murray Z. Frank, 2004, Is all that talk just noise? The information content of internet stock message boards, *The Journal of Finance* 59, 1259–1295.

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *Journal of Finance* 61, 1645–1680.

———, 2007, Investor sentiment in the stock market, *Journal of Economic Perspectives* 21, 129–157.

Chen, Hailiang, Prabuddha De, Yu Jeffrey Hu, and Byoung-Hyoun Hwang, 2014, Wisdom of crowds: The value of stock opinions transmitted through social media, *Review of Financial Studies* 27, 1367–1403.

Connolly, Robert, Chris Stivers, and Licheng Sun, 2005, Stock market uncertainty and the stock-bond return relation, *Journal of Financial and Quantitative Analysis* 40, 161–194.

Cooper, Michael, Roberto C. Gutierrez, and Allaudeen Hameed, 2004, Market states and momentum, *Journal of Finance* 59, 1345–1365.

Cooper, Michael J., John J. McConnell, and Alexei V. Ovtchinnikov, 2006, The other january effect, *Journal of Financial Economics* 82, 315–341.

Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2015, The sum of all FEARS investor sentiment and asset prices, *The Review of Financial Studies* 28, 1–32.

De Long, J. Bradford, Andrei Shleifer, Lawrence Summers, and Robert J. Waldmann, 1990, Noise trader risk in financial markets, *Journal of Political Economy* 98, 703–738.

DeMiguel, Victor, L. Garlappi, and R. Uppal, 2009, Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?, *Review of Financial Studies* 22, 1915–1953.

Gao, Lei, Yufeng Han, Sophia Zhengzi Li, and Guofu Zhou, 2015, Intraday momentum: The first half-hour return predicts the last half-hour return, Working paper, Washington University at St. Louis.

Garcia, Diego, 2013, Sentiment during recessions, *The Journal of Finance* 68, 1267–1300.

Guiso, L., P. Sapienza, and L. Zingales, 2014, Time varying risk aversion, Working paper, University of Chicago.

Heston, Steven L., Robert A. Korajczyk, and Ronnie Sadka, 2010, Intraday patterns in the cross-section of stock returns, *Journal of Finance* 65, 1369–1407.

Hirshleifer, David, and Tyler Shumway, 2003, Good day sunshine: Stock returns and the weather, *The Journal of Finance* 58, 1009–1032.

Jobson, J. D., and B. Korkie, 1981, Performance hypothesis testing with the sharpe and treynor measures, *Journal of Finance* 36, 889–908.

Lee, Charles, Andrei Shleifer, and Richard Thaler, 1991, Investors sentiment and the closed-end fund puzzle, *Journal of Finance* 46, 75–109.

Lemmon, Michael, and Evgenia Portniaguina, 2006, Consumer confidence and asset prices: Some empirical evidence, *Review of Financial Studies* 19, 1499–1529.

Loughran, Tim, and Bill McDonald, 2011, When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks, *Journal of Finance* 66, 67–97.

Memmel, C., 2003, Performance hypothesis testing with the sharpe ratio, *Finance Letters* 1, 21–23.

Newey, Whitney K., and Kenneth D. West, 1987, A simple positive semi-definite, heteroskedasticity and aurocorrelation consistent covariance matrix, *Econometrica* 51, 1233–1242.

Qiu, Lily, and Ivo Welch, 2006, Investor sentiment measures, Working paper, Brown University.

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

Welch, Ivo, and Amit Goyal, 2008, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455–1508.

Yu, Jianfeng, and Yu Yuan, 2011, Investor sentiment and the mean-variance relation, *Journal of Financial Economics* 100, 367–381.

Table I
Predictability of Intraday Stock Market Returns with Lagged Change in Investor Sentiment

This table reports results from the following predictive regressions:

$$r_{i,t} = \beta_0 + \beta_1 \Delta s_{i-1,t} + \epsilon_t, \quad i = 2, \ldots, 13,$$

where $r_{i,t}$ is the i-th half-hour return on the S&P 500 index ETF on day t , and $\Delta s_{i-1,t}$ denotes the change in investor sentiment in the $(i-1)$-th half-hour. Panels A, B, and C report results for three periods: the whole sample period, the NBER-dated Recessions from March 2001 to November 2001 and from December 2007 to June 2009, and the NBER-dated Expansions, respectively. Newey and West (1987) robust t-statistics are in parentheses and significance at the 1%, 5% , or 10% level is indicated by an ***, an ** or an *, respectively. The sample period is from January 2, 1998 to December 31, 2011.

Panel A: Full Sample Period from 1998 to 2011			
Half-hour Return Period	β_0	β_1	R^2(%)
2^{nd} Half-hour	-0.000	0.001	0.00
	(-0.14)	(0.41)	
3^{rd} Half-hour	-0.000	0.040*	0.08
	(-0.42)	(1.70)	
4^{th} Half-hour	-0.000	0.072***	0.33
	(-1.04)	(3.21)	
5^{th} Half-hour	-0.000	0.049***	0.31
	(-0.32)	(2.81)	
6^{th} Half-hour	-0.000	0.045*	0.18
	(-0.11)	(1.91)	
7^{th} Half-hour	0.000	0.036***	0.20
	(0.52)	(2.65)	
8^{th} Half-hour	0.000*	0.049*	0.12
	(1.71)	(1.71)	
9^{th} Half-hour	-0.000	0.084***	0.31
	(-0.05)	(2.92)	
10^{th} Half-hour	-0.000	0.165***	1.02
	(-0.56)	(5.46)	
11^{th} Half-hour	0.000	0.213***	1.63
	(1.59)	(6.30)	
12^{th} Half-hour	0.000***	0.212***	1.68
	(3.26)	(6.52)	
The Last Half-hour	0.000	0.269***	1.43
	(0.00)	(5.35)	

Table I
(continued)

Panel B: Business Cycles - Recession			
Half-hour Return Period	β_0	β_1	R^2(%)
2^{nd} Half-hour	-0.000	0.005	0.05
	(-0.07)	(0.54)	
3^{rd} Half-hour	0.000	0.104	0.23
	(0.48)	(1.52)	
4^{th} Half-hour	-0.000*	0.116*	0.44
	(-1.85)	(1.73)	
5^{th} Half-hour	-0.000	0.033	0.09
	(-0.82)	(0.77)	
6^{th} Half-hour	0.000	0.093	0.23
	(0.30)	(1.32)	
7^{th} Half-hour	-0.000	0.166***	1.27
	(-0.32)	(3.15)	
8^{th} Half-hour	0.000	0.152	0.54
	(0.51)	(1.55)	
9^{th} Half-hour	-0.000	0.086	0.15
	(-1.06)	(0.98)	
10^{th} Half-hour	-0.000	0.211**	0.70
	(-0.20)	(2.25)	
11^{th} Half-hour	0.001***	0.394***	2.29
	(2.95)	(2.95)	
12^{th} Half-hour	0.000	0.412***	1.79
	(1.24)	(2.98)	
The Last Half-hour	0.000	0.476**	1.86
	(0.93)	(2.55)	

Table I
(continued)

Panel C: Business Cycles - Expansion			
Half-hour Return Period	β_0	β_1	$R^2(\%)$
2^{nd} Half-hour	-0.000	0.000	0.00
	(-0.07)	(0.09)	
3^{rd} Half-hour	-0.000	0.029	0.06
	(-0.97)	(1.16)	
4^{th} Half-hour	0.000	0.059**	0.28
	(0.14)	(2.57)	
5^{th} Half-hour	0.000	0.053***	0.42
	(0.16)	(2.77)	
6^{th} Half-hour	-0.000	0.039	0.19
	(-0.31)	(1.64)	
7^{th} Half-hour	0.000	0.019	0.08
	(1.04)	(1.31)	
8^{th} Half-hour	0.000*	0.027	0.05
	(1.79)	(0.97)	
9^{th} Half-hour	0.000	0.082***	0.41
	(0.78)	(2.88)	
10^{th} Half-hour	-0.000	0.156***	1.27
	(-0.55)	(5.09)	
11^{th} Half-hour	-0.000	0.176***	1.59
	(-0.27)	(5.85)	
12^{th} Half-hour	0.000***	0.180***	2.04
	(3.37)	(6.21)	
The Last Half-hour	-0.000	0.216***	1.45
	(-0.72)	(5.16)	

Table II
Predictability of Intraday Stock Market Returns with Lagged Change in Investor Sentiment and Lagged Returns

This table reports results from the following predictive regressions:

$$r_{i,t} = \beta_0 + \beta_1 \Delta s_{i-1,t} + \beta_2 r_{1,t} + \beta_3 r_{i-1,t} + \epsilon_t, \quad i = 10, \ldots, 13,$$

where $r_{i,t}$ is the i-th half-hour return on the S&P 500 index ETF on day t , $r_{1,t}$ is the first half-hour return, and $\Delta s_{i-1,t}$ denotes the change in investor sentiment in the $(i-1)$-th half-hour. Panels A, B, and C report results for three periods: the whole sample period, the NBER-dated Recessions from March 2001 to November 2001 and from December 2007 to June 2009, and the NBER-dated Expansions, respectively. Newey and West (1987) robust t-statistics are in parentheses and significance at the 1%, 5% , or 10% level is indicated by an ***, an ** or an *, respectively. The sample period is from January 2, 1998 to December 31, 2011.

Half-hour Return Period	β_0	β_1	β_2	β_3	Adj. R^2(%)
Panel A: Full Sample Period from 1998 to 2011					
10^{th} Half-hour	-0.000	0.170***	-0.001	-0.046	1.08
	(-0.55)	(5.52)	(-0.11)	(-1.14)	
11^{th} Half-hour	0.000	0.208***	0.002	0.026	1.60
	(1.58)	(5.63)	(0.22)	(0.62)	
12^{th} Half-hour	0.000***	0.207***	0.013	0.014	1.67
	(3.13)	(6.31)	(0.74)	(0.38)	
The Last Half-hour	-0.000	0.203***	0.066***	0.101**	3.66
	(-0.49)	(4.19)	(3.67)	(2.10)	
Panel B: Business Cycles - Recession					
10^{th} Half-hour	-0.000	0.201**	0.005	0.018	0.21
	(-0.16)	(2.14)	(0.24)	(0.24)	
11^{th} Half-hour	0.001***	0.367**	0.017	0.035	2.04
	(3.03)	(2.44)	(0.75)	(0.41)	
12^{th} Half-hour	0.000	0.398***	0.017	0.002	1.36
	(1.33)	(2.82)	(0.41)	(0.04)	
The Last Half-hour	0.000	0.308*	0.096**	0.130	5.15
	(0.97)	(1.82)	(2.53)	(1.48)	
Panel C: Business Cycles - Expansion					
10^{th} Half-hour	-0.000	0.163***	-0.005	-0.091**	1.78
	(-0.43)	(5.19)	(-0.52)	(-2.22)	
11^{th} Half-hour	-0.000	0.176***	-0.006	0.019	1.53
	(-0.21)	(5.74)	(-0.58)	(0.50)	
12^{th} Half-hour	0.000***	0.176***	0.009	0.017	1.99
	(3.22)	(6.01)	(0.70)	(0.39)	
The Last Half-hour	-0.000	0.180***	0.046***	0.069*	2.63
	(-1.17)	(4.42)	(2.99)	(1.74)	

Table III
Robustness Check: Monthly Seasonality

This table reports results from the following predictive regressions:

$$r_{i,mt} = \beta_0 + \beta_1 \Delta s_{i-1,mt} + \beta_2 r_{1,mt} + \beta_3 r_{i-1,mt} + \epsilon_{mt}, \quad i = 10, \ldots, 13,$$

where $m = January, February, \ldots, December$, r_i is the i-th half-hour return on the S&P 500 index ETF, r_1 is the first half-hour return, and Δs_{i-1} denotes the change in investor sentiment in the $(i-1)$-th half-hour. Newey and West (1987) robust t-statistics are in parentheses and significance at the 1%, 5% , or 10% level is given by an ***, an ** or an *, respectively. The sample period is from January 2, 1998 to December 31, 2011.

Panel A: Predictive Regressions for the 10^{th} Half-hour					
Month	β_0	β_1	β_2	β_3	Adj. R^2(%)
January	0.000	-0.012	-0.026	0.046	-0.50
	(1.10)	(-0.13)	(-0.90)	(0.38)	
February	0.000	0.106	-0.004	-0.066	-0.65
	(0.70)	(0.94)	(-0.12)	(-0.72)	
March	-0.000	0.140	-0.014	0.007	-0.53
	(-0.23)	(1.46)	(-0.53)	(0.08)	
April	0.000	0.357***	0.004	-0.229*	7.35
	(0.16)	(2.62)	(0.14)	(-1.94)	
May	0.000	0.190**	0.008	-0.031	0.94
	(0.08)	(2.22)	(0.26)	(-0.34)	
June	-0.000	0.205*	-0.013	0.035	1.17
	(-0.38)	(1.67)	(-0.40)	(0.39)	
July	-0.000	0.246***	0.045**	-0.120	3.94
	(-0.99)	(3.36)	(1.98)	(-0.83)	
August	-0.000	0.269***	-0.013	0.125	2.91
	(-0.72)	(2.84)	(-0.64)	(1.06)	
September	-0.000	0.225**	0.029	-0.122	2.64
	(-0.28)	(2.36)	(1.55)	(-1.16)	
October	-0.000	0.415***	-0.018	-0.057	3.16
	(-0.98)	(4.05)	(-0.49)	(-0.66)	
November	0.000	0.036	-0.028	-0.019	-0.81
	(1.00)	(0.30)	(-0.58)	(-0.08)	
December	-0.000	-0.023	0.013	-0.135	0.17
	(-1.22)	(-0.56)	(0.28)	(-1.31)	

Table III
(continued)

Panel B: Predictive Regressions for the 11^{th} Half-hour					
Month	β_0	β_1	β_2	β_3	Adj. R^2(%)
January	0.000	0.066	0.095**	0.089	6.03
	(0.08)	(0.95)	(2.26)	(1.19)	
February	-0.000	0.059	0.010	0.010	-1.16
	(-1.00)	(0.45)	(0.21)	(0.09)	
March	0.000*	0.052	0.048*	0.106	1.93
	(1.86)	(0.41)	(1.80)	(1.03)	
April	-0.000*	0.383***	-0.001	0.074	7.14
	(-1.82)	(4.45)	(-0.03)	(0.71)	
May	0.000	0.192**	-0.035*	0.227*	7.27
	(0.35)	(2.47)	(-1.67)	(1.68)	
June	-0.000	0.203**	0.017	-0.071	2.07
	(-0.75)	(2.21)	(0.59)	(-1.18)	
July	0.000	0.296***	-0.031	0.080	3.12
	(0.69)	(3.02)	(-1.00)	(0.65)	
August	0.000	0.404**	-0.031	-0.135	3.33
	(0.24)	(2.46)	(-0.81)	(-1.13)	
September	0.000	0.358***	0.015	-0.109	4.07
	(1.57)	(3.06)	(0.72)	(-1.32)	
October	0.000**	0.349***	-0.024	0.077	2.58
	(2.02)	(2.72)	(-0.67)	(0.46)	
November	-0.000	0.070	0.022	0.097	0.12
	(-0.48)	(0.59)	(0.65)	(0.88)	
December	0.000	0.090	-0.063	-0.113	2.03
	(0.89)	(0.50)	(-1.61)	(-0.84)	

Table III
(continued)

Panel C: Predictive Regressions for the 12^{th} Half-hour					
Month	β_0	β_1	β_2	β_3	Adj. R^2(%)
January	-0.000	0.200*	0.004	-0.017	0.03
	(-0.21)	(1.67)	(0.09)	(-0.19)	
February	0.000**	0.065	-0.081*	-0.061	1.36
	(2.06)	(0.57)	(-1.73)	(-0.65)	
March	0.000	0.041	0.039	-0.053	-0.05
	(0.44)	(0.41)	(1.05)	(-0.55)	
April	-0.000	0.142	0.016	0.092	0.96
	(-0.49)	(1.24)	(0.46)	(0.61)	
May	0.000	0.160***	-0.045	-0.169	3.85
	(0.93)	(2.76)	(-1.29)	(-1.23)	
June	-0.000	0.076*	-0.046	0.228***	5.40
	(-1.32)	(1.92)	(-1.58)	(2.93)	
July	0.001**	0.298**	-0.040	0.035	3.22
	(2.35)	(2.05)	(-0.79)	(0.34)	
August	0.000**	0.287***	0.039	0.009	4.39
	(2.01)	(3.10)	(0.88)	(0.07)	
September	0.000*	0.298***	0.072**	0.001	7.09
	(1.66)	(2.72)	(2.16)	(0.01)	
October	0.001*	0.559***	-0.021	0.101	2.70
	(1.95)	(2.93)	(-0.25)	(0.90)	
November	0.000	0.122	0.055	-0.142	1.69
	(0.96)	(0.75)	(0.99)	(-1.33)	
December	-0.000	0.339***	0.092	0.073	7.21
	(-1.08)	(2.75)	(1.64)	(0.68)	

Table III
(continued)

Panel D: Predictive Regressions for the Last Half-hour					
Month	β_0	β_1	β_2	β_3	Adj. R^2(%)
January	0.000	0.229**	0.096***	-0.016	5.57
	(0.51)	(2.02)	(3.46)	(-0.14)	
February	-0.000	0.133	0.015	0.038	-0.49
	(-1.05)	(1.14)	(0.39)	(0.27)	
March	-0.000*	0.097	0.108***	0.202*	8.69
	(-1.96)	(0.82)	(3.17)	(1.96)	
April	0.000	0.102	0.030	-0.059	-0.03
	(1.06)	(1.25)	(0.75)	(-0.71)	
May	0.000	0.372***	0.078***	0.074	7.46
	(1.40)	(3.37)	(2.75)	(0.90)	
June	-0.000	0.203**	0.031	0.049	1.24
	(-1.29)	(2.00)	(0.95)	(0.61)	
July	-0.000	0.252*	-0.036	0.152**	4.38
	(-0.32)	(1.89)	(-1.01)	(2.36)	
August	-0.000	0.322*	0.081	0.298*	10.20
	(-1.30)	(1.87)	(1.43)	(1.94)	
September	0.000	0.230	0.020	-0.096	0.08
	(0.22)	(1.20)	(0.53)	(-0.90)	
October	0.000	0.489*	0.093	0.040	4.02
	(0.33)	(1.88)	(1.11)	(0.44)	
November	-0.000	-0.050	0.164***	0.493**	18.83
	(-0.86)	(-0.23)	(3.81)	(2.22)	
December	0.000	-0.021	0.033	-0.027	-0.68
	(0.33)	(-0.19)	(0.48)	(-0.25)	

Table IV
Robustness Check: Weekday Seasonality

This table reports results from the following predictive regressions:

$$r_{i,wt} = \beta_0 + \beta_1 \Delta s_{i-1,wt} + \beta_2 r_{1,wt} + \beta_3 r_{i-1,wt} + \epsilon_{wt}, \quad i = 10, \dots, 13,$$

where $w = Monday, Tuesday, \dots, Friday$, r_i is the i-th half-hour return on the S&P 500 index ETF, r_1 is the first half-hour return, and Δs_{i-1} denotes the change in investor sentiment in the $(i-1)$-th half-hour. Newey and West (1987) robust t-statistics are in parentheses and significance at the 1%, 5% , or 10% level is given by an ***, an ** or an *, respectively. The sample period is from January 2, 1998 to December 31, 2011.

Panel A: Predictive Regressions for the 10^{th} Half-hour					
Weekday	β_0	β_1	β_2	β_3	Adj. R^2(%)
Monday	-0.000	0.263***	0.033*	-0.057	4.23
	(-0.86)	(3.70)	(1.87)	(-0.63)	
Tuesday	-0.000	0.207***	0.004	-0.048	1.07
	(-0.40)	(3.11)	(0.18)	(-0.51)	
Wednesday	0.000	0.176**	-0.047**	0.025	1.77
	(0.02)	(2.25)	(-2.07)	(0.40)	
Thursday	-0.000	0.049	0.003	0.081	0.11
	(-0.40)	(0.79)	(0.15)	(0.97)	
Friday	-0.000	0.116**	-0.011	-0.26***	6.18
	(-0.16)	(2.55)	(-0.55)	(-4.20)	
Panel B: Predictive Regressions for the 11^{th} Half-hour					
Monday	0.000	0.066	0.019	0.050	0.70
	(0.19)	(1.08)	(0.81)	(0.64)	
Tuesday	0.000	0.150*	0.009	-0.029	0.30
	(0.12)	(1.68)	(0.24)	(-0.32)	
Wednesday	0.000	0.356***	0.027	-0.013	4.63
	(1.59)	(5.11)	(1.22)	(-0.24)	
Thursday	0.000	0.248***	0.002	0.075	2.69
	(0.01)	(2.94)	(0.07)	(1.05)	
Friday	0.000	0.238**	-0.055	0.099	3.27
	(0.97)	(2.56)	(-1.29)	(1.03)	

Table IV
(continued)

Weekday	β_0	β_1	β_2	β_3	Adj. R^2(%)
Panel C: Predictive Regressions for the 12^{th} Half-hour					
Monday	0.000**	0.200***	-0.020	0.191*	3.56
	(2.45)	(2.87)	(-0.59)	(1.87)	
Tuesday	0.000**	0.248***	0.068**	-0.063	4.24
	(1.97)	(3.34)	(2.21)	(-0.81)	
Wednesday	-0.000	0.292***	0.024	0.078	3.57
	(-0.68)	(4.10)	(0.75)	(1.08)	
Thursday	0.000	0.327***	0.025	0.050	2.76
	(1.49)	(3.26)	(0.83)	(0.41)	
Friday	0.000	0.095*	-0.048	-0.113	1.31
	(1.38)	(1.92)	(-0.75)	(-1.58)	
Panel D: Predictive Regressions for the Last Half-hour					
Monday	-0.000	0.048	0.101***	0.61	4.19
	(-0.63)	(0.36)	(3.51)	(0.65)	
Tuesday	-0.000	0.350***	0.048	0.236***	10.28
	(-1.15)	(3.86)	(1.24)	(2.94)	
Wednesday	-0.000	0.303***	0.065**	-0.035	2.66
	(-0.98)	(2.71)	(2.03)	(-0.25)	
Thursday	0.000	0.191**	0.006	0.133	2.77
	(0.45)	(2.47)	(0.17)	(1.14)	
Friday	0.000	0.079	0.089**	0.094	3.13
	(1.07)	(0.98)	(2.52)	(0.71)	

Table V
Predictability of Intraday Stock Market Returns: The Role of Macroeconomic Variables

This table reports results from the following predictive regressions:

$$r_{i,t} = \beta_0 + \beta_1 \Delta s_{i-1,t} + \beta_2 r_{1,t} + \beta_3 r_{i-1,t} + \beta_4 Term_{t-1} + \beta_5 Def_{t-1} + \beta_6 Rate_{t-1} + \epsilon_t, \quad i = 10, \ldots, 13,$$

where $r_{i,t}$ is the i-th half-hour return on the S&P 500 index ETF on day t , $r_{1,t}$ is the first half-hour return, and $\Delta s_{i-1,t}$ denotes the change in investor sentiment in the $(i-1)$-th half-hour. We also include three lagged macroeconomic variables: term spread ($Term$), default spread (Def), and short-term interest rate ($Rate$). Newey and West (1987) robust t-statistics are in parentheses and significance at the 1%, 5% , or 10% level is given by an ***, an ** or an *, respectively. The sample period is from January 1998 to December 2011.

Half-hour Return Period	β_0	β_1	β_2	β_3	β_4	β_5	β_6	Adj. R^2(%)
10^{th} Half-hour	-0.000	0.187***	-0.003	-0.051	-0.000	0.000	-0.000	1.13
	(-0.29)	(5.72)	(-0.28)	(-1.21)	(-0.01)	(0.70)	(-0.14)	
11^{th} Half-hour	-0.001	0.207***	0.006	0.020	0.000	0.001**	0.000	2.26
	(-1.24)	(5.56)	(0.56)	(0.48)	(0.10)	(2.44)	(0.36)	
12^{th} Half-hour	0.000	0.214***	0.011	0.020	0.000	0.000	0.000	1.62
	(0.20)	(6.28)	(0.56)	(0.47)	(0.02)	(0.02)	(0.35)	
The Last Half-hour	0.000	0.219***	0.062***	0.108**	-0.000	0.000	-0.000	3.69
	(0.37)	(4.31)	(3.20)	(2.20)	(-0.94)	(0.75)	(-1.17)	

Table VI
Predictability of Intraday Stock Market Returns: Alternative Measures of Investor Sentiment

This table reports results from the following predictive regressions:

$$r_{i,t} = \beta_0 + \beta_1 \Delta s_{i-1,t} + \beta_2 r_{1,t} + \beta_3 r_{i-1,t} + \beta_4 VIX_{t-1} + \beta_5 \Delta VIX_{t-1} + \beta_6 State_{t-1} + \epsilon_t, \quad i = 10, \ldots, 13,$$

where $r_{i,t}$ is the i-th half-hour return on the S&P 500 index ETF on day t , $r_{1,t}$ is the first half-hour return, and $\Delta s_{i-1,t}$ denotes the change in investor sentiment in the $(i-1)$-th half-hour. We also include three lagged values of the following three sentiment variables: CBOE's Volatility Index (VIX), change in VIX (ΔVIX), and Market State ($State$). Newey and West (1987) robust t-statistics are in parentheses and significance at the 1%, 5% , or 10% level is given by an ***, an ** or an *, respectively. The sample period is from January 1998 to December 2011.

Half-hour Return Period	β_0	β_1	β_2	β_3	β_4	β_5	β_6	Adj. R^2(%)
10^{th} Half-hour	0.000	0.179***	-0.001	-0.048	-0.000	-0.000	-0.000	1.17
	(0.28)	(5.76)	(-0.10)	(-1.17)	(-0.33)	(-1.03)	(-0.25)	
11^{th} Half-hour	-0.000	0.204***	0.003	0.027	0.000	0.000	-0.000	1.91
	(-0.92)	(5.85)	(0.29)	(0.65)	(1.39)	(0.07)	(-0.16)	
12^{th} Half-hour	-0.001*	0.185***	0.013	0.007	0.00004**	0.000	0.0003**	2.93
	(-1.87)	(5.99)	(0.74)	(0.19)	(2.02)	(1.58)	(1.98)	
The Last Half-hour	-0.000	0.166***	0.066***	0.090*	0.000	0.0002***	-0.000	4.76
	(-0.30)	(3.51)	(3.71)	(1.92)	(0.41)	(3.23)	(-0.55)	

Table VII
Out-of-Sample Forecasting Performances

This table reports the out-of-sample forecasting performance of various predictive regressions. Our initial estimation window is set to 250 days. We consider the following four metrics: (1) Out-of-sample R^2 (OOS R^2); (2) Root Meant Squared Error (RMSE); (3) Mean Absolute Error (MAE); and (4) Theil U Statistic. We calculate the forecasting errors by comparing the actual returns of r_i with the model predicted returns $\hat{r}_i$. The process is iterated until the end of sample is reached. Variations of the following regression model are used to generate forecasting errors:

$$r_{i,t} = \beta_0+\beta_1\Delta s_{i-1,t}+\beta_2 r_{1,t}+\beta_3 r_{i-1,t}+\beta_4 VIX_{t-1}+\beta_5\Delta VIX_{t-1}+\beta_6 State_{t-1}+\epsilon_t, \quad i = 10, \ldots, 13,$$

where $r_{i,t}$ is the i-th half-hour return on the S&P 500 index ETF on day t , $r_{1,t}$ is the first half-hour return, and $\Delta s_{i-1,t}$ denotes the change in investor sentiment in the $(i-1)$-th half-hour. We also include three lagged values of the following three sentiment variables: CBOE's Volatility Index (VIX), change in VIX (ΔVIX), and Market State ($State$). The sample period is from January 1998 to December 2011. The best models are shown in **bold**.

Model	OOS R^2(%)	RMSE ($\times$100)	MAE($\times$100)	Theil U Statistic
		Panel A: 10^{th} Half-hour		
Sample Mean	0.000	0.29403	**0.19841**	1.0009
Δs	**0.839**	**0.2928**	0.19853	**0.99671**
All Predictors	-0.892	0.2953	0.19976	1.0054
		Panel B: 11^{th} Half-hour		
Sample Mean	0.000	0.30437	0.20164	1.0002
Δs	**1.535**	**0.30203**	**0.20130**	**0.99254**
All Predictors	0.069	0.30427	0.20335	0.99990
		Panel C: 12^{th} Half-hour		
Sample Mean	0.000	0.37422	**0.22325**	1.0005
Δs	1.123	0.37211	0.22615	0.99483
Δs, VIX, $State$	**1.224**	**0.37192**	0.22798	**0.99432**
All Predictors	0.203	0.37384	0.22903	0.99945
		Panel D: The Last Half-hour		
Sample Mean	0.000	0.41080	0.25144	1.0004
Δs	1.152	0.40843	0.25131	0.99457
Δs, r_1,r_{12},ΔVIX	**3.920**	**0.40267**	**0.25081**	**0.98055**
All Predictors	2.865	0.40487	0.25250	0.98592

Table VIII
Evaluating the Performance of Market Timing Trading Strategies

This table reports the annualized mean returns, standard deviations, Sharpe ratio and CEQs for strategies that rely on various combinations of the four predictors to generate trading signals. The strategies enter a long (short) position in the SPY during the last half hour of a trading day if the model predicted return is positive (non-positive). The sample period is from January 1998 to December 2011 with an initial estimation window of 250 days. The risk aversion parameter γ is set to 3. The benchmark strategy generates the trading signals if the historical average return is positive and *vice versa*. We report the p-values in parentheses, which compares strategies that uses predictor(s) with the benchmark strategy.

Strategies	Mean(%)	Std Dev (%)	Sharpe Ratio	CEQ (%)
		Panel A: 10^{th} Half-hour		
Benchmark	0.6795	4.6631	0.1457	0.3533
Δs	5.0219	4.6526	1.0794**	4.6972***
			(0.0123)	(0.0000)
All Predictors	3.6743	4.6576	0.7889*	3.3489***
			(0.0845)	(0.0000)
		Panel B: 11^{th} Half-hour		
Benchmark	0.9981	4.8302	0.2066	0.64814
Δs	9.4336	4.7939	1.9678***	9.0888***
			(0.0000)	(0.0000)
All Predictors	5.8287	4.8166	1.2101***	5.4807***
			(0.0065)	(0.0000)
		Panel C: 12^{th} Half-hour		
Benchmark	2.7720	5.9352	0.46705	2.2436
Δs	9.8335	5.9054	1.6652***	9.3104***
			(0.0001)	(0.0000)
Δs, VIX, $State$	6.5353	5.9235	1.1033**	6.0090***
			(0.0147)	(0.0000)
All Predictors	5.8803	5.9262	0.9923*	5.3535***
			(0.0686)	(0.0000)
		Panel D: The Last Half-hour		
Benchmark	0.3605	6.5190	0.0553	-0.2770
Δs	8.3401	6.4978	1.2835***	7.7068***
			(0.0015)	(0.0000)
Δs, r_1,r_{12},ΔVIX	9.1903	6.4932	1.4154***	8.5579***
			(0.0011)	(0.0000)
All Predictors	8.6284	6.4963	1.3282***	7.9953***
			(0.0024)	(0.0000)

Figure 1
Means and Volatilities of Half-Hour SPY Intraday Returns and Changes in Investor Sentiment

This figure plots the means and volatilities of half-hour SPY intraday returns as well as changes in investor sentiment. The sample period is from 01/1998 to 12/2011.