All the News That's Fit to Reprint:

Do Investors React to Stale Information?

October 2010

Paul C. Tetlock*

Columbia University

Abstract

This paper tests whether stock market investors appropriately distinguish new and old information about firms. I define the staleness of a news story as its textual similarity to the previous ten stories about the same firm. I find that firms' stock returns respond less to stale news. Even so, a firm's return on the day of stale news negatively predicts its return in the following week. Individual investors trade more aggressively on news when news is stale. The subsequent return reversal is significantly larger in stocks with above-average individual investor trading activity. These results are consistent with the idea that individual investors overreact to stale information, leading to temporary movements in firms' stock prices.

* Acknowledgements: I thank Columbia University, Yale University, and the University of Texas at Austin for their research support. I am grateful to Brad Barber, Nick Barberis, John Campbell, James Choi, Martijn Cremers, Amy Dittmar, Simon Gervais, Mark Grinblatt (AFA discussant), Terry Murray, Terry Odean (NBER discussant), Chris Parsons, Matt Spiegel, Laura Starks (editor), Jeremy Stein, Philip Tetlock, Sheridan Titman, Heather Tookes, and two anonymous referees for helpful comments. I thank seminar participants at the AFA, Boston College, Columbia, MIT, NBER, Notre Dame, Rice, UC Davis, Wharton, and Yale for valuable feedback. I thank Tanya Balsky, Adam Laiacano, and the Program for Financial Studies at Columbia for excellent research assistance. All views and opinions expressed in this paper are my own and are not endorsed by the retail order data provider.

"People everywhere confuse what they read in the newspaper with news." – A.J. Liebling

This paper tests whether stock market investors appropriately distinguish new and old information about public firms. In an efficient market where firms' stock prices rapidly incorporate all value-relevant signals, new information becomes stale information almost instantly. Based on theory alone, the impact of redundant information on asset prices is unclear. The proliferation of news increases the speed and quantity of information dissemination, which could enhance informational efficiency. On the other hand, some readers of a news story may not realize the extent to which other market participants have already traded on similar past information, leading them to overreact to stale information in news.

Based on this latter argument, I hypothesize that investor overreaction to financial news increases with the staleness or redundancy of information. The central contribution of this paper is to use an extensive database on public news events to test this hypothesis and explore the mechanism behind any observed overreaction. I gauge market overreaction by the extent to which a firm's initial daily return around a news event negatively predicts its return in the week after the event. The staleness of a news story is its textual similarity to the previous ten stories about the same firm. I focus on cross-sectional variation in return reversals because there are many possible explanations for on-average return reversals.

The sequence of news events for Equitable Cos, an insurance firm with a market capitalization of $15 billion, in March of 1999 illustrates the methodology. At 5:14pm on Monday, March 1st, which I define as (one hour into) trading day $t-1$, Dow Jones (*DJ*) releases a newswire story about Equitable Cos. The story describes an SEC filing in which Equitable proposes changing its name to the empty placeholder name of "( )" until shareholders adopt another name at an upcoming meeting. On Tuesday night, part of trading day *t*, the news appears again in a very similar format when the *DJ* newswire pre-releases selected stories from the Wednesday morning Wall Street Journal (*WSJ*).[1] The headlines and lead paragraphs in the original Monday *DJ* newswire and the Tuesday *WSJ* story appear below:

DJ: Equitable Proposes Changing Its Name To (...)

The artist formerly known as Prince chose a glyph to represent his new identity. Now the insurance company about to be formerly known as the Equitable has done him one better. Until it finalizes its new moniker, it apparently wants to be known as, well ... nothing.

WSJ: Can Equitable Find Any Better Name Than '( )'?

The Artist, formerly known as Prince, chose a glyph to represent his new identity. Now the insurance company that soon may be formerly known as Equitable Cos. has gone a step further. Until it finalizes a new name, the company apparently must make do with "( )."

A simple [0,1] measure of the similarity between two texts, proposed by Jaccard (1901), is the number of unique words present in the intersection of the two texts divided by the number of unique words present in the union of the two texts. One can compute an analogous similarity measure for unique adjacent word pairings, called bigrams, rather than unique single words.[2] The 23 unique single words in the first excerpt above include: "Equitable," "propose," "artist," "finalize," and "moniker." The 22 unique single words in the second passage include "Equitable," "find," "better," and "artist." There are 28 unique single words in the union of the two paragraphs and 16 common pairings in their intersection, implying that the single-word similarity of the two first paragraphs is 16/28 = 55.2%. The rest of these two stories are even

[1] The actual WSJ story occurs on Wednesday morning, which is also part of day 0, but this story is not contained in the Dow Jones newswire archive.

[2] Before identifying unique words and bigrams, I exclude a standard list of 119 extremely common words such as "into," "so," and "that"; 42 common numbers (0 through 9 and 1978 through 2009); and 27 terms that are ubiquitous in financial news stories, such as "Dow Jones," "New York," and "newswire." I also use a standard word stemming algorithm to equate all similar forms of a word—e.g., "changing," and "changed" are both derivatives of "change."

more similar, so that their overall single-word story similarity is 79.6%, which ranks above the 99th percentile of similarity on March 2nd, 1999.[3]

Equitable's abnormal return on trading day *t* of its highly stale newswire story is -2.99%. Interestingly, Equitable's abnormal returns on trading days *t + 1* and *t + 2* are 1.76% and 1.93%, completely reversing the initial decline in its stock price. Although the return reversal after the -2.99% reaction to the highly stale story suggests that the reaction was excessive, it is difficult to draw accurate inferences based on the *ex post* performance of a single firm.

To systematically measure how markets respond to two or more possibly related news events, I examine all *DJ* newswire stories from the *DJ* archive for public US firms from November 1996 to October 2008. The market reaction to news about a firm is the firm's stock return on a trading day when *DJ* includes the firm's ticker code in the news header. I analyze whether and why the day-*t* market responses to these news events are partially reversed during days *t + 1* to *t + 5*. I use both Fama-MacBeth (1973) regression methods and calendar time portfolios to assess whether the extent of reversal depends on news staleness.

A news story's staleness is its average textual similarity to the previous ten news stories. This measure identifies news stories that contain a greater proportion of textual information that overlaps with previously known facts. Consistent with the idea that such news stories contain less new information, firms' stock return and volume reactions on news days with high average staleness are significantly smaller than their return and volume reactions on other news days. The market's initial reaction to news in the bottom staleness decile exceeds its reaction to news in the top staleness decile by 413 (75) basis points on an equal-weighted (value-weighted) basis.

---

[3] Using bigrams, such as "Equitable propose" and "propose change," instead of using single-words as the basis for computing similarity, the bigram similarity in the two excerpts would be 12/37 = 32.4% and the overall story similarity would be 63.5%. In general, the single-word and bigram similarity measures exhibit correlations exceeding 0.8, where the single-word score is higher by construction.

The first hypothesis tested here is that market reactions to news are better negative predictors of future returns when news is stale. The evidence supports this view. Equal-weighed portfolios formed on news staleness exhibit return reversals that differ by 26 basis points in the week after news. Comparing this difference in reversal to the difference in initial reaction of 413-basis-points, the market's initial distinction between stale and new news seems insufficient in the sense that it revises its initial view by another 26 / 413 = 6.3% in the following week. The value-weighted revision is larger in percentage terms (19.8%) but smaller in raw magnitude (15 bps).

In predictive regressions of firms' returns on days $t + 1$ to $t + 5$ that control for alternative explanations, the regression coefficient on a firm's day-t return decreases with the staleness of news on day t. The ability of day-t news staleness to predict post-news return reversal remains significant after controlling for weekly return reversals (Jegadeesh (1990); Lehmann (1990)), volume-induced return reversals (*e.g.*, Campbell, Grossman, and Wang (1993); Lee and Swaminathan (2000); Llorente et al. (2002)), and other variables that predict high-frequency returns and return reversals.

To examine the mechanism behind these return reversals, I focus on a subset of investors who may confuse new and old information and, as a result, actively trade on stale information. Based on abundant prior evidence in papers such as Odean (1999), Barber and Odean (2000), Barber and Odean (2008), and Barber, Odean, and Zhu (2008) that individual traders exhibit behavioral biases, one hypothesis is that individual investors are more likely to react to stale information than institutional investors. To measure the presence of both types of investors, I use a database of individual and institutional trading orders routed through a large market center from 2003 to 2007. Aggressive trading in each groups is the imbalance in buy and sell market orders for that group. I show that individual investors increase their tendencies to trade aggressively in the same direction as news on stale news days, whereas institutional investors do not. The return reversal after stale news is significantly larger in stocks with above-average individual investor trading activity on the day of the news.

This study is related to a rapidly growing area of research on financial news events. Beyond the papers cited above, recent contributions include Barber and Loeffler (1993), Busse and Green (2001), Antweiler and Frank (2006), Das and Chen (2007), Tetlock (2007), Engelberg (2008), and Fang and Peress (2009). These studies focus almost exclusively on single news events and do not consider potential interactions between news events. By contrast, this study introduces a direct textual measure of the similarity between news events, which allows for novel tests of whether investors appropriately distinguish new and old information. These tests suggest that individual investors overreact to stale information, leading to return reversals.

The few studies that explicitly consider the links between news events arrive at somewhat different conclusions from each other—e.g., Davies and Canes (1978) versus Barber and Loeffler (1993) in finance, and Hand (1990) versus Ball and Kothari (1991) in accounting. Although the data in these studies can be reconciled with the stale information hypothesis, the limited sample sizes and specific nature of the news events make it difficult to draw general conclusions. Davies and Canes (1978) find no significant reversals after their measure of stale information release, whereas Barber and Loeffler (1993) do. Hand (1991) and Ball and Kothari (1991) dispute the interpretation of the stock market's response to the announcement of 230 swap transactions, which may contain stale information. Of the four studies above, only Barber and Loeffler (1993) finds a significant return reversal, but this study does not explicitly measure the staleness of news or analyze how reversal depends on staleness.

Two more recent studies by Huberman and Regev (2001) and Gilbert et al. (2010) provide related evidence, which one can interpret as showing that the market reacts to stale information. However, neither study has a sufficiently large sample—1 event and 72 events, respectively—to demonstrate that the market reaction to news is reliably reversed or that this reversal depends on the staleness of news. In addition, none of the studies above uses data on individual investors. The evidence in this paper sheds some light on this unresolved debate by directly showing that return reversals after news events increase with the staleness of news and that individual investors trade on stale information.

This study is also related to contemporaneous work by Hanley and Hoberg (2010) and Hoberg and Phillips (2010). These studies employ measures of textual similarity in a finance context, analyzing firms' initial public offering prospectuses and their 10-K filings. A review of the literature on textual similarity outside of finance appears in Losee (1998).

The outline of the paper is as follows. Section I describes the news and financial data used in this study. Section II presents the main empirical tests of the stale information hypothesis, showing how return reversals of market reactions to news depend on staleness. Section III examines whether trading by individual or institutional investors contributes to the return reversal. Section IV discusses the implications of the stale information hypothesis. The Appendix proposes one possible theoretical framework for stale information.

I. Empirical Data and Methodology

This study uses financial news stories about publicly traded US firms in the *DJ* newswire archive from November 1996 to October 2008 to measure these firms' information environments. The *DJ* newswires are the most widely circulated financial news in the United States for institutional investors. The *DJ* newswire arguably has the most comprehensive coverage (Fang and Peress (2009)). Although *DJ* newswire articles are usually reasonably timely, most public news events coincide with the release of information that is stale to some degree. Before writing their stories, reporters often obtain facts from non-exclusive sources of information that many investors can access directly. Reliable public news data from the *DJ* archive is available after November 1996. Prior to that date, the *DJ* ticker codes exist mainly for larger firms, which appear to have been selected in a non-random way (Tetlock (2010)). Because the *DJ* ticker codes in each story determine whether a firm is mentioned, the sample period begins in November 1996 and ends in October 2008, which is the last available month.

A. Measuring Public News Events and Their Staleness

I impose three filters on the news data to enhance the precision of the staleness measures. First, I retain only stories with one or two US ticker codes and no more than three total ticker codes to ensure that the firms mentioned in the story are the main focus of the story. Second, I exclude news stories on news days in which fewer than 50 words appear because a minimum number of words is necessary to measure textual similarity with a reasonable degree of accuracy. Third, I analyze stories with a single-word (bigram) similarity to the previous ten stories of at least 5% (2%) to increase the likelihood that the story is relevant for the firm's valuation. The eligible single-word similarity sample contains over 850,000 firm-days with newswire stories about 10,187 US firms during 12-year period. Firms at the 10th, 50th, and 90th percentiles for full-sample media coverage have 40, 195, and 1175 news days out of 3,010 trading days.

Trading day t is a firm-specific public news event if at least one news story that meets the criteria above occurs within the 4pm-to-4pm market close-to-close time interval. The matching frequency for news and returns is daily because the intra-day timing of news often does not coincide with the intra-day timing of the release of information, particularly for *DJ* releases of *WSJ* stories that only occur when markets are closed. For the main analysis, I merge news stories with daily stock price data from the Center for Research on Security Prices (CRSP) and individual and institutional trades routed to a large market center.

Next, I formally define the single-word and bigram staleness measures. Consider the single-word (*Sim1*) and bigram (*Sim2*) similarities between stories j and $j - k$, where $0 < k \leq 10$, about firm i. Define the unique single-word and two-word sets for story j as $W1_j$ and $W2_j$. Using the # operator to denote the number of elements in a set, the pairwise similarities are given by:

$$Sim1_{i,j,j-k} = \frac{\#(W1_j \cap W1_{j-k})}{\#(W1_j \cup W1_{j-k})} \text{ and } Sim2_{i,j,j-k} = \frac{\#(W2_j \cap W2_{j-k})}{\#(W2_j \cup W2_{j-k})} \tag{1}$$

Both similarity measures are examples of Jaccard (1901) indexes.

The single-word and bigram staleness of story j for firm i ($Stale1_{ij}$ and $Stale2_{ij}$) are defined as the average single-word and bigram similarities to the previous 10 newswire stories:

$$Stale1_{ij} = \frac{\sum_{k=1}^{10} Sim1_{i,j,j-k}}{10} \text{ and } Stale2_{ij} = \frac{\sum_{k=1}^{10} Sim2_{i,j,j-k}}{10} \tag{2}$$

Choosing a cutoff value of $k = 10$ for the staleness measures captures most of the thematically related stories and provides sufficient textual material for the similarity computation. The main results below are qualitatively similar using k values of 1 to 5 or 6 to 10 instead of 1 to 10.

Firm i's single-word and bigram raw staleness measures on day t ($Stale1_{it}$ and $Stale2_{it}$) are the equal-weighted averages of the single-word and bigram staleness over all $j = 1, ..., J_t$ of the firm's newswire stories that occur on day t. One can interpret these two quantities as averages of single-word and bigram staleness of the firm's news on that day. Defining daily staleness using word-weighted averages does not materially affect the results.

To remove the influence of variation in the number of newswire stories and the number of unique words per story on the staleness measures, I compute residual staleness measures ($stale1_{it}$ and $stale2_{it}$) using daily cross-sectional regressions of each raw staleness measure ($Stale1_{it}$ and $Stale2_{it}$) on the log of the number of stories, the log of the average unique words per story, and the squared log of the average unique words per story. Because news content changes significantly over time, I standardize the firm-specific staleness residuals using the cross-sectional mean and standard deviation of residual staleness on each trading day. The construction of the two stale news measures requires only information known to investors by the end of day *t*.

B. Contemporaneous Relationship between News Staleness and Trading Activity

Before delving into formal tests, I estimate the contemporaneous relationship between the staleness of news and stock market trading activity. These correlation estimates in this subsection suggest that the staleness measures exhibit intuitive properties but are not appropriate for assessing whether investors overreact to stale news. Both the single-word and bigram similarity measures of the staleness of news are contemporaneously associated with lower firm-specific return volatility (correlations of -0.032 and -0.054) and lower trading volume (correlations of -0.058 and -0.089). The difference between the market's reaction to news with high (90th percentile) and low (10th percentile) staleness is statistically significant at the 1% level and economically large: return reactions are 26.7 basis points lower and volume is 37.1% lower.

These estimates come from regressions of either firm *i*'s volume or volatility on trading day *t* on the staleness of the firm's news event on the same day. The trading volume measure is abnormal turnover ($AbTurn_{it}$), which is firm *i*'s log turnover on trading day *t* minus its average log turnover on days $t - 5$ to $t - 1$. The return volatility measure is the absolute value of firm *i*'s abnormal stock return on day *t*. For simplicity, a firm's abnormal return ($AbRet_{it}$) is its raw return minus the return on the CRSP value-weighted index. Using more sophisticated benchmarks has little impact on the results because the simple market adjustment captures much of firms' systematic volatility and because the vast majority of return volatility in firm-specific news events is not explained by traditional risk factors (e.g., Tetlock, Saar-Tsechansky, and Macskassy (2008)). Furthermore, all tests below include firm size as an independent variable, which Fama and French (1992) show captures the expected return premium from market beta. The calendar time portfolio tests control for other risk-factors in returns beyond size. None of these controls materially affects the results.

Regression controls include firm *i*'s stock market variables such as log size on day $t - 1$ ($MktCap_{i,t-1}$), cumulative market-adjusted returns on days $t - 5$ to $t - 1$ ($AbRet_i$[-5,-1]), average market-adjusted volatility on days $t - 5$ to $t - 1$ ($IdVolat_i$[-5,-1]), and the log of Amihud's (2002) illiquidity measure averaged over days $t - 5$ to $t - 1$ ($Illiq_i$[-5,-1]). The daily illiquidity measure is equal to $10^6 * |Ret_{it}| / Volume_{it}$, where $Volume_{it}$ is the stock's dollar volume. Other controls include firm-specific news variables such as the log of newswire messages on day *t* (Msg_{it}), the log of average unique words per newswire on day *t* ($Words_{it}$), the fraction of newswires on day t that mention earnings-related words ($Earn_{it}$), and abnormal newswire messages on days $t - 5$ to $t - 1$ (Msg_i[-5,-1]). The earnings-related words are defined in Tetlock, Saar-Tsechansky, and Macskassy (2008) as words beginning with the stem "earn." The abnormal newswire messages variable measures media coverage in the week prior to the news event and is given by the number of messages per day in that week minus the number of messages per day in trading days $t - 60$ to $t - 6$, which is the calendar quarter preceding the news. All control variables are standardized to have zero means and unit standard deviations on each trading day to facilitate interpretation and comparison across days.

[Insert Table 1 here.]

Panel A in Table 1 summarizes the statistical properties of both raw staleness measures. It shows that the 5th, 25th, 50th, 75th, and 95th percentiles of $Stale1_{it}$ and $Stale2_{it}$ are: 5.4%, 7.0%, 9.6%, 14.0%, and 26.3%; and 2.2%, 3.5%, 5.7% 9.1%, and 21.0%, respectively. Panel B presents the average daily correlations of the raw and residual staleness measures with the key control variables. Because $stale1_{it}$ and $stale2_{it}$ are designed to be orthogonal to Msg_{it}, $Words_{it}$, and $(Words_{it})^2$, they have nearly zero univariate correlations with each of these variables. In fact, Panel B in Table 1 shows that even the raw staleness measures ($Stale1_{it}$ and $Stale2_{it}$) do not exhibit high correlations with any of the control variables. The average cross-sectional correlations between single-word raw staleness and news-related control variables are statistically significant but small: 0.117 with Msg_{it}, 0.109 with Msg_i[-5,-1], -0.213 with $Words_{it}$, and -0.089 with $Earn_{it}$. The average cross-sectional correlations between single-word raw staleness and stock market control variables are even smaller and often statistically insignificant: 0.023 with $MktCap_{i,t-1}$, 0.001 with $AbRet_i$[-5,-1]), -0.005 with $IdVolat_i$[-5,-1], -0.017 with $Illiq_i$[-5,-1], and -0.039 with $AbTurn_{it}$. These low correlations suggest that either the staleness measures are imprecise or that they capture a unique aspect of a firm's information environment.

For all firms with news on day t, I estimate daily cross-sectional Fama-MacBeth (1973) regressions to assess how news staleness on day t relates to either volume on day t or volatility on day t. The complete cross-sectional regression specifications are:

$$DepVar_{it} = a + b * StaleVar_{it} + c * Controls_{it} + \varepsilon_{it} \text{ for each trading day } t \tag{3}$$

where $DepVar_{it} = |AbRet_{it}|$ or $AbTurn_{it}$, $StaleVar_{it} = stale1_{it}$ or $stale2_{it}$, c is a 1 by 8 coefficient vector, and $Controls_{it} = [MktCap_{i,t-1}\ AbRet_i[-5,-1]\ IdVolat_i[-5,-1]\ Illiq_i[-5,-1]\ Msg_{it}\ Words_{it}\ Earn_{it}\ Msg_i[-5,-1]]^T$. The regressions include only firms with news on day t to allow the impact of control variables such as size, illiquidity, and abnormal turnover to depend on news. A shortcoming of this approach is that the coefficient estimates for days with few news events are very imprecise. For example, fewer than 50 firms have news on some trading days, which is a problem for the regression specifications that include 20 independent variables. To ensure sufficient degrees of freedom and increase the power of the tests, I restrict the sample to the 95% of trading days in which at least 100 firms have news and non-missing independent variables. To minimize market microstructure effects, I include only firms with stock prices exceeding $5 and with positive volume on days $t-1$, t, and $t+1$. The $t+1$ restriction ensures that there is no bid-ask bounce between returns on day t and days $t+2$ to $t+5$ but has little impact on the results.

[Insert Table 2 here.]

Table 2 displays the results from the cross-sectional regressions. The main finding is that both single-word and bigram staleness of news is associated with lower contemporaneous return volatility and abnormal turnover. All coefficients of interest are highly statistically and economically significant. For example, consider the impact of a +1.28 to -1.28 standard deviation change in the single-word staleness measure. If single-word staleness were normally distributed, this change would correspond to a 10th percentile to 90th percentile change. An increase in single-word staleness of 2.56 units is associated with a decrease of 26.7 basis points (2.56 * -0.104 = -0.267 percent) in return volatility and a 37.1% decrease in average daily turnover (2.56 * -0.038 / 0.262 = -0.371). In these tests, the two-word staleness results are slightly stronger but are qualitatively similar to the single-word staleness results.

The third and fourth columns in Table 2 report the regression results for "Small" firms and "Big" firms, using a small-big size cutoff equal to the median of sample firm size on day *t*. These columns reveal that single-word staleness is associated with a reduction in return volatility that is over five times stronger in small firms (-0.203) than in big firms (-0.036). One explanation for this discrepancy is that staleness is more difficult to measure in big firms because their information environments are more complex. If measured staleness is equal to true staleness plus random measurement error, the coefficient estimates in Table 2 and other regressions where staleness is an independent variable understate the true impact of stale information.

At least two sources of measurement error in staleness are larger in the complex information environments of big firms. First, measuring the textual similarity of news to recent *DJ* newswires fails to capture similarity to alternative non-*DJ* sources of information, such as analyst reports, which are abundant for big firms. Second, newswires often provide ongoing coverage of important stories about big firms. In these cases, a story's high textual similarity to the previous 10 stories may not indicate high staleness because small textual changes in an unfolding sequence of stories can have a huge impact on the market's perception of firm value. For example, the sentences "Microsoft is considering buying Yahoo! and its search engine for a price of $45 billion" and "Microsoft is ready to buy Yahoo! and its search engine for a price of $45 billion." Although the second sentence may indicate a much higher likelihood of a successful Yahoo! acquisition to a human reader, the two sentences are nearly identical according to the simple text analysis here. Even sophisticated natural language processing techniques may fail to capture such subtleties in a wide range of news stories.

There is some variation in staleness by day of the week: staleness is lowest on Monday through Thursday and highest on Friday. One explanation for the high staleness on Friday is that there is much less news overall on Fridays. Despite this low amount of news, DellaVigna and Pollet (2009) argue that readers pay less attention to *each* news story on Fridays. These patterns could occur because reporters produce less news if they expect people to ignore it. Thus, the relationship between visibility and the quantity of news is not entirely clear. The differences in staleness by day of the week are not economically large. For example, the average value of raw single-word staleness on the lowest day (Tuesday) is only 5.7% lower than the average value on the highest day (Friday).

To summarize, Table 2 shows that firms' stock return volatility decreases with the staleness of news, especially for small firms, even after controlling for many variables known to be related to return volatility. These findings suggest that news staleness is a plausible, albeit imperfect, measure of the extent to which the market has already incorporated the information in a news story. Subsequent tests explore whether investors appropriately account for the extent to which the information in news is stale.

II. Using Staleness to Predict Firms' Returns after News Events

This section presents daily Fama-MacBeth (1973) cross-sectional regressions to investigate whether the staleness of news predicts the extent of return reversal after news. These regressions control for many other variables that could predict returns after news. I focus on reversals at weekly horizons based on prior work on reversals in general, such as Jegadeesh

(1990) and Lehmann (1990), and on reversals after news in particular, such as Tetlock (2007). The dependent variable in the main regressions is raw firm returns from day $t + 2$ through day $t + 5$ ($Ret_i[2,5]$), relative to the news event that occurs on day t. The initial results conservatively exclude the return on day $t + 1$ because using adjacent formation and holding periods may induce bid-ask bounce in returns. In robustness checks, I report the results for other time horizons to assess the duration of return predictability and the possible impact of bid-ask bounce.

The two main independent variables of interest are day-t returns ($AbRet_{it}$) and interactions of staleness with day-t returns ($stale1_{it}*AbRet_{it}$ or $stale2_{it}*AbRet_{it}$) because the coefficients on these variables capture reversals of reactions to news and their dependence on (single-word or bigram) staleness. A negative (positive) coefficient on $AbRet_{it}$ indicates unconditional return reversal (continuation) of the daily reaction to news during days $t + 2$ to $t + 5$. A negative (positive) coefficient on either $stale1_{it}*AbRet_{it}$ or $stale2_{it}*AbRet_{it}$ indicates that return reversal increases (decreases) with either single-word or bigram staleness.

The control variables include several measures of stock market activity known to predict returns and the extent of return reversals at weekly frequencies, along with other measures that are likely to be correlated with staleness and could predict returns. News-related control variables in the main analysis include log newswire messages on day t (Msg_{it}), abnormal newswire messages in the prior week ($Msg_i[-5,-1]$), log unique words per newswire on day t ($Words_{it}$), and the fraction of earnings-related newswires on day t ($Earn_{it}$). To control for post-earnings announcement drift and return predictability identified in Pritamani and Singal (2001) and Tetlock (2010), the main specification includes interaction terms between earnings words and day-t returns ($Earn_{it}*AbRet_{it}$) and between log newswires and day-t returns ($Msg_{it}*AbRet_{it}$).

Stock market control variables in the main analysis include $MktCap_{i,t-1}$, $AbRet_i[-5,-1]$, $IdVolat_i[-5,-1]$, $Illiq_i[-5,-1]$, and $AbTurn_{it}$ as defined earlier. Empirical results in studies cited above, along with studies by Banz (1981), Ang et al. (2006), and Gervais, Mingelgrin, and Kaniel (2001), indicate that these variables forecast high-frequency returns. To allow for the possibility that reversals differ for small and big firms, the main specification includes an interaction term between firm size and day-t returns ($MktCap_{i,t-1}*AbRet_{it}$). Lastly, an exhaustive specification includes five more interaction terms between day-t returns and control variables: abnormal turnover, log words, prior-week abnormal newswires, prior-week return volatility, and prior-week illiquidity. As before, all control variables are standardized by trading day and interaction terms are the product of these standardized variables.

A. Cross-Sectional Regressions of Post-News Returns on News Staleness

Table 3 reports the time-series average of the daily cross-sectional coefficients and R^2 statistics for the regressions predicting firms' returns after news events during days $t + 2$ to $t + 5$. To conserve space, the table shows only the most interesting regression coefficients. The standard errors in parentheses are robust to heteroskedasticity and autocorrelation up to five days, using the Newey-West (1987) method. This procedure corrects for the three-day overlap in the holding periods of returns in adjacent daily regressions. The six cross-sectional regressions in Table 3 differ in which of the three staleness proxies they employ and in which firm and story characteristics appear as control variables. The cross-sectional regression specifications are:

$$\begin{aligned} Ret_i[2,5] = \quad & a + b_1 * AbRet_{it} + b_2 * AbRet_{it} * StaleVar_{it} + c_0 * StaleVar_{it} \\ & + c_1 * MainControls_{it} + c_2 * OtherControls_{it} + \varepsilon_{it} \text{ for each trading day } t \end{aligned} \tag{4}$$

where $StaleVar_{it} = stale1_{it}$ or $stale2_{it}$, c_1 and c_2 are 1 by 12 and 1 by 5 coefficient vectors, $MainControls_{it} = [AbRet_{it}*Msg_{it}\ Msg_{it}\ AbRet_{it}*Earn_{it}\ Earn_{it}\ AbRet_{it}*MktCap_{i,t-1}\ MktCap_{i,t-1}\ AbRet_i[-5,-1]\ AbTurn_{it}\ IdVolat_i[-5,-1]\ Illiq_i[-5,-1]\ Words_{it}\ Msg_i[-5,-1]]^{T}$ and $OtherControls_{it} = [AbRet_{it}*AbTurn_{it}\ AbRet_{it}*IdVolat_i[-5,-1]\ AbRet_{it}*Illiq_i[-5,-1]\ AbRet_{it}*Words_{it}\ AbRet_{it}*Msg_i[-5,-1]]^{T}$.

[Insert Table 3 here.]

Table 3 shows that the coefficients on $AbRet_{it}*stale1_{it}$ and $AbRet_{it}*stale2_{it}$ are negative and both statistically and economically significant. The magnitudes of these two coefficients change by a statistically immaterial amount regardless of which control variables are included in the three specifications: no controls, the main set of controls, or all controls. The coefficient units are percentages per one standard deviation change in the independent variables. The magnitude of the total return reversal of daily returns is equal to the $AbRet_{it}$ coefficient plus the sum of all the $AbRet_{it}$ interaction coefficients weighted by the value of each interaction variable.

For example, in the third column, the total return reversal of daily returns is equal to -0.049, plus the value of $stale1_{it}$ times -0.059, plus zero if all other interaction variables are equal to their zero means. Thus, when single-word staleness is two standard deviations above its mean, return reversal is 17.7 (-0.177 = -0.049 + 2 * -0.059) basis points per standard deviation of day-t returns. This means that a -1 to +1 standard deviation change in daily returns produces an average change of 17.7 * 2 = 35.2 basis points in returns on days $t + 2$ to $t + 5$. This magnitude is large compared to the regression intercepts, which range from 10.1 basis points to 11.4 basis points and can be interpreted as the average days $t + 2$ to $t + 5$ return when all independent variables are equal to their zero means. By contrast, when single-word staleness is equal to 0.83

(0.83 = -0.049 / -0.059) standard deviations below its mean, the expected return reversal is equal to zero, implying that changes in day-t returns do not predict returns on days $t + 2$ to $t + 5$.

The average daily return reversal coefficient ($AbRet_{it}$) only becomes significant when the weekly reversal control ($AbRet_i[-5,-1]$) is included. In the two no control specifications, there is no material daily return reversal (0.015 and -0.003), but there is significant daily reversal in the four specifications that control for weekly reversal (-0.049, -0.076, -0.053, and -0.093). In addition to news staleness, several other variables exhibit significant interactions with daily returns. The third and fourth regressions show that return reversal increases as newswire messages decrease ($AbRet_{it}*Msg_{it}$), as earnings-related words decrease ($AbRet_{it}*Earn_{it}$), and as firm size ($AbRet_{it}*MktCap_{i,t-1}$) increases. The last two regressions show that return reversal increases as day-t turnover decreases ($AbRet_{it}*AbTurn_{it}$), as prior-week volatility increases ($AbRet_{it}*IdVolat_i[-5,-1]$), and as prior-week illiquidity increases ($AbRet_{it}*Illiq_i[-5,-1]$). Although some of these "control" variables can be interpreted within the context of the stale information hypothesis, I do not emphasize these interpretations because the two textual similarity measures provide more direct evidence.

Most of the control variables designed to capture high-frequency stock return predictability have the expected signs and magnitudes. Unexpected results sometimes arise because these regressions assess return predictability solely on news days. The most notable results are that firms with low prior-week returns, low newswire messages, high day-t turnover, high prior-week illiquidity, low prior-week volatility, and small size experience abnormally high returns in days $t + 2$ to $t + 5$.

The six regression intercepts range from 10.1 basis points to 11.4 basis points and can be interpreted as the average days $t + 2$ to $t + 5$ stock return when all independent variables are equal to their zero means. The R^2 statistics increase from 3% to 13% to 16% as additional sets of control variables are included. Given how little the exhaustive set of controls changes the key coefficients on $AbRet_{it}*stale1_{it}$ and $AbRet_{it}*stale2_{it}$, the analyses below develop the more parsimonious results using the main set of controls.

In further analyses, I examine time variation in the staleness interaction coefficients in Table 3. The magnitude of reversal in the market's reaction to stale news does not differ significantly by day of the week, but the point estimates are largest on Fridays: -0.078 for $AbRet_{it}*stale1_{it}$ and -0.093 for $AbRet_{it}*stale2_{it}$,. The reversal coefficients also tend to be larger on days with above-average news coverage: -0.093 for $AbRet_{it}*stale1_{it}$ and -0.060 for $AbRet_{it}*stale2_{it}$. Both results are consistent with the hypothesis that reversal to stale news occurs primarily on days when investors devote less effort to processing each news story.

The next set of tests explores the duration of the return reversal by changing the time horizon of the dependent return variable. Table 4 shows the results for two one-day periods (day $t + 1$ and day $t + 2$) and for one two-week period (days $t + 2$ to $t + 10$). The point estimates suggest that staleness predicts return reversal ($AbRet_{it}*stale1_{it}$ and $AbRet_{it}*stale2_{it}$) relatively consistently throughout the two weeks after news. Comparing columns five and six in Table 4 with columns three and four in Table 3, the coefficients on $AbRet_{it}*stale1_{it}$ and $AbRet_{it}*stale2_{it}$ are slightly larger in magnitude for the days-[2,10] period (-0.069 and -0.052) than the days-[2,5] period (-0.059 and -0.046).

[Insert Table 4 here.]

Partly because the median (mean) time until a firm's next news event is just four (13.7) days, the tests begin to lose power at longer horizons, resulting in slightly larger standard errors. The one-day horizon tests do not suffer much from this issue but do not benefit from increases in precision from averaging across nearly independent time periods. The point estimates in columns one through four in Table 4 show that return reversal on days $t + 1$ and $t + 2$ is 1.0 to 1.5 basis points per standard deviation of daily returns with modest t-statistics ranging between 0.9 and 1.9. Given that slightly more reversal occurs on day $t + 2$ than day $t + 1$, bid-ask bounce on day $t + 1$ is unlikely to strengthen the stale reversal result. If bid-ask bounce has a greater impact on firms with non-stale news, it may actually weaken the main result.

Another robustness question is whether news staleness predicts return reversal in various subsamples. Table 5 displays the results from regressing $Ret_i[2,5]$ on the main set of controls in six cross-sectional and time-series subsamples. The first two columns show the coefficients in years 1996 to 2001 and years 2002 to 2008. The coefficients on $AbRet_{it}$ and $AbRet_{it}*stale1_{it}$ show that both return reversal after news and reversal after stale news diminish dramatically in the second half of the sample (-0.067 vs. -0.032 and -0.092 vs. -0.027), though the point estimates remain negative. This general reduction in return reversals could be related to the increase in institutional trading since 2001, an issue the next section explores further. The declining reversal of daily returns may also be related to Kaniel, Saar, and Titman's (2008) finding that weekly return reversal has fallen sharply in the past 20 years, particularly for large stocks.

[Insert Table 5 here.]

The third and fourth columns in Table 5 report the regression coefficients for firms with below-median (Small) and above-median (Big) market capitalization. The impact of news staleness on reversal is more than twice as large for small firms (-0.082) but remains economically meaningful for big firms (-0.038). The smaller reversal and larger standard errors for big firms could be related to the complexity of their information environments, as discussed earlier. The fifth and sixth columns show that stale reversal is more than twice as large on news days with earnings-related news, as measured by the presence of "earn*," compared to days without earnings news (-0.102 vs. -0.049). This demonstrates that the textual staleness of news is not merely a proxy for news that is irrelevant for valuation. Overall, the estimates of reversal after stale news ($AbRet_{it}*stale1_{it}$) are negative in all six subsamples and significant at the 5% level in four subsamples. The *p*-values for the hypothesis that return reversal of single-word staleness is zero in the 2002 to 2008 and large firm subsamples are 0.134 and 0.159, respectively.

B. Calendar Time Returns after News Events Sorted by Staleness and Initial Market Reaction

Based on the regressions in the previous section, the inclusion of several control variables does not materially change the estimated impact of stale information on reversal. Accordingly, in this section, I estimate the magnitude of the stale news reversal using calendar time portfolios formed on day-*t* reactions to news and the two staleness proxies. There are two main benefits of these portfolio tests. First, non-parametric sorts capture the potentially non-linear dependence of future returns on day-*t* reactions to news, whereas the regression specification imposes linearity. Second, interpreting the magnitudes of the calendar time returns is straightforward, whereas interpreting the regression coefficients requires full knowledge of the covariance matrix of the control variables and staleness proxies.

In each trading day, I determine the decile breakpoints for the two staleness interaction terms with day-*t* reactions to news ($AbRet_{it}*stale1_{it}$ and $AbRet_{it}*stale2_{it}$). Because all variables are standardized by day before computing interaction terms, the direct one-way sort on the interaction term produces results similar to a two-way sort on both variables. Using the daily decile breakpoints, I sort firms with news on day *t* into 10 portfolios based on their values of

$AbRet_{it}$*$stale1_{it}$ and $AbRet_{it}$*$stale2_{it}$. I hold each of these 10 portfolios from the beginning of day $t + b$ until the end of day $t + e$, where this interval is denoted by [b, e]. I repeat this procedure for intervals of [b, e] = [0,0], [1,1], [2,2], [2,5], [1,5], and [1,10].

When the portfolio holding period is longer than one day (i.e., $b > e$), this portfolio formation procedure generates a sequence of $b – e + 1$ portfolios, formed over the past $b – e + 1$ days, with overlapping time horizons. I apply equal weights to these portfolios to combine them into an aggregate portfolio as in, for example, Jegadeesh and Titman (1993). This procedure creates 10 such aggregate portfolios ranked by their day-t returns and staleness ($AbRet_{it}$*$stale1_{it}$ or $AbRet_{it}$*$stale2_{it}$). I form two long-short "stale momentum" portfolios that buy the stocks with $AbRet_{it}$*$stale1_{it}$ or $AbRet_{it}$*$stale2_{it}$ in the top decile and sell short the stocks with $AbRet_{it}$*$stale1_{it}$ or $AbRet_{it}$*$stale2_{it}$ in the bottom decile. These portfolios are long on firms experiencing positive news with high staleness or negative news with low staleness (*e.g.*, $AbRet_{it}$*$stale1_{it} >> 0$) and short on firms experiencing negative news with high staleness or positive news with low staleness (*e.g.*, $AbRet_{it}$*$stale1_{it} << 0$). The stale momentum portfolio has negative returns only if the reversal of news-day returns is larger for stocks with high staleness than for stocks with low staleness during the portfolio's holding period.

I construct the stale momentum portfolios using both equal weights and value weights for firms within the decile portfolios. In general, firms in the sample of news events are considerably larger than the representative publicly traded firm. Value-weighting weighs the largest firms within this sample of large firms even more heavily. I compute the risk-adjusted returns of each portfolio using a standard time series regression of portfolio returns on four return factors: the market (MKT), size (SMB), and book-to-market (HML) factors proposed in Fama and French (1992 and 1993)), and the UMD factor based on the momentum anomaly. The momentum factor represents a long-short portfolio generated by sorts of past returns over the monthly time horizons of 2 to 12 months.[4] Each news event portfolio's risk-adjusted return is the intercept in the time series regression of the portfolio's raw return on the four risk factors. Newey and West (1987) standard errors robust to heteroskedasticity and five days of serial correlation appear below the regression coefficients.

Table 6 presents the daily alphas from stale momentum portfolios at various holding period horizons. The four panels in the table differ in whether they use single-word or bigram staleness to construct the portfolios and in whether they use equal or value weights. Each column in a panel represents a different time series regression. Only Panel A explicitly shows the factor loadings on the four factors because these loadings are generally small and do not materially affect the daily alphas. The cumulative alpha is the daily alpha times the number of days in each holding period. The bottom row in each panel shows the R^2 statistics, which are all less than 1%, confirming that the risk factors do not explain the portfolio returns.

[Insert Table 6 here.]

The first notable result in Table 6 is that all stale momentum portfolio returns are negative, regardless of the portfolio holding period, firm weighting scheme, or staleness variable used in portfolio formation. The six horizons explored in the table are days [0,0], [1,1], [2,2], [2,5], [1,5], and [1,10]. The day-*t* (or [0,0]) difference in initial market reactions to stale and non-stale news provides a useful benchmark for comparing the other holding period returns. For example, column one in Panel A (Panel B) shows that the equal-weighted difference in day-*t* alphas is -4.133% (-5.219%), meaning that the market reacts much more to single-word (bigram)

[4] The daily returns of these four factors are available at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

news with low staleness. The value-weighted return difference in Panel C (Panel D) is, however, much smaller at just 74.8 (87.6) basis points.

The cumulative equal- and value-weighted stale momentum alphas on days [1,5] after the initial news are -26.2 (-27.0) basis points and -14.8 (-16.8) basis points using the single-word (bigram) staleness measure. These reversal magnitudes are not only statistically significant, but they are also economically significant. As a fraction of the initial difference in the market's reaction to single-word stale and non-stale news, the equal- and value-weighted difference in alphas on days [1,5] are equal to 6.3% and 19.8%, respectively. In other words, the market's initial assessment of the importance of stale information is revised substantially within a week. Making inferences at longer time horizons is challenging for the reasons discussed earlier.

III. Individual and Institutional Trading around Stale News

This section examines how individual and institutional investors trade around stale news events. A second set of tests analyzes whether the return reversal after stale news depends on the relative intensity of trading activity by these groups of investors. These two tests help distinguish rational and behavioral explanations for the observed return reversals after stale news.

A. Trading Behavior of Individuals and Institutions after Stale News Events

The proprietary data on individual and institutional investor trading activity come from an over-the-counter market maker in Nasdaq stocks. This market center began as a trading platform for retail broker-dealers to route their orders, but it also attracts some institutional order flow. Broker-dealers' Rule 606 reports filed under the Securities and Exchange Commission's (SEC) Regulation National Market Systems (RegNMS) reveal that most large retail brokers route significant order flow to this market center, including four of the top five online brokerages in 2005. In the quarter closest to 2005:Q1 where Rule 606 data are available for Nasdaq stocks, these four brokers route an average of 35% of their orders to this market center. Rule 606 filings indicate that most brokers routing to this market center receive small payments for directing market orders there. Such payments, between over-the-counter market makers and brokers who handle mostly retail order flow, are common.

The data include all executed market orders routed by individuals and institutions through the market center from February 2003 through December 2007. The market center classifies orders as individual or institutional based on how they are submitted and routed. There are at least two benefits to using market order data on individual and institutional traders. First, more than 99% of market orders result in executed trades. Second, traders who submit market orders are aggressively demanding liquidity and potentially exerting an impact on stock prices.

The order data are large in comparison to most other databases of individual transactions. For example, from 2003 to 2007, brokerages route over 55 million individual market orders to this market center in sample eligible Nasdaq stocks. The two brokerage databases studied in Barber, Odean, and Zhu (2009), from 1991 to 1996, include "just" 9.2 million trades and presumably far fewer market orders. The aggregate dollar value of the individual market orders here is \$787 billion compared to an aggregate traded value of \$128 billion in Barber, Odean, and Zhu (2009). The aggregate value of institutional market orders here is \$101 billion, implying that institutions comprise only 13% of the market orders. The dollar value of all market orders in the data is a significant percentage (2.1%) of total CRSP volume in Nasdaq stocks.

More importantly, the quantity of market orders routed to this market center is representative of the cross-sectional trading activity in sample eligible Nasdaq firms. To evaluate this, I compute the average monthly cross-sectional correlation in the log of total market order activity here and the log of CRSP trading activity. Activity for the market center (CRSP) is measured in either number of market orders (trades) or dollar volume of market orders (trades). The average correlations between individual market order activity here and CRSP trading activity are 0.91 and 0.86 for the number and dollar value measures, respectively. These high correlations suggest that the aggregate quantity of orders routed to the market center is representative of overall cross-sectional trading activity.

To analyze individual and institutional trading around news, this study employs three measures of market order activity for each firm i on each trading day t: individual and institutional buy-sell imbalances ($IndBSI_{it}$ and $InstBSI_{it}$), and the relative intensity of individual to institutional order activity ($IndAct_{it}$). Let the total number of shares purchased and sold by individuals (institutions) through market orders in firm i on day t be given by $IndBuy_{it}$ and $IndSell_{it}$ ($InstBuy_{it}$ and $InstSell_{it}$). Then $IndBSI_{it}$ is given by the difference between $IndBuy_{it}$ and $IndSell_{it}$, divided by their sum; and $InstBSI_{it}$ is defined analogously. Positive values of $IndBSI_{it}$ ($InstBSI_{it}$) indicate that individuals (institutions) are aggressive net buyers of shares in firm i on day t. The definition of $IndAct_{it}$ is the dollar-weighted difference between ($IndBuy_{it} + IndSell_{it}$) and ($InstBuy_{it} + InstSell_{it}$), divided by $MktCap_{it}$. Positive values of $IndAct_{it}$ indicate that more individuals than institutions are submitting aggressive orders to trade firm i's stock on day t through the market center.

Table 7 assesses how individuals and institutions trade around stale news, using the same Fama-MacBeth cross-sectional regressions as in Table 3, Table 4, and Table 5, except that market order imbalances are the dependent variable instead of stock returns. The first four columns analyze individual buy-sell imbalances (*IndBSI*) and the last four columns use institutional imbalances. The coefficient on firm *i*'s abnormal return on day-*t* ($AbRet_{it}$) measures aggressive momentum or contrarian trading. For both individuals and institutions, this coefficient is consistently positive on day *t*, implying that investors tend to aggressively trade in the same direction as returns—*e.g.*, traders submit market buy orders when positive news occurs.

Interactions terms with $AbRet_{it}$ allow aggressive momentum trading to depend on the two textual staleness measures ($stale1_{it}$ and $stale2_{it}$), newswire messages (Msg_{it}), earnings news ($Earn_{it}$), and firm size ($MktCap_{i,t-1}$). Columns one and two show that *individual* buy-sell imbalances increase by statistically and economically large amounts on day *t* when there is positive stale news as measured by either single-word or bigram staleness. The coefficient on $AbRet_{it}*stale1_{it}$ ($AbRet_{it}*stale2_{it}$) shows that a one-standard deviation increase in staleness is associated with an increase in news-momentum trading of 0.900% (1.045%) by individuals. In other words, individuals increase their tendencies to trade in the same direction as the news on day *t* if the news is stale. By contrast, these same two coefficients in columns five and six show that institutional buy-sell imbalances on day *t* are unaffected by the direction of stale news on day *t*. One can reject the hypothesis that the $AbRet_{it}*stale1_{it}$ or $AbRet_{it}*stale2_{it}$ coefficients for individuals and institutions are the same at the 1% level.

The coefficient on $AbRet_{it}*stale1_{it}$ in columns three and four show that individual imbalances no longer depend on stale news after day $t + 1$. The -0.195 coefficient on $AbRet_{it}*stale1_{it}$ in column six provides suggestive (*t*-statistic of just -1.48) evidence that institutions trade against stale news on day $t + 1$. If so, this institutional trading could expedite the market's return reversal after stale news, but the tests are not sufficiently powerful to definitively confirm or reject this speculation. Still, the results in Table 7 establish that individuals aggressively trade on stale news mainly on days t and $t + 1$, whereas institutions seem to trade on the opposite direction, particularly on day $t + 1$. This evidence is consistent with the hypothesis that individual traders exert temporary buying pressure on stocks with positive stale news on day t and that this pressure subsides soon after day $t + 1$. This timing coincides closely with the timing of the observed return reversal after stale news.

B. Return Reversals after Stale News Sorted by Relative Intensity of Individual Trading Activity

This subsection explicitly examines how the return reversal after stale news depends on the relative intensity of individual and institutional trading activity, as measured by the $IndAct_{it}$ variable. The stale information hypothesis suggests that the impact of staleness on reversals should increase with trading activity from irrational investors who confuse old and new information. The results in Table 7 suggest that individuals could play this role.

Table 8 repeats the return reversal analysis in Table 5 for the subsamples of firms in which the relative intensity of individual trading is either above or below its median on each trading day. This analysis is restricted to February 2003 to December 2007 because this is the period in which market center data are available. The reader should interpret these results with some caution because many traders, excluding market makers at broker dealers, do not have real-time access to data on the relative trading intensity of individuals and institutions. Thus, the return predictability in this table does not correspond to implementable trading strategy returns for the traders without these data. Nevertheless, these tests provide intriguing evidence on whether individuals exert price pressure on stocks experiencing stale news.

[Insert Table 8 here.]

The significantly negative coefficients on $AbRet_{it}*stale1_{it}$ and $AbRet_{it}*stale2_{it}$ in columns one and three in Table 8 show that the return reversal after stale news is quite strong in stocks with above-median individual trading activity. This finding is notable because the stale reversal results are weakest in the 2003 to 2008 subperiod (see Table 5). For stocks with below-median individual trading activity, however, the point estimates suggest that there is no reversal after stale news, as measured by either staleness variable. Despite the smaller samples in this analysis, one can reject at the 5% level the hypothesis that return reversal after stale news is equal in the above-median and below-median individual investor activity subsamples.

IV. Concluding Thoughts

This paper presents evidence consistent with the hypothesis that individual investors overreact to stale information about publicly traded firms. News-event returns are smaller and partially reverse when news stories overlap more with past information, based on two textual similarity measures. The impact of news staleness on return reversal remains strong after controlling for a wide range of variables known to forecast future returns. Staleness predicts increases in return reversals in many cross-sectional subsamples and time periods, and at several time horizons ranging from two days to two weeks.

The return reversal after stale news is distinct from previously documented reversals and momentum, such as the weekly return reversal, volume-induced return reversal, and post-earnings announcement drift. The reversal after stale news does not seem to be driven by microstructure biases, such as bid-ask bounce, given that most of the reversal occurs more than one day after the stale news. The fact that reversal after stale news occurs even within the group of stories that focuses on firms' earnings indicates that the textual staleness of news is not merely a proxy for news that is irrelevant for valuation.

The impact of staleness on return reversals is significantly greater in stocks with above-median individual investor trading activity. The market center data show that individual investors increase their tendencies to aggressively trade on news when news is stale. The implication is that individual investors sometimes fail to distinguish between old and new information in news.

The comprehensive nature of the news events studied here suggests that the stale information processing bias is quite general and has the potential to explain other empirical anomalies in stock returns. The appendix to this paper proposes a theory in which this cognitive bias affects equilibrium asset prices. The model considers the sequential release of two pieces of information: one signal (s_1) consisting of pure new information, followed by a second "impure" signal ($s_1 + s_2$) consisting of both new information (s_2) and stale information (s_1).

This simple model initially predicts return momentum as the first signal (s_1) elicits two similar reactions—when it is initially released and when it released again—followed by return reversal that corrects investors' overreaction to stale information.[5] This paper focuses on returns after the market reaction to the follow-up news event ($s_1 + s_2$), which elicits an unambiguous reversal. Yet one could also explore whether there is any positive correlation in the market reactions to successive news events—i.e., s_1 and $s_1 + s_2$. Chan (2003) provides possibly related evidence that return momentum occurs only in firms with public news events, and return reversals occur in firms without news. Future research could test the return momentum implications of the stale information hypothesis by applying the distinction between stale news and other news to the Chan (2003) analysis.

The results here suggest that one role of financial media is to transmit stale news to a subset of investors who unwittingly make prices less efficient in the short run. Methodologically, this paper shows that one can use the content of news stories to quantify the staleness of information transmitted by news providers. Researchers can explore other dimensions of information content, such as the evolution of particular news topics over time, using similarity measures analogous to staleness. Developing improved measures of firms' and investors' information environments is necessary for testing numerous predictions from asset pricing theory and market microstructure. The findings have the potential to inform broad policy questions such as whether active trading around news events improves market efficiency in the long run.

[5] I thank Sheridan Titman for helpful discussions of this point.

Appendix: Modeling Stale Information

This appendix outlines a model to suggest one mechanism for the stale information hypothesis. The model shares features with theories of overreaction and underreaction that are based on investors' individual decision making errors (e.g., Barberis, Shleifer, and Vishny (1998); Daniel, Hirshleifer, and Subrahmanyam (1998); Brunnermeier (2005)), and those that are based on investor heterogeneity (e.g., Hong and Stein (1999)). Although the stale information hypothesis and the theory in Hong and Stein (1999) both emphasize the importance of the information environment, these two theories still make two distinct predictions, which I discuss below. When describing stock returns, the term overreaction refers to reversals. When describing stock price levels, overreaction refers to a price response to a signal that exceeds the change in fundamental value.

The main result of the model is that the release of an informative signal elicits overreaction as irrational investors receive both new and old information concurrently. The key assumption is that irrational investors' perceptions of their signal conflate old and new information, implying that they react to old information (e.g., DeMarzo, Vayanos, and Zwiebel (2003)). Because rational investors anticipate that irrational investors will overreact, rational investors trade intensely on the signal that will soon be re-released to irrational investors. This triggers an initial overreaction in prices that occurs even before irrational investors receive the stale signal. Recognizing that the initial release of a signal that will soon be stale is difficult to measure empirically, I emphasize the subsequent overreaction in prices and its relationship to observable variables.

Formally, suppose there are two types of investors: one is completely rational and the other is imperfectly rational. The investors with bounded rationality are present in measure m, whereas the rational investors are present in measure $1 - m$. Both types of investors have negative exponential utility functions that possess the convenient constant absolute risk aversion (CARA) property. I focus on the rational expectations equilibrium in which all individual traders are atomistic price takers.

There is a single asset that pays a normally distributed liquidating dividend d, where $d = s_1 + s_2 + s_3$ with $s_i \sim N(0, \sigma_i^2)$ for $i = 1, 2, 3$, and the s_i terms are independent. I assume that the liquidating dividend is gradually revealed to all investors in three periods, so that investors observe the signals s_1 in period 1, $s_1 + s_2$ in period 2, and $s_1 + s_2 + s_3$ in period 3. The two investor types differ only in the way that they process the signal in period 2, which contains both old information (s_1) and new information (s_2). Rational investors perfectly separate the two types of information. They correctly perceive only s_2 as new information, and completely disregard the old signal s_1. By contrast, irrational investors perceive $ks_1 + s_2$ as new information, where $0 < k < 1$, implying that they also partially react to old information. The k parameter captures the extent to which irrational investors overreact to stale information—i.e., $k = 0$ corresponds to a rational investor. A more general model could include an additional parameter that measures the extent of underreaction to new information (s_2).

In each period, including an initial period 0, both types of investors set their asset demands to maximize their CARA utility functions based on all information available to them. To simplify the exposition of the equilibrium, I assume investors myopically maximize expected utility of next-period wealth. I also make two assumptions to suppress the influence of risk aversion on asset prices. First, the rational and irrational investors' risk aversion parameters are equal. Second, the single asset is available in zero net supply. These assumptions enable me to focus on how traders' expectations affect prices. They do not affect the qualitative results because the fraction of irrational investors remains a free parameter. Lastly, I normalize the market interest rate to zero.

One can solve for the equilibrium asset prices using the traditional backward induction approach.[6] The market clearing prices in both periods are:

$$p_2 = (1 + m \cdot k)s_1 + s_2, \tag{A1}$$

$$p_1 = [1 + m \cdot (1 - m) \cdot k] \cdot s_1 \tag{A2}$$

In period 0, prices are equal to zero because zero is the common prior mean for all investors, and all investors have the same background information.

Now I compare the market prices above to the prices that would prevail in the limiting case of $m = 0$ in which all investors are rational Bayesians with unlimited cognitive resources. Equation (A1) would become $p_2 = s_1 + s_2$, and Equation (A2) would be $p_1 = s_1$ if no irrational investors participated. When $m > 0$, prices in period 1 overreact to the signal s_1 relative to the benchmark $m = 0$ case. This initial overreaction persists in period 2, and is reinforced by the overreaction from irrational investors. Empirically, it is difficult to distinguish these two sources of overreaction, particularly if the signals in periods 1 and 2 arrive at nearly the same time. The root cause of both overreactions is that irrational investors confuse stale and new information.

Researchers commonly interpret return reversals as empirical manifestations of overreaction. Indeed, most price changes in this model are negatively autocorrelated. Two empirically relevant quantities are the covariances between the two news-event returns—around

[6] For simplicity and realism, I assume that irrational investors are unaware of their own perceptual errors. The main results in the model do not require this assumption.

the release of s_1 and $s_1 + s_2$—and subsequent returns. Using Equations (A1) and (A2), one obtains these expressions for the news-event return reversals after s_1 and $s_1 + s_2$, respectively:

$$Cov(d - p_2, p_2 - p_1) = -m^3k^2\sigma_1^2 \leq 0 \quad \text{(A3)}$$

$$Cov(d - p_2, p_1 - p_0) = -mk[1 + m(1 - m)k]\sigma_1^2 \leq 0 \quad \text{(A4)}$$

Equations (A3) and (A4) show that both anticipated and unanticipated overreaction to stale information lead to return reversals after stale information is released in period 2. The main implication for the empirical work is that return reversals are likely to be larger when there is abundant recent information (σ_1^2), particularly if this old information resembles new information. The model's second implication is that the return reversal after the second release of s_1 should be greater when more irrational investors are present—i.e., as m increases.

A final interesting implication—not explored in this paper—is that the two price responses to the signals are positively autocorrelated because they represent overreaction to similar underlying information (s_1):

$$Cov(p_2 - p_1, p_1 - p_0) = m^2k[1 + m(1 - m)k]\sigma_1^2 \geq 0 \quad \text{(A5)}$$

Equation (A5) implies that there is return momentum in the two news-event returns. This occurs before the partial return reversal of both news-event returns. In this paper, I focus only on the unambiguous return reversals that occur after the release of stale information, leaving tests of the positive correlation between the news-event returns of successive news releases for future work.

Despite the similarities in the stale information hypothesis and the theory in Hong and Stein (1999), the empirical predictions of the two models are somewhat distinct. Both models feature agents who respond to news events and ignore the information in market prices. A key difference between the two models is the information diffusion process, which Hong and Stein (1999) suppose is a sequence of pure innovations in signals and I model as the arrival of two potentially related signals. This difference generates overreaction in this model and underreaction in Hong and Stein (1999). Nevertheless, many of the comparative statics of the two models are similar if one is willing to consider an increased return reversal as equivalent to reduced return momentum. For example, greater analyst coverage reduces return momentum in Hong and Stein (1999), and increases return reversals in the stale information model.

Fortunately, relative to the predictions made in Hong and Stein (1999), the current model does deliver at least two unique comparative statics. First and foremost, the stale information model draws an explicit link between traders' reactions to successive signals, predicting that return reversals will increase after stale news events. By contrast, the Hong and Stein (1999) model does not make an obvious prediction. Second, an increase in the fraction of irrational traders increases the magnitude of return reversal in the stale information model, whereas it increases return momentum in Hong and Stein (1999). The empirical tests in Sections II and III examine these two unique predictions as well as the predictions that both models make.

References

Amihud, Yakov, 2002, Illiquidity and stock returns cross-section and time-series effects, *Journal of Financial Markets* 5, 31-56.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259-299.

Antweiler, Werner, and Murray Z. Frank, 2006, Do U.S. stock markets typically overreact to corporate news stories? Working paper, University of British Columbia.

Ball, Ray J., and S. P. Kothari, 1991, Security returns around earnings announcements, *The Accounting Review* 66, 718-738.

Banz, Rolf W., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3-18.

Barber, Brad M., and Douglas Loeffler, 1993, The 'Dartboard' column: second-hand information and price pressure, *Journal of Financial and Quantitative Analysis* 28, 273-284.

Barber, Brad M., and Terrance Odean, 2000, Trading is hazardous to your wealth: the common stock investment performance of individual investors, *Journal of Finance* 55, 773-806.

Barber, Brad M., and Terrance Odean, 2008, All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors, *Review of Financial Studies* 21, 785-818.

Barber, Brad M., Terrance Odean, and Ning Zhu, 2008, Do retail trades move markets? *Review of Financial Studies* 22, 1-36.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.

Brunnermeier, Markus K., 2005, Information leakage and market efficiency, *Review of Financial Studies* 18, 417-457.

Busse, Jeffrey A. and Clifton T. Green, 2003, Market efficiency in real time, *Journal of Financial Economics* 65, 415-437.

Campbell, John Y., Sanford J. Grossman, and Jiang Wang, 1993, Trading volume and serial correlation in returns, *Quarterly Journal of Economics* 108, 905-939.

Chan, Wesley S., 2003, Stock price reaction to news and no-news: drift and reversal after headlines, *Journal of Financial Economics* 70, 223-260.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and overreactions, *Journal of Finance* 53, 1839-1886.

Das, Sanjiv R., and Mike Y. Chen, 2007, Yahoo! for Amazon: sentiment extraction from small talk on the web, *Management Science* 53, 1375-1388.

Davies, Peter Lloyd., and Michael Canes, 1978, Stock prices and the publication of second-hand information, *Journal of Business* 51, 43-56.

DellaVigna, Stefano, and Joshua M. Pollet, 2009, Investor inattention and Friday earnings announcements, *Journal of Finance* 64, 709-749.

DeMarzo, Peter M., Dimitri Vayanos, and Jeffrey Zwiebel, 2003, Persuasion bias, social influence, and unidimensional opinions, *Quarterly Journal of Economics* 118, 909-968.

Engelberg, Joseph, 2008, Costly information processing: Evidence from earnings announcements, Working Paper, University of North Carolina.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, Return, and Equilibrium: Empirical tests, *Journal of Political Economy* 81, 607-636.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns of stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Fang, Lily, and Joel Peress, 2009, Media coverage and the cross-section of expected returns, *Journal of Finance* 64, 2023-2052.

Gervais, Simon, Ron Kaniel, and Dan H. Mingelgrin, 2001, The high-volume return premium, *Journal of Finance* 56, 877-919.

Gilbert, Thomas, Shimon Kogan, Lars Lochstoer, and Ataman Ozyildirim, 2010, Investor inattention and the market impact of summary statistics, Working Paper, University of California, Berkeley.

Hand, John R.M., 1990, A test of the extended functional fixation hypothesis, *The Accounting Review* 65, 740-753.

Hand, John R.M., 1991, Extended functional fixation and security returns around earnings announcements: a reply to Ball and Kothari, *Accounting Review* 66, 739-746.

Hanley, Kathleen, and Gerard Hoberg, 2010, The information content of IPO prospectuses, *Review Financial Studies* 23, 2821-2864.

Hoberg, Gerard, and Gordon M. Phillips, 2010, Product market synergies and competition in mergers and acquisitions: A text-based analysis, *Review of Financial Studies* 23, 3773-3811.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *Journal of Finance* 54, 2143-2184.

Huberman, Gur, and Tomer Regev, 2001, Contagious speculation and a cure for cancer: a nonevent that made stock prices soar, *Journal of Finance* 56, 387-396.

Jaccard, Paul, 1901, Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines, *Bulletin de la Société Vaudoise des Sciences Naturelles* 37, 241-272.

Jegadeesh, Narasimhan, 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881-898.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Kaniel, Ron, Gideon Saar, and Sheridan Titman, 2008, Individual investor trading and stock returns, *Journal of Finance* 63, 273-310.

Lee, Charles M. C., and Swaminathan, Bhaskaran, 2000, Price momentum and trading volume, *Journal of Finance* 55, 2017-2069.

Lehmann, Bruce N., 1990, Fads, martingales, and market efficiency, *Quarterly Journal of Economics* 105, 1-28.

Llorente, Guillermo, Roni Michaely, Gideon Saar, and Jiang Wang, 2002, Dynamic volume-return relation of individual stocks, *Review of Financial Studies* 15, 1005-1047.

Losee, Robert M., 1998, *Text Retrieval and Filtering: Analytic Models of Performance.* Kluwer Academic Publishers, Boston, Massachusetts.

Newey, Whitney K., and Kenneth D. West, 1987, A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix estimator, *Econometrica* 55, 703-708.

Odean, Terrance, 1999, Do investors trade too much? *American Economic Review* 89, 1279-1298.

Pritamani, Mahesh, and Vijay Singal, 2001, Return predictability following large price changes and information releases, *Journal of Banking and Finance* 25, 631-656.

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

Tetlock, Paul C., Maytal Saar-Tsechansky, and Sofus Macskassy, 2008, More than words: quantifying language to measure firms' fundamentals, *Journal of Finance* 63, 1437-1467.

Tetlock, Paul C., 2010, Does public financial news resolve asymmetric information? *Review of Financial Studies* 23, 3520-3557.

Table 1: Cross-Sectional Distributions and Correlations for the Two Staleness Measures

Panel A in Table 1 summarizes the cross-sectional statistical properties of the single-word and bigram raw and residual textual staleness measures ($Stale1_{it}$ and $Stale2_{it}$; and $stale1_{it}$ and $stale2_{it}$). For each measure, it shows the average of the daily cross-sectional mean, standard deviation, 5^{th}, 25^{th}, 50^{th}, 75^{th}, and 95^{th} percentiles. Both residual staleness measures are orthoganalized with respect to three news story characteristics and standardized by day as explained in the text. Panel B presents the daily average of the correlations of each of the four staleness measures with several control variables. Some control variables measure characteristics of news stories, including newswire messages (Msg_{it}), prior-week abnormal newswire messages ($Msg_i[-5,-1]$), words in the news ($Words_{it}$), squared words ($(Words_{it})^2$), and earnings-related news ($Earn_{it}$). Other control variables measure firm characteristics, including firm size ($MktCap_{i,t-1}$), abnormal turnover ($AbTurn_{it}$), prior-week idiosyncratic volatility ($IdVolat_i[-5,-1]$), prior-week illiquidity ($Illiq_i[-5,-1]$), and cumulative abnormal returns in the past week ($AbRet_i[-5,-1]$). See text for further details. In Panel B, the symbols **, *, and $^{+}$ denote statistical significance at the 1%, 5%, and 10% levels, respectively. Significance is based on Newey-West (1987) standard errors that are robust to heteroskedasticity and five days of autocorrelation.

Panel A: Distributions of the Raw and Residual Staleness Measures

	$Stale1_{it}$	$Stale2_{it}$	$stale1_{it}$	$stale2_{it}$
Mean	0.117	0.077	0.000	0.000
Standard Deviation	0.068	0.064	1.000	1.000
5th Percentile	0.058	0.028	-1.033	-0.969
10th Percentile	0.063	0.031	-0.882	-0.827
25th Percentile	0.075	0.039	-0.606	-0.580
50th Percentile	0.097	0.056	-0.230	-0.236
75th Percentile	0.133	0.090	0.306	0.284
90th Percentile	0.194	0.150	1.110	1.118
95th Percentile	0.251	0.206	1.867	1.920

Panel B: Correlations of Firm and News Characteristics with the Staleness Measures

	$Stale1_{it}$	$Stale2_{it}$	$stale1_{it}$	$stale2_{it}$
Msg_{it}	0.117^{**}	0.044^{**}	0.000	-0.006^{**}
$Words_{it}$	-0.213^{**}	-0.130^{**}	0.000	0.002^{**}
$(Words_{it})^2$	-0.006	0.000	0.000	0.001^{**}
$Earn_{it}$	-0.089^{**}	-0.120^{**}	-0.057^{**}	-0.089^{**}
$AbRet_{it}$	-0.107^{**}	-0.007^{**}	-0.003^{**}	-0.002
$\|AbRet_{it}\|$	-0.019^{**}	-0.055^{**}	-0.032^{**}	-0.054^{**}
$Abturn_{it}$	-0.039^{**}	-0.090^{**}	-0.058^{**}	-0.089^{**}
$MktCap_{i,t-1}$	0.023^{**}	0.055^{**}	0.004	0.040^{**}
$Msg_i[-5,-1]$	0.109^{**}	0.120^{**}	0.117^{**}	0.116^{**}
$IdVolat_i[-5,-1]$	0.002	-0.005	0.016^{**}	0.003
$Illiq_i[-5,-1]$	-0.017^{**}	-0.052^{**}	-0.006^{*}	-0.041^{**}

Table 2: Relating Absolute Returns and Trading Volume to News Staleness

Table 2 displays results from daily cross-sectional Fama-MacBeth (1973) regressions of either absolute abnormal returns ($|AbRet_{it}|$) or abnormal turnover ($AbTurn_{it}$) during day t on news and control variables known by day t. The table reports the time-series average of the coefficients, R^2 statistics, and value of the dependent variable. The single-word and bigram measures of textual staleness are $stale1_{it}$ and $stale2_{it}$. Control variables include newswire messages (Msg_{it}), prior-week abnormal newswire messages ($Msg_i[-5,-1]$), words in the news ($Words_{it}$), earnings-related news ($Earn_{it}$), firm size ($MktCap_{i,t-1}$), abnormal turnover ($AbTurn_{it}$), prior-week idiosyncratic volatility ($IdVolat_i[-5,-1]$), prior-week illiquidity ($Illiq_i[-5,-1]$), and cumulative abnormal returns in the past week ($AbRet_i[-5,-1]$). See text for further details. The regression intercept is the average value of the dependent variable on day t. Columns three and four in the table separately examine subsamples of below-median and above-median firm size. I standardize all independent variables by day so that the coefficient units are basis points per standard deviation increase in the independent variables. Newey-West (1987) standard errors robust to heteroskedasticity and five days of autocorrelation appear in parentheses. All coefficients below are statistically significant at the 1% level.

Dependent Variable	$\|AbRet_{it}\|$	$\|AbRet_{it}\|$	$\|AbRet_{it}\|$	$\|AbRet_{it}\|$	$AbTurn_{it}$	$AbTurn_{it}$
Firms Included	All	All	Small	Big	All	All
$stale1_{it}$	-0.104		-0.203	-0.036	-0.038	
	(0.007)		(0.012)	(0.005)	(0.002)	
$stale2_{it}$		-0.143				-0.048
		(0.008)				(0.002)
Msg_{it}	1.111	1.046	1.740	0.536	0.227	0.217
	(0.015)	(0.016)	(0.030)	(0.011)	(0.003)	(0.003)
$Msg_i[-5,-1]$	-0.168	-0.150	-0.188	-0.092	-0.039	-0.037
	(0.007)	(0.007)	(0.010)	(0.005)	(0.001)	(0.001)
$Words_{it}$	0.108	0.085	0.396	-0.005	0.000	-0.005
	(0.007)	(0.007)	(0.016)	(0.005)	(0.002)	(0.002)
$Earn_{it}$	0.092	0.080	0.040	0.135	0.056	0.050
	(0.007)	(0.007)	(0.011)	(0.006)	(0.001)	(0.001)
$MktCap_{i,t-1}$	-1.221	-1.152	-1.767	-0.789	-0.160	-0.152
	(0.018)	(0.018)	(0.042)	(0.015)	(0.003)	(0.003)
$IdVolat_i[-5,-1]$	0.836	0.776	0.828	0.805	-0.014	-0.018
	(0.021)	(0.020)	(0.023)	(0.028)	(0.002)	(0.002)
$Illiq_i[-5,-1]$	-0.244	-0.216	-0.387	-1.383	0.060	0.058
	(0.011)	(0.012)	(0.016)	(0.122)	(0.003)	(0.003)
$AbRet_i[-5,-1]$	-0.185	-0.159	-0.197	-0.151	-0.018	-0.017
	(0.013)	(0.014)	(0.015)	(0.013)	(0.002)	(0.002)
Average Value	2.894	2.733	2.614	1.935	0.262	0.239
	(0.043)	(0.043)	(0.050)	(0.070)	(0.006)	(0.007)
Trading Days	2853	2652	2367	2238	2853	2652
Obs per Day	333	305	197	178	333	305
R^2	20.7%	20.9%	20.2%	21.2%	18.8%	18.6%

Table 3: Predicting Return Reversals after News Using Staleness and Control Variables

Table 3 presents results from daily cross-sectional Fama-MacBeth (1973) regressions of raw post-news firm returns from days *t + 2* through *t + 5* on news-related variables and other control variables known by day *t*. The table reports the time-series average of the coefficients, return from days *t + 2* through *t* + 5, and R^2 statistics. The independent variable for the firm's abnormal return on day-*t* ($AbRet_{it}$) measures return reversal. The interaction terms with $AbRet_{it}$ allow reversal to depend on the two textual staleness measures ($stale1_{it}$ and $stale2_{it}$), newswire messages (Msg_{it}), earnings-related news ($Earn_{it}$), firm size ($MktCap_{i,t-1}$), abnormal turnover ($AbTurn_{it}$), prior-week idiosyncratic volatility ($IdVolat_i[-5,-1]$), prior-week illiquidity ($Illiq_i[-5,-1]$), words in the news ($Words_{it}$), and prior-week abnormal newswire messages ($Msg_i[-5,-1]$). Each regression controls for the direct effects of all included interaction term variables. The other control variable is the firm's cumulative abnormal return in the past week ($AbRet_i[-5,-1]$). See text for variable construction details. I standardize all independent variables by day so that the coefficient units are basis points per standard deviation increase in the independent variables. Newey-West (1987) standard errors that are robust to heteroskedasticity and five days of autocorrelation appear in parentheses. The symbols **, *, and $^{+}$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	$Ret_i[2,5]$	$Ret_i[2,5]$	$Ret_i[2,5]$	$Ret_i[2,5]$	$Ret_i[2,5]$	$Ret_i[2,5]$
Controls Included	None	None	Main	Main	All	All
$AbRet_{it}$	0.015	-0.003	-0.049*	-0.076**	-0.053*	-0.093**
	(0.019)	(0.020)	(0.022)	(0.024)	(0.024)	(0.026)
$AbRet_{it}*stale1_{it}$	-0.057**		-0.059**		-0.066**	
	(0.016)		(0.016)		(0.017)	
$stale1_{it}$	-0.024^{+}		-0.007		-0.007	
	(0.014)		(0.012)		(0.012)	
$AbRet_{it}*stale2_{it}$		-0.053**		-0.046**		-0.047**
		(0.016)		(0.017)		(0.018)
$stale2_{it}$		-0.029^{+}		-0.025^{+}		-0.021
		(0.016)		(0.014)		(0.014)
$AbRet_{it}*Msg_{it}$			0.057**	0.069**	0.023	0.027
			(0.014)	(0.015)	(0.016)	(0.018)
Msg_{it}			-0.033**	-0.018	-0.030*	-0.017
			(0.012)	(0.013)	(0.012)	(0.013)
$AbRet_{it}*Earn_{it}$			0.032*	0.021	0.032*	0.017
			(0.013)	(0.015)	(0.014)	(0.017)
$Earn_{it}$			0.001	-0.002	0.001	-0.002
			(0.011)	(0.011)	(0.011)	(0.011)
$AbRet_{it}*MktCap_{i,t-1}$			-0.065**	-0.063**	-0.122**	-0.114**
			(0.018)	(0.020)	(0.027)	(0.028)
$MktCap_{i,t-1}$			-0.115**	-0.103**	-0.116**	-0.108**
			(0.029)	(0.030)	(0.029)	(0.030)
$AbRet_i[-5,-1]$			-0.060**	-0.060**	-0.059**	-0.061**
			(0.022)	(0.023)	(0.023)	(0.023)
$AbTurn_{it}$			0.060**	0.054**	0.055**	0.053**
			(0.011)	(0.011)	(0.011)	(0.011)
$IdVolat_i[-5,-1]$			-0.067*	-0.044	-0.060^{+}	-0.044
			(0.030)	(0.031)	(0.033)	(0.033)
$Illiq_i[-5,-1]$			-0.156**	-0.141**	-0.157**	-0.144**

			(0.023)	(0.023)	(0.023)	(0.024)
$Words_{it}$			-0.030*	-0.031**	-0.028*	-0.029*
			(0.012)	(0.012)	(0.012)	(0.012)
$Msg_i[-5,-1]$			0.006	0.013	0.005	0.014
			(0.010)	(0.010)	(0.010)	(0.010)
$AbRet_{it}*AbTurn_{it}$					0.038**	0.045**
					(0.014)	(0.015)
$AbRet_{it}*IdVolat_i[-5,-1]$					-0.092**	-0.076**
					(0.015)	(0.016)
$AbRet_{it}*Illiq_i[-5,-1]$					-0.045*	-0.051**
					(0.018)	(0.020)
$AbRet_{it}*Words_{it}$					0.006	-0.007
					(0.016)	(0.017)
$AbRet_{it}* Msg_i[-5,-1]$					-0.007	-0.011
					(0.014)	(0.016)
Average $Ret_i[2,5]$	0.101	0.108	0.112	0.114	0.107	0.109
	(0.086)	(0.086)	(0.087)	(0.088)	(0.087)	(0.088)
Trading Days	2863	2661	2853	2652	2853	2652
Observations per Day	337	309	333	305	333	305
R^2	2.7%	2.9%	12.3%	13.0%	15.8%	16.8%

Table 4: Predicting Return Reversals at Different Time Horizons Using Staleness

Table 4 presents daily cross-sectional Fama-MacBeth (1973) regressions of raw firm returns from days $t + b$ through $t + e$ on news variables and control variables known by day t. The three [b, e] intervals in the columns are [1,1], [2,2] and [2,10]. The table reports the time-series average of the regression coefficients, returns on days [b, e], and R^2 statistics. The coefficient on the firm's abnormal return on day t ($AbRet_{it}$) measures return reversal. Interaction terms with $AbRet_{it}$ allow reversal to depend on the two textual staleness measures ($stale1_{it}$ and $stale2_{it}$), newswire messages (Msg_{it}), earnings news ($Earn_{it}$), and firm size ($MktCap_{i,t-1}$). Other controls include the direct effects of these variables, weekly returns ($AbRet_i[-5,-1]$), turnover ($AbTurn_{it}$), weekly volatility ($IdVolat_i[-5,-1]$), weekly illiquidity ($Illiq_i[-5,-1]$), words in the news ($Words_{it}$), and weekly newswire messages ($Msg_i[-5,-1]$). The coefficient units are basis points per standard deviation of the independent variables. Newey-West (1987) standard errors robust to heteroskedasticity and five days of autocorrelation appear in parentheses. The symbols **, *, and $^{+}$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	$Ret_i[1,1]$	$Ret_i[1,1]$	$Ret_i[2,2]$	$Ret_i[2,2]$	$Ret_i[2,10]$	$Ret_i[2,10]$
Controls Included	Main	Main	Main	Main	Main	Main
$AbRet_{it}$	0.041**	0.025^{+}	-0.043**	-0.049**	-0.042	-0.078*
	(0.013)	(0.013)	(0.012)	(0.012)	(0.032)	(0.034)
$AbRet_{it}$*$stale1_{it}$	-0.012		-0.015^{+}		-0.069**	
	(0.010)		(0.008)		(0.022)	
$stale1_{it}$	-0.009^{+}		-0.008		-0.031	
	(0.005)		(0.005)		(0.021)	
$AbRet_{it}$*$stale2_{it}$		-0.010		-0.014		-0.052*
		(0.011)		(0.009)		(0.024)
$stale2_{it}$		-0.010^{+}		-0.010^{+}		-0.048^{+}
		(0.005)		(0.005)		(0.025)
$AbRet_{it}$*Msg_{it}	0.038**	0.048**	0.023**	0.032**	0.088**	0.101**
	(0.009)	(0.009)	(0.007)	(0.007)	(0.018)	(0.019)
Msg_{it}	-0.029**	-0.031**	-0.021**	-0.014**	-0.051**	-0.032
	(0.005)	(0.006)	(0.005)	(0.005)	(0.017)	(0.019)
$AbRet_{it}$*$Earn_{it}$	0.043**	0.050**	0.015*	0.012	0.048**	0.043*
	(0.009)	(0.010)	(0.007)	(0.008)	(0.018)	(0.020)
$Earn_{it}$	0.010*	0.003	-0.011**	-0.008	0.032^{+}	0.039*
	(0.005)	(0.005)	(0.004)	(0.005)	(0.018)	(0.018)
$AbRet_{it}$*$MktCap_{i,t-1}$	-0.004	-0.010	-0.021*	-0.018^{+}	-0.066**	-0.057
	(0.011)	(0.011)	(0.010)	(0.010)	(0.026)	(0.028)
$MktCap_{i,t-1}$	-0.017^{+}	-0.013	-0.013	-0.006	-0.182**	-0.163**
	(0.010)	(0.010)	(0.009)	(0.010)	(0.052)	(0.052)
$AbRet_i[-5,-1]$	-0.087**	-0.083**	-0.039**	-0.036**	-0.022	-0.017
	(0.010)	(0.011)	(0.009)	(0.009)	(0.039)	(0.040)
Average Return	0.042^{+}	0.039	0.026	0.022	0.254	0.265
	(0.025)	(0.025)	(0.024)	(0.025)	(0.159)	(0.162)
Trading Days	2853	2652	2853	2652	2853	2652
Obs per Day	333	305	333	305	333	305
R^2	13.6%	14.3%	12.2%	13.0%	11.9%	12.6%

Table 5: Predicting Return Reversals Using Staleness in Different Subsamples

Table 5 presents daily cross-sectional Fama-MacBeth (1973) regressions of raw firm returns from days $t + 2$ through $t + 5$ on news variables and control variables measured prior to day t. The six columns show different subsamples: years 1996 to 2001, and 2002 to 2008; above-average firm size ($MktCap_{i,t-1}$), and below-average size; news days with earnings-related words as measured by $Earn_{it}$, and news days without earnings news. The table reports the time-series average of the regression coefficients, returns on days $t + 2$ through $t + 5$, and R^2 statistics. The coefficient on the firm's abnormal return on day t ($AbRet_{it}$) measures return reversal. Interaction terms with $AbRet_{it}$ allow reversal to depend on the single-word staleness measure ($stale1_{it}$), newswire messages (Msg_{it}), earnings news ($Earn_{it}$), and firm size ($MktCap_{i,t-1}$). Other controls include the direct effects of these variables, weekly returns ($AbRet_i[-5,-1]$), turnover ($AbTurn_{it}$), weekly volatility ($IdVolat_i[-5,-1]$), weekly illiquidity ($Illiq_i[-5,-1]$), words in the news ($Words_{it}$), and weekly newswire messages ($Msg_i[-5,-1]$). The coefficient units are basis points per standard deviation of the independent variables. Newey-West (1987) standard errors robust to heteroskedasticity and five days of autocorrelation appear in parentheses. The symbols **, *, and $^{+}$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	$Ret_i[2,5]$	$Ret_i[2,5]$	$Ret_i[2,5]$	$Ret_i[2,5]$	$Ret_i[2,5]$	$Ret_i[2,5]$
Subsample	96 to 01	02 to 08	Small	Big	Earnings	No Earnings
$AbRet_{it}$	-0.067^{+}	-0.032	-0.008	-0.010	0.026	-0.112**
	(0.037)	(0.024)	(0.039)	(0.044)	(0.042)	(0.027)
$AbRet_{it}*stale1_{it}$	-0.092**	-0.027	-0.082**	-0.038	-0.102*	-0.049*
	(0.026)	(0.018)	(0.020)	(0.027)	(0.043)	(0.022)
$stale1_{it}$	-0.028	0.014	-0.007	-0.013	0.035	-0.013
	(0.022)	(0.010)	(0.017)	(0.014)	(0.032)	(0.013)
$AbRet_{it}*Msg_{it}$	0.088**	0.026^{+}	0.046**	0.120**	0.035	0.043^{+}
	(0.024)	(0.014)	(0.016)	(0.024)	(0.025)	(0.022)
Msg_{it}	-0.053**	-0.013	-0.086**	-0.026	-0.014	-0.040**
	(0.020)	(0.012)	(0.020)	(0.014)	(0.024)	(0.015)
$AbRet_{it}*Earn_{it}$	0.030	0.033**	0.008	0.046^{+}		
	(0.024)	(0.013)	(0.017)	(0.027)		
$Earn_{it}$	0.002	-0.001	0.029	-0.028		
	(0.020)	(0.009)	(0.015)	(0.013)		
$AbRet_{it}*MktCap_{i,t-1}$	-0.048	-0.083**	-0.013	-0.208**	-0.046	-0.095**
	(0.032)	(0.017)	(0.035)	(0.044)	(0.029)	(0.023)
$MktCap_{i,t-1}$	-0.139**	-0.090**	-0.438**	-0.037	-0.105	-0.091**
	(0.052)	(0.028)	(0.062)	(0.034)	(0.051)	(0.033)
$AbRet_i[-5,-1]$	-0.101*	-0.020	-0.054*	-0.056	-0.038	-0.087**
	(0.040)	(0.020)	(0.024)	(0.035)	(0.038)	(0.026)
Average $Ret_i[2,5]$	0.134	0.091	-0.159	-0.340*	0.097	0.112
	(0.137)	(0.106)	(0.100)	(0.159)	(0.117)	(0.092)
Trading Days	1418	1435	2367	2238	925	2504
Obs per Day	273	391	197	178	165	257
R^2	13.8%	10.8%	14.1%	17.7%	15.8%	12.9%

Table 6: Returns of Calendar Time Portfolios Formed on Reactions to Stale News

Table 6 displays the risk factor loadings and daily alphas of stale momentum portfolios. These portfolios are formed each day by buying firms in the top decile of either $AbRet_{it}*stale1_{it}$ or $AbRet_{it}*stale2_{it}$ and selling firms in the bottom decile of $AbRet_{it}*stale1_{it}$ or $AbRet_{it}*stale2_{it}$. The four panels below show the daily equal- and value-weighted returns for the reversal portfolios using the two (single-word or bigram) staleness measures. The columns show reversal portfolios held for different time horizons after news events on day *t*: days [0,0], [1,1], [2,2], [2,5], [1,5], and [1,10]. Each daily alpha is the intercept from a time series regression of portfolio returns on four factors: the market (MKT), size (SMB), book-to-market (HML), and momentum (UMD) factors available on Kenneth French's web site. Cumulative alphas equal daily alphas times days in the holding period. Newey and West (1987) standard errors robust to heteroskedasticity and five days of autocorrelation appear in parentheses below the coefficients. The symbols **, *, and $^{+}$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Equal-weighted Returns, Formed on Top minus Bottom Decile of $AbRet_{it}*stale1_{it}$						
Holding Period	[0,0]	[1,1]	[2,2]	[2,5]	[1,5]	[1,10]
MKT	-0.079	0.001	-0.019	0.007	0.011	-0.002
	(0.057)	(0.026)	(0.023)	(0.014)	(0.042)	(0.011)
SMB	-0.333**	-0.020	0.017	0.027	0.020	-0.002
	(0.118)	(0.056)	(0.040)	(0.025)	(0.023)	(0.015)
HML	-0.156	0.031	0.027	0.030	0.034	0.026
	(0.125)	(0.054)	(0.045)	(0.026)	(0.026)	(0.021)
UMD	-0.036	0.037	0.027	0.033^{+}	0.043**	0.034**
	(0.081)	(0.034)	(0.030)	(0.018)	(0.016)	(0.013)
Daily Alpha	-4.133**	-0.089**	-0.059**	-0.043**	-0.052**	-0.030**
	(0.100)	(0.026)	(0.021)	(0.011)	(0.011)	(0.008)
Cumulative Alpha	-4.133**	-0.089**	-0.059**	-0.175**	-0.262**	-0.301**
	(0.100)	(0.026)	(0.021)	(0.045)	(0.055)	(0.085)
Trading Days	2863	2861	2859	2947	2956	2980
R^2	0.3%	0.1%	0.1%	0.2%	0.3%	0.3%
Panel B: Equal-weighted Returns, Formed on Top minus Bottom Decile of $AbRet_{it}*stale2_{it}$						
Holding Period	[0,0]	[1,1]	[2,2]	[2,5]	[1,5]	[1,10]
Cumulative Alpha	-5.219**	-0.140**	-0.033	-0.129**	-0.270**	-0.210*
	(0.103)	(0.027)	(0.024)	(0.050)	(0.065)	(0.107)
Trading Days	2662	2661	2659	2830	2851	2911
R^2	0.2%	0.1%	0.3%	0.2%	0.1%	0.2%
Panel C: Value-weighted Returns, Formed on Top minus Bottom Decile of $AbRet_{it}*stale1_{it}$						
Holding Period	[0,0]	[1,1]	[2,2]	[2,5]	[1,5]	[1,10]
Cumulative Alpha	-0.748**	-0.057^{+}	-0.110**	-0.069	-0.148^{+}	-0.186^{+}
	(0.092)	(0.032)	(0.033)	(0.076)	(0.086)	(0.116)
R^2	0.2%	0.2%	0.1%	0.0%	0.1%	0.1%
Panel D: Value-weighted Returns, Formed on Top minus Bottom Decile of $AbRet_{it}*stale2_{it}$						
Holding Period	[0,0]	[1,1]	[2,2]	[2,5]	[1,5]	[1,10]
Cumulative Alpha	-0.876**	-0.062^{+}	-0.125**	-0.153*	-0.168**	-0.144
	(0.089)	(0.033)	(0.032)	(0.070)	(0.64)	(0.128)
R^2	0.2%	0.1%	0.2%	0.5%	0.4%	0.2%

Table 7: Individual and Institutional Market Order Imbalances After News

Table 7 presents daily cross-sectional Fama-MacBeth (1973) regressions of individual (*IndBSI*) and institutional market order imbalances (*InstBSI*) from days $t + b$ through $t + e$ on news and control variables. The three columns show [b, e] intervals of [0,0], [1,1], and [2,5]. The table reports the time-series average of the regression coefficients, imbalances on days [b, e], and R^2 statistics. The coefficient on the firm's abnormal return on day t ($AbRet_{it}$) measures aggressive momentum or contrarian trading. Interaction terms with $AbRet_{it}$ allow momentum trading to depend on the two textual staleness measures ($stale1_{it}$ and $stale2_{it}$), newswire messages (Msg_{it}), earnings news ($Earn_{it}$), and firm size ($MktCap_{i,t-1}$). The complete set of independent variables is the same as in Table 4 and Table 5. See these tables for details. The coefficient units are basis points per standard deviation of the independent variables. Newey-West (1987) standard errors robust to heteroskedasticity and five days of autocorrelation appear in parentheses. The symbols **, *, and $^{+}$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	*IndBSI*	*IndBSI*	*IndBSI*	*IndBSI*	*InstImb*	*InstBSI*	*InstBSI*	*InstBSI*
Individual Activity	[0,0]	[0,0]	[1,1]	[2,5]	[0,0]	[0,0]	[1,1]	[2,5]
Controls Included	Main	Main	Main	Main	Main	Main	Main	Main
$AbRet_{it}$	4.137**	3.825**	1.266**	-0.023	1.339**	1.078**	0.459*	0.021
	(0.219)	(0.215)	(0.199)	(0.126)	(0.179)	(0.174)	(0.172)	(0.090)
$AbRet_{it}$*$stale1_{it}$	0.900**		0.212	-0.134	-0.051		-0.195	0.077
	(0.196)		(0.182)	(0.107)	(0.138)		(0.132)	(0.071)
$stale1_{it}$	0.176		0.195^{+}	0.479**	0.335**		0.149^{+}	0.239**
	(0.145)		(0.120)	(0.087)	(0.086)		(0.080)	(0.051)
$AbRet_{it}$*$stale2_{it}$		1.045**				0.049		
		(0.227)				(0.143)		
$stale2_{it}$		0.224				0.341**		
		(0.153)				(0.090)		
$AbRet_{it}$*Msg_{it}	-3.029**	-2.651**	-1.255**	-0.372**	-0.566**	-0.463**	-0.279**	0.045
	(0.137)	(0.144)	(0.114)	(0.073)	(0.113)	(0.114)	(0.104)	(0.049)
Msg_{it}	-0.406**	-0.046	-0.971**	-0.725**	-0.320**	-0.288**	-0.364**	-0.161**
	(0.146)	(0.148)	(0.134)	(0.087)	(0.106)	(0.106)	(0.098)	(0.054)
$AbRet_{it}$*$Earn_{it}$	-0.372**	-0.407**	-0.115	0.173*	-0.072	-0.119	0.036	0.125*
	(0.126)	(0.145)	(0.140)	(0.084)	(0.137)	(0.140)	(0.112)	(0.056)
$Earn_{it}$	0.215	0.053	0.225^{+}	0.316**	0.026	0.065	-0.074	0.010
	(0.137)	(0.150)	(0.129)	(0.080)	(0.086)	(0.098)	(0.085)	(0.046)
$AbRet_{it}$*$MktCap_{i,t-1}$	-1.330**	-1.322**	-0.217	-0.130	0.702**	0.444**	0.278^{+}	0.044
	(0.174)	(0.184)	(0.175)	(0.103)	(0.132)	(0.144)	(0.145)	(0.071)
$MktCap_{i,t-1}$	-2.843**	-3.231**	-2.084**	-2.262**	-1.327**	-1.431**	-1.385**	-1.636**
	(0.346)	(0.368)	(0.368)	(0.320)	(0.271)	(0.267)	(0.248)	(0.217)
$AbRet_i[-5,-1]$	-0.441**	-0.307*	-0.774**	-0.443**	0.029	0.048	0.065	0.024
	(0.141)	(0.148)	(0.133)	(0.083)	(0.100)	(0.110)	(0.091)	(0.059)
Average Imbalance	-7.379**	-8.118**	-7.343**	-7.611**	0.087	-0.138	-0.005	0.049
	(0.480)	(0.454)	(0.394)	(0.376)	(0.228)	(0.229)	(0.209)	(0.178)
Trading Days	1148	1084	1148	1148	1148	1084	1148	1148
Obs per Day	301	275	301	301	301	275	301	301
R^2	7.4%	7.9%	7.3%	8.1%	7.9%	8.3%	7.5%	7.9%

Table 8: Predicting Return Reversals Using Staleness and Individual Investor Activity

Table 8 presents daily cross-sectional Fama-MacBeth (1973) regressions of raw firm returns from days $t + 2$ through $t + 5$ on news variables and control variables known by day t. The columns show subsamples with above-median and below-median individual investor trading activity ($IndAct_{it}$). The table reports the time-series average of the regression coefficients, returns on days $t + 2$ through $t + 5$, and R^2 statistics. The coefficient on the firm's abnormal return on day t ($AbRet_{it}$) measures return reversal. Interaction terms with $AbRet_{it}$ allow reversal to depend on the two textual staleness measures ($stale1_{it}$ and $stale2_{it}$), newswire messages (Msg_{it}), earnings news ($Earn_{it}$), and firm size ($MktCap_{i,t-1}$). The complete set of independent variables is the same as in Table 4 and Table 5. See these tables for details. The coefficient units are basis points per standard deviation of the independent variables. Newey-West (1987) standard errors robust to heteroskedasticity and five days of autocorrelation appear in parentheses. The symbols **, *, and $^{+}$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	$Ret_i[2,5]$	$Ret_i[2,5]$	$Ret_i[2,5]$	$Ret_i[2,5]$
Individual Activity	High	Low	High	Low
Controls Included	Main	Main	Main	Main
$AbRet_{it}$	-0.044	-0.035	-0.069*	-0.050
	(0.030)	(0.034)	(0.032)	(0.035)
$AbRet_{it}*stale1_{it}$	-0.063*	0.002		
	(0.026)	(0.037)		
$stale1_{it}$	0.026	-0.004		
	(0.018)	(0.014)		
$AbRet_{it}*stale2_{it}$			-0.091**	0.020
			(0.027)	(0.040)
$stale2_{it}$			0.033	-0.013
			(0.018)	(0.015)
$AbRet_{it}*Msg_{it}$	0.011	0.035	0.016	0.032
	(0.018)	(0.025)	(0.020)	(0.026)
Msg_{it}	-0.013	-0.018	-0.002	-0.016
	(0.019)	(0.012)	(0.019)	(0.013)
$AbRet_{it}*Earn_{it}$	0.021	0.046^{+}	0.018	0.054*
	(0.020)	(0.027)	(0.021)	(0.027)
$Earn_{it}$	-0.002	-0.002	-0.012	-0.004
	(0.018)	(0.013)	(0.020)	(0.014)
$AbRet_{it}*MktCap_{i,t-1}$	-0.105**	-0.082**	-0.113**	-0.079**
	(0.028)	(0.027)	(0.027)	(0.030)
$MktCap_{i,t-1}$	-0.078^{+}	-0.064*	-0.089**	-0.066*
	(0.041)	(0.028)	(0.042)	(0.030)
$AbRet_i[-5,-1]$	-0.001	-0.021	0.003	-0.031
	(0.024)	(0.028)	(0.025)	(0.030)
Average Imbalance	0.142	0.186^{+}	0.120	0.193^{+}
	(0.107)	(0.096)	(0.107)	(0.098)
Trading Days	967	970	912	919
Obs per Day	165	165	149	149
R^2	15.2%	18.0%	16.3%	19.7%