The Power of Prices: Information, Trade, and Salient Returns[a]

Fabian Brunner and Michael Ungeheuer[b]

November 2019

Abstract

We show that a considerable part of information acquisition and trading is driven by salient returns, above and beyond the effects of underlying news shocks. To establish causality, we first estimate the effect of overnight earnings announcements on information acquisition before the market opens versus after. Larger surprises are associated with more information acquisition only after the open, in line with salient returns as a cause. Second, we exploit *Wall Street Journal* stock rankings to show that prominently placed salient returns drive information acquisition and trading. Last, we document that stocks experiencing salient returns are particularly mispriced, in line with return-induced noise trading as a major moderator of anomalies.

Keywords: Salience, information placement, returns, information acquisition, trading activity, mispricing.

JEL Classification Numbers: G12, G14.

[a] We are grateful for helpful comments from Peter Bossaerts, Zhi Da, Ralf Elsas, Samuel Hartzmark, Alexander Hillert, Zwetelina Iliewa, Lena Jaroszek, Thomas Johann, Alexander Klos, Kasper Meisner Nielsen, Albert Menkveld, Christoph Merkle, Terrance Odean, Stefan Ruenzi, Nic Schaub, Stefanie Schraeder, and Mark Seasholes, as well as seminar participants at the University of Mannheim, Aalto University, and Ludwig Maximilian University of Munich. We would particularly like to thank Anja Kunzmann for her contributions to an early version of the project. We appreciate the technical support from Antti Lehtinen and Miko Niemenmaa, as well as research assistance from Otso Ärmänen and Samuel Taylor. Some of the material in Subsection 4.3 is based on an earlier working paper by Ungeheuer titled "Stock Returns and the Cross-Section of Investor Attention."

[b] Fabian Brunner: University of Mannheim, brunner@bwl.uni-mannheim.de. Michael Ungeheuer: Aalto University, michael.ungeheuer@aalto.fi.

The Power of Prices: Information, Trade, and Salient Returns

Abstract

We show that a considerable part of information acquisition and trading is driven by salient returns, above and beyond the effects of underlying news shocks. To establish causality, we first estimate the effect of overnight earnings announcements on information acquisition before the market opens versus after. Larger surprises are associated with more information acquisition only after the open, in line with salient returns as a cause. Second, we exploit *Wall Street Journal* stock rankings to show that prominently placed salient returns drive information acquisition and trading. Last, we document that stocks experiencing salient returns are particularly mispriced, in line with return-induced noise trading as a major moderator of anomalies.

Keywords: Salience, information placement, returns, information acquisition, trading activity, mispricing.

JEL Classification Numbers: G12, G14.

1 Introduction

Figure 1 reports the relation between daily stock returns and subsequent information acquisition (SEC EDGAR downloads, Panel A) and trading (CRSP turnover, Panel B).

[Insert Figure 1 about here.]

Information acquisition and trading are roughly proportional to absolute deviations of stock returns from the day's median. A one percentage point higher magnitude of daily returns is associated with an increase of more than 2% in abnormal information acquisition, as well as more than 4% in abnormal trading activity. Is this strong association driven by underlying information shocks (like earnings surprises) that drive returns, as well as information acquisition and trading? Or do investors also react directly to salient returns? After all, returns (relative price changes) are an omnipresent feature of financial markets. They are placed prominently by exchanges (e.g., NYSE and NASDAQ), financial information intermediaries (e.g., Bloomberg, Reuters), and the news media (e.g., as part of newspaper articles on the *WSJ* website), as documented in Figure 2.

[Insert Figure 2 about here.]

We provide evidence in line with the latter conjecture: Prominently placed salient returns drive information acquisition and trading, above and beyond a direct effect of underlying information shocks. This moderating role of observed returns provides a piece to the puzzle of the high levels of costly information acquisition and trading in financial markets (Milgrom and Stokey, 1982; Grossman and Stiglitz, 1980). We use two settings to show a causal effect of return salience on information acquisition and trading. First, we study information acquisition following overnight earnings announcements. In this context, we can compare the information acquisition activity for firms with large versus small absolute earnings surprises in the hours before and after the market opens using a difference-in-difference methodology. Only after the market opens and salient returns become available do large surprise stocks exhibit higher information acquisition activity. Hence, the availability of a salient return causally increases information acquisition activity beyond underlying information shocks.

Second, we exploit daily gainer and decliner rankings in the *Wall Street Journal* (*WSJ*) as a particularly prominent placement of salient returns. We show that stocks ranked in the

WSJ experience significantly more subsequent information acquisition and trading activity, compared to stocks with equally extreme returns that are not ranked in the *WSJ*. Hence, we provide strong evidence for a causal role of the placement of returns in determining information acquisition and trading activity.

Last, we establish a connection between return-induced activity and mispricing. Under the assumption that a considerable part of return-induced trading reflects the systematic activity of noise traders, one would expect amplified mispricing after stocks experience prominent salient returns (Grossman and Stiglitz, 1980; Kyle, 1985). We use Stambaugh et al. (2012)'s mispricing score to show that mispricing is indeed amplified after return-induced trading, even in a universe of large, liquid stocks with high institutional ownership. While part of this effect might be caused by the arbitrage risk of stocks with salient returns (Stambaugh et al., 2012), our finding provides a different view of mispricing, via the ultimate cause, systematic noise trading, instead of subsequent impediments to liquidity-provision for noise traders.

To analyze the effect of returns on information acquisition and trading activity, we focus on U.S. common stocks traded on major exchanges after excluding micro-caps and penny-stocks for the period 2007 to 2016. Our information acquisition measures are based on the number of downloads of firm specific information from the SEC's EDGAR platform (for more sophisticated investors) and the number of Wikipedia firm page views (for less sophisticated investors). For trading activity, we consider both total turnover and retail turnover (as in Boehmer et al., 2019).

As a point of departure, we analyze the reaction of information acquisition after the market closes (after 4:00 PM) to realized returns since the previous market close. In line with the visual impression from Figure 1's Panel A, we find that stocks experience around 2.5% more information acquisition on EDGAR per percentage point return deviation.

To provide causal evidence for an effect of returns beyond underlying information shocks, we analyze effects around overnight earnings announcements. Specifically, we focus on firm-days with earnings announcements before 7:30 AM. We then compare stocks with large absolute earnings surprises to stocks with small absolute surprises based on ex-ante analyst estimates. Large surprises, near-instantaneously transmitted to investors via newswires, provide an information shock well before the market opens *without a contemporaneous salient return*. As returns become available at the open (9:30 AM), firms with large absolute earnings surprises predictably experience large absolute returns. This event lets us separate the effect of underlying information shocks from the effect of salient returns in a difference-in-difference setting.

We find that directly before the market opens (but after the announcement of the earnings surprise), firms with large earnings surprises do not experience larger levels of information acquisition than firms with small earnings surprises. Both, the group of large and small surprise firms experience high levels of information acquisition well before the market opens, but there is no significant difference between the groups. However, starting in the hour directly after the market opens, the large surprise group of firms experiences a relative increase in information acquisition. Results for Wikipedia firm page views are qualitatively the same, indicating robustness for less sophisticated investors. Hence, our difference-in-difference results are consistent with prominently placed salient returns as an important driver of the patterns in information acquisition we observe in Figure 1 beyond the impact of underlying information shocks.

We also compare the intraday dynamics of information acquisition and trading activity after overnight earnings announcements. They mirror each other, in line with trading similarly responding to observed salient returns. Both abnormal information acquisition and abnormal trading activity for large surprise firms accumulate steadily after the market opens. In particular, there is a relatively constant abnormal level of information acquisition for large surprise firms until 4:00 PM when the market closes. Similarly, abnormal trading activity for large surprise stocks remains steady throughout the day. These steady levels of abnormal information acquisition and trading are hard to reconcile with a complete materialization of the underlying cause (the overnight earnings news) when the market opens. However, salient returns are prominently displayed throughout the day after the market opens, and thus provide a simple and consistent explanation for intradaily patterns in information acquisition and trading activity after overnight earnings surprises.

Our difference-in-difference results are robust when excluding downloads of contemporaneously added EDGAR filings and stocks ranked as daily winners and losers in the media (Kumar et al., 2019). They are much weaker for algorithmic downloads from EDGAR ('robots'), in line with humans as the driver behind return-induced information acquisition. Our findings hold across firms of different size, with ex-ante varying levels of analyst coverage, information acquisition activity or institutional ownership, and across different days, with varying levels of implied volatility (VIX) or previous month market returns (up and down markets).

As a second identification strategy for the impact of prominently placed returns, we analyze gainer and decliner rankings in the *Wall Street Journal* (*WSJ*). The *WSJ* provides a daily list of stocks with the highest and lowest returns. These rankings provide a particularly prominent placement of salient returns. The stock universe used by the *WSJ* does not include all stocks of our CRSP universe. We can therefore find a control group of stocks with similarly extreme returns, which are not ranked in the *WSJ*. This methodology allows us to estimate the marginal effect of a prominent *WSJ* placement of a return, keeping the magnitude of the return constant.

We find a highly significant effect of the prominent placement of returns in the ranking on information acquisition as well as trading activity. A day's top five gainers (bottom five decliners) experience 30% (32%) more subsequent downloads on EDGAR and 26% (18%) more next-day trading activity compared to stocks with equally extreme returns but no prominent placement in the *WSJ*. We find qualitatively similar results for Wikipedia firm page views and retail trading activity, indicating that prominence in the *WSJ* also affects less sophisticated investors. The reactions to *WSJ* gainer and decliner rankings provide strong support for prominently placed salient returns as a major moderator between underlying information shocks and information acquisition, as well as trading activity.

To analyze asset pricing implications of return-induced trading, we focus on U.S. common stocks traded on major exchanges and large firms (above the 2^{nd} NYSE decile), after excluding penny-stocks (with prices below $5) for the period from August 1965 to 2015. Our measure of daily return-induced activity is based on our well-identified findings for information acquisition and trading after salient returns. We construct a monthly proxy for return-induced noise trading by averaging this daily measure. In our asset pricing tests, we then analyze the interaction between this measure for return-induced trading and Stambaugh et al. (2012)'s mispricing score in predicting abnormal returns.

We find that stocks, which are prone to mispricing according to the Stambaugh et al. (2012) score, exhibit significantly stronger abnormal returns after experiencing return-induced noise trading. This effect is statistically highly significant (t-statistics beyond 5), alleviating concerns about multiple testing (Harvey, 2017). It is also economically large. The difference in anomalous returns between mispricing quintiles one and five increases from 34 basis points per month for stocks with weak return-induced trading to 177 basis points per month for stocks with strong return-induced trading. The interaction between the mispricing score and return-induced trading remains significant even for the largest, most liquid stocks, with high institutional ownership and low short-sale constraints. It is also robust to controlling for a large battery of factor returns and firm characteristics. Hence, we provide strong support for amplified mispricing after return-induced noise trading.

We discuss our contribution to the literature in Section 2. Section 3 introduces the data used. In Section 4, we report our results on the effect of prominently placed salient returns on information acquisition and trading activity. We then explore the effect of return-induced trading on mispricing in Section 5. Section 6 concludes.

2 Related Literature

By analyzing the effect of salient returns on information acquisition and trading, we contribute to two broad and central research questions. What drives the high levels of costly information acquisition and trading activity we observe in financial markets? And how important are salience and placement of information in financial markets?

First, we contribute to the literature on the determinants of costly information acquisition and overall trading activity. The high levels of both in stock markets[1] are puzzling from the standpoint of early theory built on rational traders and efficient markets (Milgrom and Stokey, 1982). To generate large levels of trading activity, subsequent models assume the existence of noise traders with associated market inefficiencies (Grossman and Stiglitz, 1980) or specific behavioral biases that can cause investors to trade, like overconfidence (Harrison and Kreps, 1978; Scheinkman and Xiong, 2003; Daniel and Hirshleifer, 2015). Alternative models based on heterogeneous shocks to preferences, heterogeneous agents and rebalancing, or life-cycle considerations cannot generate observed levels of trading activity.

We contribute to this literature by showing the key role of prominently placed salient stock returns. They can explain large parts of the high levels of information acquisition and trading after information shocks. In other words, stocks with extreme information shocks but without prominently placed salient returns do not exhibit comparably large levels of

[1] E.g., Daniel and Hirshleifer (2015) report, that in 2014, the total dollar trade in the largest 500 U.S. stocks was $29.5 trillion, i.e., nearly double the US GDP. French (2008) estimates that "society's capitalized cost of price discovery is at least 10% of current market capitalization."

abnormal stock turnover and information acquisition.

One interpretation of this finding is that the direction of noise trader attention towards salient returns is a key driver of costly information acquisition and trading (Grossman and Stiglitz, 1980; Kyle, 1985). The activity of overconfident investors might be a driving force behind excessive trading, but it is likely to be restricted to stocks with salient returns (or similar prominent placements of information), which are actually on the radar of investors. A plausible consequence of increased noise trading after salient returns is stronger mispricing due to *systematic* noise trading. In line with this hypothesis, we find strong evidence that stocks, which are prone to mispricing ex-ante, exhibit particularly anomalous returns after such return-induced trading.

Next to the noise trader view of this association, there is a complementary, friction-based view. Stambaugh et al. (2012) argue that idiosyncratic volatility, which is driven by salient returns, is a proxy for arbitrage risk and find that stocks with higher arbitrage risk exhibit stronger mispricing. In their model, they explicitly assume that the demand of noise traders is independent of the salience of returns. In this study, we provide evidence against this assumption: Salient, i.e., large-magnitude returns drive trading activity, in line with return-induced noise trading. However, our results do not preclude arbitrage risk as a driver of mispricing. They rather provide a different view of mispricing: We analyze variation in the ultimate cause of mispricing, systematic noise trading, instead of the subsequent impediments to liquidity-provision for noise traders (such as illiquidity, short-sale constraints, or Stambaugh et al. (2012)'s arbitrage risk).[2]

Another, complementary interpretation of return-induced trading is that price changes themselves reveal the information that is relevant for investors. Building on Hayek (1945), Bossaerts et al. (2019) analyze asset markets when valuation is a computationally complex problem. They argue that prices—as opposed to underlying information shocks (or "data")—provide a signal about how to optimize portfolios. Their experimental evidence supports that prices can help investors acquire knowledge, i.e., optimization-relevant information. Extreme returns in stock markets might thus themselves drive subsequent acquisition of information relevant to improve investors' portfolio choices. This interpretation is consistent

[2] Empirically, arbitrage risk, as measured by Stambaugh et al. (2012), and return-induced trading, as documented by us, are highly similar. E.g., the correlation between idiosyncratic volatility and our measure for return-induced trading is approximately 0.8 and it is likely that even remaining differences are correlated with both arbitrage risk and noise trading.

with our finding that investors strongly react to prominently placed stock returns, holding the underlying new data (earnings surprise or magnitude of return shock) constant.[3] The interpretation of price changes as valuable signals about decision-relevant information in a computationally complex world does not preclude the above interpretation, that a significant number of noise traders persistently react to salient returns based on limited attention and behavioral biases.

Second, we contribute to the literature on how salience and the prominent placement of information affect investor decision making. This recent literature presents well-identified evidence for economically large effects of information placement on trading activity. As an illustration, Engelberg and Parsons (2011) disentangle the release of earnings announcements from their placement in newspapers and find that stock trading activity reacts strongly to newspaper coverage beyond the effect of the underlying announcement. Kaniel and Parham (2015) show that mutual fund flows react strongly to placement in a prominent *Wall Street Journal* fund ranking, beyond the effect of mutual fund performance. In a similar vein, there is evidence that fund flows to past winners (Berk and van Binsbergen, 2015; Barber et al., 2016) are driven by prominently placed performance indicators, like Morningstar Ratings, instead of underlying fund performance (Evans and Sun, 2018; Ben-David et al., 2019).

Some of this literature documents effects of salience on information efficiency, both negative effects (mostly for individual investor attention) and positive effects (mostly for institutional investor attention). On the one hand, there is evidence that individual investors trade too heavily (Odean, 1999; Barber and Odean, 2000) and particularly buy stocks in response to attention-grabbing salient information (Barber and Odean, 2008), including extreme daily returns. Da et al. (2011) provide asset pricing results in line with attention-induced price impact and reversal patterns, i.e., a negative effect of salience on market efficiency. Building on these insights, (Kumar et al., 2019) provide evidence that prominent daily winner and loser rankings lead to attention-induced upward price pressure and a subsequent reversal that can explain the puzzling underperformance of stocks with extreme daily returns (Ang et al.,

[3]It might also explain the high levels of trading activity after earnings announcements (Kim and Verrecchia, 1994; Crego, 2018). Some traders might learn more from the resulting returns than others, leading to more asymmetric information and trading. Maybe returns also provide a signal to inattentive, well-informed investors, telling them it makes sense to search information and provide liquidity for these stocks. In line with an attention-based explanation, Hendershott et al. (2018) find that pricing errors can persist for a long time when some investors are inattentive.

2006). There is also evidence that institutional investors trade as a reaction to salient returns and consequently underperform, potentially due to the price impact of their systematic trading decisions (Hartzmark, 2014; Akepanidtaworn et al., 2019).

On the other hand, Fedyk (2018) provides causal evidence that prominent placement of information on Bloomberg terminals drives up institutional investor trading. This increase in institutional trading activity immediately after the prominent placement of information seems to increase market efficiency by speeding up the pricing of new information. Ben-Rephael et al. (2017) also provide evidence that institutional attention towards news facilitates price adjustment. Consistently, Schaub (2018) shows that delayed dissemination of earnings news by data providers like First Call (a precursor of I/B/E/S) is associated with delayed reactions of financial markets.

We contribute to this literature by demonstrating that the placement and salience of returns causally affects information acquisition and trading activity. Some of these reactions are likely driven by better-informed, sophisticated investors and contribute to information efficiency. However, we also provide asset pricing evidence in line with return-induced noise trading amplifying anomalous returns for stocks that are prone to mispricing.

Prior work mostly focuses on specific contexts where predictable price effects are likely (e.g., mispricing due to retail buying pressure or more efficient pricing due to institutional informed trading). For most of our analysis (up to our asset pricing tests), we deliberately choose a broad scope that allows us to better study the puzzling overall levels of costly information acquisition and trading activity after salient returns, as illustrated in Figure 1. Stock returns are omnipresent in financial markets. They are not just prominently placed in one particular newspaper or information service, but systematically prominent throughout all kinds of information services including newspapers (like the *Wall Street Journal*), financial news providers (like Bloomberg), and stock exchanges (like the NYSE), see Figure 2. This abundance of returns makes them a plausibly highly potent stimulus for the activity of *all* investors, both institutional and individual, as well as more and less sophisticated. As returns usually reflect contemporaneous information shocks, it is an empirical challenge to measure the direct effects of returns themselves. To the best of our knowledge, we are the first to disentangle reactions to information shocks from reactions to prominently placed stock returns.

3 Data

The main variables of interest in this study are stock returns, information acquisition for stocks or firms, and trading activity in stocks.

Our primary source of data for the information acquisition of investors are log-files for the usage of company filings on the SEC EDGAR platform. The log-files contain a partially anonymized version of the IP address of the user, a time-stamp of the request as well as the identifier for the filing being requested from the server. Internet Appendix E explains in greater detail how we prepare and clean this dataset.

The Edgar platform is not the only provider of information on listed companies. We are however convinced that the Edgar platform is a relevant source of information for market participants. The platform provides high quality and non-manipulated company information in standardized form with immediate access at no cost. A considerable amount of literature documents the importance of the platform for markets and market participants in various contexts.[4]

We only include EDGAR activity that likely derives from human read-and-process usage of filings. The classification of human users is based on their daily revealed downloading behavior on the IP address level. Internet Appendix E provides more information on the user classification and corresponding criteria.

We also rely on information acquisition measured as the number of page views of Wikipedia firm pages to capture the information retrieval of less sophisticated investors. Focke et al. provide a more detailed explanation of these data.

We use stock and company information from CRSP and Compustat. We include common shares for listed companies (CRSP "shrcd" 10 and 11) traded on NYSE, AMEX or NASDAQ (CRSP "exchcd" 1, 2, and 3). Intraday stock variables such as prices and shares traded are from Algoseek, a commercial provider of minute bar data for the consolidated tape of actual trades for the US. The sample period is determined by the availability of EDGAR log-files and intraday data from Algoseek. While the SEC provides some log-files going back to 2003, there several missing and damaged files, especially in 2005 and 2006. Therefore, we include data from January 2007 to August 2016 in the sample, where the EDGAR data is of reliably

[4]These studies include Chen et al. (2018), Drake et al. (2015), Gibbons et al. (2018), Lee et al. (2015) and Li and Sun (2019).

high quality.

To mitigate the impact of small and illiquid stocks on our results, we exclude observations if the market capitalization one month ago was smaller than the first decile of all stocks listed on the NYSE or if the price one month ago was below five dollars. We also exclude days where the exchanges have special trading hours (before public holidays) as these are confounding our empirical approach.

We measure the information acquisition of investors per company and day as the number of downloads of filings for this company on the EDGAR platform or Wikipedia page views. To minimize the impact of contemporary return developments on the results, we only consider activity when the exchanges are closed from 4:00 PM until 9:00 AM of the following day. We measure abnormal information acquisition ("AEdgar" and "AWiki") as the logarithmic difference of downloads today and downloads on the same day of the week of the previous week. This approach limits the impact of varying levels of downloads for different companies, firm-specific time trends in download levels, as well as firm-specific intraweekly seasonality.

Both close-to-close (4:00 PM to 4:00 PM) as well as returns over different time horizons are generated from Algoseek data and adjusted for splits and dividends. In line with existing work on intraday market data, we employ a filter for data errors in the intraday prices. If a stock exhibits an absolute logarithmic return in excess of three within one hour and this return is reversed by at least 90% considering the next available price within six hours, we remove the outlier price from the sample.[5] We use the last price available within each hour for our return calculations. As the open price, we use the value weighted price of all transactions from 9:30 until 10:00 AM following Lou et al. (2018).

Total turnover is calculated as the number of shares traded relative to the number of shares outstanding. Abnormal turnover ("ATurnover") is obtained by taking the logarithmic difference of turnover today and turnover on the same day of the previous week.

Beyond total turnover, we also investigate turnover clearly attributable to retail investors. We use the procedure described in Boehmer et al. (2019) based on trade level TAQ data to identify sub-penny price improved off-exchange trades that capture the largest part of retail trading activity.

We also add data on the announcement of tangible news. Specifically, we consider 8-K, 10-

[5] We validate the Algoseek data by comparing the constructed close-to-close returns with the daily holding period returns provided in CRSP. For our sample, the correlation is 0.991

K and 10-Q filing dates obtained from the SEC EDGAR master index file, quarterly earnings announcement days from Compustat and IBES and dividend declaration days from CRSP. Last but not least, we add media coverage dates. A day is a media coverage day if there is at least one article on the company in a national newspaper. As national newspapers, we use the New York Times, USA Today, the Washington Post and the *Wall Street Journal*, available via Nexis.

For the difference-in-difference test based on overnight earnings announcements, we augment the data with earnings announcement information from IBES. We obtain quarterly earnings along with the announcement dates and time-stamps from the IBES unadjusted actuals file. For the difference-in-difference tests, we include announcements made from 8:00 PM on the previous day to 7:30 AM on the current day to exclude confounding effects from previous day after-market returns and to make sure earnings information is available before we begin to measure information acquisition at 7:30 AM.[6] We use ex-ante analysts' estimates to quantify to what extent news is provided in the earnings announcement. Estimates are retrieved from the IBES unadjusted detail file. We only include estimates entered or updated within the 30 days prior to the actual announcement and only keep the most recent estimate per analyst. We generate standardized unexpected earnings ("SUE") as the difference between actual earnings-per-share (EPS) and the median split-adjusted analyst EPS forecast standardized by the share price five days before the announcement. We generate an indicator variable for announcements with large absolute surprises ("Large Surprise"). The indicator takes on a value of one if SUE is below the 30th or above the 70th percentile compared to all announcements in the previous 180 calendar days.

For the test based on *WSJ* rankings, we obtain the historical ranking from 2007 to 2017 from the *WSJ* website. We use the ranking based on the composite universe, that include stocks traded on NYSE, AMEX and Nasdaq.[7]

An overview of the variables used in this study and more information on their construction can be found in Internet Appendix D. Table 1 provides summary statistics on the full sample (Panel A) and the overnight earnings announcements sub-sample (Panel B). In Panel C,

[6] Figure A2 in the Internet Appendix provides an overview of the general timing of earnings announcements over the course of the day.

[7] In 2019, the *WSJ* has considerably modified the market data websites. Since the update, historical gainer/decliner rankings are unfortunately no longer available. The contemporary daily rankings continue to exist in very similar form on the new website.

we compare the groups of companies with large and small earnings surprises used in the difference-in-difference test along several dimensions. Section C.1 of the internet appendix provides a discussion of the summary statistics.

[Insert Table 1 about here.]

4 Results

We now discuss our findings in detail. In Subsection 4.1, we present stylized facts on the relation of daily stock returns and subsequent levels of information acquisition, as well as overall and retail trading. Subsection 4.2 provides causal evidence for the effect of salient returns on information acquisition, based on a difference-in-difference test with overnight earnings announcements. We also discuss the dynamics of intraday trading after overnight earnings surprises that resemble the dynamics of information acquisition. In Subsection 4.3, we use stocks with returns prominently placed returns in the *WSJ* rankings and a control group of stocks with similar returns to identify an economically strong effect of the placement of returns on information acquisition and trading. Subsection 4.4 presents additional results based on subsample splits along the cross-section or time-series.

4.1 Stylized Facts

As a motivation, we report the association between daily stock returns and subsequent information acquisition, as well as trading activity in Figure 1. The figure is based on a portfolio sort into 100 percentile portfolios by daily returns. We then compute portfolio-level statistics as time series averages of daily equally-weighted portfolio-level observations. In Panel A, we report each portfolio's subsequent (i.e. after returns have realized) abnormal level of information acquisition from Edgar. Clearly, information acquisition is positively associated with higher magnitudes of daily stock returns. For example, downloads are 25% to 30% higher than usual for the most extreme percentile portfolios and around 5% higher than usual in percentiles 5 and 96. Note that due to our definition of abnormal information acquisition, firm-specific differences in download levels and intraweekly seasonalities do not affect these estimates. We report analogous results for Wikipedia firm page views in the Internet Appendix, see Figure A1 Panel A. Downloads from EDGAR are likely driven by more sophisticated, mostly institutional investors who follow the market closely. In contrast, Wikipedia firm page views are likely driven by less sophisticated, mostly retail investors, who arguably do not follow daily stock returns as closely. However, the association between daily stock returns and subsequent abnormal Wikipedia firm page views is still strong, with increases in information acquisition of 10% to 15% for the most extreme percentile portfolios. Overall, we find that higher magnitudes of daily stock returns are followed by strong increases in information acquisition.

In Panel B of Figure 1, we plot abnormal turnover on the next trading day after the realized return for each return portfolio. As for information acquisition, we observe a clear positive association between daily stock returns and subsequent activity. Trading volumes are 50% higher than usual for the most extreme percentile portfolios and 8% to 10% higher than usual in percentiles 5 and 96. Results for abnormal retail trading activity are qualitatively and quantitatively similar to overall trading activity (see Internet Appendix, Figure A1 Panel B). Hence, both overall and retail trading activity is strongly related to the magnitude of previously realized returns.

It seems that both information acquisition and trading activity are roughly proportional to absolute average portfolio returns (that are for comparison reported in the Internet Appendix in Figure A1 Panel C). This observations motivates our study. Do only underlying information shocks cause returns, information acquisition and trading? Or can prominently placed and salient stock returns by themselves drive information acquisition and trading?

There is a pervasive convention of reporting returns starting from the previous market close. To explore the role of the salience of a return on information acquisition, we check whether returns measured from different starting points (e.g., intradaily, open-to-close returns) have marginal predictive power for subsequent information acquisition. We find that conventionally reported returns from the previous close are the dominant determinants of subsequent information acquisition and trading. Intraday returns have little marginal predictability although they capture more recent information shocks. This result seems hard to reconcile with the conjecture that the strong association between return magnitudes and subsequent information acquisition is driven only by underlying information shocks but highlights the role of the salience of returns. The results of this analysis are discussed in section C.2 of the internet appendix in detail.

4.2 Difference-in-Difference Around Earnings Announcements

In this section, we present a cleanly identified effect of prominently placed salient stock returns on information acquisition, holding the underlying information shock constant. Our identification strategy relies on overnight earnings surprises.

Large absolute earnings surprises are predictably associated with extreme subsequent returns relative to small surprises. However, these predictable large returns only receive a prominent placement when the market opens at 9:30 AM, whereas the earnings surprises we use for identification are released and available to investors by 7:30 AM at the latest. Such overnight announcements are near-instantaneously made available to investors via newswires.[8] Figure A2 in the Internet Appendix reports the overall distribution of earnings announcements over an EDT-day. Most earnings announcements are made in the two hours after the market closes or in the four hours before it opens. We use announcements made after 8:00 PM on the previous day, to avoid announcements that might affect after-hours trading, which results in returns as well. We additionally restrict announcements to those before 7:30 AM to end up with a clean pre-open period for difference-in-difference tests.[9] Figure A3 in the Internet Appendix reports the cumulative number of large and small earnings announcements between 8:00 PM and 7:30 AM, which we use in our tests.

If the association of information acquisition with extreme returns, which we document in Figure 1 and Table B2, is driven only by underlying information shocks, we would expect a reaction to large earnings surprises relative to small surprises already before the market opens. An additional reaction to large earnings surprises directly after the market opens is plausibly caused by salient returns, not underlying information shocks. As earnings announcements are also easy to anticipate and information about earnings surprises is relatively cheap to acquire, the comparison of pre- and post-open information acquisition for large versus small surprise firms provides a particularly rigorous test for the hypothesis that investors react to returns, not just underlying information shocks.

[8] See Bradley et al. (2014) for an analysis of the timing of newswires relative to earnings announcements and analyst recommendations. Usually newswires seem to be out *before* the I/B/E/S-timestamps, which we use. Such early information releases are unproblematic for our identification strategy.

[9] Pre-market trading between 4:00 AM and 9:30 AM should work against our difference-in-difference test, as it would lead to pre-open availability of returns in addition to the information shock. The difference-in-difference estimate should thus rather be biased down if such pre-market trading leads to prominently placed returns for many stocks.

We report regression results for this simple difference-in-difference test, based on the two hours before and after the market opens at 9:30 AM, in Table 2. Panel A reports results for EDGAR downloads and Panel B for Wikipedia firm page views. For EDGAR downloads, Specification (1) of Panel A shows that there is no significant difference in information acquisition before the market opens between firms with large and small surprises, as measured by the coefficient of the large surprise dummy. However, information acquisition for firms with large surprises jumps by 9.4% relative to small surprise firms when the market opens at 9:30 AM, as indicated by the coefficient of the large surprise dummy interacted with the post-open dummy. This effect is not just economically large, but also statistically significant at the 1% level with a t-statistic of 4.98. Interestingly, the overall level of information acquisition is significantly lower in the two hours after the market opens relative to before, as indicated by the -39.0% coefficient of the post-open dummy. This estimate indicates that overnight earnings announcements receive the most significant abnormal information acquisition right after the announcement and before the open. Importantly, this pre-open information acquisition is not related to the strength of the earnings surprise.[10] Specification (2) of Table 2's Panel A adds stock × announcement day fixed effects to Specification (1), thus directly comparing post-open information acquisition for each announcement with pre-open information acquisition. This regression necessarily leads to the same difference-in-difference estimate, but does not allow for an estimate of pre-open abnormal information acquisition via the the large surprise dummy (which does not vary within stock × announcement day).

[Insert Table 2 about here.]

To directly estimate the economic significance of the relation between stock returns and information acquisition, we next run a two-stage regression. In the first stage, we construct the predicted return of following an earnings announcement as the average return of all earnings announcement return in the same surprise decile over the past 180 day. In the second stage

[10] We report the overall level of abnormal information acquisition for small- and large-surprise earnings announcements (not just the difference between the two) in Figure A4 in Internet Appendix A. This figure clearly shows that market participants are highly active in acquiring information via EDGAR (Panel A) and Wikipedia (Panel B) for firms with earnings announcements long before the market opens. However, despite the availability of the earnings surprise, strong-surprise firms only experience more information acquisition than small-surprise firms *after* the market opens. Figure Figure A4 also supports the parallel trends assumption, in addition to the statistical tests in Panel C of Table 1.

we run our difference-in-difference regression with these predicted returns instead of the large surprise dummy, using a continuous rather than binary treatment. We thus get an estimate for the increase in information acquisition per percentage point of return. Specifications (3) and (4) of Table 2 are otherwise analogous to Specifications (1) and (2). Estimates are in line with a reaction of 3.8% more information acquisition per percentage point absolute stock return. Again, the difference-in-difference estimate is statistically significant at the 1%-level, with a t-statistic of 5.18.[11]

In Panel B of Table 2, we report analogous results for information acquisition via Wikipedia. Again, effect sizes are somewhat lower than for EDGAR downloads, but they are still highly significant. In particular, large surprise firms receive a 3.4% increase in Wikipedia page views after the market opens relative to small surprise firms (Specifications (1) and (2)), which is statistically significant at the 1% level with a t-stat of 2.68). This is consistent with a 1.2% increase in Wikipedia page views per percentage point return magnitude (Specifications (3) and (4)). Hence, both sophisticated investors using EDGAR and less sophisticated investors acquiring information via Wikipedia firm pages significantly react to the avilability of salient stock returns when the market opens, even if the underlying information shock happened hours before.

To show that the difference-in-difference results are not driven by the selection of the two-hour periods before and after the market opens in Table 2 (7:30 to 9:30 AM and 9:30 to 11:30 AM for EDGAR, 7:00-9:00 AM and 10:00-12:00 for Wikipedia), we report hourly effects for the entire 24 hours, starting from the previous day at 8:00 PM in Figure 3. Panel A reports hourly abnormal information acquisition for large early-morning earnings surprise firms via EDGAR and Panel B via Wikipedia. Although the underlying information shock realizes before 7:30 AM in the morning, downloads only significantly and persistently increase after the market opens. Specifically, abnormal information acquisition jumps when the market opens, but remains relatively stable around zero before the market opens and then stable above zero until the market closes. There is a small significantly positive effect between 5:00 and 6:00 AM, in line with few investors reacting to the underlying information shock (both

[11]We can also gauge economic significance via comparing coefficients from Specifications (1) and (2) to predictable large surprise returns: Predictable returns of large surprise firms are higher by about 2.17% than returns for small surprise firms, see Panel C of Table 1. Hence, the 9.4% difference-in-difference from Specification (1) for abnormal information acquisition after the market opens suggests approximately 4.3% more information acquisition via EDGAR per percentage point more extreme return.

in Panels A and B). However, the main driver of information acquisition for large surprise firms is intraday activity, starting when the market opens. Thus, our difference-in-difference test provides strong support for returns, as opposed to underlying information shocks as the main driver of information acquisition for firms with large absolute returns.

[Insert Figure 3 about here.]

We run a few additional robustness tests. First, we check that information acquisition via EDGAR after earnings announcements is not purely driven by recent filings (potentially including filings, that are directly related to the announcement). In Table B4 of the Internet Appendix, we re-run our analysis from Table 2 and exclude downloads for filings made on the announcement day or the day before. Results remain statistically highly significant and qualitatively similar. As expected, the economic significance decreases by around one third. Second, we check that the effect is not driven exclusively by firms appearing in online rankings of daily winners and losers (Kumar et al., 2019). Table B5 in Internet Appendix B reports results from Table 2 after excluding the stocks in the most extreme percentiles of the cross-sectional return distribution, which are likely to appear in such rankings. In particular sophisticated investors using EDGAR react similarly strongly to surprises that do not result in a ranking (Panel A of Table B5). This is consistent with institutional investors observing returns beyond prominent daily winner and loser rankings. Reactions of less sophisticated investors using Wikipedia decrease by around one third when ranked stocks are excluded, becoming statistically insignificant, which indicates that they rely much more on such highly prominent placements of daily return rankings than sophisticated EDGAR users. Third, we analyze algorithmic EDGAR downloads in Table B6 of the Internet Appendix, instead of the human downloads we focus on for our main analysis. Algorithms might be programmed to download filings from EDGAR after large earnings surprises or extreme returns. Our evidence indicates that algorithms do react to large surprises directly with an increase in algorithmic EDGAR downloads of 6.7% (t-statistic 2.36) even before the market opens. The post-open difference-in-difference of 1.2% is economically smaller than the human reaction to the market open and statistically insignificant. Hence, it seems that algorithms are programmed to react to large underlying information shocks rather than absolute stock returns.

Overall, our difference-in-difference results provide clear causal evidence in support of prominently placed returns as a driver of human information acquisition on top of information shocks. The stability of the difference over the entire trading day further supports the interpretation of prominently placed salient returns as a stimulus. Intuitively, the underlying information shock should have a fading effect as the associated overnight newswire becomes more and more distant over the trading day. However, the strong-minus-weak earnings surprise effect remains similar from open-to-close, in line with the persistent stimulus through observed salient returns throughout the trading day.

It is not possible to run a difference-in-difference test for trading activity, because—per construction of the test—there is no trading activity before the market opens. However, we can analyze the intradaily dynamics of abnormal trading activity for firms with large and small earnings surprises. We can then compare intradaily trading to intradaily information acquisition, as reported in Figure 3. We next show that intradaily trading activity after overnight earnings surprises matches the intradaily information acquisition patterns we find above.

As a first result, Table 3 shows that turnover is significantly higher on the trading day directly after the overnight announcement for large- relative to small-surprise stocks. Specification (1) shows that large surprise firms experience around 16% more trading activity than small surprise firms. Taking into account stock and month fixed effects in Specification (2) increases the effect size to a 22% increase in turnover of large- relative to small-surprise firms. These turnover effects between large- and small-surprise firms translate to economic effects of 4% to 9% more turnover per percentage point absolute return. While these regressions suffer from endogeneity issues, results are consistent with the effect sizes we observe in Figure 1.

[Insert Table 3 about here.]

If the higher level of trading activity we observe for large-surprise firms is driven by prominently placed salient returns, we would expect this effect to be present and relatively stable throughout the trading day, as the extreme returns of large-surprise firms remain salient throughout the day. This kind of non-vanishing difference between large- and small-surprise firms would also be consistent with the stable intradaily pattern we observe in Figure 3 for abnormal information acquisition via EDGAR and Wikipedia. On the other hand, if the abnormal trading activity for large-surprise firms is driven by overnight information shocks, we would expect this effect to vanish quickly after the market opens, as the stimulus, i.e., the newswire about the earnings surprise, becomes more and more distant.

Figure 4 is analogous to Figure 3 and shows that intradaily trading effects indeed mirror those of information acquisition. This is unexpected if the underlying overnight information shock (earnings surprise) were to explain trading activity. All information we use to identify large versus small surprise firms is announced before 7:30 AM in the morning, so that most trading as a reaction to this shock should occur directly after the market opens. However, the non-vanishing, stable intradaily pattern is consistent with persistently prominently placed salient close-to-X returns driving trading activity.

[Insert Figure 4 about here.]

In summary, while a clean difference-in-difference test is not possible for trading activity, the intraday dynamics of trading activity seem more in line with prominently placed salient returns as a stimulus, as opposed to underlying overnight information shocks.

4.3 Varying the Prominence of Salient Returns: *WSJ* Rankings

In this section, we exploit the inclusion of certain stocks in the *WSJ* gainer/decliner ranking as a shock to the availability of returns. Since at least 2007, the *WSJ* publishes a ranking of the top and bottom 100 "Gainers" and "Decliners" on its website. The gainers and decliners are sorted according to their daily percentage return. A similar ranking is featured in the daily print edition of the *WSJ*. Figure 5 features a screenshot from the *WSJ* website and print edition with the respective rankings. The rankings provide no fundamental information on the included stocks but return and volume data.

Crucially for our identification strategy, the universe of securities used in the *WSJ* does not coincide with the universe of CRSP stocks used in this study. In addition to common stocks, the *WSJ* rankings include certain ADRs and ETFs.[12] Due to the different universes used to construct the *WSJ* ranking, we are able to isolate variation in the salience of a

[12]The *WSJ* also seems to impose certain filters with regard to the trading volume, price and listing status of securities.

return via the ranking after controlling for the general magnitude of the return itself. This comparison of prominently placed *WSJ* gainers and decliners to less prominently placed stocks with equal returns enables us to identify an economically strong causal effect of salient returns on information acquisition *and* trading activity. However, a drawback of this identification strategy is that we can only estimate the marginal effect of one specific (particularly prominent) placement of daily returns in the *WSJ*. Stocks in our control group are likely to also be prominently placed in other information sources, such as rankings in different newspapers or on different websites. Thus, we will likely underestimate the true placement effect of returns.

Empirically, we estimate regressions with information acquisition after close as the dependent variable on indicator variables for the stock being ranked in the top or bottom ranks 1 to 5, 6 to 10 and 11 to 15 of the *WSJ* gainer/decliner rankings. To isolate the pure placement effect from variation in returns, we carefully control for underlying return movements. Specifically, we include indicator variables for each rank in the top and bottom 200 stock returns with ranks being formed daily in our sample. Moreover, we include 100 indicators for each percentile rank of the daily returns in our sample as well as the for each percentile rank of returns considering the whole sample period. The regressions also include stock and day fixed effects as well as five lags of the dependent variable.

We validate this empirical approach in table B10 of the internet appendix. To do so, we regress contemporaneous on the ranking indicators. In column (1), before controlling for returns using the three sets of indicator variables, we can see that placement in the ranking is strongly correlated with large absolute returns by construction. For example, stocks ranked in the top 5 spots of the gainer ranking on average exhibit a return of about 28.4%. In column (2) however, the ranking indicators do not exhibit (economically or statistically) significant predictability for the magnitude of the returns of ranked stocks after we control for returns using the three groups of indicator variables. Hence, the regression estimates on the ranking variable should now capture a placement or salience effect of the return in the ranking beyond the specific magnitude of the underlying return.

Table 4 presents the results for the information acquisition variables based on Edgar downloads (columns (1) and (2)) and Wikipedia page views (columns (3) and (4)). Focusing on Edgar in column (1), we can see that abnormal information acquisition is strongly associated with the ranking status of the stock. For example, if a stock is ranked as one of the top 5 gainers, abnormal information acquisition from Edgar amounts to 74.2% on average. The magnitude decreases the further down a stock is included in the ranking, to about 41.1% for stock that are featured in the gainer ranking on spots 11 to 15. A symmetric pattern is observable for stocks in the decliner ranking. Note that these estimates subsume any effects from the returns themselves, underlying information shocks as well as any potential placement effects of the returns in the rankings.

In column (2), we include the return controls to isolate the latter effect purely related to the placement of a return in the ranking. Notably, economically and statistically significant coefficients on the *WSJ* ranking indicators remain after controlling for the magnitude of the underlying returns. For the top 5 gainers, the response to the placement in the ranking after controlling for return variation in abnormal information acquisition on Edgar is about 29.6%. Hence, investors causally react to the placement of a return in the ranking with an increase in information acquisition. This placement effect account for about 40% of the entire effect of returns on information acquisition for the top 5 gainers. The placement effect decreases to about 18.5% and 8.3% for ranks 6 to 10 and 11 to 15, consistent with the largest salience effects relating to the very extreme spots in the rankings. Accordingly, the relative contribution of the placement effect to the overall response to returns decreases from 40% to 20% comparing ranks 1 to 5 with ranks 11 to 15. A symmetric and very similar pattern can be observed for the placement in the decliner rankings.

Columns (3) and (4) show the results of a similar analysis, but use information acquisition from Wikipedia as the dependent variable. Even before controlling for return variation, the coefficients on the dummies are lower compared to the results based on the Edgar proxy. This result is consistent with Edgar users, who are relatively more sophisticated, showing a higher sensitivity to returns and information shocks generally. The placement effect measured by the coefficients on the ranking dummies after controlling for returns in column four is significant only for the top 10 gainers and the top 5 decliners. Again, this results is consistent with less sophisticated Wikipedia users only reacting to very extreme placement stimuli, such as the very top and bottom spots in the *WSJ* rankings.[13]

[13] We further explore this idea with respect to the relation of information acquisition and returns generally by estimating spline regressions allowing for a non-linear relation of returns and information acquisition. The results are presented in Table B9 and Figure A5 of the internet appendix. Consistent with the *WSJ* ranking results, we find that more sophisticated Edgar users react to a wide range of returns, even if they are of small magnitude. In contrast, Wikipedia users almost exclusively react to very extreme returns, and almost

For this test, we can check whether similar conclusions hold for two other prominently used information acquisition measures: Bloomberg news search and Google search volume. When relying on these proxies, we obtain qualitatively similar results. The results are reported in Table B11 of the internet appendix.

This setting based on *WSJ* rankings allows us to identify effects of salient returns on trading on top of information acquisition, thereby complementing the difference-in-difference approach used in the previous section for a clean identification of information acquisition effects. We implement an analysis of trading activity by using total abnormal turnover on the next trading day and retail based abnormal turnover on the next trading day as the dependent variable in table 5.

Columns (1) and (2) reveal very similar patterns in the response of next day trading to the inclusion of a return in the *WSJ* rankings compared to information acquisition. Before controlling for returns (column (1)), abnormal turnover strongly relates to the ranking status of a return. As an example, a stock ranked in the top 5 spots of the gainer ranking exhibits abnormal turnover of about 82.6% on the next trading day. The ranking effect decreases the further down a return is included in the ranking list. A symmetric pattern can be observed for the stocks included in the decliner ranking. The coefficients in column (2) show the placement effect of returns in the rankings after including the return controls, thus isolating the placement effect from underlying information shocks. Even if the return variation is accounted for, economically and statistically significant placement effects of returns in the rankings remain. For example, after a stock is included in the top 5 gainer ranking, the stock exhibits abnormal trading volume of 26.1% that is attributed to the pure placement effect of the return. Hence the placement of a return in the ranking causes both more information acquisition after the close and more trading activity on the next trading day. The placement effect decreases the further down a stock is included in the ranking. For ranks 6 to 10, this effect amounts to 13.4% on average, and for ranks 11 to 15 to 6.8% on average. Correspondingly, the relative contribution of the placement effect to the total reaction to returns decreases from 32% for the very top ranking spots to about 19% for the stock ranked in place 11 to 15. Very similar patterns can be observed if the decliner instead of the gainer ranking is considered.

Columns (3) and (4) report the results for abnormal retail turnover on the next trading day.

only if this return is very extreme considering the daily cross-section of returns.

The coefficients reveal a seemingly higher association between being included in the ranking and retail trading as compared to total trading. For the top 5 gainers, retail trading increases by about 124.5% on the day after the stock is included in the ranking. This reaction is about 50% stronger than the reaction in total trading. One interpretation of these results is that retail investors in general react stronger with trades as a response to past returns compared to other investors. The coefficients in columns (4) show that the relative importance of the pure placement effect in the rankings after controlling for the magnitude of the returns is of similar magnitude as for the results concerning general trading. Focusing on gainers, the relative contribution of the placement effect to the total return effect ranges between 29% for the top five ranks and 17% for ranks 11 to 15. Similar patterns are present for the decliner ranking.

Overall, the results in this section show that not only information acquisition, but also trading activity is increased by how prominently a return is placed for investors. This placement effect operates beyond a basic relation of information acquisition, trading and returns due to underlying information shocks.

4.4 Robustness: Variation Across Firms and Days

In this section, we analyze variation in the results of the difference-in-difference test and ranking results across firms and over time. For the difference-in-difference results from Table 2, we construct subsets of firms to analyze cross-sectional stability by splitting the sample according to a given characteristic lagged by one month. In Table B7 of the Internet Appendix, we report results for these median-splits.

Specifications (1) and (2) of Panel A split firms based on market capitalization at the cross-sectional median. Note that the "small" firms in our sample are still relatively large, as we require a price above 5$ and a market capitalization above the 1$^{\text{st}}$ NYSE-percentile in the sample construction. The increase in information acquisition for large-surprise firms after the open remains highly significant at 7.4% for the smaller half of firms, while we estimate only a 4.3% increase of information acquisition for larger firms with marginal significance. The smaller estimate for large firms is consistent with higher monitoring activity of market participants for the pre-scheduled earnings announcements of large companies in general and not just after a return is observed.

To analyze the effect of ex-ante visibility in more detail, we next split by analyst coverage in the previous month. In order to avoid a confounding correlation with firm size, we split by *residual* analyst coverage, similar to Nagel (2005)'s orthogonalization for institutional ownership. Therefore, we regress analyst coverage on firm size in last month's cross-section and use the residual from that regression to split the sample. Specifications (3) and (4) of Table B7 Panel A show the results of this split. In line with the above discussion on ex-ante visibility and firm size, low-coverage firms exhibit a stronger 11.2% effect of large- vs. small-surprise firms at the market open compared to the 7.6% for high-coverage firms

Specifications (5) and (6) split by the residual level of EDGAR downloads for a firm in the previous month. As for residual analyst coverage, the less visible firms experience a much stronger effect due to returns, while ex-ante visible firms experience a weaker (but still statistically significant) effect.

In a last split presented in columns (7) and (8), we analyze firms with varying levels of residual institutional ownership. This split is motivated by the typical investors using EDGAR, who are likely to be sophisticated and often institutional investors. For such investors, who follow firms more closely throughout earnings announcement days, one could expect overall effects to be stronger. Indeed, we find that high institutional ownership firms with large-surprises experience a 10.9% increase in information acquisition after the market opens, whereas low institutional ownership firms only experience a 7.7% effect (both significant at the 5%-level). Hence, the effect of prominently placed salient returns on information acquisition from Edgar is particularly relevant for firms with sophisticated, institutional investors.

Panel B of Table B7 reports the results of similar analyses for the Wikipedia based information acquisition variable. Generally, difference between the sub-sample are less pronounced for the Wikipedia based results as compared to the Edgar based results.

In Table B8 of the Internet Appendix, we report the variation of our difference-in-difference effect over time. In Specifications (1) and (2) we split by the median of the VIX into low- and high-uncertainty times. The effect of large-surprise returns at the open is particularly strong for high-uncertainty times, which might be driven by more extreme stock returns for large surprises in these times or a greater focus on the processing of market instead of stock specific news. However, effects are also statistically significant in low-uncertainty times. In Specifications (3) and (4) we split by the sign of the previous month's market return. The effect of large-surprise returns at the open is stable at about 9% across the two subsamples.

Overall, we find that the information acquisition effect of returns on top of the earnings surprise information shock is particularly strong for firms that are less visible and less closely followed ex-ante. It is slightly stronger for firms with high institutional ownership and in times of high uncertainty. In any case we observe a positive effect of returns on information acquisition, making the difference-in-difference results stable.

We also consider subsample splits with respect to firm size for the analyses of information acquisition and turnover based on *WSJ* rankings. Again, we split the sample based on the median market cap in the sample one month prior to the observation. This split is also partially informative with respect to endogeneity concerns regarding the selection of stocks eligible to be included in the *WSJ* rankings. Potentially, the set of stocks eligible for inclusion has higher visibility compared to non-eligible stocks. It is important to remember in this context, that all our dependent variables are based on log-changes, such that general differences in visibility cannot explain the effects on changes that we observe in the regressions. Beyond this point, it is not clear that the response in the change to a ranking event should be larger for more visible firms. One could even argue that the marginal effect on the changes for more visible firms could be smaller due to the generally higher baseline level of information acquisition and trading for firms with high visibility.

Table B12 of the internet appendix presents the results of the subsample analysis of information acquisition. Considering AEdgar in columns (1) and (2), we find an economically similar reaction to being included in the ranking for both large and small firms. The same holds true for AWiki in columns (3) and (4), although the split results in a lack of power and mostly insignificant coefficients for the Wikipedia based results.

Table B13 of the internet appendix presents the results of the subsample analysis of trading activity. Both concerning total turnover (columns (1) and (2)) and retail turnover (columns (3) and (4)), we find an economically meaningful turnover reaction to a return being included in the *WSJ* ranking. For the gainer ranking, the effects are slightly larger for larger companies, while for the decliner ranking we observe stronger effects for small firms.

The findings for both information acquisition and trading suggest that potential differences in visibility (as proxied by firm size) between ranking eligible and non-eligible stocks are unlikely to be the driver of the main findings. Overall, we find no strong systematic relation between firm size and the marginal effect of the ranking. Some estimates even point at a negative relation between firm size and the marginal effect of the ranking, suggesting that we might underestimate the true ranking effect due to the eligibility criteria set by the *WSJ*.

5 Implications for Mispricing

In this section, we explore the implications of return-induced trading for mispricing. Above, we show that both information acquisition and trading activity are driven by salient returns beyond the effects of underlying information shocks. It seems plausible that a substantial part of return-induced activity is due to attention-constrained noise traders. This does not imply price pressure and mispricing. Noise trading for a specific stock might be truly idiosyncratic and balanced, with similar levels of uninformed buying and selling. However, one would expect that behavioral biases that tend to drive noise traders to buy or sell certain stocks are systematic.[14] When the level of noise trading increases, associated price pressure is likely to increase as well. Increased price pressure for stocks, which are prone to mispricing ex-ante, should thus amplify anomalous returns.[15]

To explore this idea we test if stocks, which are prone to mispricing ex-ante exhibit particularly strong anomalous returns after salient returns. We rely on the mispricing score of Stambaugh et al. (2012), a compound score based on eleven well-established anomalies ("SYY Score"). A higher SYY score indicates that a stock is usually overpriced and will on average underperform going forward, whereas a low SYY score indicates that a stock is rather underpriced and will on average outperform going forward. We derive a measure of the salience effects of returns by relying on the stylized facts documented in Figure 1. Each day, we sort stocks into 100 percentiles based on their returns. For each return percentile, we assign the abnormal level of information acquisition from Figure 1, as a measure for return-

[14] One example for such systematic behavioral factors is media-induced positive or negative sentiment (see, e.g., Tetlock, 2007; Garcia, 2012).

[15] We are deliberately not restricting ourselves to individual investors. Institutional investors may also react to salient returns, even if it is detrimental to their performance (see, e.g., Akepanidtaworn et al., 2019). For individual investors, one might expect particularly strong effects on uninformed *buying* when returns are salient. The theoretical motivation behind attention-induced buying pressure and overpricing goes back to Lintner (1969), Miller (1977), and Mayshar (1983), who analyze the effects of heterogeneous beliefs combined with short-sale constraints on asset prices. Individual investors are particularly short-sale constrained. Barber and Odean (2008) add that individual investors' search problem is greater for buying than selling decisions, so that attention leads to buying pressure even for investors who already own a stock. See Kumar et al. (2019) for empirical evidence of return-induced overpricing.

induced activity. To arrive at a monthly measure for return-induced activity ("PoP"), we average across all days in a month. If salient returns amplify mispricing via increased noise trading activity, we expect more overpricing in high SYY score stocks and thus a stronger underperformance after high levels of PoP. By the same argument, we expect to find more underpricing in low SYY score stocks and thus a stronger outperformance after high levels of PoP. In short, we expect a negative interaction effect between PoP and the SYY score.

Table 6 presents the results of dependent double-sorts of stocks along the SYY score and PoP. Each month, all stocks are first sorted into quintile portfolios based on the SYY score. Within each SYY portfolios, stocks are sorted by PoP. Thus, this double sort allows us to compare the effect of the SYY mispricing score on abnormal returns at different levels of return-induced trading, PoP. Panel A presents the returns of value weighted portfolios. Consistent with the findings of Stambaugh et al. (2012), we can see that high SYY score (overpriced) stocks underperform low SYY score (underpriced) stocks. This negative difference for the returns of overpriced minus underpriced portfolio holds for all quintiles of the PoP measure. However, the difference is not constant across the PoP quintiles. As expected, the return spread increases considerably with PoP, from -34 bps per month for the low PoP quintile to -177 bps per month for the high PoP quintile. The difference of the mispricing spread of -143 bps is shown in the last column and is highly statistically significant (t-statistic of -5.57). Hence, a simple portfolio sort support the notion that return-induced trading amplifies mispricing in the stock market.

In the remaining rows of Panel A, we show that the significant interaction effect between return-induced trading and mispricing is robust when controlling for factor returns of the Carhart (1997) and Fama and French (2015) models. Very similar results are obtained for equal-weighted instead of value-weighted returns in Panel B of Table 6. In fact, the spreads are even slightly larger for the value-weighted returns, indicating that illiquidity and short sale constraints, which correlate strongly with firm size are not a major driver of our effect.

We report alphas from a large number of other factor models in Table B15 of the Internet Appendix. Results are robust to controlling for the CAPM 1-factor (1F) and Fama and French (1993) 3-factor (3F) model, the Hou et al. (2015) Q-model, the short- and long-term reversal factors from Kenneth French's data library, the Frazzini and Pedersen (2014) betting-against-beta factor, factors based on idiosyncratic volatility (IVol, Ang et al., 2006) and maximum daily returns (MAX, Bali et al., 2011), the Pástor and Stambaugh (2003)

(PS) and the Sadka (2006) (fixed-transitory and variable-permanent) systematic liquidity factors, the Stambaugh and Yuan (2017) mispricing factors, the Hirshleifer and Jiang (2010) undervalued-minus-overvalued factor, the Asness et al. (2019) quality-minus-junk factor, and the Novy-Marx (2013) profitability factor. Irrespective of the specific factor model, we find uniformly strong evidence of a large interaction between return-induced activity and mispricing that never falls below 0.66% per month. t-statistics range from 3.47 up to 8.00.

We visualize the large moderating effect of salient returns on the abnormal returns of low- over high-SYY score stocks in Figure 6. Both lines report the cumulative returns of a value-weighted long-short strategy based on SYY score quintiles. Each month the portfolio includes the bottom 20% of stocks with the lowest SYY score in the long leg and the top 20% of stocks with the highest SYY score in the short leg. However, we split each leg further in two groups based on PoP. The dashed blue line depicts the returns of the long-short SYY strategy that only includes the 40% of stocks with the highest PoP ("Mispricing Portfolio Hi PoP"). The solid black line depicts the returns of the long-short SYY strategy that only includes the 40% of stocks with the lowest PoP ("Mispricing Portfolio Lo PoP"). In line with the portfolio double sort, we can see that mispricing effects are amplified in the high- relative to the low-PoP portfolio. This is consistent with return-induced noise trading as an amplifier of mispricing. The high PoP strategy accumulates returns of more than 800% over 50 years (16% per year), while the cumulative returns of the low PoP portfolio "only" sum to less than 300% (6% per year).

Table 7 reports the results of Fama and MacBeth (1973) regressions of individual stock returns on the SYY score, the PoP measure, as well as the interaction term of the two. SYY and PoP are standardized to exhibit zero mean and unit standard deviation. In all specifications, we confirm the findings of Stambaugh et al. (2012) and find that a higher mispricing score is associated with lower returns. The base effect of Pop is also negative and particularly significant when we control for short-term reversal. A negative effect for stocks with high levels of return salience is in line with attention-induced buying for attention grabbing stocks (see Kumar et al., 2019). More importantly, we again find a significant negative interaction between PoP and the SYY score, as for portfolio sorts (t-statistic of -7.60 in Specification (2)). This indicates that stocks, which are more prone to mispricing exhibit particularly abnormal returns when return-induced trading is high. As an illustration of the economic magnitude of this effect based on Specification (2), consider the implied change in monthly returns for a one standard deviation change in the SYY score at different levels of PoP. For low levels of PoP (one standard deviation below the mean), a one-standard deviation increase in the SYY score implies a relatively moderate subsequent underperformance of $-30 - (-18) = -8$ bps per month. In contrast, for high levels of PoP (one standard deviation above the mean), the one-standard-deviation effect of SYY is amplified six-fold to $-30 + (-18) = -48$ bps per month.

The negative interaction term between the SYY score and PoP is hardly affected by the inclusion of additional known predictors of returns. In Specification (3), we control for a stock's β, the logarithm of its size and its book to market ratio, last year's return (momentum), last month's return (short-term reversal), and the previous two years' returns (long-term reversal).[16] The coefficients of all control variables are as expected. In particular, small firms, value stocks, last year's winners, last month's losers, and long-term losers exhibit higher returns. Motivated by Fama and French (2015) we add operating profitability and asset growth as control variables in Specification (4). As expected profitable firms exhibit higher future returns, while firms with strong asset growth exhibit lower returns. Our main results remain largely unaffected. Gervais et al. (2001) find that trading activity is related to future stock returns. Our analysis of mispricing is motivated by return-induced trading activity, so that controlling for trading activity might influence our results. Thus, in Specification (5), we add last month's level and change in turnover. As in Gervais et al. (2001), stocks with increasing trading activity exhibit higher future returns (the 'high-volume return premium'), while a high turnover level is related to lower future returns, which is consistent with high turnover stocks being more liquid and delivering lower returns. However, both effects are statistically insignificant within our highly liquid stock universe. More importantly, controlling for these effects does not affect our main findings either. Finally, to make sure results are not driven by the salience of industry-returns or small- vs large-firm returns in particular months we add Fama/French-48 industry dummies and NYSE-size-decile dummies in Specification (6). Additionally, we also include exchange dummies. This does not change our estimates for the interaction effect between return-induced trading PoP and the SYY mispricing score either. Hence, our results are robust to controlling for common firm-specific characteristics.

Panel B of Table 7 provides further robustness checks. Our asset pricing results concerning the interaction of PoP and the SYY score hold after the exclusion of daily winners and

[16] All variables are defined in detail in Appendix D.

losers from the sample as shown in Specification (1) (Kumar et al., 2019). Specifications (2) to (4) study the effect of return salience on mispricing via PoP for different samples based on residual institutional ownership (Nagel, 2005). Specification (2) shows that the baseline results holds for the sample period for which institutional (13f) ownership data is available, 1980-2015. Next, we split the sample into to subsample based on the cross-sectional median of institutional ownership. As expected, we find stronger mispricing (the coefficient of SYY) in the low institutional ownership sample, combined with a stronger interaction term between SYY and PoP, see Specification (3). Mispricing and its interaction with return-induced trading are much weaker in the high institutional ownership subsample, see Specification (4). Importantly, the interaction effect between PoP and SYY remains statistically significant (t-statistic -2.69). A one-standard-deviation increase in the SSY score is associated with a decrease in returns of -10 bps for low-PoP stocks and -30 bps for high-PoP stocks, indicating a robust strong interaction between return-induced trading and mispricing, even at high levels of institutional ownership within our stock universe of large, liquid firms. The attenuating effect of high institutional ownership can be driven by both, (i) lower short sale constraints for stocks at institutional ownership with many lendable shares, as well as (ii) increased trading of institutional investors, who are less prone to systematic behavioral biases and thus less likely to cause buy-sell imbalances. Last, Specifications (5) and (6) show that the interaction between return-induced trading and mispricing is significant within both, the first (1965-1990) and the second (1991-2015) half of our sample period.

In Table B16 of the Internet Appendix, we provide additional robustness checks. Our results hold when skipping one month between the measurement and prediction period (avoiding microstructure issues), for alternative measures of return-induced activity, when employing different data requirements (price, size, or exchange filters) or return adjustments (DGTW- or industry-adjustments). We also provide sub-sample splits based on size and liquidity measures. As for institutional ownership and mispricing effects are attenuated for large firms, whereas illiquidity proxies (Amihud, 2002; Corwin and Schultz, 2012) do not qualitatively change results. Last, as documented in Lou et al. (2018), abnormal returns flip signs between overnight and intraday returns, in line with clientele-effects (e.g., individual investors are particularly likely to trade overnight, whereas institutional investors are particularly likely to trade intraday).

As discussed in Section 2, our return salience proxy PoP is highly correlated (at around

0.8) with idiosyncratic volatility, Stambaugh et al. (2012)'s proxy for arbitrage risk. It is hard to distinguish this proxy for arbitrage risk and the salient returns that we link to noise-trading. Empirically, it is likely that the remaining differences between the proxies PoP (based on cross-sectional return ranks) and idiosyncratic volatility (based on time-series volatility of returns) still measure both arbitrage risk *and* return saliency and thus increased noise trading. Stambaugh et al. (2012) similarly find that historically more salient returns amplify abnormal returns of stocks that are prone to mispricing. However, they interpret this as an effect of arbitrage risk, whereas our analysis is motivated by the demand of noise traders. In their model, Stambaugh et al. (2012) explicitly assume noise trading to be exogenous. In Section 4, we provide strong evidence rejecting this assumption: Salient, extreme returns drive trading activity, in line with significant return-induced noise trading. However, we do not claim that arbitrage risk does not (also) drive mispricing. We rather provide a complementary view of mispricing, by analyzing variation in the ultimate cause of price distortions, systematic noise trading, instead of the subsequent limits to arbitrage (such as illiquidity, short-sale constraints, or Stambaugh et al. (2012)'s arbitrage risk).

6 Conclusion and Implications

We provide evidence that information acquisition and trading activity in response to information shocks are largely driven by prominently placed salient returns. First, large overnight earnings surprises strongly predict higher levels of downloads from the SEC's EDGAR platform and Wikipedia firm page views only after salient returns become available when the market opens. The intradaily dynamics of trading activity for stocks with large earnings surprises are consistent with the patterns of information acquisition. Second, we find that stocks with extreme returns that receive prominent placements in the *WSJ* daily gainer and decliner ranking experience significantly higher information acquisition and trading activity than stocks with equally large return shocks but a less prominent placement. This powerful, moderating role of prominently placed salient returns between information shocks and financial markets might help to understand the high levels of costly information acquisition and trading we observe in stock markets (Milgrom and Stokey, 1982).

One interpretation of our finding is that market prices provide valuable signals to investors about what information they need to acquire to optimize their portfolio (Bossaerts et al., 2019). Another, complementary interpretation is that attention-constrained investors are attracted to such highly visible stocks, which leads to trading between uninformed noise traders and informed liquidity providers. In line with this interpretation, we find that stocks that are prone to mispricing exhibit particularly anomalous returns after return-induced trading activity, consistent with Grossman and Stiglitz (1980) and Kyle (1985).

The finding that prominently placed salient stock returns themselves affect information acquisition and trading is particularly important for empirical work in financial economics. It is common practice to use historical stock returns to measure expected risk (e.g., betas or volatility, Fama and MacBeth, 1973; Ang et al., 2006) or investors' preferences (e.g., cumulative prospect theory values, Barberis et al., 2016). If some of these historical stock returns are more prominently placed and salient, thus causing particularly high levels of information acquisition and trading, this might confound or explain results in the existing empirical literature. As an illustration, Kumar et al. (2019) build on our insight that prominently placed daily winner and loser rankings drive investor attention and show that the attention-induced overpricing of these stocks provides a simple explanation for the high prices and low returns of high idiosyncratic volatility stocks (Ang et al., 2006).

The powerful impact of salient price changes might also affect strategic information releases. If managers anticipate the attention received by stocks with extreme returns, they might strategically release information to stay under the radar (e.g., bundling positive and negative information shocks on one day, so that the effects cancel out) or to stand out relative to other firms (e.g., separating information into a negative release on one day and a positive release later, so that daily returns are extreme for both announcements). Managers might even overstate and later retract information to stand out and attract trading activity. There is some evidence of such strategic information releases by managers. Segal and Segal (2016) use Form 8-K filings to analyze strategic reporting of news. They find that managers time news and bundle positive and negative news to avoid investor attention.

The return-induced effects we analyze are arguably particularly important in public stock markets, as opposed to less liquid private markets. For privately held companies, the effects of information shocks on information acquisition and trading are likely to be much smaller, as there are no prominently placed high-frequency returns. Thus, any positive and negative side effects of the regular information acquisition and trading shocks experienced by publicly traded companies will likely affect the decision of companies to go public or private.

More generally, the availability of prices is a feature of market-based solutions. Meloso et al. (2009) compare a market-based solution for promoting knowledge advances and the conventional solution, patents (temporary monopolies). In a series of experiments, they find that markets, via providing priced shares in potential discoveries, better promote the generation and spread of knowledge than patents. While we document that return-induced activity is associated with amplified anomalous returns for stocks prone to mispricing, this does not preclude that overall, the positive effects of widely available market prices, through providing incentives to generate and share knowledge, dominate.

References

Akepanidtaworn, Klakow, Rick Di Mascio, Alex Imas, and Lawrence Schmidt, 2019, Selling Fast and Buying Slow: Heuristics and Trading Performance of Institutional Investors.

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, *The Journal of Finance* 61, 259–299.

Asness, Clifford S., Andrea Frazzini, and Lasse Heje Pedersen, 2019, Quality minus junk, *Review of Accounting Studies* 24, 34–112.

Bali, Turan G., Nusret Cakici, and Robert F. Whitelaw, 2011, Maxing out: Stocks as lotteries and the cross-section of expected returns, *Journal of Financial Economics* 99, 427–446.

Barber, Brad M., Xing Huang, and Terrance Odean, 2016, Which Factors Matter to Investors? Evidence from Mutual Fund Flows, *Review of Financial Studies* 29, 2600–2642.

Barber, Brad M., and Terrance Odean, 2000, Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors, *The 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, *The Review of Financial Studies* 21, 785–818.

Barberis, Nicholas, Abhiroop Mukherjee, and Baolian Wang, 2016, Prospect Theory and Stock Returns: An Empirical Test, *Review of Financial Studies* 29, 3068–3107.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song, 2019, What Do Mutual Fund Investors Really Care About?

Ben-Rephael, Azi, Zhi Da, and Ryan D. Israelsen, 2017, It depends on where you search: Institutional investor attention and underreaction to news, *Review of Financial Studies* 30, 3009–3047.

Berk, Jonathan B., and Jules H. van Binsbergen, 2015, Assessing Asset Pricing Models Using Revealed Preference, *Journal of Financial Economics* 119, 1–23.

Boehmer, Ekkehart, Charles M. Jones, Xiaoyan Zhang, and Xinran Zhang, 2019, Tracking Retail Investor Activity.

Bossaerts, Peter, Elizabeth Bowman, Felix Fattinger, Shijie Huang, Carsten Murawski, Shireen Tang, and Nitin Yadav, 2019, Asset Pricing under Computational Complexity.

Bradley, Daniel, Jonathan Clarke, Suzanne Lee, and Chayawat Ornthanalai, 2014, Are Analysts' Recommendations Informative? Intraday Evidence on the Impact of Time Stamp Delays, *Journal of Finance* 69, 645–673.

Carhart, Mark M., 1997, On Persistence in Mutual Fund Performance, *The Journal of Finance* 52, 57–82.

Chen, Huaizhi, Lauren Cohen, Dong Lou, and Christopher J Malloy, 2018, IQ from IP : Simplifying Search in Portfolio Choice, *Working Paper* .

Corwin, Shane A., and Paul Schultz, 2012, A Simple Way to Estimate Bid-Ask Spreads from Daily High and Low Prices, *The Journal of Finance* 67, 719–759.

Crego, Julio, 2018, Why Does Public News Augment Information Asymmetries?

Da, Zhi, Joseph E. Engelberg, and Pengjie Gao, 2011, In Search of Attention, *The Journal of Finance* 66, 1461–1499.

Daniel, Kent, and David Hirshleifer, 2015, Overconfident Investors, Predictable Returns, and Excessive Trading, *Journal of Economic Perspectives* 29, 61–88.

Drake, Michael S., Darren T. Roulstone, and Jacob R. Thornock, 2015, The Determinants and Consequences of Information Acquisition via EDGAR, *Contemporary Accounting Research* 32, 1128–1161.

Engelberg, Joseph E., and Christopher A. Parsons, 2011, The Causal Impact of Media in Financial Markets, *The Journal of Finance* 66, 67–97.

Evans, Richard B., and Yang Sun, 2018, Models or Stars: The Role of Asset Pricing Models and Heuristics in Investor Risk Adjustment.

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

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

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

Fedyk, Anastassia, 2018, Front Page News: The Effect of News Positioning on Financial Markets.

Frazzini, Andrea, and Lasse Heje Pedersen, 2014, Betting Against Beta, *Journal of Financial Economics* 111, 1–25.

French, Kenneth R., 2008, Presidential Address: The Cost of Active Investing, *The Journal of Finance* 63, 1537–1573.

Garcia, Diego, 2012, Sentiment during recessions, *The Journal of Finance* .

Gervais, Simon, Ron Kaniel, and Dan H. Mingelgrin, 2001, The High-Volume Return Premium, *The Journal of Finance* 56, 877–919.

Gibbons, Brian, Peter Iliev, and Jonathan Kalodimos, 2018, Analyst Information Acquisition via EDGAR, *Working paper* .

Grossman, Sanford J., and Joseph E. Stiglitz, 1980, On the Impossibility of Informationally Efficient Markets, *American Economic Review* 70, 393–408.

Harrison, Michael J., and David M. Kreps, 1978, Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations, *The Quarterly Journal of Economics* 92, 323–336.

Hartzmark, Samuel M., 2014, The Worst, the Best, Ignoring All the Rest: The Rank Effect and Trading Behavior, *Review of Financial Studies* 28, 1024–1059.

Harvey, Campbell R., 2017, Presidential Address: The Scientific Outlook in Financial Economics, *The Journal of Finance* 72, 1399–1440.

Hayek, Friedrich, 1945, The Use of Knowledge in Society, in Randall S. Kroszner, and Louis Putterman, eds., *The American Economic Review*, volume 35, 519–530 (Cambridge University Press, Cambridge).

Hendershott, Terrence, Albert J. Menkveld, Remy Praz, and Mark S. Seasholes, 2018, Asset Price Dynamics with Limited Attention.

Hirshleifer, David, and Danling Jiang, 2010, A Financing-Based Misvaluation Factor and the Cross-Section of Expected Returns, *Review of Financial Studies* 23, 3401–3436.

Hou, Kewei, Chen Yue, and Lu Zhang, 2015, Digesting Anomalies: An Investment Approach, *Review of Financial Studies* 28, 650–705.

Kaniel, Ron, and Robert Parham, 2015, WSJ Category Kings - the impact of media attention on consumer and mutual fund investment decisions.

Kim, Oliver, and Robert E. Verrecchia, 1994, Market liquidity and volume around earnings announcements, *Journal of Accounting and Economics* 17, 41–67.

Kumar, Alok, Stefan Ruenzi, and Michael Ungeheuer, 2019, Daily Winners and Losers.

Kyle, Albert S., 1985, Continuous Auctions and Insider Trading, *Econometrica* 53, 1315–1335.

Lee, Charles M.C., Paul Ma, and Charles C.Y. Wang, 2015, Search-based peer firms: Aggregating investor perceptions through internet co-searches, *Journal of Financial Economics* 116, 410–431.

Li, Frank Weikai, and Chengzhu Sun, 2019, Information Acquisition and Expected Returns: Evidence from EDGAR Search Traffic, *Working Paper* .

Lintner, John, 1969, The Aggregation of Investor's Diverse Judgments and Preferences in Purely Competitive Security Markets, *Journal of Financial and Quantitative Analysis* 4, 347–400.

Lou, Dong, Christopher Polk, and Spyros Skouras, 2018, A Tug of War : Overnight Versus Intraday Expected Returns, *Journal of Financial Economics* Forthcomin.

Mayshar, Joram, 1983, On Divergence of Opinion and Imperfections in Capital Markets, *American Economic Review* 73, 114–128.

Meloso, Debrah, Jernej Copic, and Peter Bossaerts, 2009, Promoting Intellectual Discovery: Patents Versus Markets, *Science* 323, 1335–1339.

Milgrom, Paul, and Nancy Stokey, 1982, Information, Trade and Common Knowledge, *Journal of Economic Theory* 26, 17–27.

Miller, Edward M., 1977, Risk, Uncertainty, and Divergence of Opinion, *Journal of Finance* 32, 1151–1168.

Nagel, Stefan, 2005, Short sales, institutional investors and the cross-section of stock returns, *Journal of Financial Economics* 78, 277–309.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1–28.

Odean, Terrance, 1999, Do Investors Trade Too Much?, *American Economic Review* 89, 1279–1298.

Pástor, Luboš, and Robert F. Stambaugh, 2003, Risk and Expected Stock Returns, *Journal of Political Economy* 111, 642–685.

Ryans, James P., 2017, Using the EDGAR Log File Data Set, *Working Paper* .

Sadka, Ronnie, 2006, Momentum and Post-Earnings-Announcement Drift Anomalies: The Role of Liquidity Risk, *Journal of Financial Economics* 80, 309–349.

Schaub, Nic, 2018, The Role of Data Providers as Information Intermediaries, *Journal of Financial and Quantitative Analysis* 53, 1805–1838.

Scheinkman, José A., and Wei Xiong, 2003, Overconfidence and Speculative Bubbles, *Journal of Political Economy* 111, 1183–1220.

Segal, Benjamin, and Dan Segal, 2016, Are managers strategic in reporting non-earnings news? Evidence on timing and news bundling, *Review of Accounting Studies* 21, 1–42.

Stambaugh, Robert F., Jianfeng Yu, and Yu Yuan, 2012, The short of it: Investor sentiment and anomalies, *Journal of Financial Economics* 104, 288–302.

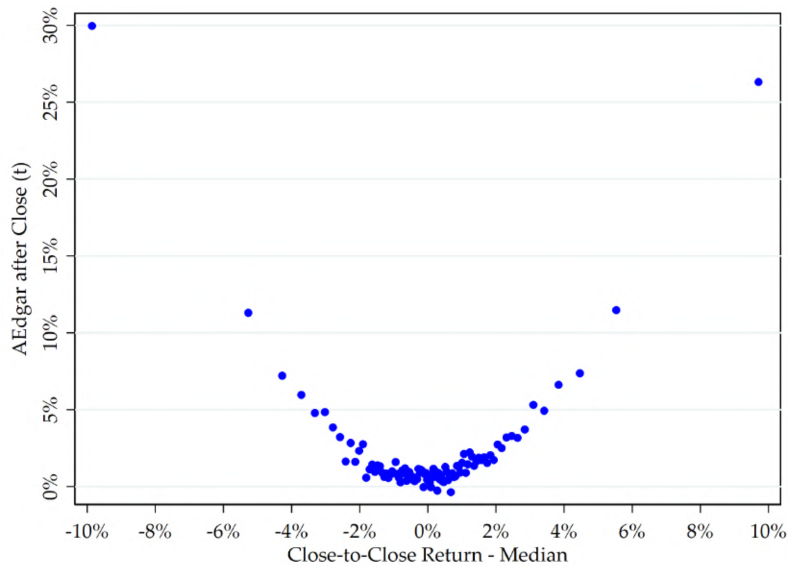
Stambaugh, Robert F., and Yu Yuan, 2017, Mispricing Factors, *Review of Financial Studies* 30, 1270–1315.

Tetlock, Paul C., 2007, Giving Content to Investor Sentiment: The Role of Media in the Stock Market, *Journal of Finance* 62, 1139–1168.

Figure 1: Information Acquisition, Trading and Close to Close Returns

This figure shows average abnormal EDGAR downloads and abnormal turnover following returns in Panels A and B. Each trading day, stocks are sorted into 100 portfolios based on their median adjusted return. Then, within each portfolio, variables are averaged within the cross-section of included stocks and in a second step averaged within the time-series. EDGAR downloads are measured after the market closes and before the next open. Turnover is measured on the next trading day. Hence, Panels A and B do not represent contemporaneous associations. Variables are defined in Internet Appendix D.

Panel A: Information Acquisition after the Market Closes



Panel B: Total Turnover on the Next Trading Day

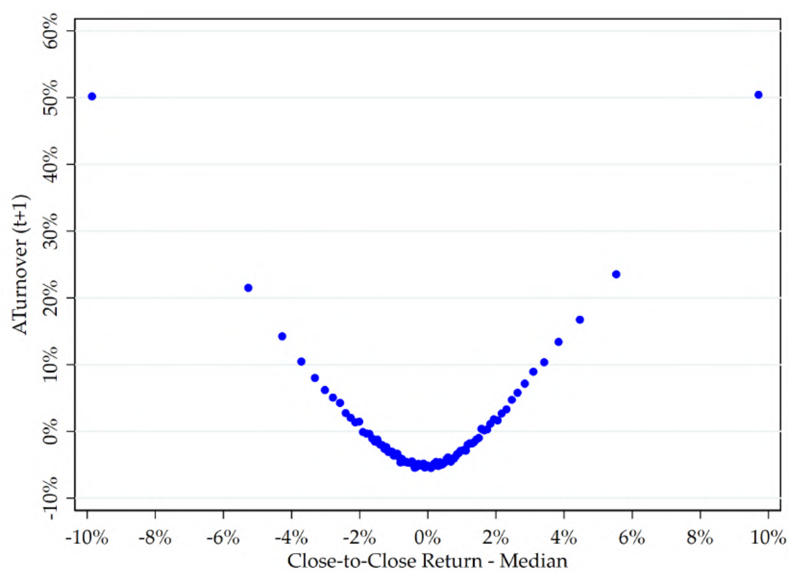
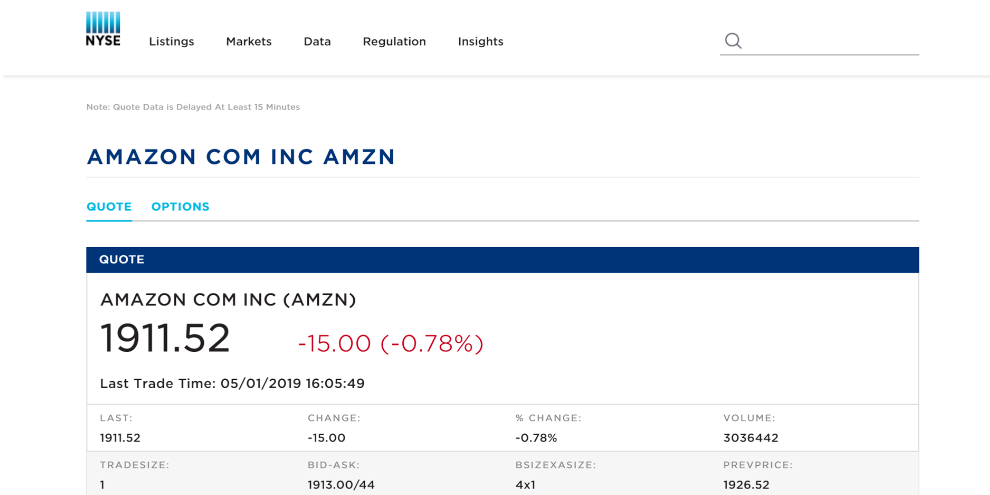


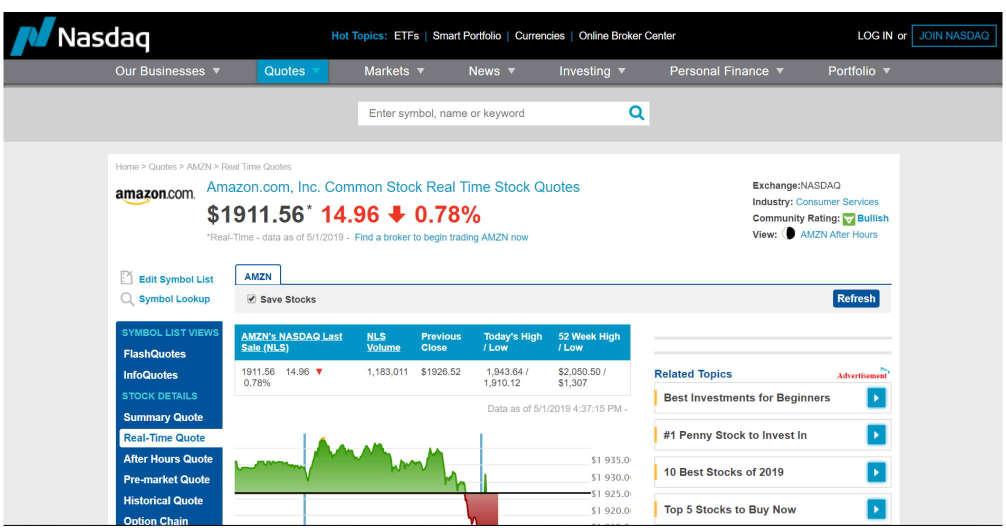
Figure 2: Examples of Salient Daily Returns

This figure shows examples of daily returns that are prominently placed by stock exchanges (NYSE and NASDAQ), financial news providers (Bloomberg, Reuters, Google Finance, Yahoo! Finance), newspapers on their websites (Wall Street Journal and Financial Times) and in their online articles (Wall Street Journal article). The examples are all for Amazon's daily close-to-close return on May 1st 2019.

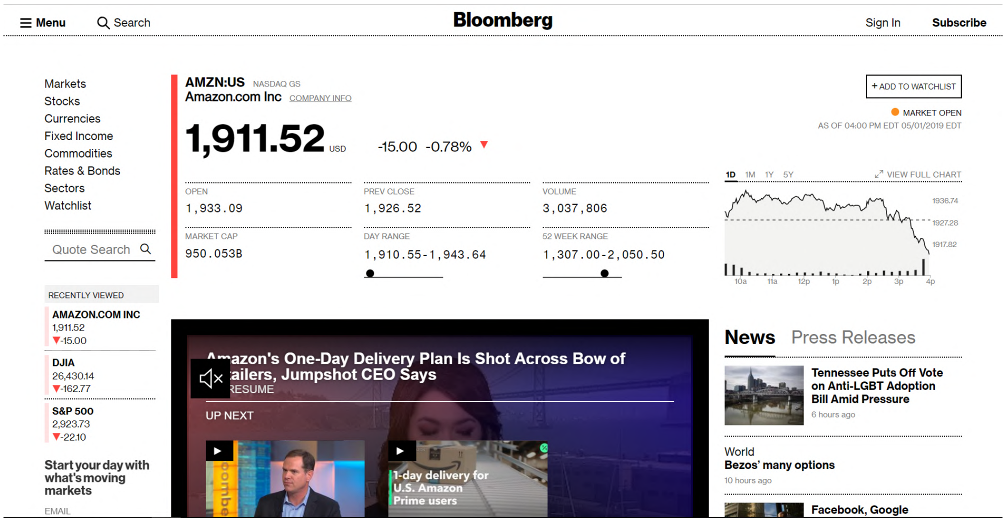
Panel A: New York Stock Exchange



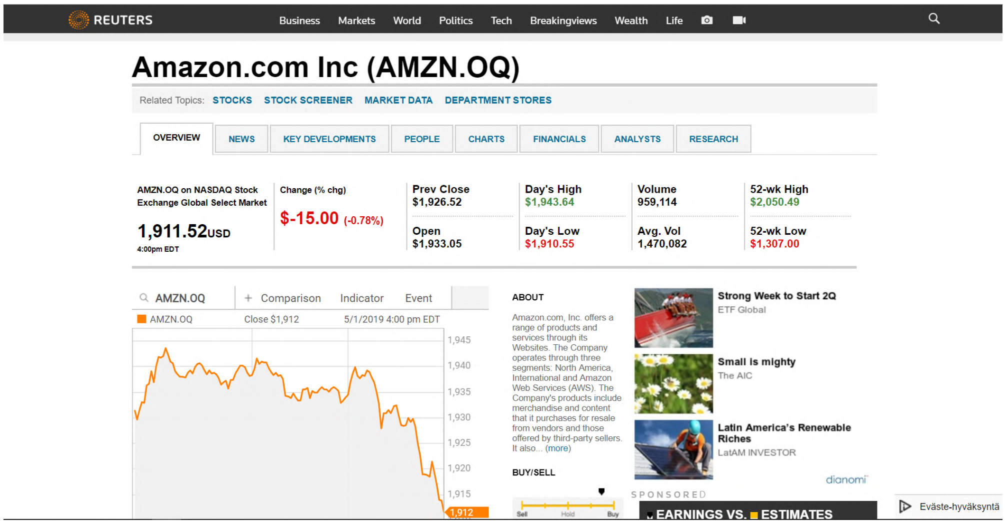
Panel B: NASDAQ



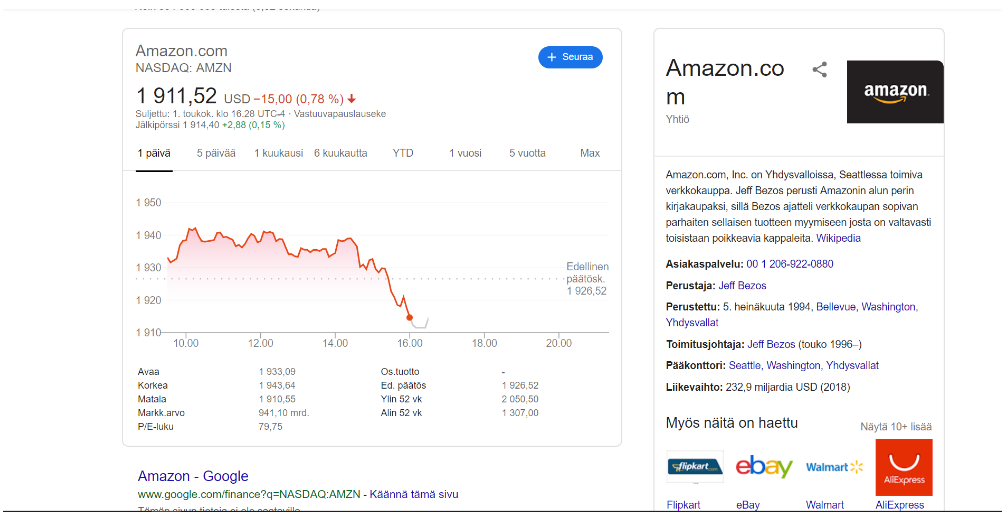
Panel C: Bloomberg



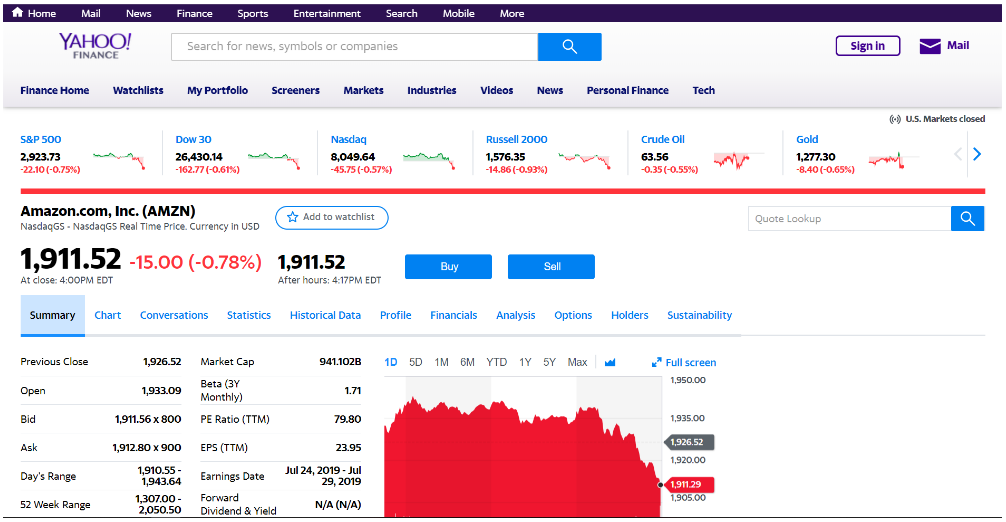
Panel D: Reuters



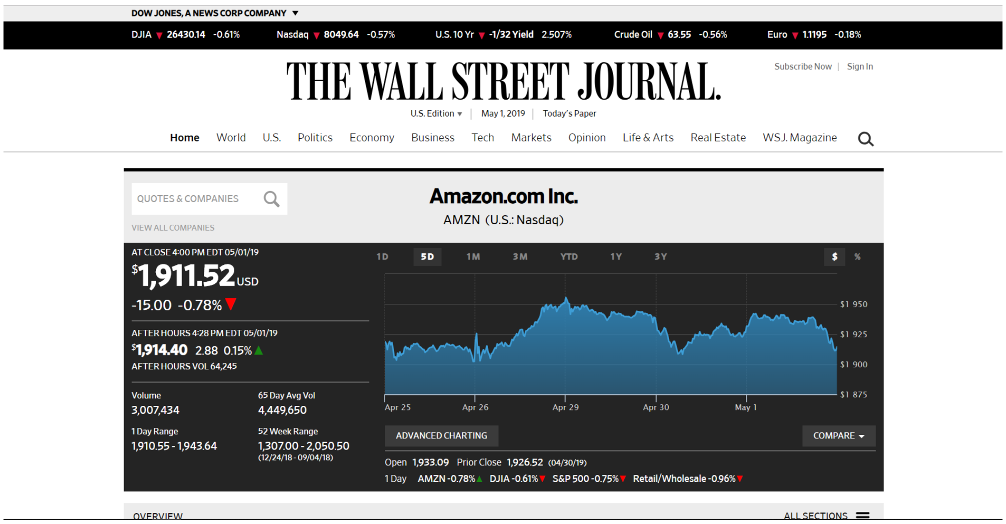
Panel E: Google Finance



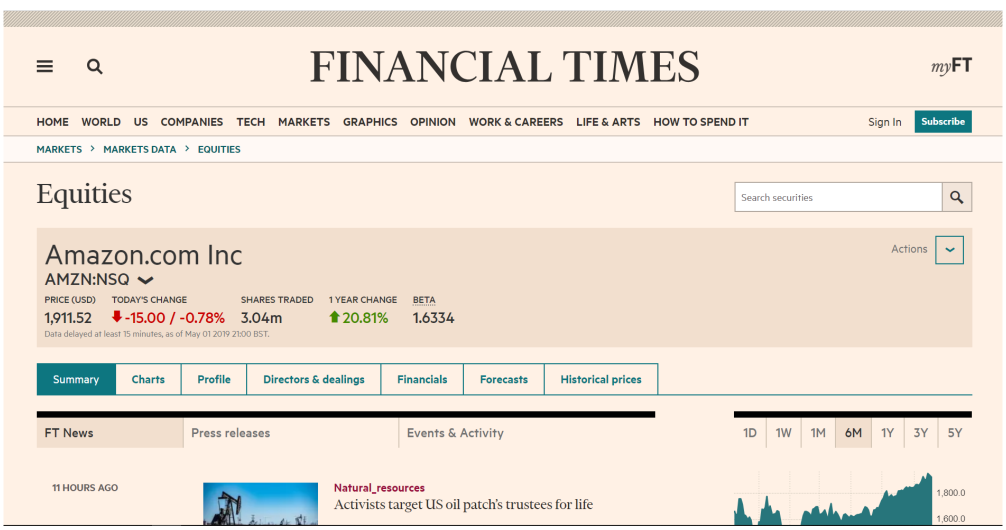
Panel F: Yahoo! Finance



Panel G: Wall Street Journal



Panel H: Financial Times



Panel I: WSJ Article from April 25^{th}, accessed on May 1^{st}, with Amazon's May 1^{st} return

By Yoree Koh ⊕

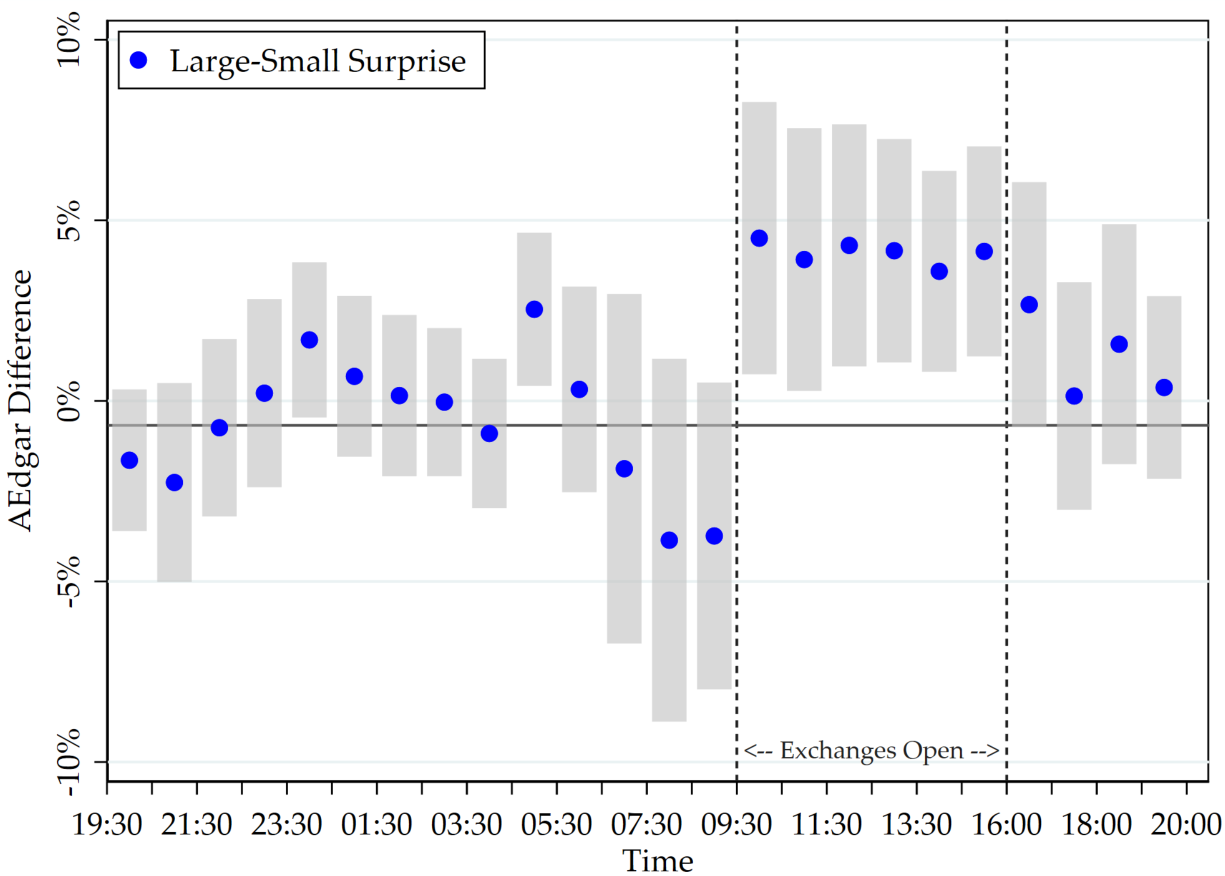
Updated April 25, 2019 08:39 p.m. EDT

Amazon.com Inc. AMZN -0.78% ▼ notched a best-ever $3.56 billion quarterly profit as it continued to lean on higher margin businesses and put a lid on costs.

Figure 3: Information Acquisition after Overnight Earnings Announcements with Large and Small Surprises

This figure shows the difference in abnormal information acquisition (AEdgar in Panel A, AWiki in Panel B) following overnight earnings announcements with small and large absolute earnings surprises. Surprises are based on SUE calculated from analysts' forecasts. Large surprises are surprises where SUE is smaller than the 30th or larger than the 70th percentile based on earnings announcements in the past 180 days. AEdgar/AWiki is measured hourly based on the log of downloads/page views differenced with the log of downloads/page views for the same hour and the same company one week before the announcement. The blue dots show the estimated difference for large versus small surprises and the gray area depicts the 95% confidence interval associated with the difference. The gray dashed lines indicate the beginning and end of trading hours. See text and Internet Appendix D for additional information.

Panel A: AEdgar



Panel B: AWiki

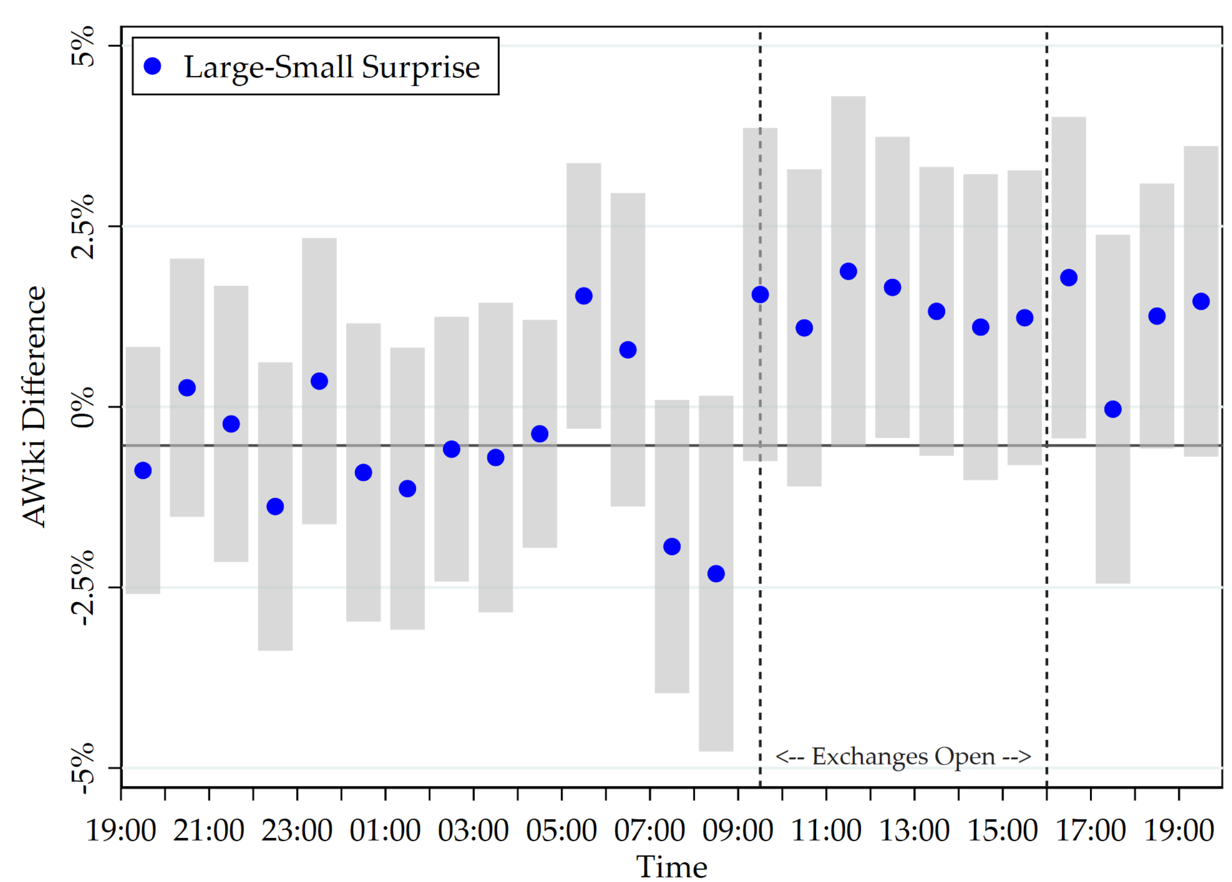


Figure 4: Turnover after Overnight Earnings Announcements with Large and Small Surprises

This figure shows the difference in abnormal turnover (ATurnover) following overnight earnings announcements with small and large absolute earnings surprises. Surprises are based on SUE calculated from analysts' forecasts. Large surprises are surprises where SUE is smaller than the 30th or larger than the 70th percentile based on earnings announcements in the past 180 days. ATurnover is measured hourly based on the log of turnover differenced with the log of turnover for the same hour and the same company one week before the announcement. The blue dots show the estimated difference for large versus small surprises and the gray area depicts the 95% confidence interval associated with the difference. See text and Internet Appendix D for additional information.

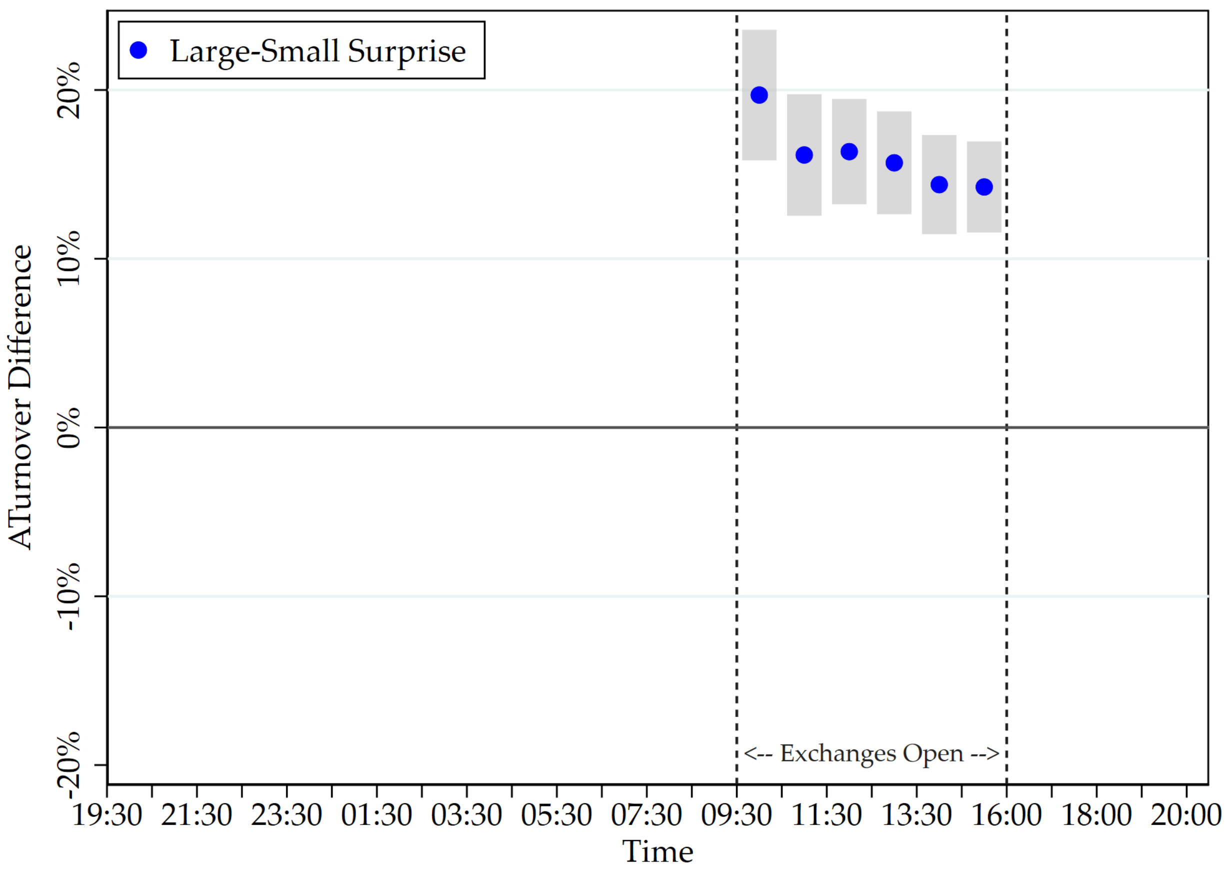


Figure 5: Wall Street Journal Gainer and Decliner Rankings

This figure exhibits Wall Street Journal Gainer and Decliner rankings of returns from both the website (Panel A) as well as the print edition (Panel B). These rankings are used in section 4.3 for identification purposes.

Panel A: WSJ Website (November 3rd, 2016)

Gainers (Roll over for charts and headlines) 5:02 pm EDT 11/03/16

NYSE | **Nasdaq** | **Arca** | **Composite**

Issue	Price	Chg	% Chg	Volume
Inteliquent (IQNT)	22.58	5.84	34.89	5,639,533
MetaldynePerform (MPG)	19.20	4.90	34.27	4,287,811
TechnicalComms (TCCO)	2.95	0.65	28.26	1,842,305
EnviroStar (EVI)	10.40	1.85	21.64	108,862
EnerNOC (ENOC)	5.80	0.95	19.59	297,150

See all Gainers Get this by E-mail

Decliners (Roll over for charts and headlines) 5:02 pm EDT 11/03/16

NYSE | **Nasdaq** | **Arca** | **Composite**

Issue	Price	Chg	% Chg	Volume
TrilliumTherap (TRIL)	6.85	-7.75	-53.08	2,552,000
AACHoldings (AAC)	7.80	-8.09	-50.91	3,734,902
DiplomatPharmacy (DPLO)	12.95	-9.43	-42.14	16,961,196
SequentialBrands (SQBG)	4.80	-2.43	-33.61	2,570,952
Fitbit (FIT)	8.51	-4.30	-33.57	72,700,405

See all Decliners Get this by E-mail

Panel B: WSJ Print Edition (April 5th, 2016)

Percentage Gainers...

Company	Symbol	Latest Session Close	Net chg	% chg	52-Week High	Low	% chg
OncoCyte	OCX	6.07	1.85	43.84	10.24	2.45	...
Virgin America	VA	55.11	16.21	41.67	55.43	26.30	88.0
Sky Solar Holdings ADR	SKYS	5.90	1.64	38.50	12.00	1.12	-50.3
Ruckus Wireless	RKUS	13.24	3.24	32.40	13.50	7.25	5.3
Transcontinental Realty	TCI	12.03	2.12	21.39	14.75	8.05	10.3
USMD Holdings	USMD	12.16	1.81	17.49	13.59	6.50	24.7
Unico American	UNAM	11.00	1.61	17.20	13.76	8.15	...
Edwards Lifesciences	EW	105.08	15.16	16.86	107.90	61.38	50.3
Flexion Therapeutics	FLXN	10.95	1.53	16.24	29.09	7.56	-52.2
Sorrento Therapeutics	SRNE	6.30	0.85	15.60	26.80	4.25	-46.0
Genocea Biosciences	GNCA	6.83	0.88	14.79	16.18	2.56	-39.2
MediciNova	MNOV	9.00	1.10	13.92	9.37	2.62	164.7
NantKwest	NK	9.72	1.09	12.63	38.48	6.10	...
Global Blood Therapeutics	GBT	18.14	2.01	12.46	57.00	12.24	...
SteadyMed	STDY	2.89	0.32	12.45	11.14	2.00	-66.3

Percentage Losers

Company	Symbol	Latest Session Close	Net chg	% chg	52-Week High	Low	% chg
Vericel	VCEL	3.72	-2.31	-38.31	6.69	1.69	1.9
Great Basin Scientific	GBSN	4.46	-2.05	-31.49	12810.00	3.69	-99.9
Staffing 360 Solutions	STAF	3.61	-0.94	-20.64	10.24	1.80	33.7
iRadimed	IRMD	15.51	-3.85	-19.89	33.25	14.54	0.1
Natus Medical	BABY	31.84	-7.80	-19.68	51.05	29.34	-19.7
Smith Wesson Hldg	SWHC	22.78	-4.98	-17.94	30.44	12.72	76.2
Direxion Brazil Bull 3X	BRZU	70.90	-11.88	-14.35	339.06	26.40	-71.5
ConforMIS	CFMS	10.28	-1.71	-14.26	26.93	7.56	...
ARC Group Worldwide	ARCW	2.17	-0.36	-14.23	8.44	1.05	-64.7
Brocade Comm Systems	BRCD	9.19	-1.45	-13.63	12.88	7.40	-22.8
GeoPark	GPRK	2.55	-0.35	-12.07	5.73	2.45	-39.6
Clovis Oncology	CLVS	17.26	-2.11	-10.89	116.75	16.78	-74.9
Naked Brand Group	NAKD	2.00	-0.23	-10.31	6.74	1.79	-52.4
Cartesian Inc.	CRTN	2.00	-0.23	-10.31	4.40	1.81	-44.0
ProShs Ultra MSCI Brazil	UBR	36.05	-4.08	-10.17	87.00	17.61	-48.9

Figure 6: Mispricing, Pop, and Returns

This figure plots the cumulative long-short trading strategy returns of two portfolios based on the SYY mispricing score of Stambaugh et al. (2012). Both portfolios go long each month in the stocks with the 20% lowest mispricing scores and short in the stocks with the 20% highest mispricing scores. For the low PoP portfolio (dashed blue line), only stocks with PoP in the lowest 40% are included. For the high PoP portfolio (solid black line), only stocks with PoP in the highest 40% are included. The stocks in the portfolios are value-weighted. PoP is the return salience measure based on daily returns, see the main text for details.

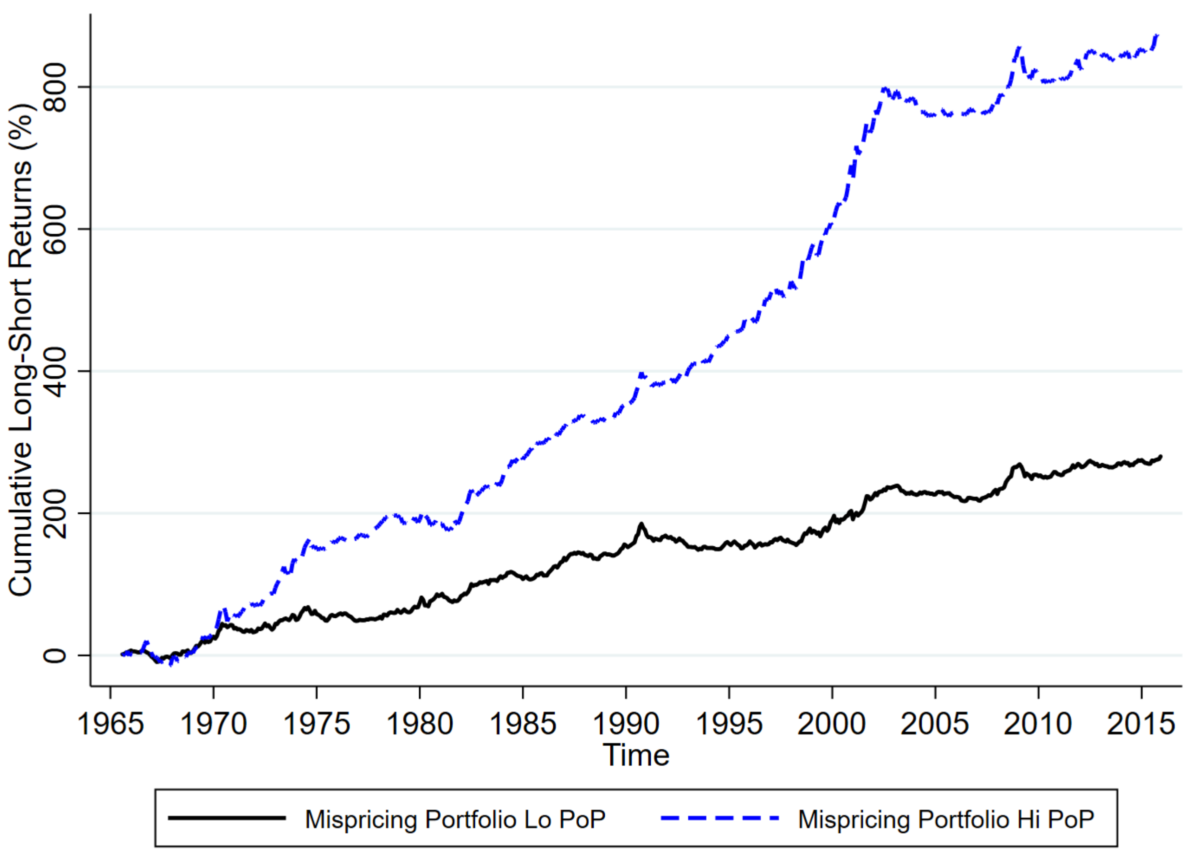


Table 1: Descriptive Statistics

This table provides summary statistics on the information acquisition, trading and return variables used in the analyses. Panel A provides summary statistics on the sample that includes all stock days. Panel B provides summary statistics on the overnight earnings announcement sample used in the difference-in-difference tests. Panel C provides summary statistics for this sample conditional on the absolute earnings surprise group, along with the difference in mean between the two groups and a corresponding two-sample t-test of the null hypothesis of equal means. The large surprise group entails the top and bottom 30% of earnings surprises and the remaining 40% form the small surprise group. All variables are defined in Internet Appendix D. ***, ** and * represent significance at the 1%-, 5%- and 10%-level, respectively.

Panel A: Summary Statistics (Full Sample - All Stock-Days)

	count	mean	sd	p1	p25	p50	p75	p99
Edgar after Close	5085434	24.83	90.03	0.00	4.00	11.00	26.00	212.00
AEdgar after Close	5085434	0.02	1.03	-2.48	-0.62	0.00	0.69	2.60
Wiki after Close	3900795	196.36	1459.07	0.00	6.00	31.00	110.00	2429.00
AWiki after Close	3890153	0.01	0.59	-1.54	-0.17	0.00	0.18	1.61
Turnover [%]	5392011	0.98	1.59	0.01	0.35	0.63	1.13	5.88
ATurnover	5317287	0.01	0.68	-1.66	-0.39	-0.01	0.39	1.84
RetailTurnover [%]	5340611	0.07	0.28	0.00	0.01	0.03	0.06	0.60
ARetailTurnover	5080852	0.01	0.97	-2.48	-0.55	-0.01	0.55	2.62
$\text{Close}_{\text{t-1}}$ to Close_{t} Return	5274042	-0.01	3.01	-8.39	-1.18	0.00	1.20	8.17
Open_{t} to Close_{t} Return	5288605	0.01	2.32	-6.62	-0.94	0.03	0.98	6.54
12pm_{t} to Close_{t} Return	5316020	0.02	1.66	-4.67	-0.62	0.01	0.67	4.77
3pm_{t} to Close_{t} Return	5322468	0.02	1.02	-2.72	-0.34	0.00	0.36	3.02
$\text{Close}_{\text{t-1}}$ to Open_{t} Return	5230996	-0.02	1.86	-4.67	-0.63	-0.00	0.61	4.55
$\text{Close}_{\text{t-1}}$ to 12pm_{t} Return	5257173	-0.03	2.48	-6.67	-0.98	0.00	0.96	6.41
$\text{Close}_{\text{t-1}}$ to 3pm_{t} Return	5260317	-0.03	2.82	-7.84	-1.12	0.00	1.12	7.37
$\text{12pm}_{\text{t-1}}$ to Close_{t} Return	5299808	0.01	3.39	-9.54	-1.33	0.04	1.39	9.19
8-K Filing	5393174	0.05	0.21	0.00	0.00	0.00	0.00	1.00
10-K or 10-Q Filing	5393174	0.02	0.12	0.00	0.00	0.00	0.00	1.00
Earnings Announcement	5393174	0.02	0.12	0.00	0.00	0.00	0.00	1.00
Dividend Announcement	5393174	0.01	0.09	0.00	0.00	0.00	0.00	0.00
Media Coverage	5393174	0.01	0.09	0.00	0.00	0.00	0.00	0.00

Panel B: Summary Statistics (Overnight Earnings Announcements)

	N	Mean	StdDev	P1	P25	P50	P75	P99
$Close_{t\text{-}1}$ to $Close_t$ Return [%]	22691	-0.14	7.23	-23.29	-3.23	0.09	3.48	17.29
$Close_{t\text{-}1}$ to $Open_t$ Return [%]	22654	-0.01	5.70	-18.32	-2.44	0.12	2.82	14.02
Turnover [%]	22849	2.67	3.25	0.11	0.91	1.69	3.19	15.83
Edgar (Preceeding Month)	22849	1435.09	2604.04	0.00	430.00	851.00	1602.00	10349.00
Edgar (Previous Day)	22849	59.31	108.98	0.00	14.00	33.00	69.00	452.00
Edgar (EA Day)	22849	114.71	170.87	0.00	32.00	71.00	139.00	733.00
Edgar Pre	22849	19.18	31.88	0.00	3.00	10.00	24.00	141.00
Edgar Post	22849	22.48	34.77	0.00	5.00	13.00	28.00	151.00
AEdgar (Preceeding Month)	22849	0.01	0.45	-1.06	-0.23	-0.01	0.21	1.48
AEdgar (Previous Day)	22849	0.28	0.96	-1.99	-0.25	0.18	0.77	3.00
AEdgar (EA Day)	22849	0.71	0.82	-1.34	0.21	0.69	1.17	2.94
AEdgar Pre	22849	1.30	1.16	-1.50	0.51	1.34	2.08	3.87
AEdgar Post	22849	0.96	1.09	-1.73	0.18	0.96	1.69	3.47
Wiki (Preceeding Month)	17749	8412.36	19245.24	0.00	533.00	2172.00	7474.00	94377.00
Wiki (Previous Day)	17285	317.16	782.19	0.00	21.00	83.00	282.00	3529.00
Wiki (EA Day)	17285	356.32	807.54	0.00	25.00	99.00	329.00	3791.00
Wiki Pre	17257	16.32	39.40	0.00	0.00	4.00	15.00	185.00
Wiki Post	17257	21.11	47.55	0.00	1.00	6.00	20.00	227.00
AWiki (Preceeding Month)	17591	-0.08	0.89	-6.02	-0.10	0.00	0.13	1.22
AWiki (Previous Day)	17282	0.01	0.44	-1.28	-0.13	0.00	0.15	1.20
AWiki (EA Day)	17282	0.12	0.47	-1.10	-0.04	0.08	0.28	1.54
AWiki Pre	17254	0.16	0.71	-1.61	-0.07	0.00	0.56	2.08
AWiki Post	17254	0.16	0.69	-1.61	-0.10	0.00	0.51	2.08
Large Surprise	22849	0.59	0.49	0.00	0.00	1.00	1.00	1.00
SUE [%]	22849	0.03	1.02	-3.51	-0.03	0.05	0.21	2.41
Predicted Return	21726	-0.00	0.03	-0.06	-0.02	0.00	0.02	0.05
Number of Analysts	22849	7.01	5.37	1.00	3.00	6.00	10.00	24.00
Market Cap [bio]	22843	10.83	26.54	0.24	0.92	2.55	8.51	155.81
InstOwn [%]	22849	77.28	27.17	0.39	67.67	83.14	93.33	120.28

Panel C: Summary Statistics (Overnight Earnings Announcements by Absolute Surprise Group)

	Large Surprise		Small Surprise		Difference	
	Mean	StdDev	Mean	StdDev	Delta	t-Stat
|SUE| [%]	0.62	1.18	0.07	0.06	0.55***	(53.67)
|$\text{Close}_{\text{t-1}}$ to Close_{t} Return| [%]	5.64	5.94	3.88	4.07	1.76***	(26.47)
|$\text{Close}_{\text{t-1}}$ to Open_{t} Return| [%]	4.47	4.67	3.02	3.22	1.45***	(27.62)
|Predicted Return| [%]	3.01	1.11	0.84	0.47	2.17***	(195.64)
Turnover [%]	3.00	3.63	2.20	2.53	0.79***	(19.45)
AEdgar (Preceeding Month)	0.01	0.45	0.01	0.45	0.00	(0.04)
AEdgar (Previous Day)	0.27	0.97	0.30	0.94	-0.03**	(-2.00)
AEdgar Pre	1.29	1.17	1.31	1.13	-0.02	(-1.24)
AEdgar Post	0.99	1.10	0.92	1.07	0.07***	(5.13)
AWiki (Preceeding Month)	-0.08	0.92	-0.08	0.86	0.01	(0.49)
AWiki (Previous Day)	0.01	0.45	0.01	0.43	0.00	(0.15)
AWiki Pre	0.15	0.71	0.18	0.71	-0.02**	(-2.11)
AWiki Post	0.17	0.69	0.16	0.68	0.01	(1.04)
Market Cap [bio]	7.40	19.91	15.68	33.15	-8.28***	(-21.70)
Number of Analysts	6.32	5.17	7.98	5.49	-1.65***	(-22.99)
InstOwn [%]	76.56	26.46	78.30	28.10	-1.74***	(-4.71)
Observations	13380		9469		22849	

Table 2: Information Acquisition and Returns - Difference-in-Difference Test using Overnight Earnings Announcements

This table provides the results of difference-in-difference tests of abnormal information acquisition (AEdgar in Panel A, AWiki in Panel B) for overnight earnings announcements around the opening of the exchanges at 9:30 AM. Large Surprise is an indicator that takes on a value of one if SUE is smaller than the 30th or larger than the 70th percentile based on earnings announcements in the past 180 days. For each announcement, we measure information acquisition for two windows: 7:30 AM to 9:30 AM and 9:30 AM to 11:30 AM for AEdgar and 8:00am to 09:00am and 10:00am to 11:00am for AWiki. Post-Open is an indicator that take on a value of one for the latter window. Predicted return is the average announcement day return of companies with similarly surprising announcements over the previous 180 days. All variables are defined in appendix D. Standard errors are two-way clustered by stock and month. ***, ** and * represent significance at the 1%-, 5%- and 10%-level, respectively.

Panel A: AEdgar

	(1) AEdgar	(2) AEdgar	(3) AEdgar	(4) AEdgar
Large Surprise × Post-Open	0.094*** (4.98)	0.094*** (4.98)		
Large Surprise	-0.019 (-0.80)			
|Predicted Return|× Post-Open			3.812*** (5.18)	3.812*** (5.18)
|Predicted Return|			-1.455 (-1.48)	
Post-Open	-0.390*** (-19.57)	-0.390*** (-19.57)	-0.428*** (-19.61)	-0.428*** (-19.61)
Constant	1.308*** (36.00)		1.354*** (34.83)	
Stock x Announcement FE	No	Yes	No	Yes
Observations	45698	45698	43452	43452
Adjusted R2	0.022	0.355	0.024	0.357

Panel B: AWiki

	(1) AWiki	(2) AWiki	(3) AWiki	(4) AWiki
Large Surprise × Post-Open	0.034***	0.034***		
	(2.68)	(2.68)		
Large Surprise	-0.023*			
	(-1.83)			
|Predicted Return|× Post-Open			1.229***	1.229***
			(3.01)	(3.01)
|Predicted Return|			-1.221***	
			(-2.65)	
Post-Open	-0.021**	-0.021**	-0.027**	-0.027**
	(-2.07)	(-2.07)	(-2.26)	(-2.26)
Constant	0.178***		0.190***	
	(12.69)		(11.81)	
Stock x Announcement FE	No	Yes	No	Yes
Observations	34508	34508	34508	34508
Adjusted R2	0.000	0.265	0.000	0.265

Table 3: Turnover and Overnight Earnings Announcements

This tables shows the results of OLS regressions of abnormal turnover on a dummy for large earnings surprises or the predicted return for the sample used in Table 2. All variables are defined in Internet Appendix D. Standard errors are two-way clustered by stock and month. ***, ** and * represent significance at the 1%-, 5%- and 10%-level, respectively.

	(1) ATurnover	(2) ATurnover	(3) ATurnover	(4) ATurnover
Large Surprise	0.161***	0.216***		
	(9.29)	(14.71)		
|Predicted Return|			3.988***	8.719***
			(3.76)	(11.43)
Constant	1.130***		1.139***	
	(49.40)		(47.24)	
Stock FE	No	Yes	No	Yes
Year-Month FE	No	Yes	No	Yes
Observations	22370	21885	21292	20819
Adjusted R2	0.008	0.242	0.004	0.243

Table 4: WSJ Ranks and Information Acquisition

This tables shows the results of OLS regressions of abnormal information acquisition (AEdgar/AWiki) on dummies according to the stock being ranked among the top or bottom 15 stocks in the Wall Street Journal Composite Gainer/Decliner ranking. The regressions include fixed effects for the top and bottom 200 daily ranks of the CRSP universe in our sample (400 effects), for each percentile rank in the daily cross-section of returns of the CRSP universe in our sample (100 effects) and for each whole sample percentile rank of the CRSP universe in our sample (100 effects) as indicated below the coefficients. All variables are defined in Internet Appendix D. Standard errors are two-way clustered by stock and month. ***, ** and * represent significance at the 1%-, 5%- and 10%-level, respectively.

	AEdgar after Close (t)		AWiki after Close (t)	
	(1)	(2)	(3)	(4)
WSJ Gainer 1-5	0.742***	0.296***	0.666***	0.171***
	(26.66)	(4.95)	(24.16)	(3.66)
WSJ Gainer 6-10	0.522***	0.185***	0.363***	0.076**
	(23.81)	(4.27)	(20.65)	(2.36)
WSJ Gainer 11-15	0.411***	0.083***	0.278***	0.029
	(18.91)	(2.68)	(14.87)	(1.24)
WSJ Decliner 11-15	0.425***	0.078***	0.214***	0.030
	(20.36)	(2.69)	(13.07)	(1.38)
WSJ Decliner 6-10	0.546***	0.171***	0.250***	0.029
	(27.12)	(4.63)	(16.10)	(1.10)
WSJ Decliner 1-5	0.858***	0.323***	0.448***	0.086**
	(39.14)	(6.31)	(23.56)	(2.21)
CRSP Top/Bottom 200 FE	No	Yes	No	Yes
CRSP Return Percentile FE	No	Yes	No	Yes
CRSP Return Group FE	No	Yes	No	Yes
5 Lags of Dependent	Yes	Yes	Yes	Yes
Stock and Day FE	Yes	Yes	Yes	Yes
Observations	3279467	3279349	2761254	2760631
Adjusted R2	0.228	0.229	0.405	0.406

Table 5: WSJ Ranks and Other Outcomes

This tables shows the results of OLS regressions of abnormal turnover and abnormal retail turnover on dummies according to the stock being ranked among the top or bottom 15 stocks in the Wall Street Journal Composite Gainer/Decliner ranking. The regressions include fixed effects for the top and bottom 200 daily ranks of the CRSP universe in our sample (400 effects), for each percentile rank in the daily cross-section of returns of the CRSP universe in our sample (100 effects) and for each whole sample percentile rank of the CRSP universe in our sample (100 effects) as indicated below the coefficients. All variables are defined in Internet Appendix D. Standard errors are two-way clustered by stock and month. ***, ** and * represent significance at the 1%-, 5%- and 10%-level, respectively.

	ATurnover (t+1)		ARetailTurnover (t+1)	
	(1)	(2)	(3)	(4)
WSJ Gainer 1-5	0.826***	0.261***	1.245***	0.364***
	(38.39)	(5.74)	(42.96)	(5.64)
WSJ Gainer 6-10	0.466***	0.134***	0.855***	0.216***
	(29.70)	(4.20)	(38.56)	(4.40)
WSJ Gainer 11-15	0.359***	0.068***	0.687***	0.118***
	(23.94)	(2.98)	(30.02)	(3.45)
WSJ Decliner 11-15	0.375***	0.054**	0.739***	0.020
	(25.70)	(2.50)	(33.73)	(0.63)
WSJ Decliner 6-10	0.535***	0.180***	0.980***	0.188***
	(30.12)	(5.60)	(43.55)	(4.05)
WSJ Decliner 1-5	0.775***	0.216***	1.430***	0.274***
	(44.46)	(4.76)	(56.61)	(3.97)
CRSP Top/Bottom 200 FE	No	Yes	No	Yes
CRSP Return Percentile FE	No	Yes	No	Yes
CRSP Return Group FE	No	Yes	No	Yes
5 Lags of Dependent	Yes	Yes	Yes	Yes
Stock and Day FE	Yes	Yes	Yes	Yes
Observations	3467229	3467229	2724346	2724344
Adjusted R2	0.327	0.331	0.262	0.269

Table 6: Dependent Double Sorts

In this table, we report the average excess returns of dependent bivariate sorts. All stocks are first sorted on lagged SYY score into quintile bins and then within each bin sorted by lagged PoP into quintile. PoP is the return salience measure based on daily returns, see the main text for details. SYY score is the mispricing score of Stambaugh et al. (2012). For definitions of other variables, see Appendix D. Alphas are estimated as the intercept estimates of time-series regression of long-short portfolio on factor models. The sample covers all $\geq$ \$5 U.S. common stocks traded on the NYSE, AMEX and NASDAQ with market cap higher than the 2nd NYSE decile from 1963 to 2015. t-statistics are based on Newey-West standard errors with one lag and are reported in parentheses.

Panel A: Value Weighted Returns

	PoP					
SYY Score	Low	2	3	4	High	5-1
Underpriced	0.72%	0.74%	0.71%	0.95%	0.94%	0.22%
2	0.68%	0.63%	0.52%	0.47%	0.67%	-0.01%
3	0.63%	0.57%	0.54%	0.49%	0.52%	-0.12%
4	0.54%	0.37%	0.43%	0.47%	0.04%	-0.51%
Overpriced	0.38%	0.15%	0.06%	-0.16%	-0.83%	-1.21%
5-1	**-0.34%**	**-0.59%**	**-0.65%**	**-1.12%**	**-1.77%**	**-1.43%**
	(-2.74)	(-3.86)	(-3.41)	(-4.87)	(-6.78)	(-5.57)
Carhart alpha	**-0.28%**	**-0.49%**	**-0.41%**	**-0.92%**	**-1.54%**	**-1.25%**
	(-2.33)	(-3.59)	(-2.63)	(-5.04)	(-7.59)	(-5.45)
FF5 alpha	**-0.27%**	**-0.46%**	**-0.46%**	**-0.98%**	**-1.66%**	**-1.39%**
	(-2.11)	(-3.28)	(-2.64)	(-4.73)	(-7.00)	(-5.71)

Panel B: Equal Weighted Returns

	PoP					
SYY Score	Low	2	3	4	High	5-1
Underpriced	0.94%	0.91%	0.93%	1.14%	1.00%	0.06%
2	0.85%	0.90%	0.79%	0.88%	0.80%	-0.05%
3	0.78%	0.70%	0.83%	0.74%	0.62%	-0.15%
4	0.62%	0.65%	0.67%	0.69%	0.38%	-0.24%
Overpriced	0.53%	0.33%	0.23%	0.00%	-0.67%	-1.20%
5-1	**-0.41%**	**-0.58%**	**-0.70%**	**-1.14%**	**-1.67%**	**-1.25%**
	(-4.36)	(-5.38)	(-5.06)	(-6.12)	(-8.18)	(-6.34)
Carhart alpha	**-0.34%**	**-0.50%**	**-0.52%**	**-0.89%**	**-1.41%**	**-1.07%**
	(-3.54)	(-5.16)	(-4.89)	(-7.50)	(-9.16)	(-6.28)
FF5 alpha	**-0.27%**	**-0.41%**	**-0.48%**	**-0.94%**	**-1.41%**	**-1.14%**
	(-3.04)	(-4.56)	(-4.24)	(-6.08)	(-8.29)	(-6.64)

Table 7: Fama-MacBeth Regressions

In this table, we report results from Fama and MacBeth (1973) regressions of this month's return on stock characteristics available at the end of last month. PoP is the return salience measure based on daily returns, see the main text for details. SYY score is the mispricing score of Stambaugh et al. (2012). For definitions of other variables, see Appendix D. The sample covers all ≥ \$5 U.S. common stocks traded on the NYSE, AMEX and NASDAQ with market cap higher than the 2nd NYSE decile from 1963 to 2015. t-statistics are based on Newey-West standard errors with one lag and are reported in parentheses. ***, **, and * indicate significance at the 1%-, 5%- and 10%-level, respectively.

Panel A

	(1)	(2)	(3)	(4)	(5)	(6)
SYY Score	-0.0029***	-0.0030***	-0.0025***	-0.0019***	-0.0025***	-0.0025***
	(-8.65)	(-8.50)	(-9.44)	(-6.21)	(-9.33)	(-10.15)
SYY Score × PoP		**-0.0018***	**-0.0016***	**-0.0015***	**-0.0016***	**-0.0017***
		(-7.60)	**(-7.26)**	**(-6.54)**	**(-7.19)**	**(-8.19)**
PoP	-0.0019**	-0.0014*	-0.0021***	-0.0022***	-0.0020***	-0.0023***
	(-2.36)	(-1.74)	(-5.24)	(-5.22)	(-5.42)	(-6.21)
Beta			0.0021	0.0014	0.0023*	0.0019
			(1.54)	(1.11)	(1.81)	(1.67)
ln(Size)			-0.0013***	-0.0013***	-0.0013***	0.0000
			(-3.68)	(-3.68)	(-3.87)	(0.04)
ln(BTM)			0.0015**	0.0016**	0.0016***	0.0023***
			(2.55)	(2.28)	(2.68)	(4.59)
$\text{ret}_{t-12,t-2}$			0.0076***	0.008***	0.0078***	0.0066***
			(4.60)	(4.91)	(4.66)	(4.66)
$\text{ret}_{t-1,t-1}$			-0.0344***	-0.0356***	-0.035***	-0.0498***
			(-8.01)	(-8.36)	(-8.20)	(-12.42)
$\text{ret}_{t-36,t-13}$			-0.0001	-0.0002	-0.0001	0.0002
			(-0.09)	(-0.31)	(-0.18)	(0.32)
Profitability (FF)				0.0055***		
				(2.64)		
Investments (FF)				-0.0023*		
				(-1.92)		
ln(Turnover)					-0.0004	
					(-0.83)	
Δln(Turnover)					0.0007	
					(1.55)	
FF48 Industry FE	No	No	No	No	No	Yes
Size FE	No	No	No	No	No	Yes
Exchange FE	No	No	No	No	No	Yes
Months	605	605	605	605	605	605
Avg. N	1584	1584	1418	1026	1378	1401
Avg. R2	0.030	0.033	0.097	0.100	0.104	0.192
Years	1965-2015	1965-2015	1965-2015	1965-2015	1965-2015	1965-2015

Panel B

	(1) Excl. DWL	(2) 13f Sample	(3) Low IO	(4) High IO	(5) 1965-1990	(6) 1991-2015
SYY Score	-0.0027***	-0.0026***	-0.0036***	-0.002***	-0.0023***	-0.0026***
	(-9.01)	(-8.07)	(-8.92)	(-5.10)	(-7.59)	(-6.15)
SYY Score × PoP	**-0.0017***	**-0.0019***	**-0.0027***	**-0.0010***	**-0.0012***	**-0.0019***
	(-4.87)	(-6.94)	(-7.05)	(-2.69)	(-4.66)	(-5.62)
PoP	-0.0006	-0.002***	-0.0024***	-0.0018***	-0.0027***	-0.0016**
	(-0.89)	(-3.97)	(-3.79)	(-3.06)	(-5.96)	(-2.33)
Beta	0.0019	0.0019	0.0027	0.0012	0.0027	0.0016
	(1.39)	(1.11)	(1.44)	(0.70)	(1.49)	(0.76)
ln(Size)	-0.0011***	-0.0012***	-0.0011***	-0.0012***	-0.0014***	-0.0012**
	(-3.40)	(-2.99)	(-2.80)	(-2.83)	(-2.68)	(-2.53)
ln(BTM)	0.0015**	0.0015**	0.0019**	0.0011	0.0021**	0.0010
	(2.56)	(2.19)	(2.53)	(1.57)	(2.33)	(1.22)
$\text{ret}_{t-12,t-2}$	0.0079***	0.0049**	0.0034*	0.0064***	0.0113***	0.004*
	(4.66)	(2.56)	(1.67)	(3.14)	(4.9)	(1.67)
$\text{ret}_{t-1,t-1}$	-0.0506***	-0.0248***	-0.0172***	-0.0337***	-0.0517***	-0.0169**
	(-10.71)	(-4.86)	(-3.06)	(-6.20)	(-8.83)	(-2.82)
$\text{ret}_{t-36,t-13}$	0.0004	-0.000	-0.0003	0.0003	0.0004	-0.0005
	(0.68)	(-0.05)	(-0.33)	(0.38)	(0.35)	(-0.63)
Months	605	420	420	420	305	300
Avg. N	1252	1416	651	765	1201	1638
Avg. R2	0.099	0.091	0.115	0.088	0.104	0.090
Years	1965- 2015	1980- 2015	1980- 2015	1980- 2015	1965- 1990	1991- 2015

Internet Appendix to
The Power of Prices:
Information, Trade, and Salient Returns

Abstract

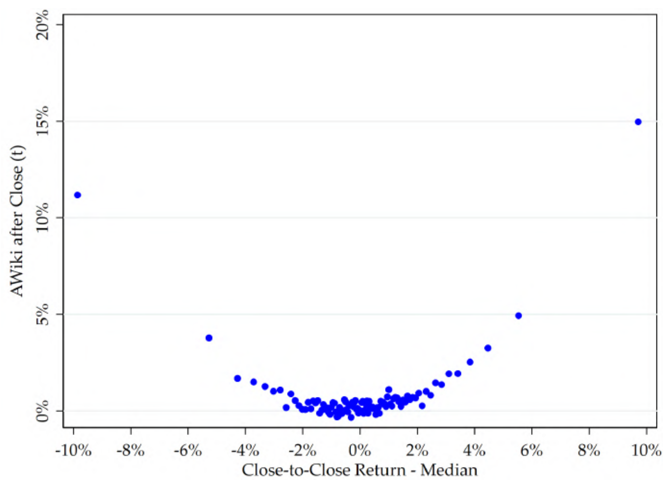
The Internet Appendix consists of six sections. Appendices A and B contain additional figures and tables. Appendix C contains discussions of additional empirical results. In Appendix D, we define the main variables and explain how they are constructed. In Appendix E, we provide information on how the SEC log-file dataset is cleaned and prepared and how users are classified.

A Additional Figures

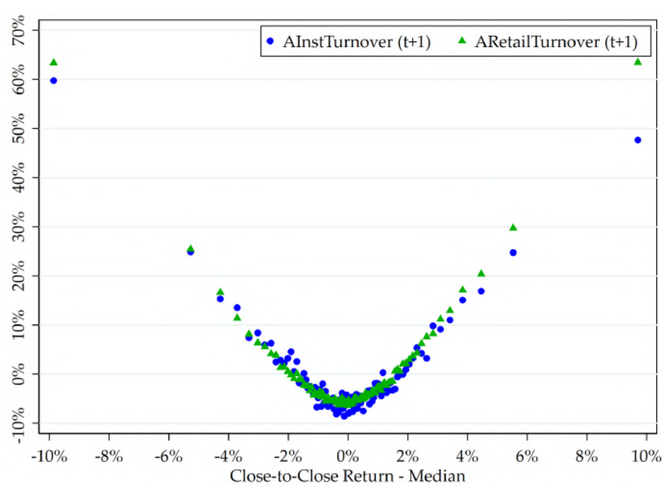
Figure A1: Information Acquisition, Trading and Close to Close Returns

This figure shows average abnormal Wiki pageviews and abnormal turnover by trader type in Panels A and B. Each trading day, stocks are sorted into 100 portfolios based on their median adjusted return. Then, within each portfolio, variables are averaged within the cross-section of included stocks and in a second step averaged within the time-series. Wiki pageviews are measured after the market closes and before the next open. Retail and institutional turnover are measured on the next trading day. Hence, Panels A and B do not represent contemporaneous associations. Panel C shows the average absolute median adjusted return for each portfolio. Variables are defined in Internet Appendix D.

Panel A: Information Acquisition (AWiki) after the Market Closes



Panel B: Turnover on the Next Trading Day By Trader Type



Panel C: Close to Close Returns

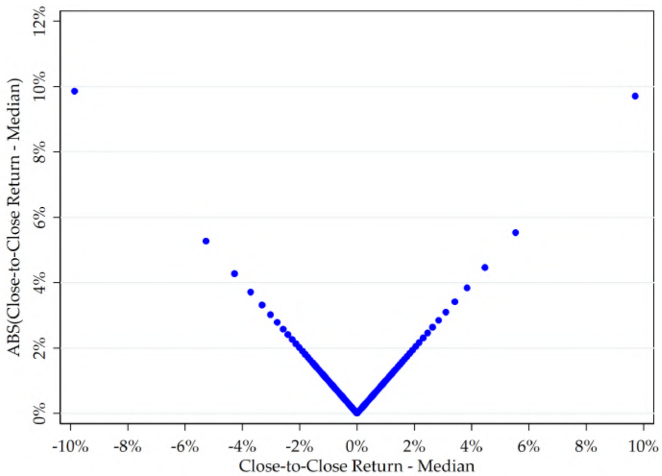


Figure A2: Timing of Earnings Announcements

This figure provides the frequency of earnings announcements in the sample for every five minute interval of the day based on IBES time-stamps. The sample construction is explained in detail in Section 3 of the main text. Earnings announcement cluster heavily in the first two hours after the market closes and in the 3.5 hours before the market opens.

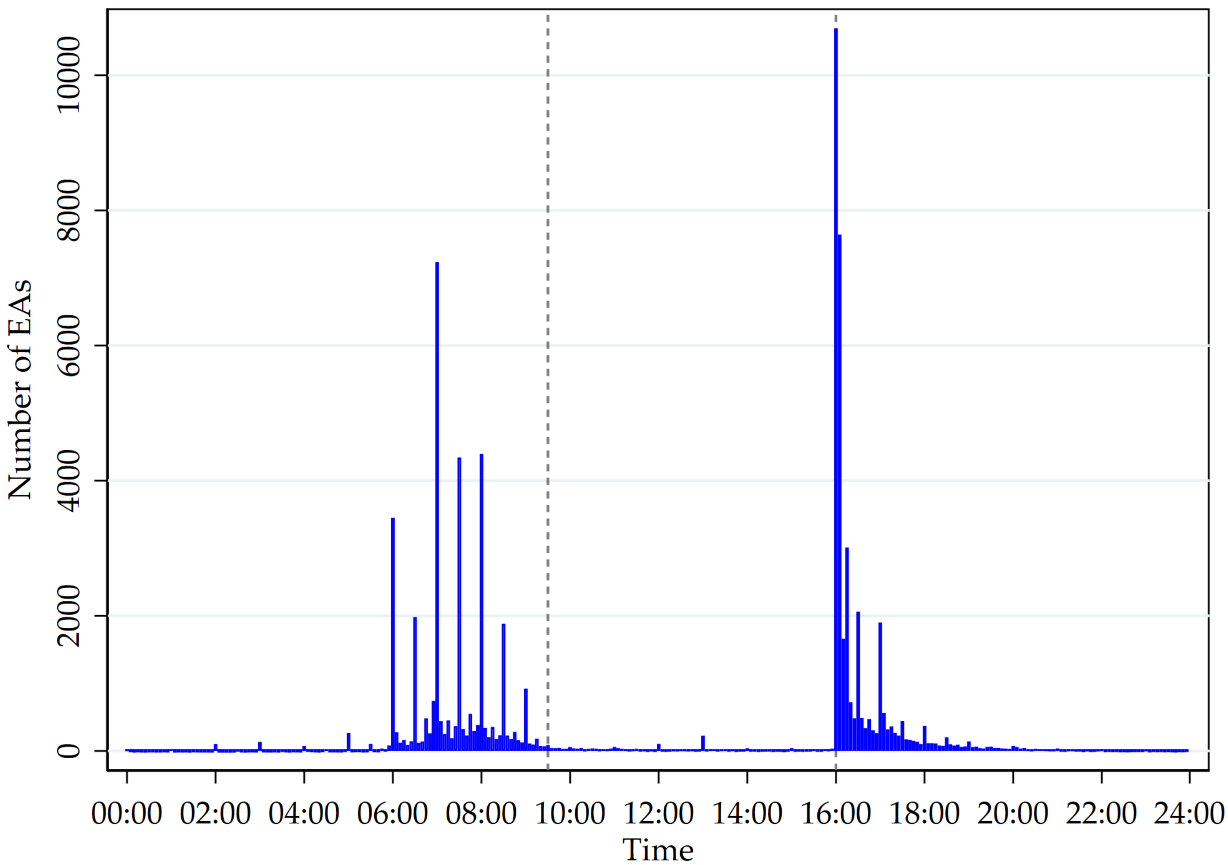


Figure A3: Cumulative Number of Earnings Announcements by Time

This figure shows the cumulative number of overnight earnings announcement for the announcements included in the sample for the difference-in-difference tests. To be included, the announcement needs to take place between 8:00 PM on the previous day and 7:30 AM in the morning. Firms with small and large surprised do exhibit very similar timing profiles.

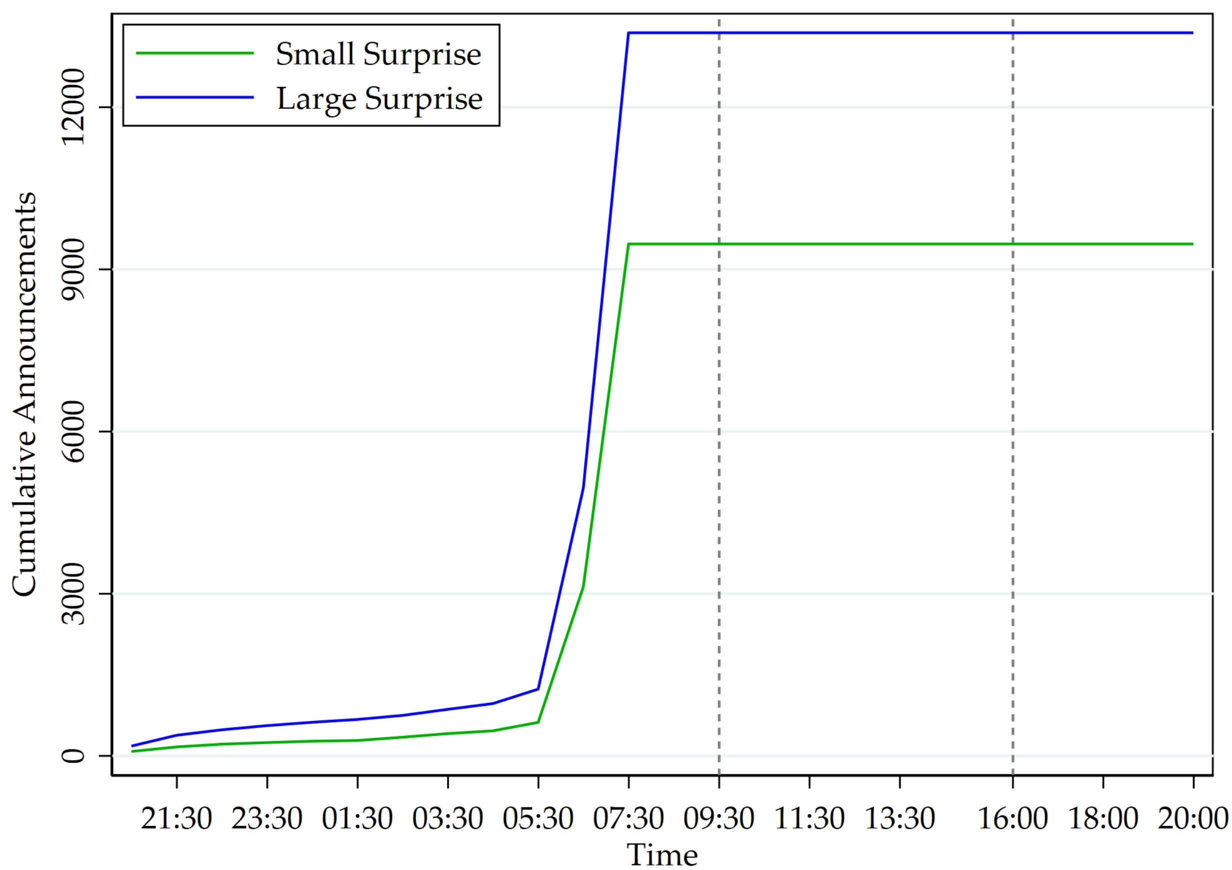
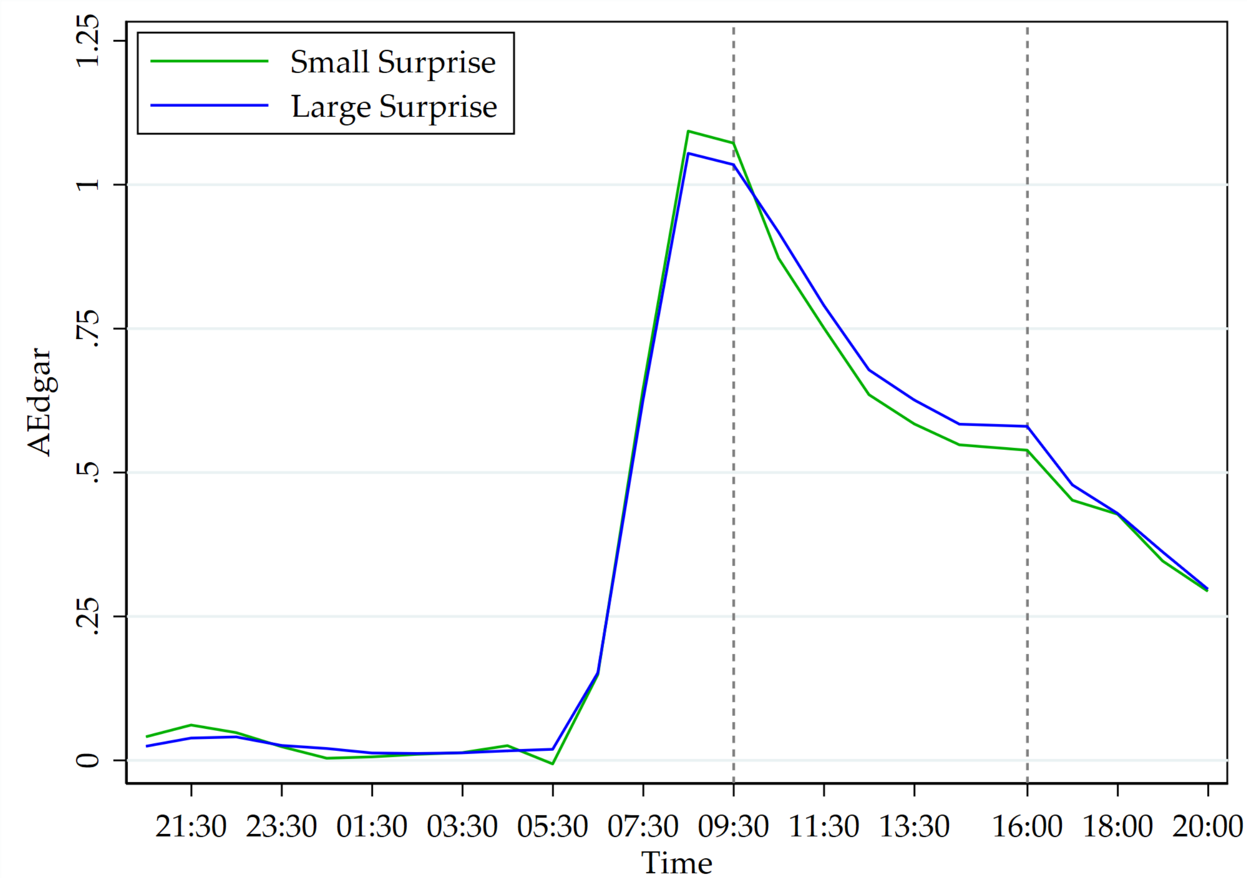


Figure A4: Abnormal Edgar for Large and Small Surprises

This figure shows hourly abnormal information acquisition (AEdgar in Panel A, AWiki in Panel B) for the overnight earnings announcement sample conditional on the magnitude of the absolute earnings surprise. To be included, the announcement needs to take place between 8:00 PM on the previous day and 7:30 AM in the morning. The grey lines indicate the opening and closing of the exchanges considered.

Panel A: AEdgar



Panel B: AWiki

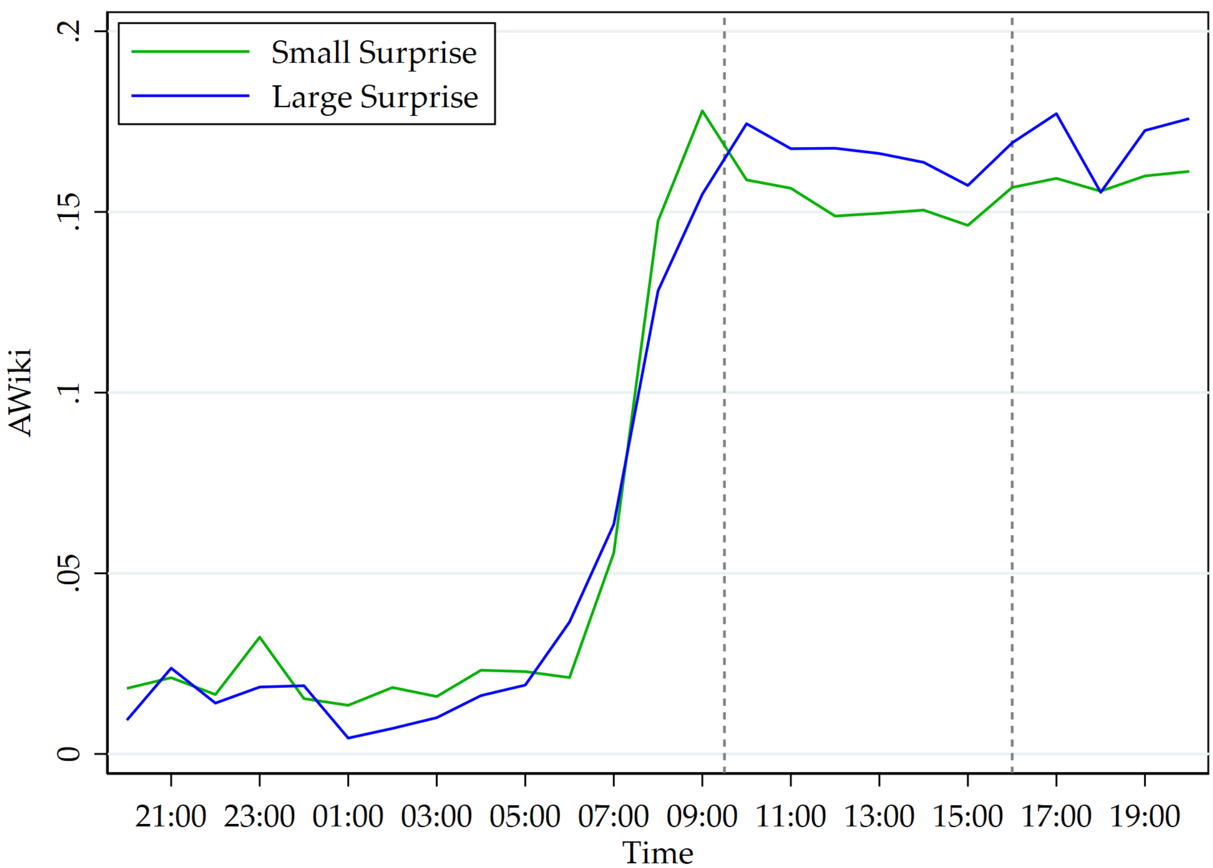
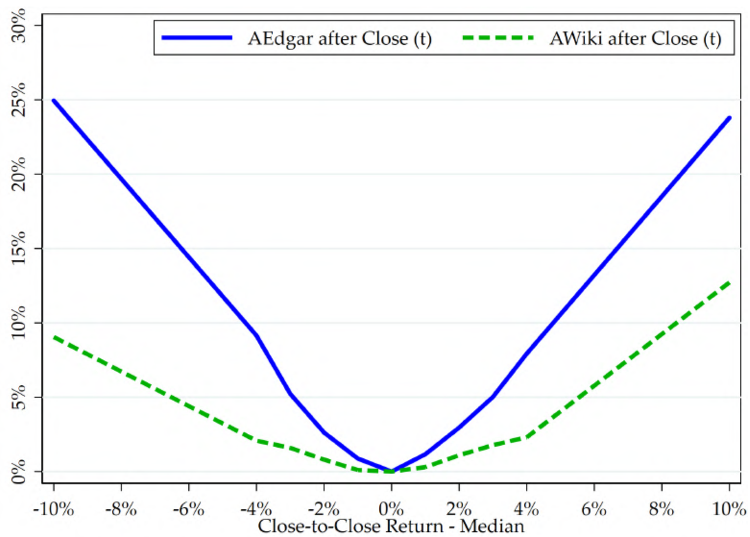


Figure A5: Spline Regressions of Information Acquisition

This figure shows the implied conditional mean function from the results of linear spline regressions of abnormal information acquisition (AEdgar in Panel A, AWiki in Panel B) on median adjusted close-to-close returns. The corresponding regression results can be found in table B9 of the internet appendix. In panel A we use spline terms based on continuous returns with knots at every integer percentage in the interval [-4,4]. In panel B, we use spline terms based on cross-sectional return percentile ranks with knots at the 5th, 10th, 50th, 90th and 95th percentile.

Panel A: Specification based on Returns



Panel B: Specification based on Percentile-Ranks

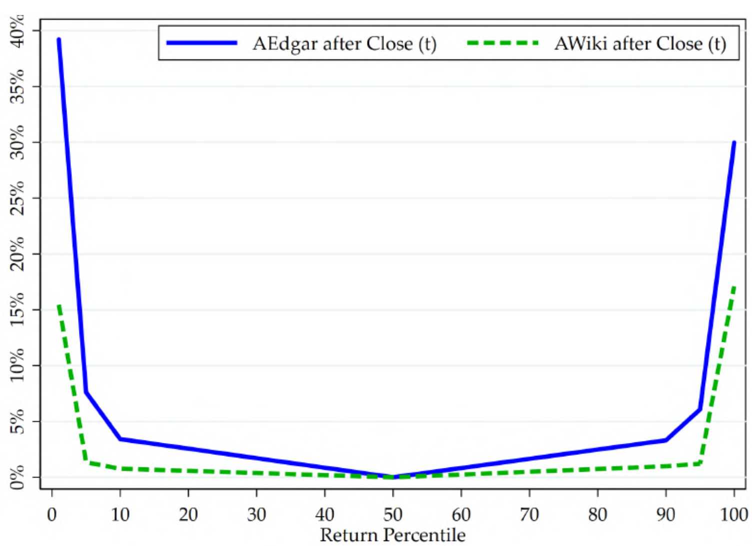
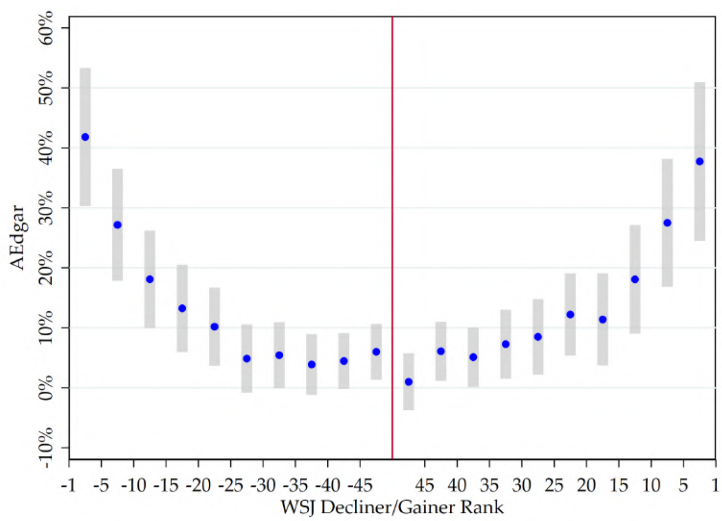


Figure A6: WSJ Ranks and Information Acquisition

This figure shows the regression coefficients of regressions of abnormal information acquisition after close (AEdgar in Panel A, AWiki in Panel B) on top and bottom 50 WSJ Gainer/Decliner composite ranking indicators (by groups of 5 ranks). The regressions include fixed effects for the top and bottom 200 daily ranks of the CRSP universe in our sample (400 effects), for each percentile rank in the daily cross-section of returns of the CRSP universe in our sample (100 effects) and for each whole sample percentile rank of the CRSP universe in our sample (100 effects) as well as stock and day fixed effects and five lags of the dependent variable. Positive rank labels indicate gainers, while negative rank labels indicate decliners. The blue dot marks the point estimate, while the grey bars represent the 95% confidence interval. Standard errors are clustered by stock and day.

Panel A: AEdgar



Panel B: AWiki

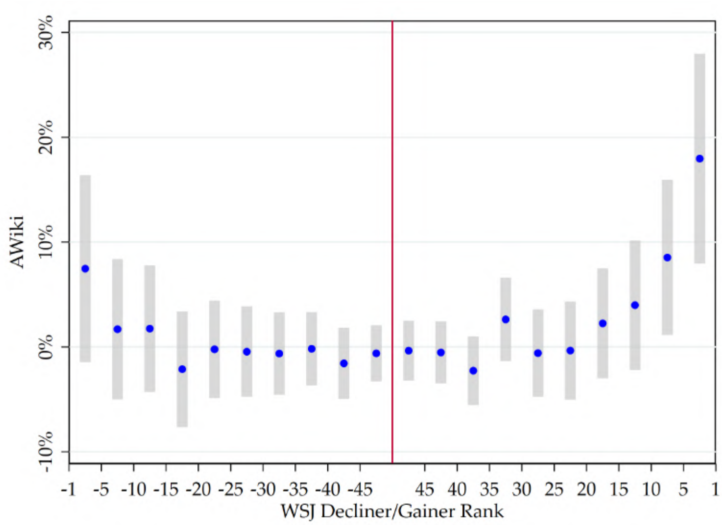
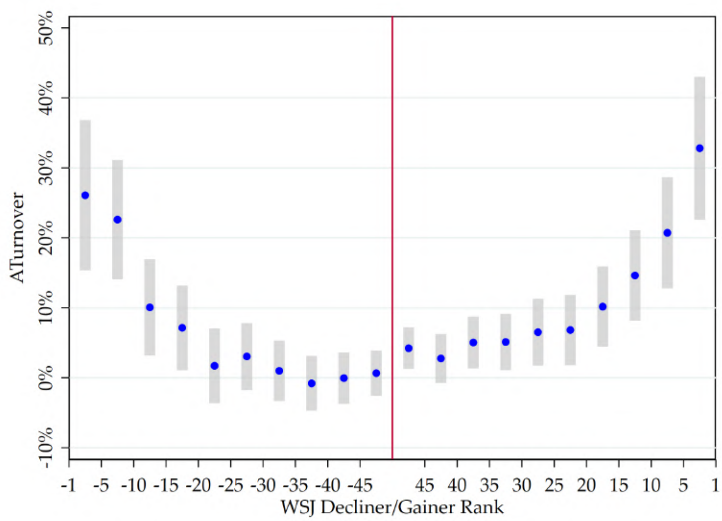


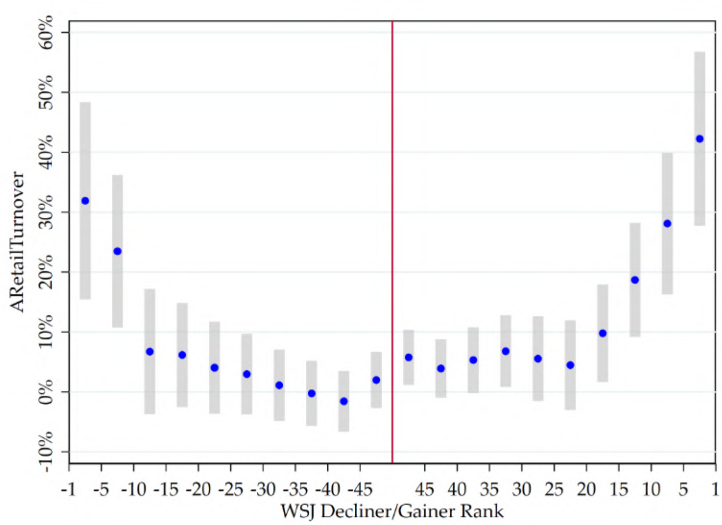
Figure A7: WSJ Ranks and Turnover

This figure shows the regression coefficients of regressions of abnormal turnover on the next trading day (Total turnover in Panel A, retail turnover in Panel B) on top and bottom 50 WSJ Gainer/Decliner composite ranking indicators (by groups of 5 ranks). The regressions include fixed effects for the top and bottom 200 daily ranks of the CRSP universe in our sample (400 effects), for each percentile rank in the daily cross-section of returns of the CRSP universe in our sample (100 effects) and for each whole sample percentile rank of the CRSP universe in our sample (100 effects) as well as stock and day fixed effects and five lags of the dependent variable. Positive rank labels indicate gainers, while negative rank labels indicate decliners. The blue dot marks the point estimate, while the grey bars represent the 95% confidence interval. Standard errors are clustered by stock and day.

Panel A: Total Turnover



Panel B: Retail Turnover



B Additional Tables

Table B1: Correlations

This table provides Pearson correlations for the variables included in the full sample. All variables are defined in Internet Appendix D.

	AEdgar after Close	ATurnover	ABS($Close_{t-1}$ to $Close_t$ Return-Median)	ABS($Open_t$ to $Close_t$ Return-Median)	ABS($Close_{t-1}$ to $Open_t$ Return-Median)
AEdgar after Close	1.00				
ATurnover	0.09	1.00			
ABS($Close_{t-1}$ to $Close_t$ Return-Median)	0.05	0.27	1.00		
ABS($Open_t$ to $Close_t$ Return-Median)	0.03	0.20	0.73	1.00	
ABS($Close_{t-1}$ to $Open_t$ Return-Median)	0.05	0.24	0.67	0.30	1.00

	AEdgar after Close	ATurnover	8-K Filing	10-K or 10-Q Filing	Earnings Announcement	Dividend Announcement	Media Coverage
AEdgar after Close	1.00						
ATurnover	0.09	1.00					
8-K Filing	0.12	0.09	1.00				
10-K or 10-Q Filing	0.10	0.04	0.13	1.00			
Earnings Announcement	0.06	0.18	0.29	0.21	1.00		
Dividend Announcement	0.02	0.02	0.09	0.02	0.06	1.00	
Media Coverage	0.01	0.02	0.02	0.01	0.02	0.00	1.00

	$Close_{t-1}$ to $Close_t$ Return	$Open_t$ to $Close_t$ Return	$Close_{t-1}$ to $Open_t$ Return
$Close_{t-1}$ to $Close_t$ Return	1.00		
$Open_t$ to $Close_t$ Return	0.79	1.00	
$Close_{t-1}$ to $Open_t$ Return	0.64	0.04	1.00

	$Close_{t-1}$ to $Close_t$ Return-Median	$Open_t$ to $Close_t$ Return-Median	$Close_{t-1}$ to $Open_t$ Return-Median
$Close_{t-1}$ to $Close_t$ Return-Median	1.00		
$Open_t$ to $Close_t$ Return-Median	0.77	1.00	
$Close_{t-1}$ to $Open_t$ Return-Median	0.65	0.01	1.00

Table B2: Information Acquisition and Returns

This table presents the results of OLS regressions of abnormal information acquisition (measured from 4:00 PM to 9:00 AM on the following days; AEdgar in Panel A, AWiki in Panel B) on absolute returns calculated for different time periods. Close to close returns are measured from 4:00 PM yesterday to 4:00 PM today, while alternative returns are calculated using an alternative start or end hour. All returns are centered by the daily cross-sectional median. All variables are defined in Internet Appendix D. Standard errors are two-way clustered by stock and day. ***, ** and * represent significance at the 1%-, 5%- and 10%-level, respectively.

Panel A: AEdgar after Close_t

		X=Open		X=12pm		X=3pm	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\lvert\text{Close}_{t-1}$ to Close_t Ret$\rvert$	2.445***	2.564***	1.557***	2.330***	1.500***	2.347***	1.503***
	(54.98)	(50.20)	(38.78)	(55.04)	(30.95)	(55.70)	(21.24)
$\lvert\text{X}_t$ to Close_t Ret$\rvert$		-0.219***		0.560***		1.279***	
		(-4.00)		(8.28)		(10.27)	
$\lvert\text{Close}_{t-1}$ to X_t Ret$\rvert$			2.052***		1.361***		1.060***
			(31.32)		(26.11)		(15.25)
5 Lags of Dependent	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock and Day FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4948540	4905375	4905375	4929088	4929088	4934488	4934488
Adjusted R2	0.217	0.217	0.218	0.217	0.218	0.217	0.217

	X=Open_t				X=12pm_{t-1}	
	(8)	(9)	(10)	(11)	(12)	(13)
$\lvert\text{Close}_{t-1}$ to Close_t Ret$\rvert$				1.143***	1.942***	1.831***
				(17.62)	(31.49)	(32.69)
$\lvert$X to Close_t Ret$\rvert$	2.237***		1.570***	0.648***	0.569***	0.579***
	(40.46)		(30.40)	(8.54)	(12.23)	(13.78)
$\lvert\text{Close}_{t-1}$ to X_t Ret$\rvert$		3.410***	3.072***	2.271***		
		(50.03)	(46.56)	(29.66)		
5 Lags of Dependent	Yes	Yes	Yes	Yes	No	Yes
Stock and Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Lagged Close to Close	No	No	No	No	Yes	Yes
Observations	4962739	4909370	4905375	4905375	4882444	4844585
Adjusted R2	0.214	0.217	0.218	0.218	0.053	0.220

Panel B: AWiki after Close_t

		X=Open		X=12pm		X=3pm	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\lvert \text{Close}_{t-1}$ to Close_t Ret$\rvert$	1.078***	1.330***	0.488***	1.095***	0.491***	1.071***	0.466***
	(28.50)	(30.97)	(16.52)	(29.19)	(15.01)	(29.10)	(10.71)
$\lvert X_t$ to Close_t Ret$\rvert$		-0.525***		-0.071*		0.088	
		(-13.49)		(-1.75)		(1.22)	
$\lvert \text{Close}_{t-1}$ to X_t Ret$\rvert$			1.286***		0.822***		0.677***
			(24.80)		(21.65)		(14.78)
5 Lags of Dependent	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock and Day FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3811771	3789130	3789130	3801869	3801869	3804775	3804775
Adjusted R2	0.465	0.467	0.467	0.466	0.466	0.466	0.466

	$X=\text{Open}_t$				$X=\text{12pm}_{t-1}$	
	(8)	(9)	(10)	(11)	(12)	(13)
$\lvert \text{Close}_{t-1}$ to Close_t Ret$\rvert$				0.465***	0.909***	0.827***
				(12.28)	(21.21)	(21.38)
$\lvert X$ to Close_t Ret$\rvert$	0.778***		0.415***	0.035	0.233***	0.244***
	(18.52)		(11.95)	(0.79)	(8.11)	(9.67)
$\lvert \text{Close}_{t-1}$ to X_t Ret$\rvert$		1.714***	1.639***	1.299***		
		(30.92)	(30.74)	(22.99)		
5 Lags of Dependent	Yes	Yes	Yes	Yes	No	Yes
Stock and Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Lagged Close to Close	No	No	No	No	Yes	Yes
Observations	3827930	3789959	3789130	3789130	3757411	3744378
Adjusted R2	0.463	0.467	0.467	0.467	0.331	0.470

Table B3: Tangible versus Non-tangible News Days

This table presents the results of OLS regressions of abnormal information acquisition (measured from 4:00 PM to 9:00 AM on the following days; AEdgar in Panel A, AWiki in Panel B) on absolute returns calculated for different time periods. Close to close returns are measured from 4:00 PM yesterday to 4:00 PM today, while alternative returns are calculated using an 9:30 AM (Open) as start or end hour. Columns one to three include only observations with companies that announced earnings, for which a form 8-K, 10-K or 10-Q was filed, who declared a dividend or who received media coverage in a national newspaper on day t or $t-1$. Columns four to six include all other days. All returns are centered by the daily cross-sectional median. All variables are defined in Internet Appendix D. Standard errors are two-way clustered by stock and day. ***, ** and * represent significance at the 1%-, 5%- and 10%-level, respectively.

Panel A: AEdgar after Close$_t$

	Days with Tangible News			Days without Tangible News		
	(1)	(2)	(3)	(4)	(5)	(6)
$\lvert$Close$_{t-1}$ to Close$_t$ Ret$\rvert$	1.995***	1.793***	1.472***	1.206***	1.170***	1.064***
	(38.80)	(31.51)	(22.97)	(32.80)	(23.66)	(27.20)
$\lvert$Open$_t$ to Close$_t$ Ret$\rvert$		0.622***			0.073	
		(7.10)			(1.29)	
$\lvert$Close$_{t-1}$ to Open$_t$ Ret$\rvert$			0.801***			0.590***
			(9.47)			(9.41)
5 Lags of Dependent	Yes	Yes	Yes	Yes	Yes	Yes
Stock and Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	706320	701755	701755	4242085	4203446	4203446
Adjusted R2	0.252	0.252	0.252	0.217	0.217	0.217

Panel B: AWiki after Close$_t$

	Days with Tangible News			Days without Tangible News		
	(1)	(2)	(3)	(4)	(5)	(6)
$\lvert$Close$_{t-1}$ to Close$_t$ Ret$\rvert$	1.412***	1.450***	0.674***	0.646***	0.892***	0.355***
	(29.89)	(27.72)	(14.45)	(17.53)	(18.04)	(11.77)
$\lvert$Open$_t$ to Close$_t$ Ret$\rvert$		-0.120			-0.416***	
		(-1.64)			(-9.46)	
$\lvert$Close$_{t-1}$ to Open$_t$ Ret$\rvert$			1.115***			0.956***
			(16.17)			(15.91)
5 Lags of Dependent	Yes	Yes	Yes	Yes	Yes	Yes
Stock and Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	528352	526017	526017	3283359	3263049	3263049
Adjusted R2	0.474	0.475	0.476	0.467	0.468	0.468

Table B4: Robustness - Excluding Recent Filings (Day t and $t-1$)

This table provides the results of difference-in-difference tests of abnormal EDGAR downloads for overnight earnings announcements around the opening of the exchanges at 9:30 AM. In the construction of the AEdgar variable, we exclude downloads for filings made on the earnings announcement day or the previous day. Large Surprise is an indicator that takes on a value of one if SUE is smaller than the 30th or larger than the 70th percentile based on earnings announcements in the past 180 days. For each announcement, we measure in information acquisition for two windows: 7:30 AM to 9:30 AM and 9:30 AM to 11:30 AM. Post-Open is an indicator that take on a value of one for the latter window. Predicted return is the average announcement day return of companies with similarly surprising announcements over the previous 180 days. All variables are defined in Internet Appendix D. Standard errors are two-way clustered by stock and month. ***, ** and * represent significance at the 1%-, 5%- and 10%-level, respectively.

	(1) AEdgar	(2) AEdgar	(3) AEdgar	(4) AEdgar
Large Surprise × Post-Open	0.064*** (3.46)	0.064*** (3.46)		
Large Surprise	-0.040** (-2.04)			
|Predicted Return|× Post-Open			2.528*** (4.00)	2.528*** (4.00)
|Predicted Return|			-2.092*** (-2.88)	
Post-Open	-0.300*** (-17.11)	-0.300*** (-17.11)	-0.326*** (-17.65)	-0.326*** (-17.65)
Constant	0.732*** (27.73)		0.770*** (27.03)	
Stock x Announcement FE	No	Yes	No	Yes
Observations	45698	45698	43452	43452
Adjusted R2	0.015	0.236	0.016	0.238

Table B5: Robustness - Excluding Daily Winners and Losers

This table provides the results of difference-in-difference tests of abnormal information acquisition (AEdgar in Panel A, AWiki in Panel B) for overnight earnings announcements around the opening of the exchanges at 9:30 AM as in table 2. We exclude companies which rank in the 1st or 99th percentile of returns (in the whole cross-section of traded stocks) on the earnings announcement day. All variables are defined in Internet Appendix D. Standard errors are two-way clustered by stock and month. ***, ** and * represent significance at the 1%-, 5%- and 10%-level, respectively.

Panel A: AEdgar

	(1) AEdgar	(2) AEdgar	(3) AEdgar	(4) AEdgar
Large Surprise × Post-Open	0.083*** (4.32)	0.083*** (4.32)		
Large Surprise	-0.036 (-1.44)			
|Predicted Return|× Post-Open			3.445*** (4.22)	3.445*** (4.22)
|Predicted Return|			-2.119** (-2.09)	
Post-Open	-0.391*** (-19.35)	-0.391*** (-19.35)	-0.426*** (-18.82)	-0.426*** (-18.82)
Constant	1.290*** (36.16)		1.338*** (34.92)	
Stock x Announcement FE	No	Yes	No	Yes
Observations	40324	40324	38334	38334
Adjusted R2	0.024	0.347	0.025	0.348

Panel B: AWiki

	(1) AWiki	(2) AWiki	(3) AWiki	(4) AWiki
Large Surprise × Post-Open	0.022	0.022		
	(1.63)	(1.63)		
Large Surprise	-0.026**			
	(-2.03)			
|Predicted Return|× Post-Open			0.797*	0.797*
			(1.84)	(1.84)
|Predicted Return|			-1.213**	
			(-2.45)	
Post-Open	-0.025**	-0.025**	-0.028**	-0.028**
	(-2.35)	(-2.35)	(-2.28)	(-2.28)
Constant	0.173***		0.182***	
	(11.84)		(11.19)	
Stock x Announcement FE	No	Yes	No	Yes
Observations	30512	30512	30512	30512
Adjusted R2	0.000	0.256	0.000	0.256

Table B6: Overnight Earnings Announcements and Information Acquisition of Robots

This table provides the results of difference-in-difference tests of abnormal EDGAR downloads by robots/algorithms for overnight earnings announcements around the opening of the exchanges at 9:30 AM. Large Surprise is an indicator that takes on a value of one if SUE is smaller than the 30th or larger than the 70th percentile based on earnings announcements in the past 180 days. For each announcement, we measure in information acquisition for two windows: 7:30 AM to 9:30 AM and 9:30 AM to 11:30 AM. Post-Open is an indicator that take on a value of one for the latter window. All variables are defined in Internet Appendix D. Standard errors are two-way clustered by stock and month. ***, ** and * represent significance at the 1%-, 5%- and 10%-level, respectively.

	(1) AEdgar (Robot)	(2) AEdgar (Robot)	(3) AEdgar (Robot)	(4) AEdgar (Robot)
Large Surprise × Post-Open	0.012 (0.46)	0.012 (0.46)		
Large Surprise	0.067** (2.36)			
|Predicted Return|× Post-Open			0.521 (0.52)	0.521 (0.52)
|Predicted Return|			2.529* (1.86)	
Post-Open	-0.311*** (-11.17)	-0.311*** (-11.17)	-0.320*** (-10.00)	-0.320*** (-10.00)
Constant	1.151*** (25.90)		1.160*** (22.36)	
Stock x Announcement FE	No	Yes	No	Yes
Observations	45698	45698	43452	43452
Adjusted R2	0.012	0.430	0.013	0.430

Table B7: Median Splits: Cross-Section

This table provides the results of difference-in-difference tests of abnormal information acquisition (AEdgar in Panel A, AWiki in Panel B) for overnight earnings announcements around the opening of the exchanges at 9:30 AM. We split the sample along the median of several variables in two sub-samples: one-month lagged market capitalization, residual analyst coverage, residual EDGAR downloads and residual institutional ownership. Residual variables are obtained from a cross-sectional regressions of the base variable on log market cap in the previous month. All variables are defined in Internet Appendix D. Standard errors are two-way clustered by stock and month. ***, ** and * represent significance at the 1%-, 5%- and 10%-level, respectively.

Panel A: AEdgar

	Market Cap		Res. Analyst Coverage	
	(1) Low	(2) High	(3) Low	(4) High
Large Surprise	-0.004	0.043	-0.020	-0.019
	(-0.13)	(1.34)	(-0.75)	(-0.56)
Post-Open	-0.289***	-0.451***	-0.363***	-0.416***
	(-11.75)	(-17.67)	(-16.17)	(-16.38)
Large Surprise × Post-Open	0.074***	0.043*	0.112***	0.076***
	(3.25)	(1.79)	(4.60)	(2.84)
Constant	1.202***	1.374***	1.249***	1.367***
	(27.40)	(34.55)	(33.94)	(32.14)
Observations	22786	22912	22786	22912
Adjusted R2	0.011	0.038	0.018	0.027

	Res. Downloads		Res. Institutional Ownership	
	(5) Low	(6) High	(7) Low	(8) High
Large Surprise	-0.072**	0.007	-0.008	-0.025
	(-2.32)	(0.23)	(-0.27)	(-0.78)
Post-Open	-0.361***	-0.425***	-0.351***	-0.426***
	(-15.53)	(-17.53)	(-13.84)	(-16.67)
Large Surprise × Post-Open	0.153***	0.053**	0.077***	0.109***
	(6.37)	(2.09)	(2.82)	(4.29)
Constant	1.253***	1.376***	1.208***	1.405***
	(29.82)	(35.65)	(31.29)	(32.69)
Observations	22786	22912	22786	22912
Adjusted R2	0.016	0.030	0.019	0.026

Panel B: AWiki

	Market Cap		Res. Analyst Coverage	
	(1) Low	(2) High	(3) Low	(4) High
Large Surprise	0.003	-0.003	-0.027*	-0.020
	(0.19)	(-0.15)	(-1.71)	(-1.08)
Post-Open	0.010	-0.039***	-0.018	-0.025*
	(0.53)	(-3.35)	(-1.22)	(-1.93)
Large Surprise × Post-Open	0.027	0.022	0.029	0.039**
	(1.29)	(1.12)	(1.57)	(2.11)
Constant	0.098***	0.221***	0.159***	0.197***
	(6.76)	(11.74)	(8.92)	(12.73)
Observations	15906	18602	17086	17422
Adjusted R2	0.000	0.000	0.000	0.000

	Res. Downloads		Res. Institutional Ownership	
	(5) Low	(6) High	(7) Low	(8) High
Large Surprise	-0.043**	-0.010	-0.033*	-0.014
	(-2.51)	(-0.61)	(-1.88)	(-0.80)
Post-Open	-0.012	-0.032**	-0.018	-0.025*
	(-0.79)	(-2.23)	(-1.20)	(-1.80)
Large Surprise × Post-Open	0.026	0.044***	0.033*	0.035*
	(1.19)	(2.76)	(1.78)	(1.96)
Constant	0.169***	0.188***	0.184***	0.172***
	(11.11)	(10.25)	(11.11)	(10.46)
Observations	16878	17630	16956	17552
Adjusted R2	0.000	0.000	0.000	-0.000

Table B8: Median Splits: Time-Series

This table provides the results of difference-in-difference tests of abnormal information acquisition (AEdgar in Panel A, AWiki in Panel B) for overnight earnings announcements around the opening of the exchanges at 9:30 AM. We split the sample along the median of several variables in two sub-samples: VIX and previous month market return. All variables are defined in Internet Appendix D. Standard errors are two-way clustered by stock and month. ***, ** and * represent significance at the 1%-, 5%- and 10%-level, respectively.

Panel A: AEdgar

	VIX		Market Return (Previous Month)	
	(1) Low	(2) High	(3) Low	(4) High
Large Surprise	0.013	-0.047	-0.029	-0.007
	(0.44)	(-1.50)	(-1.00)	(-0.23)
Post-Open	-0.400***	-0.379***	-0.403***	-0.376***
	(-15.99)	(-14.10)	(-14.90)	(-14.10)
Large Surprise × Post-Open	0.059***	0.126***	0.091***	0.096***
	(2.67)	(4.62)	(3.15)	(4.03)
Constant	1.340***	1.275***	1.340***	1.276***
	(27.96)	(28.21)	(32.92)	(23.67)
Observations	22810	22888	22514	23184
Adjusted R2	0.027	0.018	0.024	0.020

Panel B: AWiki

	VIX		Market Return (Previous Month)	
	(1) Low	(2) High	(3) Low	(4) High
Large Surprise	-0.008	-0.035**	-0.048**	0.004
	(-0.46)	(-2.34)	(-2.67)	(0.23)
Post-Open	-0.016	-0.027*	-0.031**	-0.011
	(-1.05)	(-1.77)	(-2.13)	(-0.77)
Large Surprise × Post-Open	0.026	0.042**	0.049***	0.018
	(1.38)	(2.32)	(2.78)	(1.04)
Constant	0.189***	0.168***	0.195***	0.160***
	(10.11)	(8.71)	(9.83)	(9.14)
Observations	16582	17926	17798	16710
Adjusted R2	-0.000	0.000	0.000	-0.000

Table B9: Spline Regressions of Information Acquisition on Returns

This table shows the results of linear spline regressions of abnormal information acquisition (AEdgar and AWiki) on median adjusted close-to-close returns. The corresponding regression results are visualized in figure A5. In columns 1 and 2, we use spline terms based on continuous returns with knots at every integer percentage in the interval [-4,4]. In columns 3 and 4, we use spline terms based on cross-sectional return percentile ranks with knots at the 5th, 10th, 50th, 90th and 95th percentile. The regressions include stock and day fixed effects as well as five lags of the dependent variable. All variables are defined in Internet Appendix D. Standard errors are two-way clustered by stock and month. ***, ** and * represent significance at the 1%-, 5%- and 10%-level, respectively.

	(1) AEdgar	(2) AWiki	(3) AEdgar	(4) AWiki
Return Percentile Ranks:				
Ranks 96-100	4.778***	3.179***		
	(35.05)	(29.44)		
Ranks 91-95	0.555***	0.043		
	(8.01)	(1.01)		
Ranks 51-90	0.083***	0.025***		
	(18.65)	(7.18)		
Ranks 11-50	-0.085***	-0.019***		
	(-18.53)	(-6.01)		
Ranks 6-10	-0.837***	-0.114**		
	(-11.01)	(-2.47)		
Ranks 1-5	-7.900***	-3.529***		
	(-39.62)	(-24.18)		
Return Continuous:				
Return>4%			2.643***	1.735***
			(23.76)	(19.23)
3%<Return<4%			2.906***	0.529**
			(6.70)	(2.00)
2%<Return<3%			2.063***	0.677***
			(5.48)	(3.08)
1%<Return<2%			1.790***	0.800***
			(7.09)	(5.69)
0%<Return<1%			1.170***	0.302**
			(6.40)	(2.45)
-1%<Return<0%			-0.878***	-0.104
			(-4.87)	(-0.90)
-2%<Return<-1%			-1.748***	-0.702***
			(-6.76)	(-4.92)
-3%<Return<-2%			-2.603***	-0.790***
			(-6.84)	(-3.54)
-4%<Return<-3%			-3.921***	-0.480*
			(-8.92)	(-1.70)
Return<-4%			-2.632***	-1.162***
			(-30.94)	(-18.04)
5 Lags of Dependent	Yes	Yes	Yes	Yes
Stock and Day FE	Yes	Yes	Yes	Yes
Observations	4948540	3811771	4948540	3811771
Adjusted R2	0.217	0.465	0.217	0.466

Table B10: WSJ Ranks and Returns

This tables shows the results of OLS regressions of close-to-close returns on dummies according to the stock being ranked among the top or bottom 15 stocks in the Wall Street Journal Composite Gainer/Decliner ranking. The regressions include fixed effects for the top and bottom 200 daily ranks of the CRSP universe in our sample (400 effects), for each percentile rank in the daily cross-section of returns of the CRSP universe in our sample (100 effects) and for each whole sample percentile rank of the CRSP universe in our sample (100 effects) as indicated below the coefficients. All variables are defined in Internet Appendix D. Standard errors are two-way clustered by stock and month. ***, ** and * represent significance at the 1%-, 5%- and 10%-level, respectively.

	(1)	(2)
WSJ Gainer 1-5	0.284***	-0.008
	(55.42)	(-0.39)
WSJ Gainer 6-10	0.161***	-0.007
	(148.10)	(-0.81)
WSJ Gainer 11-15	0.131***	-0.005
	(156.57)	(-1.36)
WSJ Decliner 11-15	-0.113***	0.002*
	(-111.22)	(1.72)
WSJ Decliner 6-10	-0.139***	0.002
	(-101.28)	(0.72)
WSJ Decliner 1-5	-0.225***	-0.006
	(-81.18)	(-1.17)
CRSP Top/Bottom 200 FE	No	Yes
CRSP Return Percentile FE	No	Yes
CRSP Return Group FE	No	Yes
5 Lags of Dependent	Yes	Yes
Stock and Day FE	Yes	Yes
Observations	3494287	3494287
Adjusted R2	0.414	0.953

Table B11: WSJ Ranks and Bloomberg/Google Information Acquisition

This tables shows the results of OLS regressions of abnormal information acquisition (measured using Bloomberg AIAC and abnormal Google Search Volume AGoogle) on dummies according to the stock being ranked among the top or bottom 15 stocks in the Wall Street Journal Composite Gainer/Decliner ranking. The regressions include fixed effects for the top and bottom 200 daily ranks of the CRSP universe in our sample (400 effects), for each percentile rank in the daily cross-section of returns of the CRSP universe in our sample (100 effects) and for each whole sample percentile rank of the CRSP universe in our sample (100 effects) as indicated below the coefficients. All variables are defined in Internet Appendix D. Standard errors are two-way clustered by stock and month. ***, ** and * represent significance at the 1%-, 5%- and 10%-level, respectively.

	AIAC (t+1)		AGoogle (t+1)	
	(1)	(2)	(3)	(4)
WSJ Gainer 1-5	0.622***	0.228***	0.569***	0.166**
	(15.28)	(2.64)	(17.10)	(2.32)
WSJ Gainer 6-10	0.526***	0.135**	0.394***	0.191***
	(14.91)	(2.18)	(13.46)	(3.77)
WSJ Gainer 11-15	0.461***	0.075	0.293***	0.136***
	(12.90)	(1.50)	(10.01)	(3.57)
WSJ Decliner 11-15	0.554***	0.090*	0.191***	0.046
	(17.00)	(1.86)	(7.08)	(1.29)
WSJ Decliner 6-10	0.551***	0.081	0.268***	0.090*
	(17.73)	(1.35)	(9.62)	(1.89)
WSJ Decliner 1-5	0.645***	0.132	0.464***	0.126*
	(21.01)	(1.64)	(15.97)	(1.78)
CRSP Top/Bottom 200 FE	No	Yes	No	Yes
CRSP Return Percentile FE	No	Yes	No	Yes
CRSP Return Group FE	No	Yes	No	Yes
Stock and Day FE	Yes	Yes	Yes	Yes
Observations	2022578	2022508	3336513	3334823
Adjusted R2	0.034	0.037	0.002	0.002

Table B12: WSJ Ranks, Firm Size and Information Acquisition

This tables shows the results of OLS regressions of abnormal information acquisition (measured using Edgar downloads and Wikipedia) on dummies according to the stock being ranked among the top or bottom 15 stocks in the Wall Street Journal Composite Gainer/Decliner ranking. We split the entire sample in two samples based on the stock markets capitalization one month ago. The regressions include fixed effects for the top and bottom 200 daily ranks of the CRSP universe in our sample (400 effects), for each percentile rank in the daily cross-section of returns of the CRSP universe in our sample (100 effects) and for each whole sample percentile rank of the CRSP universe in our sample (100 effects) as indicated below the coefficients. All variables are defined in Internet Appendix D. Standard errors are two-way clustered by stock and month. ***, ** and * represent significance at the 1%-, 5%- and 10%-level, respectively.

	AEdgar		AWiki	
	(1) Large	(2) Small	(3) Large	(4) Small
WSJ Gainer 1-5	0.250**	0.230***	0.290***	0.069
	(2.07)	(3.27)	(3.25)	(1.23)
WSJ Gainer 6-10	0.149*	0.151***	0.098	0.013
	(1.88)	(3.11)	(1.53)	(0.34)
WSJ Gainer 11-15	0.113*	0.047	0.015	0.014
	(1.96)	(1.28)	(0.37)	(0.48)
WSJ Decliner 11-15	0.081	0.075**	0.017	0.021
	(1.63)	(2.04)	(0.52)	(0.75)
WSJ Decliner 6-10	0.133**	0.167***	0.050	-0.009
	(2.04)	(3.52)	(1.15)	(-0.25)
WSJ Decliner 1-5	0.178**	0.302***	0.010	0.076
	(1.98)	(4.47)	(0.15)	(1.44)
CRSP Top/Bottom 200 FE	Yes	Yes	Yes	Yes
CRSP Return Percentile FE	Yes	Yes	Yes	Yes
CRSP Return Group FE	Yes	Yes	Yes	Yes
Stock and Day FE	Yes	Yes	Yes	Yes
5 Lags of Dependent	Yes	Yes	Yes	Yes
Observations	1634330	1625104	1506329	1242363
Adjusted R2	0.246	0.218	0.509	0.323

Table B13: WSJ Ranks, Firm Size and Trading

This tables shows the results of OLS regressions of abnormal turnover and abnormal retail turnover on dummies according to the stock being ranked among the top or bottom 15 stocks in the Wall Street Journal Composite Gainer/Decliner ranking. We split the entire sample in two samples based on the stock markets capitalization one month ago. The regressions include fixed effects for the top and bottom 200 daily ranks of the CRSP universe in our sample (400 effects), for each percentile rank in the daily cross-section of returns of the CRSP universe in our sample (100 effects) and for each whole sample percentile rank of the CRSP universe in our sample (100 effects) as indicated below the coefficients. All variables are defined in Internet Appendix D. Standard errors are two-way clustered by stock and month. ***, ** and * represent significance at the 1%-, 5%- and 10%-level, respectively.

	ATurnover (t+1)		ARetailTurnover (t+1)	
	(1) Large	(2) Small	(3) Large	(4) Small
WSJ Gainer 1-5	0.324***	0.254***	0.363***	0.355***
	(4.09)	(4.70)	(2.97)	(4.59)
WSJ Gainer 6-10	0.200***	0.114***	0.146	0.232***
	(3.37)	(3.19)	(1.40)	(4.17)
WSJ Gainer 11-15	0.087**	0.063**	0.126**	0.103***
	(2.37)	(2.35)	(1.97)	(2.62)
WSJ Decliner 11-15	0.033	0.065**	-0.037	0.033
	(1.01)	(2.31)	(-0.72)	(0.82)
WSJ Decliner 6-10	0.096**	0.208***	0.069	0.203***
	(2.01)	(5.06)	(1.01)	(3.51)
WSJ Decliner 1-5	0.176***	0.207***	0.128	0.269***
	(2.88)	(3.57)	(1.23)	(3.23)
CRSP Top/Bottom 200 FE	Yes	Yes	Yes	Yes
CRSP Return Percentile FE	Yes	Yes	Yes	Yes
CRSP Return Group FE	Yes	Yes	Yes	Yes
Stock and Day FE	Yes	Yes	Yes	Yes
5 Lags of Dependent	Yes	Yes	Yes	Yes
Observations	1723960	1728917	1380751	1333851
Adjusted R2	0.373	0.315	0.290	0.263

Table B14: Company Events and Information Acquisition

This tables shows the results of OLS regressions of abnormal information acquisition (AEdgar in Panel A, AWiki in Panel B) on dummies for several corporate events. All variables are defined in Internet Appendix D. Standard errors are two-way clustered by stock and month. ***, ** and * represent significance at the 1%-, 5%- and 10%-level, respectively.

Panel A: AEdgar

	(1) AEdgar	(2) AEdgar	(3) AEdgar	(4) AEdgar	(5) AEdgar	(6) AEdgar
8-K Filing	0.561***					0.496***
	(116.36)					(111.36)
10-K or 10-Q Filing		0.822***				0.701***
		(71.95)				(67.14)
Earnings Announcement			0.493***			0.098***
			(73.10)			(17.23)
Dividend Announcement				0.217***		0.086***
				(24.55)		(13.32)
Media Coverage					0.104***	0.080***
					(15.03)	(13.23)
5 Lags of Dependent	Yes	Yes	Yes	Yes	Yes	Yes
Stock and Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5035665	5035665	5035665	5035665	5035665	5035665
Adjusted R2	0.227	0.223	0.216	0.213	0.213	0.235

Panel B: AWiki

	(1) AWiki	(2) AWiki	(3) AWiki	(4) AWiki	(5) AWiki	(6) AWiki
8-K Filing	0.077***					0.048***
	(33.60)					(29.05)
10-K or 10-Q Filing		0.053***				0.010***
		(13.95)				(3.92)
Earnings Announcement			0.182***			0.147***
			(28.70)			(25.86)
Dividend Announcement				0.047***		0.028***
				(11.44)		(8.40)
Media Coverage					0.062***	0.042***
					(10.90)	(8.50)
5 Lags of Dependent	Yes	Yes	Yes	Yes	Yes	Yes
Stock and Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3686063	3686063	3686063	3686063	3686063	3868785
Adjusted R2	0.323	0.322	0.323	0.322	0.322	0.463

Table B15: Additional Factor Models

In this table, we report alphas of the difference in the long-short portfolio build on the SYY score for high and low PoP. PoP is the return salience measure based on daily returns, see the main text for details. SYY score is the mispricing score of Stambaugh et al. (2012). For definitions of other variables, see Appendix D. Alphas are estimated as the intercept estimates of time-series regression of long-short portfolio on factor models. The sample covers all ≥ $5 U.S. common stocks traded on the NYSE, AMEX and NASDAQ with market cap higher than the 2nd NYSE decile from 1963 to 2015. t-statistics are based on Newey-West standard errors with one lag and are reported in parentheses.

Factor Model	Value-Weighted	Equal-Weighted	N Months	Available
none	-1.43%	-1.25%	605	8/1965 - 12/2015
	(-5.57)	(-6.34)		
CAPM 1F	-1.67%	-1.45%	605	8/1965 - 12/2015
	(-7.04)	(-8)		
Fama-French 3F	-1.58%	-1.38%	605	8/1965 - 12/2015
	(-6.69)	(-8.15)		
Carhart 4F	-1.25%	-1.07%	605	8/1965 - 12/2015
	(-5.45)	(-6.28)		
Fama-French 5F	-1.39%	-1.14%	605	8/1965 - 12/2015
	(-5.71)	(-6.64)		
Q-model	-1.37%	-1.02%	504	7/1972 - 12/2013
	(-4.97)	(-5.11)		
Carhart 4F + Short-/Long-term Reversal	-1.28%	-1.05%	605	8/1965 - 12/2015
	(-5.33)	(-5.9)		
Carhart 4F + Betting-Against-Beta	-1.24%	-0.94%	605	8/1965 - 12/2015
	(-5.15)	(-5.48)		
Carhart 4F + IVol Q5-Q1 factor	-1.03%	-0.86%	605	8/1965 - 12/2015
	(-4.51)	(-5.22)		
Carhart 4F + MAX Q5-Q1 factor	-1.15%	-0.94%	605	8/1965 - 12/2015
	(-4.99)	(-5.62)		
Carhart 4F + PS Liquidity factor	-1.24%	-1.05%	576	1/1968 - 12/2015
	(-5.16)	(-5.95)		
Carhart 4F + Sadka Liquidity factor	-1.65%	-1.35%	357	4/1983 - 12/2012
	(-5.15)	(-5.88)		
Carhart 4F + Stambaugh 2F	-0.96%	-0.76%	605	8/1965 - 12/2015
	(-3.94)	(-4.24)		
Carhart 4F + Stambaugh mgmt	-1.10%	-0.94%	605	8/1965 - 12/2015
	(-4.64)	(-5.16)		
Carhart 4F + Stambaugh perf	-1.12%	-0.89%	605	8/1965 - 12/2015
	(-4.74)	(-5.32)		
Carhart 4F + Undervalued-Minus-Overvalued	-1.18%	-0.92%	510	7/1972 - 12/2014
	(-4.45)	(-4.85)		
Carhart 4F + Quality-Minus-Junk	-0.93%	-0.66%	605	8/1965 - 12/2015
	(-4.07)	(-4.01)		
Carhart 4F + Novy-Marx Profitability	-0.87%	-0.66%	569	7/1963 - 12/2012
	(-3.51)	(-3.47)		

Table B16: Fama-MacBeth Regressions - Additional Specification and Robustness

In this table, we report results from Fama and MacBeth (1973) regressions of this month's return on stock characteristics available at the end of last month. PoP is the return salience measure based on daily returns, see the main text for details. SYY score is the mispricing score of Stambaugh et al. (2012). For definitions of other variables, see Appendix D. The sample covers all ≥ $5 U.S. common stocks traded on the NYSE, AMEX and NASDAQ with market cap higher than the 2nd NYSE decile from 1963 to 2015. t-statistics are based on Newey-West standard errors with one lag and are reported in parentheses. ***, **, and * indicate significance at the 1%-, 5%- and 10%-level, respectively.

Panel A

	(1) 1-month Gap	(2) Wiki based	(3) IVol based	(4) Close-to-close Return	(5) Close-to-open Return	(6) Open-to-close Return
SYY Score	-0.0022***	-0.0025***	-0.0025***	-0.0038***	0.0019***	-0.0058***
	(-8.39)	(-9.54)	(-9.72)	(-7.73)	(9.59)	(-10.93)
SYY Score × PoP	**-0.0010*** **	**-0.0014*** **	**-0.0017*** **	**-0.0028*** **	**0.0010*** **	**-0.0037*** **
	(-5.1)	**(-6.51)**	**(-7.61)**	**(-7.51)**	**(4.12)**	**(-9.16)**
PoP	-0.0018***	-0.0021***	-0.0018***	-0.0045***	0.005***	-0.0094***
	(-4.32)	(-5.55)	(-4.28)	(-6.19)	(10.86)	(-12.3)
Beta	0.0015	0.0019	0.0016	-0.0027	0.0044***	-0.0072***
	(1.13)	(1.36)	(1.19)	(-1.22)	(5.24)	(-3.65)
ln(Size)	-0.001***	-0.0012***	-0.0012***	-0.0001	0.0022***	-0.0024***
	(-3.02)	(-3.58)	(-3.54)	(-0.26)	(9.8)	(-4.92)
ln(BTM)	0.0013**	0.0016***	0.0016***	0.0017**	0.0009***	0.0008
	(2.16)	(2.66)	(2.66)	(2.11)	(3.44)	(1)
$ret_{t-12,t-2}$	0.0069***	0.0077***	0.0074***	0.0031	0.0126***	-0.0095***
	(4.26)	(4.62)	(4.47)	(1.27)	(15.07)	(-4.13)
$ret_{t-1,t-1}$	0.0068*	-0.0317***	-0.0352***	-0.0081	-0.0296***	0.0215***
	(1.86)	(-7.18)	(-8.12)	(-1.28)	(-8.38)	(3.5)
$ret_{t-36,t-13}$	-0.0002	0	-0.0001	0.0002	0.0023***	-0.0022***
	(-0.39)	(-0.07)	(-0.11)	(0.22)	(7.07)	(-2.86)
Months	604	605	605	283	283	283
Avg. N	1379	1418	1418	1648	1648	1648
Avg. R2	0.092	0.096	0.096	0.090	0.052	0.097
Years	1965-2015	1965-2015	1965-2015	1992-2015	1992-2015	1992-2015

Panel B

	(1) Price>1	(2) No Size Filter	(3) No NASDAQ	(4) DGTW adjusted	(5) Industry adjusted
SYY Score	-0.0025***	-0.0032***	-0.0022***	-0.0025***	-0.0021***
	(-9.51)	(-12.71)	(-7.99)	(-8.68)	(-8.02)
SYY Score × PoP	**-0.0016*** **	**-0.0018*** **	**-0.0012*** **	**-0.0017*** **	**-0.0016*** **
	(-7.16)	**(-9.39)**	**(-4.87)**	**(-6.78)**	**(-7.64)**
PoP	-0.0022***	-0.0027***	-0.002***	-0.0022***	-0.0021***
	(-5.26)	(-7.52)	(-4.88)	(-4.98)	(-5.66)
Beta	0.0022	0.0021	0.0019	0.0015	0.0017**
	(1.56)	(1.57)	(1.41)	(1.1)	(2.03)
ln(Size)	-0.0013***	-0.0012***	-0.0011***	-0.0007***	-0.0011***
	(-3.67)	(-3.93)	(-3.33)	(-4.26)	(-4.08)
ln(BTM)	0.0016**	0.0018***	0.0019***	0.0006	0.0017***
	(2.58)	(3.2)	(3.42)	(1.33)	(4.28)
$\text{ret}_{t-12,t-2}$	0.0077***	0.0085***	0.0071***	0.0025*	0.0063***
	(4.65)	(6)	(3.75)	(1.72)	(4.85)
$\text{ret}_{t-1,t-1}$	-0.0345***	-0.0402***	-0.034***	-0.0317***	-0.0436***
	(-7.97)	(-10.28)	(-7.5)	(-6.95)	(-11.53)
$\text{ret}_{t-36,t-13}$	0	-0.0002	0.0005	0.0001	0.0001
	(-0.05)	(-0.33)	(0.74)	(0.09)	(0.29)
Months	605	605	605	450	605
Avg. N	1421	2455	1028	1499	1401
Avg. R2	0.097	0.072	0.103	0.052	0.059
Years	1965-2015	1965-2015	1965-2015	1965-2015	1965-2015

Panel C

	(1) Low Size	(2) High Size	(3) Low Amihud	(4) High Amihud	(5) Low Corwin	(6) High Corwin
SYY Score	-0.0025***	-0.0023***	-0.0029***	-0.0022***	-0.0025***	-0.0025***
	(-8.28)	(-7.66)	(-9.66)	(-7.34)	(-7.46)	(-7.61)
SYY Score × PoP	**-0.0017*** **	**-0.0011*** **	**-0.0018*** **	**-0.0015*** **	**-0.0015*** **	**-0.0015*** **
	(-6.83)	**(-3.32)**	**(-6.08)**	**(-5.26)**	**(-4.25)**	**(-5.71)**
PoP	-0.0023***	-0.002***	-0.0025***	-0.0017***	-0.0011**	-0.0028***
	(-5.08)	(-3.78)	(-4.83)	(-3.77)	(-2.25)	(-6.51)
Beta	0.0026*	0.0017	0.002	0.0019	0.0016	0.0027*
	(1.86)	(1.16)	(1.3)	(1.46)	(1.17)	(1.86)
ln(Size)	-0.0009	-0.001***	-0.0013***	-0.0018***	-0.0011***	-0.0014***
	(-1.46)	(-3.07)	(-3.58)	(-3.32)	(-3.27)	(-3.42)
ln(BTM)	0.0013*	0.0017***	0.0016***	0.0016**	0.0014**	0.0019**
	(1.88)	(2.86)	(2.66)	(2.29)	(2.33)	(2.54)
$\text{ret}_{t-12,t-2}$	0.0077***	0.0074***	0.0072***	0.0083***	0.0075***	0.0083***
	(4.93)	(3.78)	(3.66)	(5.29)	(4.25)	(4.76)
$\text{ret}_{t-1,t-1}$	-0.0347***	-0.0349***	-0.0284***	-0.0408***	-0.042***	-0.0332***
	(-7.68)	(-7.08)	(-5.92)	(-9.04)	(-9.55)	(-6.99)
$\text{ret}_{t-36,t-13}$	-0.0006	0.0007	0.0001	-0.0003	0.0012*	-0.0008
	(-0.93)	(0.86)	(0.18)	(-0.36)	(1.74)	(-1.08)
Months	605	605	605	605	594	594
Avg. N	651	767	746	690	754	661
Avg. R2	0.086	0.123	0.128	0.087	0.098	0.098
Years	1965-2015	1965-2015	1965-2015	1965-2015	1965-2015	1965-2015

C Discussion of Additional Tables and Figures

C.1 Discussion of Summary Statistics

The summary statistics are presented in Table 1. For the full sample presented in Panel A, we find that on average, there are about 25 downloads by human users on the EDGAR platform in the 12 hours following the closure of the exchanges at 4:00 PM. For about 92% of the stock-days in the sample, we observe at least one human download on EDGAR (not tabulated). The abnormal transformation AEdgar of the information acquisition variable exhibits close to zero mean and unit standard-deviation. Looking at turnover, 0.98% of the shares outstanding are traded for the average stock-day in the sample and about 0.07% of the shares outstanding are traded by retail investors as identified from trade data following Boehmer et al. (2019). The abnormal transformations of the turnover variables exhibit close to zero mean.

Turning to returns, average daily close-to-close returns are very close to zero. Alternative returns calculated using a different initial price today exhibit a similar mean but reduced standard-deviation. As an example, consider the noon-to-close returns: the standard-deviation is reduced by about 1.3 percentage points compared to close-to-close returns. Nevertheless, there remains non-trivial variation in returns over intraday horizons, as the standard-deviation of noon-to-close returns still amounts to 1.66%.

In Panel B, we provide summary statistics on the overnight earnings announcements sample used for the difference-in-difference tests. Earnings announcements are one of the major systematic drivers of stock returns and valuations, and we can see corresponding patterns in the summary statistics. The standard-deviation of close-to-close returns increases to more than 7%, which is more than twice the standard-deviation of returns in the full sample. On an earnings announcement day, there are on average 115 downloads from the Edgar platform (Edgar (EA Day)) , a considerable increase with respect to the unconditional average of 25 downloads. Average AEdgar is correspondingly high at more than 70% on earnings announcement days (AEdgar (EA Day)).

In Panel C, we compare the groups of companies with large and small earnings surprises along several dimensions. The first four columns tabulate mean and standard-deviation by earnings surprise group and the last two columns provide the difference in means and an associated two-sample t-test with the null hypothesis of equal means. We can see that the construction of the two groups results in meaningful variation in absolute SUE, returns and turnover, as indicated by the significant differences in the means.

The key identifying assumption of our difference-in-difference approach is that firms with large earnings surprises would behave like the firms with small earnings surprises if they did not exhibit a large absolute earnings surprise (parallel trends assumption). To validate the parallel trends assumption, we provide a test for the difference in AEdgar in the month and day before the actual earnings announcement. We do not detect a significant difference in the level of AEdgar before the earnings announcements days (or even on earnings announcement days before markets open) which is consistent with parallel trends in the absence of treatment.

C.2 Return Conventions Tests

To explore whether the higher levels information acquisition after more extreme stock returns are driven by prominently placed stock returns or underlying information shocks, we exploit the pervasive convention of quoting daily returns measured from the previous day's close. If returns themselves drive information effects, then such conventional returns should drive information acquisition, whereas unconventional returns measured from a different starting point, e.g., from the market open, should have a smaller marginal effect on information acquisition.

In Table B2, we report results for a panel regression of abnormal EDGAR downloads (Panel A) and Wikipedia firm page views (Panel B) on absolute deviations of stock returns from the day's median stock return.[17] We include stock and day fixed effects and cluster standard errors in both dimensions. We also include five lags of the dependent variable to account for potential reverse causality and firm-specific persistent deviations from the long-run average level of information acquisition. Panel A's Specification (1) confirms that a higher magnitude of daily returns is associated with higher levels of information acquisition. The effect is economically highly significant and consistent with Figure 1 Panel A: 1 percentage point more extreme returns are associated with around 2.5% higher levels of abnormal EDGAR

[17]The evidence in Figure 1 suggests that absolute returns are a good approximation for the relation and are easy to interpret. Results are not affected when we use spline regressions or quadratic specifications that take into account possible convexity (i.e. a stronger reaction to very extreme returns). These results are currently unreported but available from the authors.

downloads. In Specification (2), we show that this effect on information acquisition after the market close is driven by conventionally prominent close-to-close returns, whereas unconventional open-to-close returns exhibit a much lower, negative coefficient. This negative marginal association should be surprising if one expects underlying information shocks to directly drive information acquisition. After all, extreme open-to-close returns indicate a more recent information shock than extreme close-to-close returns. Additionally, open-to-close returns exhibit a higher magnitude than close-to-open returns (see standard-deviations in Panel A of Table 1), so that extremeness of close-to-open returns cannot explain results. However, the dominance of conventional close-to-close returns in predicting information acquisition after the market closes is consistent with prominently placed salient returns driving our results in Specification (1).

[Insert Table B2 about here.]

To be clear, this regression faces some endogeneity issues, even if only prominently placed returns—measured from the previous close—drive information acquisition. E.g., the negative sign of the open-to-close returns in Specification (2) might be driven by a downward omitted variable bias. In particular, holding absolute close-to-close returns constant, a more extreme open-to-close returns tends to be associated with a less extreme close-to-open return. Hence, a positive effect of the omitted absolute close-to-open return should lead to a downward bias in the included absolute open-to-close return coefficient. Indeed, despite being afflicted by the same downward omitted variable bias, Specification (3) shows that overnight close-to-open returns exhibit such a positive effect when added to Specification (1). This strong positive effect is again in line with the convention to display returns from the previous day's market close to the current point in time: The overnight return is prominently displayed when the market opens and can thus plausibly attract information acquisition. Therefore, despite some clear endogeneity issues, these initial results are more consistent with prominently placed returns than recent information shocks as a driver of information acquisition.

Specifications (4) to (7) of Panel A show that the above results are not dependent on the usage of the market open (at 9:30 AM) to cut returns into conventionally reported versus unconventional returns. Specifications (4) and (5) move the cutoff to noon. Specifications (6) and (7) move the cutoff even further to 3:00 PM, i.e., one hour before the market closes. Moving the cutoff point closer to the starting point for our close-to-open information acquisition period has two important effects. First, the returns become less extreme in the x-to-close return window (see decreasing standard-deviations in Panel A of Table 1), which might decrease x-to-close effects if extreme returns are particularly influential (see convexity in Panel A of Figure 1). Second, extreme noon-to-close or 3:00 PM-to-close returns indicate that shocks happened more recently than for extreme close-to-close returns. If the reaction to shocks is particularly strong directly after the shock, a later cutoff point might increase x-to-close effects. In line with recent shocks having some positive effect, the coefficient of x-to-close returns turns positive as the cutoff point is moved from the market open-to-3:00 PM. However, the close-to-close return coefficient remains economically most significant in all specifications.

Specifications (9) to (11) of Panel A show that our result is robust for the four other possible combinations of (conventional) close-to-close, close-to-open, and (unconventional) open-to-close returns. That is, close-to-open or close-to-close returns are systematically more powerful in explaining subsequent information acquisition than open-to-close returns, despite the recency of X-to-close returns and underlying information shocks. To continue the above example of omitted variable bias with respect to Specification (2), Specification (11) shows that including both absolute close-to-open and open-to-close returns on top of absolute close-to-close returns turns the unintuitive negative open-to-close coefficient positive, in line with the above omitted variable bias explanation. Close-to-close and Close-to-open return however remain the strongest drivers of information acquisition.

In Specification (12) of Panel A, we show that the dominance of close-to-close returns in explaining subsequent information acquisition is not caused by the fact that close-to-close returns enclose sub-period returns used as competing explanatory variables (and thus exhibit a larger standard-deviation, see Panel A of Table 1). In particular, moving the starting point of the alternative x-to-close return period beyond yesterday's close, to noon yesterday, does not affect our conclusion that close-to-close returns are most powerful in predicting subsequent information acquisition. Despite the now much larger standard-deviation of x-to-close returns (see Panel A of Table 1), close-to-close returns are the dominant predictor of subsequent information acquisition in Specifications (12) and (13). In Specification (12) we leave out the lagged dependent variables, as they are clearly correlated with yesterday's returns and might thus explain a lower coefficient for x-to-close returns. Specification (13) shows that their inclusion only slightly affects the result. Close-to-close returns consistently remain the most powerful predictor of subsequent information acquisition.[18]

In Panel B of Table B2, we report analogous results using Wikipedia firm page views to proxy for information acquisition instead of EDGAR downloads. While effect sizes are smaller, conventional close-to-x returns are again most powerful in predicting information acquisition via Wikipedia, in line with results in Panel A. That is, our results seem to hold for both sophisticated investors' information acquisition on EDGAR and less sophisticated investors' information acquisition via Wikipedia.

In sum, Table B2 demonstrates that conventional returns from close-to-close explain the positive association between return magnitudes and information acquisition better than unconventional x-to-close returns. One could argue that acquiring information after observed returns is reasonable because underlying information shocks are often hard to identify. Hence, it might be cheaper to only acquire information once large absolute returns indicate an information shock, and particularly cheap to use conventionally reported close-to-x returns as indicators (as opposed to unconventional subperiod returns). If investors actually follow such a deliberate information acquisition heuristic, one would expect that our effect becomes weaker on days with easily identifiable information shocks. Contrary to this hypothesis, Table B3 shows that the dominance of close-to-close returns in predicting information acquisition remains strong when we restrict ourselves to days where there is an easily available, tangible information shock (Specifications 1 to 3) as opposed to no such information shock (Specifications 4 to 6). We report results for EDGAR downloads in Panel A and results for Wikipedia firm page views in Panel B. Tangible news days are defined as days with 8-K, 10-K, or 10-Q filings, earnings announcements and dividend declarations, or newspaper coverage in a national newspaper. On those days, one might expect our results to be weaker if they are driven by the difficulty of obtaining the information on whatever drives returns. However,

[18]Note that we do not run these regressions on conventional and unconventional returns with trading activity on the subsequent trading day as a dependent variable. This is because we expect much larger endogeneity issues for trading activity on the trading day after extreme returns. E.g., returns on a day after extreme stock returns are likely to be affected by these extreme returns through bid-ask-bounce or, more generally, short-term reversal. Such microstructure effects are much stronger after extreme open-to-close returns than after extreme close-to-open returns on the previous day (due to the recency of underlying shocks). This effect would bias up the reaction of next-day trading activity to previous-day open-to-close returns. For information acquisition after the market close, however, there are no such contemporaneous confounding returns.

conventionally reported close-to-open returns remain more powerful than open-to-close returns in predicting information acquisition on tangible news days (see Specifications 1 to 3), as well as days without such tangible news (see Specifications 4 to 6).[19]

[Insert Table B3 about here.]

Hence, we provide evidence in favor of conventionally prominently placed returns as major drivers of information acquisition after extreme returns. Unconventional returns that similarly measure large information shocks and are cheap to compute for any investor do not exhibit a comparable association with subsequent information acquisition. Results are even stronger on days with tangible news, suggesting that the close-to-close return effect is not explained by a heuristic to learn about underlying information shocks when acquiring information is difficult or costly.

[19]This test is about the relative effects of close-to-close returns versus open-to-close or close-to-open returns within each panel. The consistently stronger absolute effects in Specifications 1-3 compared to Specifications 4-6 might be driven by the higher return volatility on tangible news days together with the convexity of the relation between returns and information acquisition, see Panel A of Figure 1, which we do not model here.

D Overview of Variables

The main variables are defined in the table below.

Variable Name	Description
SEC EDGAR Information Acquisition Variables	
Edgar	SEC EDGAR Downloads. The number of EDGAR downloads by users exhibiting human user behavior (robots are excluded). The classification of users is described in Internet Appendix E. For the full sample, we measure downloads from 4pm to 9am on the following day. For the overnight earnings announcement sample, we measures downloads for the intervals 7.30am to 9.30am and 9.30am to 11.30am.
AEdgar	Abnormal SEC EDGAR Views. The log-difference of one plus Edgar today and one plus Edgar one week ago.
Wikipedia Information Acquisition Variables	
Wiki	Wikipedia page views. The number of Wikipedia page views of a firm's Wikipedia article. For the full sample, we measure page views from 4pm to 9am on the following day. For the overnight earnings announcement sample, we measure page views for the intervals 8.00am to 9.00am and 10.00am to 11.00am.
AWiki	Abnormal Wikipedia page views. The log-difference of one plus Wiki today and one plus Wiki one week ago.
Other Information Acquisition Variables	
AIAC	Abnormal Bloomberg news search activity. The difference of the continuous bloomberg news search score tranformation today and one week ago.
AGoogle	Abnormal Google search volume. The log-difference of one plus Google search volume today and one plus Google search volume one week ago. Google search volume is based on search activity for the tickers of stocks. We exclude tickers that are words included in The Online Plain Text English Dictionary.
Turnover Variables	
Turnover	Turnover. Shares traded (from Algoseek) divided by shares outstanding (CRSP "shrout").
ATurnover	Abnormal Turnover. The log-difference of Turnover today and Turnover one week ago. If the same day of the week in the previous week was not a trading day, we use the week before for the difference.
RetailTurnover	Retail Turnover. Shares traded by retailers devided by shares outstanding. Shares traded by retailers are identified from TAQ trade level data based on exchange identifier and sub-penny price improvements following the procedure of Boehmer et al. (2019).
ARetailTurnover	Abnormal Retail Turnover. The log-difference of RetailTurnover today and RetailTurnover one week ago.

Continued on next page

Variable Name	Description
Return Variables	
All price data is sourced from Algoseek. All returns are adjusted for stock splits and dividends (from CRSP). Price data is cleaned for data errors, see main text for details. Returns are centered with the cross-sectional median each day.	
Close_{t-1} to Close_t Return	Return from yesterday 4pm to today 4pm.
Open_t to Close_t Return	Return from the opening of the exchange today to 4pm today. The open price is the volume weighted transaction price of trades until 10am.
12pm_t to Close_t Return	Return from 12pm today to 4pm today.
3pm_t to Close_t Return	Return from 3pm today to 4pm today.
Open_{t-1} to Close_t Return	Return from the opening of the exchange yesterday to 4pm today. The open price is the volume weighted transaction price of trades until 10am.
12pm_{t-1} to Close_t Return	Return from 12pm yesterday to 4pm today.
3pm_{t-1} to Close_t Return	Return from 3pm yesterday to 4pm today.
Tangible News Variables	
8-K Filing	8-K Indicator. Dummy that takes on a value of one if the company filed an 8-K filing with the SEC today or yesterday. Filings dates are obtained form the master index file.
10-K or 10-Q Filing	10-K or 10-Q Indicator. Dummy that takes on a value of one if the company filed an 10-K or 10-Q filing with the SEC today or yesterday. Filings dates are obtained form the master index file.
Earnings Announcement	Earnings Announcement Indicator. Dummy that takes on a value of one if the company announced earnings today or yesterday. Quarterly earnings announcement dates are obtained from Compustat.
Dividend Announcement	Dividend Announcement Indicator. Dummy that takes on a value of one if the company announced a dividend today or yesterday. Dividend announcement dates are obtained from CRSP (declaration date).
Media Coverage	Media Coverage Indicator. Dummy that takes on a value of one if the company is covered in a national newspaper today. The newspapers included are the New York Times, USA Today, the Washington Post and the Wall Street Journal. Newspaper data are obtained from Nexis.

Continued on next page

Variable Name	Description
Earnings Announcement Variables	
SUE	Standardized Unexpected Earnings. SUE is calculated as the difference between actual EPS and the median split-adjusted analyst EPS forecast standardized by the share price five days before the announcement. Actual EPS data is from IBES unadjusted acutals file. Estimates are from unadjusted detail file. To be included, in the sample, we require at least one valid analyst forecast. A forecast is valid if it has been entered or updated in the 30 days prior to the earnings announcement. We only keep the most recent forecast per analyst.
Large Surprise	Large Surprise Dummy. Indicator that take on a value of one if SUE is smaller than the 30th or larger than the 70th percentile of all earnings announcements in the previous 180 days.
Post-Open	Post-Open Dummy. Indicator that take on a value of one if exchanges are open for the considered time interval.
Predicted Return	Predicted Overnight Close to Open Return. The predicted return is the average return of all earnings announcements in the same SUE group over the past 180 days. We split firms in ten SUE groups using the distribution of SUE in the 180 days before each earnings announcement.
Post-Open	Post-Open Dummy. Indicator that take on a value of one if exchanges are open for the considered time interval.
Number of analysts	Number of Analysts. The number of valid forecasts used to compute SUE.
Other Variables	
InstOwn	Institutional Ownership. Fraction of shares outstanding held by institutional investors. Based on Thomson Reuters S-34 database, retrieved via WRDS TR 13-F stock ownership tool.
Market Cap	Market capitalization. The product of price and shares outstanding.
WSJ Gainer X-Y	WSJ Gainer Ranking Indicator. Takes on a value of one if a stock is ranked in the WSJ gainer ranking (covering stocks with the largest returns) in ranks X to Y. WSJ rankings are scraped from the WSJ website.
WSJ Decliner X-Y	WSJ Decliner Ranking Indicator. Takes on a value of one if a stock is ranked in the WSJ decliner ranking (covering stocks with the smallest returns) in ranks X to Y. WSJ rankings are scraped from the WSJ website.
Variables used in Asset Pricing Tests	
PoP	Measure of Return Salience. Obtained by averaging the daily salience within a month. The daily salience score is the level of abnormal information acquisition for the return percentile depicted in Figure 1. See main text for details.
SYY Score	Mispricing Score of Stambaugh et al. (2012). The score is based on 11 stock market anomalies to infer whether a stock is currently underpriced or overpriced. Obtained from Roberst Stambaugh's website.
ln(Size)	Natural logarithm of market capitalization.
ln(BTM)	Natural logarithm of the book-to-market ratio.

Continued on next page

Variable Name	Description
$\mathrm{ret}_{\text{t-12,2}}$	Momentum. Return over the previous twelve months excluding the most recent month.
$\mathrm{ret}_{\text{t-1,t-1}}$	Short-term reversal. Return over the most recent month.
$\mathrm{ret}_{\text{t-36,t-13}}$	Long-term reversal. Return over the previous 36 months excluding the most recent 12 months.
Profitability (FF)	Profitability measure of Fama and French (2015).
Investments (FF)	Investments measure of Fama and French (2015).
ln(Turnover)	Natural logarithm of turnover (shares traded over shares outstanding).
Δln(Turnover)	First-difference of ln(Turnover).

E SEC EDGAR Log File Data

The SEC provides data on the usage of filings available via the EDGAR platform in the log file dataset.[20] The files contain a timestamp in second precision, a partly anonymized IP address of the requester as well as the unique SEC identifier for the filing requested ("accession key"). We match the filings to companies based on the company identifiers assigned by the SEC ("CIK") in the master index file. If a filing is related to multiple companies (e.g. takeover related filing), we record a download of such a filing for both CIKs. We exclude requests for index pages, that just show the filing's contents without providing information and we exclude requests with http response codes indicating a technical error.

E.1 Classifying Human vs Robot Users

For our main results, we only rely on download requests that derive from IP addresses that show behavior consistent with a human downloading and looking at the filings. Therefore, we classify IP addresses into humans and robots based on their downloading behavior following Ryans (2017). In line with his suggestion, an IP address is classified as a human user if and only if all three of the following criteria are fulfilled:

1. The IP address does not issue more than 25 requests in any single minute.
2. The IP address does not issue requests for filings associated with three distinct companies in any single minute.
3. The IP address does not issue more than 500 requests during the day.

We use this classification each day for all IP addresses that issued at least one request on that day. Ryans (2017) provides a detailed explanation of the criteria and why they likely outperform other classification algorithms used in the prior literature.

[20]The dataset is available at https://www.sec.gov/dera/data/edgar-log-file-data-set.html