Pre-earnings Announcement Over-extrapolation*

Aytekin Ertan, LBS

Stephen A. Karolyi, Carnegie Mellon

Peter W. Kelly, Yale

Robert Stoumbos, Yale

March 10, 2015

Abstract

We uncover evidence that individual investors over-extrapolate from past returns. Specifically, we study the relation between previous earnings announcement returns and investor beliefs regarding subsequent earnings announcement returns. Consistent with models of over-extrapolation, we find that investors become overly optimistic about the firm's future prospects after positive earnings announcement returns. This leads to predictable patterns in returns both before and after earnings announcements: If a firm is in the top decile of our extrapolated return measure, we expect a 5-day pre-earnings announcement return 27 basis points more than all other 5-day pre-earnings announcement returns, and a 5-day post-earnings announcement return 23 basis points lower than all other 5-day post-earnings announcement returns.

*We are grateful for helpful comments from Samuel Hartzmark, Jake Thomas, and Frank Zhang. Peter Kelly acknowledges support from a Whitebox Advisors fellowship.

1 Introduction

In their seminal paper, Kahneman and Tversky (1974) highlight that experimental subjects tend to ignore the laws of probability - instead, they assess likelihoods by the degree to which an event reflects the salient characteristics of a specific class. The authors refer to this tendency as the representativeness heuristic. A number of authors (e.g., Barberis et al., 1998) have suggested that investors, guided by the representativeness heuristic, over-extrapolate, or draw strong conclusions from small samples of data. In this paper, we explore this argument by studying behavior around earnings announcements.

An investor using the representativeness heuristic may misclassify a stock that has consistently beaten earnings expectations as one that will *always* beat earnings expectations. We predict that individual investors will "bet" on this classification. Namely, if a firm has had a string of great earnings announcement returns, the investor will bet on, or purchase, the firm's stock shortly before the next earnings announcement. To test this, we first look at how individual investor behavior responds to our extrapolated return measure, a measure based on past earnings-announcement returns. We analyze household trading data from a large discount brokerage and find that past earnings announcement returns are a strong determinant of individual investor purchase decisions in the period leading up to the next earnings announcement. Specifically, if a firm is in the top decile of our extrapolated return measure, we predict that the total value of purchases in the 5-day pre-earnings announcement period will be 56 percent higher (t=2.94) than for a firm not in the top decile.[1,2]

We next examine whether this buying behavior leads to predictable patterns in asset prices. The prediction is that over-extrapolation of previous positive earnings announcement returns will motivate the investor to purchase shares of the stock before the next earnings announcement. This

[1] We focus on earnings announcement returns as our primary measure of whether a firm beats expectations. We use this measure as it does not rely on an expectations measure—that is, we do not have to proxy for what market expectations are. We also think it is plausible that earnings announcement returns are the most salient characteristic.

[2] Clearly, this result conjectures a setting in which investors know earnings announcement dates in advance. There is reason to believe this is the case: (i) Scheduled earnings announcement dates are publicly accessible (e.g., `http://www.zacks.com/earnings/earnings-calendar`), and (ii) under normal circumstances, a majority of companies announce earnings on the same day (e.g., third Thursday of March) every year.

will lead to a rise in the stock price, and overpricing, before the next earnings announcement.[3] Then, on the earnings announcement date, information is revealed, and expectations, on average, are not met. The stock price will start to revert back to its "fundamental value". To summarize, we expect to see predictable patterns in returns both before and after such earnings announcements—specifically, a high value for our extrapolated return measure should predict excess positive returns before the earnings announcement, and a reversal after.

We confirm our prediction in the data and find that the return in the 5 trading days before the earnings announcement period is strongly positively associated with earnings announcement returns of previous quarters. Namely, if a firm is in the top decile of our extrapolated return measure, we expect to see a 5-day pre-earnings announcement return 27 basis points higher (t=3.16) than all other 5-day pre-earnings announcement returns. Consistent with the idea that investors place "bets" shortly before the earnings announcement date, we find that being in the top decile only predicts a 30-day pre-earnings announcement return 17 basis points higher (t=0.73).[4]

We also find evidence that this is an overreaction. Being in the top decile of our extrapolated return measure is associated with a 5-day post-earnings announcement return 23 basis points lower than all other post-earnings announcement returns (t=-3.01). The reversal extends to longer horizons. If a firm is in the top decile of our extrapolated return measure, this predicts a 30-day post-earnings announcement return 47 basis points lower than all other post-earnings announcement periods (t=3.16).

We strengthen our analysis to consider other possible extrapolation behavior. We first consider fundamentals extrapolation. Do investors extrapolate earnings growth? Do investors think that firms that have previously beaten earnings estimates are likely to continue to beat earnings estimates? Using fundamental-level measures, we find that any fundamental extrapolation seems to be driven by our correlated extrapolated return measure. We also consider whether investors over-extrapolate earnings management behavior. Does our result stem from investors' over-anticipation

[3] We argue that this "mispricing" will not be corrected by appealing to the limits to arbitrage literature. See, for example, Shleifer and Vishny (1997).

[4] Notice that this finding also rules out concerns regarding a post-earnings-announcement-drift overhang driving our results. Namely, if our high-extrapolated-returns bucket simply captured firms with the best earnings performance in the previous quarter, we would observe larger, not smaller, returns over 30-day window preceding earnings announcement.

of earnings management? We find evidence that this is not the case.

Our paper provides individual-level and market-level evidence that investors over-extrapolate. By examining pre-earnings and post-earnings announcement returns, we show that this belief formation impacts asset prices. As such, we provide evidence that the survey evidence and theoretical work on over-extrapolation should be considered.[5] We also add to the literature on short-term reversals. So and Wang (2014) document a six-fold increase in short-term return reversals during earnings announcement periods; we highlight a driver of the price movement before the announcement.

Section 2 reviews the literature. Section 3 details the data used. In Section 4, we look at the relation between individuals' investment decisions and previous earnings announcement returns. We present, in Section 5, evidence that over-extrapolation leads to predictable patterns in returns both before and after earnings announcements. Section 6 concludes.

2 Literature Review

The representativeness heuristic refers to the tendency of individuals to ignore the laws of probability and determine likelihoods by the extent to which an event reflects the characteristics of the process that generates it or a parent population. (Kahneman and Tversky (1974)) This heuristic leads to seemingly contradictory phenomena: the gambler's fallacy — the erroneous belief that random sequences should exhibit systematic reversals — and the hot-hand effect — the belief that a successful event is more likely following a sequence of successful events than after an unsuccessful event. When individuals know the true distribution, we see evidence of the gambler's fallacy. Individuals expect small random sequences to reflect the overall distribution; as such, subjects anticipate "balancing out" so that small sequences better reflect the true distribution. We see evidence of the hot-hand effect when the true distribution is unknown; individuals believe that recent random sequences reflect the true distribution. After a sequence of successful outcomes, individuals believe that the generating process is one that delivers successful outcomes; they

[5] Greenwood and Shleifer (2014)review the survey evidence and find strong evidence of over-extrapolation. Barberis et al. (2015) develop a model consistent with this survey evidence that captures a number of features of actual prices and returns, such as excess volatility and the negative correlation of returns at long horizons.

anticipate future successful outcomes. (Gilovich et al. (1985); Rabin and Vayanos (2010)) The hot-hand effect can also be termed over-extrapolation: individuals infer too much from small samples of data.

There is a large literature that examines whether investors, analysts, and firm management exhibit over-extrapolation. The theoretical paper that is most tied to our work is Barberis et al. (1998). The authors consider a model where investors use the representativeness heuristic and also exhibit conservatism—the slow updating of models in the face of new evidence. In their model, investors' expectations are based on previous earnings. The authors show that the model helps explains a number of known pricing anomalies including momentum and long-term reversions. Camerer (1987) finds that investors over-extrapolate in an experimental market setting. Bloomfield and Hales (2002) also offer experimental evidence that investors over-extrapolate—the authors find evidence consistent with the model of Barberis et al. (1998). Frieder (2008) finds that net buying by small investors increases as the number of consecutive earnings surprises increases and that this purchase activity is negatively correlated with future returns. The evidence whether sell-side analysts over-extrapolate is mixed. De Bondt and Thaler (1990) find that analysts overreact to new information, while Teoh and Wong (2002) uncover evidence that analysts underreact to new information. Easterwood and Nutt (1999) reconcile these findings by showing that analysts underreact to negative information and overreact to positive information. Finally, Greenwood and Hanson (2015) offer evidence that management over-extrapolate from recent earnings information. The authors show that firms in the shipping industry overpay for ships and overinvest during boom periods.

In our paper, we find that investors extrapolate based on past earnings announcement returns. As such, our work is tied to the earlier literature on return extrapolation. There is ample survey evidence that investors hold extrapolative return expectations. They believe stock prices will rise after they have previously risen and will fall after they have previously fallen; Greenwood and Shleifer (2014) review data from six different surveys and find that expectations from different surveys are strongly correlated with one another and with past returns.[6] Barberis et al. (2015)

[6]Earlier survey evidence includes Vissing-Jorgensen (2004), Amronin and Sharpe (2014), and Bacchetta et al.

develop a model consistent with this survey evidence that captures a number of features of actual prices and returns, such as excess volatility and the negative correlation of returns at long horizons.

A central prediction of our paper is that over-extrapolation leads to predictability in returns around earnings announcements. Specifically, we test that previous earnings announcement returns are positively correlated with this period's pre-earnings announcement returns and negatively correlated with this period's post-earnings announcement returns. This investigation is related to a number of other empirical facts regarding earnings announcement returns. Barth et al. (1999) and Myers et al. (2007) find that price to earnings multiples fall at the termination of a positive earnings streak. Ball and Brown (1968) and Foster et al. (1984) note that even after earnings are announced, abnormal returns continue to be positive for firms with good earnings and continue to be negative for firms with bad earnings, the post earnings-announcement drift.[7] Millian (2015) argues that investors have overreacted to the post-earnings announcement drift phenomena; he finds significantly negative autocorrelation between a firm's earnings announcement news and next period's earnings announcement return for easy-to-arbitrage firms in the 1996–2010 period. So and Wang (2014) document a six-fold increase in short-term return reversals during earnings announcement periods. The authors, who do not cite the source of the initial price increase or decrease, attribute the strength of these reversals to the increased inventory risks that the market maker faces.

We attribute the earnings-related anomaly we find to over-extrapolation; we are not the first paper to find that behavioral factors can drive earnings announcement returns. Chang et al. (2014) show that investors are surprised by the unsurprising; firms whose earnings are historically larger in one quarter of the year have higher returns when those earnings are typically announced. The authors attribute this anomaly to investors overweighting the earnings of recent quarters. Loh and Warachka (2012) find that investors underreact to the continuation of earnings streaks. That is, they find that the post-earnings announcement drift is especially strong when the earnings surprise extends a streak. In contrast to our finding, the authors find that their result is consistent with

(2009).

[7]Bernard and Thomas (1989) show that it is difficult to reconcile this result with a risk based explanation. In their subsequent paper, Bernard and Thomas (1990) argue that the autocorrelation in earnings surprises helps explain the post-earnings announcement drift.

the gambler's fallacy—investors believe that trends require an immediate balancing by the opposite outcome.

3 Data

The data for our main sample comes from IBES, CRSP, and Compustat. It includes all US firm-quarters from 1993 to 2013. We drop observations before 1993 because of a potential mismatch between IBES's calculation method of realized and forecasted earnings in the early 1990s (Cohen et al. (2007)). For each firm-quarter, we calculate consensus analyst forecasts by limiting our attention to the most recent, individual (analyst-level), split-unadjusted forecasts (Diether et al. (2002)) that are confirmed reliable at most 90 days before the earnings announcement.[8] We remove firm-quarters that are missing this quarter's or the previous quarter's earnings surprise, measured as the IBES actual earnings per share minus the median of individual forecasts. Firm-quarters with a share price less than $5 at the beginning of the quarter are also excluded to mitigate confounding microstructure effects.

We investigate investor buying behavior using data from a large discount brokerage. This data set has been used extensively (e.g. Odean (1998), Strahilevitz et al. (2011), Hartzmark (2014), etc.) and includes information on 78,000 households trades from 1991 through 1996. We use information of the dollar value of purchase and the dollar value of sales, to calculate total buying, total selling, and net purchases. We restrict our sample to individual and joint-tenancy accounts to focus on individual investors, as opposed to corporations, partnerships, and pension trusts.[9] As explained in section 4, we do this in order to provide evidence that our results are not driven by excessive institutional trading on the post-earnings-announcement drift strategy (Millian (2015)).

Stock return data come from CRSP daily and monthly security files. We calculate returns surrounding this quarter using various windows. We calculate pre-earnings-announcement (post-earnings-announcement) returns as the 5-day and 30-day returns leading up to (following) two

[8] We determine this date using *revdats*, defined in IBES as the most recent date that an estimate was confirmed as accurate. We identify earnings-announcement date as the earlier of IBES earnings announcement date (*actdats*) and Compustat earnings announcement date (*rdq*).

[9] The results are similar if we include the entire sample from the discount brokerage.

days before (after) the earnings announcement. Similarly, earnings announcement returns are calculated over the 3-day trading window surrounding the earnings announcement. We use earnings announcement returns of *previous* quarter(s) to construct our independent variable of interest. This variable, which we argue captures investor extrapolation, is equal to a weighted-average return of the past eight earnings announcement returns. Specifically, we weight more recent quarters more heavily as they may be more salient to the investor (Barberis et al. (2015)).

In order to construct a vector of controls we collect financial data from Compustat Fundamentals quarterly, CRSP, and IBES. These controls include size, the book-to-market ratio, profitability, earnings surprise, forecast dispersion, trailing stock performance, volatility and trading volume.[10] Our final sample consists of 132,533 firm-quarters for which we can estimate all variables including extrapolated returns, which requires the past 8 earnings announcement returns.[11]

Table 1 presents descriptive statistics for our sample.

4 Individual Investor Buying Behavior

Our hypothesis is that investors over-extrapolate from past-earnings announcement returns. We predict that this belief formation will guide their future purchase behavior—namely, we expect to see investors "bet" on earnings by purchasing shares of a stock with a high extrapolated return in the period shortly before the next earnings announcement. We test our prediction by looking at data from a large discount brokerage.

We first create an extrapolated return measure:

$$\text{Extrapolated Return}_{i,t} = \sum_{j=1}^{8} \frac{1}{j} R_{t-j} \tag{1}$$

Our extrapolated return measure considers returns from the past 8 earnings announcements. Consistent with the idea that more recent events may be more salient, we weight more recent

[10] Definitions of variables of interest and control variables are detailed in Appendix A. Continuous independent variables are Winsorized at the 1% level.

[11] By comparison, our initial IBES sample has 195,691 firm-quarters with a non-missing earnings surprise, forecast dispersion, and at least a $5 share price. We are able to estimate our dependent and control variables for 169,721 observations.

earnings-announcement periods more heavily.[12]

We then look to see whether individual investor behavior is influenced by this behavior; that is, we look for evidence that investors are placing their "bets" shortly before the next earnings announcement. To examine this, we consider a regression of the following form:

$$\begin{aligned}\log(\text{Total Buying})_{i,t} = \quad & \beta_0 + \beta_1 \text{Test Period}_{i,t} + \beta_2 \text{Top Decile Extrapolated Return}_{i,t} \\ & + \beta_3 \text{Test Period * Top Decile Extrapolated Return}_{i,t} + \\ & \gamma_1 \text{Calendar Quarter}_t + \gamma_2 \text{Firm Identifier}_i + \epsilon_{i,t}\end{aligned}$$

This test represents a sort of difference-in-differences design. To calculate the dependent variable, *log(Total Buying)*$_{i,t}$, we determine the total dollar value of purchases made by individuals in our data set and take the log of that value. We restrict our sample to two periods. The first consists of the 5 days immediately prior to the earnings announcement period. This represents the test period, similar to the "post" period in a traditional difference-in-differences. The second period is our control period, similar to the "pre" period in a traditional difference-in-differences. It consists of the 41-45 days before the earnings announcement period, so chosen in order to have the same number of days as the treatment period and to be far away from any earnings announcements.[13] *Test Period*$_{i,t}$ is a dummy variable that equals one if the observation is from the 5-day period before the earnings announcement period. For each calendar quarter, we sort the aforementioned extrapolated return variable into deciles and construct a dummy variable, *Top Decile Extrapolated Return*$_{i,t}$ that equals one if the firm falls into the top decile of our extrapolated return measure. In our difference-in-differences test, firm-quarters with an extrapolated return in the top decile comprise our treatment group, and the remaining firm-quarters comprise our control group.

We are particularly interested in the interaction term, *Top Decile Extrapolated Return*Test Period*$_{i,t}$. The coefficient on this interaction term, β_3, tells us how much the increase in purchases for top-decile firms, in the period right before the earnings announcement, is greater than it is for other firms. This difference-in-differences design isolates purchases in anticipation of good earnings-

[12] We have also used an equal-weighted measure and have found similar results.

[13] We have considered other control periods and found similar results.

announcement performance for top-decile firms from normal changes in purchase behavior before earnings announcements, and from any general differences between purchases of top-decile firms and other firms. We predict that β_3 will be positive as we expect more purchases in the period right before earnings for firms that have a high extrapolated return measure.

It is possible that some firms may have more purchases before earnings announcements, and happen to have high extrapolated returns during our sample. To address this possible concern, we include firm fixed effects. We also include time fixed effects, $Calendar\ Quarter_t$, to capture quarters where there may be particularly high buying pressure during the test periods. We cluster standard errors by firm as, by definition, there is overlap within-firm in the extrapolated return measure.

We present the results in Table 3 - we find results in support of our prediction that $\beta_3 > 0$. Specifically, if a firm is in the top decile of our extrapolated return measure, we predict that the total value of purchases in the 5-day pre-earnings announcement period will be 56 percent higher (t=2.94) than for a firm not in the top decile. If we replace the dependent variable with total selling, we find that firms in the top decile are associated with 11 percent less selling during the test period than all other firms, but the result is not statistically significant (t=0.59).

We interpret this as strong evidence that our results are distinct from the story proposed by Millian (2015). Millian (2015) argues that investors have overreacted to the post-earnings announcement drift phenomenon; he finds significant negative autocorrelation between a firm's earnings announcement news and next period's earnings announcement return for easy-to-arbitrage firms in the 1996-2010 period. Finding significant evidence that *individual investors* respond to our extrapolated return measure suggests that this result is not driven by institutional investor overreaction to the post-earnings announcement phenomenon.

5 Earnings Announcement Returns

Having established that individual investor purchasing behavior is positively correlated with past earnings announcement returns in the period immediately leading up to the next earnings

announcement date, we next examine whether this has any asset pricing implications. Specifically, we expect that if there is excess buying pressure in the period right before earnings (and this excess buying pressure is the result of over-extrapolation) then we should see a rise in the price of the stock before earnings and a fall in the stock price after earnings are announced. Specifically, we estimate a regression of the following form:

$$\text{Previous 5-Day Return}_{i,t} = \beta_0 + \beta_1 \text{Extrapolated Return}_{i,t} + \gamma_1 \text{Controls} + \gamma_2 \text{Industry-Quarter}_{i,t} + \epsilon_{i,t}$$

The prediction is that the coefficient on *Extrapolated Return*$_{i,t}$, defined as before, will be positive and statistically significant. That is, a high extrapolated return should be associated with higher pre-earnings announcement returns. We include a number of different controls that could plausibly predict earnings announcement (and potentially, pre-earnings-announcement) returns as well. We include controls for known predictors of returns like size, the book-to-market ratio, and momentum. It is well known that earnings are auto-correlated (Bernard and Thomas (1990)). Therefore, we include controls for last quarter's earnings surprise and the decile rank of the last quarter's earning surprise. More directly, we control for the current period's earning surprise, earnings growth, accruals, and a dummy that equals one if this period saw negative earnings. We also include a control equal to the average of four-quarters ago earnings and eight-quarters ago earnings as Chang et al. (2014) show that previous earnings in the same season, can predict this periods' earnings announcement return. We also include controls for the volume of trades, the stock price volatility, and the standard deviation of individual analyst earnings estimates.[14] It is possible that the pre-earnings announcement returns are high in industry-quarters where extrapolated returns are high. To address this, we include industry-quarter dummies. We cluster standard errors by calendar quarter as there are likely subsets of firms within each quarter that are hit by common shocks which affect the extrapolated return and the pre-earnings announcement return.

[14] We also require the dispersion of analyst estimates to be greater than zero.

We present the results in Table 5: We find evidence consistent with our prediction. Specifically, we estimate $\beta_1 = 0.0346$ ($t = 3.29$), which implies that a one-standard deviation in our extrapolated return measure is associated with a 5-day pre-earnings announcement return 13 basis points higher. Consistent with the idea that investors are purchasing shares shortly before earnings for firms with high extrapolated returns: When we replace the left-hand side variable with the previous 30-day return $\beta_1 = 0.0013$ ($t = 0.08$), the coefficient on *Extrapolated Return*$_{i,t}$ is no longer economically or statistically significant. Importantly, we find evidence of a reversal. If we replace the left-hand side variable with the 5-day post-earnings announcement return, we estimate $\beta_1 = -0.0313$ ($t = 3.34$) - opposite in sign and nearly equal in magnitude to the 5-day pre-earnings increase in price.

For ease of interpretation, we also consider regressions where we replace the *Extrapolated Return*$_{i,t}$ variable with a dummy variable, *Top Decile Extrapolated Return*$_{i,t}$ that equals one if the firm falls into the top decile of our extrapolated return measure. We find that being in the top decile of our extrapolated return measure is associated with a 5-day pre-earnings announcement return 27 basis points higher (t=3.16) than all firms outside the top decile. We also find evidence of a reversal - being in the top decile of our extrapolated return measure is associated with a 5-day pre-earnings announcement return 23 basis points lower (t=-3.01) than all firms outside the top decile. We also add a dummy variable, *Bottom Decile Extrapolated Return*$_{i,t}$ that equals one if the firm falls into the bottom decile of our extrapolated return measure. We find that being in the bottom decile of our extrapolated return measure is associated with a 5-day pre-earnings announcement return 19 basis points lower (t=-2.11).

One alternative interpretation of our result is that the predictable run-up in price and subsequent reversal comes from firms that manipulate earnings. Investors may expect firms with a history of earnings manipulation to continue to manipulate earnings in the future. Therefore, we would expect to see a run-up in prices before earnings and a large decrease when the firm is finally forced to stop manipulating earnings. However, we interpret the top decile results as evidence against such a story - it is unlikely that firms that have manipulated earnings to just beat forecasts will be in the top decile of our extrapolated return measure.

6 Conclusion

Our paper provides individual-level and market-level evidence that investors over-extrapolate. We first examine individual investor trades using data from a large discount brokerage and uncover support for the notion that investors' purchase decisions are driven by extrapolative beliefs. We then find market-level evidence consistent with extrapolative beliefs. Specifically, we find strong evidence of a run-up in prices before earnings announcements for firms that have historically had good earnings announcement returns and a reversal once the information is released. In conclusion, we provide empirical evidence that over-extrapolation is prevalent and thus the survey evidence and theoretical work on over-extrapolation should be taken seriously in future studies of financial markets.

References

Amronin, G. and S. Sharpe (2014). From the horse's mouth: Economic conditions and investor expectations of risk and return. *Management Science 60*(4), 845–866.

Bacchetta, P., E. Mertens, and E. van Wincoop (2009). Predictability in financial markets: What do survey expectations tell us? *Journal of International Money and Finance*, 406–426.

Ball, R. and P. Brown (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research*, 159–178.

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer (2015). X-capm: An extrapolative capital asset pricing model. *Journal of Financial Economics 115*(1), 1–24.

Barberis, N., A. Shleifer, and R. Vishny (1998). A model of investor sentiment. *Journal of Financial Economics 49*, 307–343.

Barth, M., J. Elliot, and M. Finn (1999). Market rewards associated with patterns of increasing earnings. *Journal of Accounting Research 37*(2), 387–413.

Bernard, V. and J. Thomas (1989). Post-earnings-announcement drift: delayed price response or risk premium? *Journal of Accounting Research*, 1–36.

Bernard, V. and J. Thomas (1990). Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics 13*(4), 305–340.

Bloomfield, R. and J. Hales (2002). Predicting the next step of a random walk: experimental evidence of regime shifting beliefs. *Journal of Financial Economics 65*(3), 397–414.

Camerer, C. (1987). Do biases in probability judgment matter in markets? experimental evidence. *The American Economics Review*, 981–997.

Chang, T., S. Hartzmark, D. Solomon, and E. Soltes (2014). Being surprised by the unsurprising: Earnings seasonality and stock returns. *Available at SSRN 2460166*.

Cohen, D., R. Hann, and M. Ogneva (2007). Another look at gaap versus the street: an empirical assessment of measurement error bias. *Review of Accounting Studies 12*(2-3), 271–303.

De Bondt, W. and R. Thaler (1990). Do security analysts overreact. *The American Economic Review*, 52–57.

Diether, K., C. Malloy, and A. Scherbina (2002). Differences of opinion and the cross section of stock returns. *The Journal of Finance 57*(5), 2113–2141.

Easterwood, J. and S. Nutt (1999). Inefficiency in analysts' earnings forecasts: Systematic misreaction or systematic optimism. *The Journal of Finance 54*(5), 1777–1797.

Foster, G., C. Olsen, and T. Shevlin (1984). Earnings releases, anomalies, and the behavior of security returns. *Accounting Review*, 574–603.

Frieder, L. (2008). Investor and price response to patterns in earnings surprises. *Journal of Financial Markets 11*(3), 259–283.

Gilovich, V., R. Vallone, and A. Tversky (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology 17*, 295–314.

Greenwood, R. and S. Hanson (2015). Waves in ship prices and investment. *Quarterly Journal of Economics 130*(1), 55–109.

Greenwood, R. and A. Shleifer (2014). Expectations of returns and expected returns. *Review of Financial Studies 27*(3), 714–746.

Hartzmark, S. (2014). The worst, the best, ignoring all the rest: The rank effect and trading behavior. *Review of Financial Studies*.

Kahneman, D. and A. Tversky (1974). Judgment under uncertainty: heuristics and biases. *Science 185*, 1124–1131.

Loh, R. and M. Warachka (2012). Streaks in earnings surprises and the cross-section of stock returns. *Management Science 58*(7), 1305–1321.

Millian, J. A. (2015). Unsophisticated arbitrageurs and market efficiency: overreacting to a history of underreaction? *Journal of Accounting Research 00*, 0–0.

Myers, J., L. Myers, and D. Skinner (2007). Earnings momentum and earnings management. *Journal of Accounting, Auditing, & Finance 22*(2), 249–284.

Odean, T. (1998). Are investors reluctant to realize their losses? *Journal of Finance 53*(5), 1775–1798.

Rabin, M. and D. Vayanos (2010). The gambler's and hot-hand fallacies: Theory and applications. *The Review of Economic Studies 77*(2), 730–778.

Shleifer, A. and R. Vishny (1997). The limits of arbitrage. *The Journal of Finance 52*(1), 35–55.

So, E. C. and S. Wang (2014). News-driven return reversals: Liquidity provision ahead of earnings announcements. *Journal of Financial Economics 114*, 20–35.

Strahilevitz, M., T. Odean, and B. Barber (2011). Once burned, twice shy: How naive learning, counterfactuals, and regret affect the repurchase of stocks previously sold. *Journal of Marketing Research 48*, 102–120.

Teoh, S. and T. Wong (2002). Why new issues and high-accrual firms underperform: The role of analysts' credulity. *Review of Financial Studies 15*(3), 869–900.

Vissing-Jorgensen, A. (2004). Perspectives on behavioral finance: Does 'irrationality' disappear with wealth? evidence from expectations and actions. *NBER Macroeconomics Annual 18*.

Appendix A: Variable Definitions

Variable	Calculation	Source
Pre-announcement returns	Return accum. over [–30, –1] or [–5, –1]	CRSP
Announcement returns	Return accum. over [–1, +1]	CRSP
Post-announcement returns	Return accum. over [+1, +5] or [+1, +30]	CRSP
Extrapolated returns	Inverse weighted average of past *announcement returns* from quarter –8 to –1	CRSP
Past consecutive beatings	Number of consecutive previous quarters with *ann. returns* > 0 (capped at 8)	
Earnings surprise	Actual EPS minus the most recent consensus analyst forecast, unscaled, unadjusted	IBES
Forecast dispersion	Standard deviation of individual analyst forecasts	IBES
Staleness	Average of the difference between earnings ann. date and analyst forecast date	
Market capitalization	$CSHOQ \times PRCCQ$	Compustat
Operating cash flows	*OANCFY*, converted into quarterly levels	Compustat
Accruals	*(IBQ – Operating cash flows) / (l.ATQ)*	Compustat
Book-to-market	$CEQQ$ / $(CSHOQ \times PRCCQ)$	Compustat
Loss indicator	1 if $IBQ < 0$, and 0 otherwise	Compustat
Leverage	*(DLCQ + DLTTQ) / ATQ*	Compustat
Volume	6-month average of *VOL/SHROUT* leading up to the earn. announcement	CRSP
Volatility	6-month standard deviation of stock returns *RET* leading up to earn. announcement	CRSP
Momentum	12-month average of monthly *RET* leading up to earn. announcement	CRSP

Variables are as of the beginning of the quarter, unless stated otherwise.

Adjusted signifies conversion of year-to-date (cash flow) items into quarterly figures.

Table 1: Descriptive Statistics

	Count	Mean	Standard Deviation	Min	25%	50%	75%	Max
Extrapolated Return	133445	.0060933	.0379137	-.2954939	-.0131431	.0046592	.0244768	.5875545
Sell-Side Surprise (cents)	194750	1.042757	164.305	-5572	-1	1	4	55400
Sell-Side Coverage	194750	7.374773	5.739028	2	3	6	10	51
Investor Surprise (cents)	6056	.3110551	28.519	-1187	-2.5	.999999	4.5	403.5
No. of Investor Forecasts	6102	5.879056	11.56811	1	1	3	7	373
% of Optimistic Investors	6102	76.52508	34.43155	0	63.63636	100	100	100
Size	193095	4962.572	18349.87	2.780375	352.055	939.2238	2900.571	626550.4
Total Assets	193969	9288.881	60989.78	3.884	304.832	1051.915	3819.834	2439494
Book-to-Market	192374	.4981665	.3920302	-11.68505	.2593319	.4325754	.6607085	21.06497

Table 2: Pearson Correlations

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1)	Extrapolated Return	1.00								
(2)	Previous 10-day Return	0.02	1.00							
		(0.00)								
(3)	Previous 30-day Return	0.03	0.54	1.00						
		(0.00)	(0.00)							
(4)	Earnings Announcement Return	0.01	-0.05	-0.02	1.00					
		(0.00)	(0.00)	(0.00)						
(5)	Forward 10-day Return	-0.04	-0.02	-0.03	0.02	1.00				
		(0.00)	(0.00)	(0.00)	(0.00)					
(6)	Forward 30-day Return	-0.03	-0.00	-0.01	0.04	0.57	1.00			
		(0.00)	(0.41)	(0.00)	(0.00)	(0.00)				
(7)	% of Optimistic Investors	0.16	0.02	0.03	-0.02	0.02	0.04	1.00		
		(0.00)	(0.06)	(0.01)	(0.08)	(0.21)	(0.00)			
(8)	Investor Surprise	-0.02	0.02	0.04	0.12	0.01	-0.00	-0.05	1.00	
		(0.17)	(0.17)	(0.00)	(0.00)	(0.29)	(0.83)	(0.00)		
(9)	Sell-Side Surprise	0.03	0.01	0.01	0.02	0.00	0.00	0.04	0.68	1.00
		(0.00)	(0.00)	(0.00)	(0.00)	(0.05)	(0.37)	(0.00)	(0.00)	

Table 3: Individual Trades

	(1) Log(Net Purchase)	(2) Log(Total Purchases)	(3) Log(Total Sales)
Top Decile Extrapolated Return	-0.360	0.189	0.412***
	(-1.32)	(1.20)	(2.89)
Test Period	0.097	0.099***	0.020
	(1.55)	(2.76)	(0.53)
Top Decile Extrapolated Return*Test Period	1.09***	0.56***	-0.11
	(3.08)	(2.94)	(-0.59)
Constant	0.041	1.902***	1.938***
	(0.30)	(21.99)	(23.11)
Observations	31,652	31,652	31,652
Calendar Quarter Dummies	YES	YES	YES
Firm FE	YES	YES	YES
Robust t-statistics in parentheses			
*** p<0.01, ** p<0.05, * p<0.1			

Top Decile Extrapolated Return is a dummy variable that equals one if the firm is in the top decile of the extrapolated return measure. The extrapolated return measure is equal to $\sum_{j=1}^{10} \frac{1}{j} R_{t-j}$, where R_{t-j} is the earnings announcement return from period $t-j$. *Test Period* is a dummy variable that equals one if the observation is based on the 5-day period before the earnings announcement. The left-hand-side variable is the log value of total purchases and the log value of total sales in the second and third column, respectively. The left-hand-side variable, in the first column, equals the log value of the absolute value of total purchases minus total sales if total purchases is greater than total sales and the negative of that value if total purchases is less than total sales.

Table 4: 5-Day Return Predictability

	(1)	(2)	(3)	(4)	(5)	(6)
	Previous 5-day Return	Previous 5-day Return	Previous 5-day Return	Previous 5-day Return	Previous 5-day Return	Previous 5-day Return
Top Decile Extrapolated Return	0.42***	0.31***	0.40***	0.30***	0.27**	0.26***
	(4.26)	(3.72)	(3.97)	(3.16)	(3.09)	(3.03)
Bottom Decile Extrapolated Return	-0.06	-0.11	0.03	-0.05	-0.26**	-0.19**
	(-0.61)	(-1.23)	(0.30)	(-0.58)	(-2.06)	(-2.11)
Surprise			2.04***	1.77***	1.92***	1.67***
			(6.05)	(9.10)	(6.05)	(7.94)
Lagged Surprise			-0.659	-0.365	-0.625*	-0.435*
			(-1.56)	(-1.47)	(-1.77)	(-1.73)
PEAD Rank			0.221	0.170	0.272*	0.220**
			(1.38)	(1.58)	(1.73)	(2.03)
Loss Firm			-0.001	-0.001	-0.004***	-0.002**
			(-0.91)	(-1.54)	(-3.35)	(-2.24)
Seasonal Return			0.642	0.101	0.925	0.312
			(1.08)	(0.24)	(1.50)	(0.71)
Size					1.97	1.42
					(0.53)	(0.59)
Book-to-Market					-1.41	-1.53
					(-0.16)	(-0.36)
Earnings Growth					1.46**	1.07***
					(2.23)	(2.80)
Accruals					-0.106	-0.094
					(-1.13)	(-1.32)
Momentum					-0.075*	-0.041*
					(-1.90)	(-1.77)
Volatility					0.051**	0.024
					(2.13)	(1.64)
Volume					0.431	0.536*
					(0.87)	(1.80)
Dispersion					-0.002	-0.008*
					(-0.28)	(-1.95)
Constant	0.003**	0.003***	0.002	0.002	-0.004	-0.001
	(2.24)	(36.23)	(1.20)	(3.85)	(-0.98)	(-0.60)
Observations	185,219	185,171	144,197	144,173	132,557	132,533
Industry-Quarter FE	NO	YES	NO	YES	NO	YES

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The controls are as follows: Surprise and Lagged Surprise are the sell-side earnings announcement surprises from the current and previous quarters, respectively, where surprise is measured as the median difference between actual EPS and the analyst forecast divided by 10000. PEAD Rank is the decile that the firm's earnings surprise was in during the previous calendar quarter divided by 1000. Loss Firm is an indicator that equals 1 if earnings before extraordinary items is less than zero in the current quarter. Seasonal Return is the average of the firm's earnings-announcement return 4 quarters ago and 8 quarters ago divided by 100. Size is the logarithm of the market value of equity as of the beginning of the quarter divided by 10000. Book to Market is the logarithm of the beginning-of-the-quarter book-to-market ratio divided by 10000. Earnings growth is the growth in earnings year over year divided by 1000. Accruals is equal to the quantity of earnings before extraordinary items minus the cash flow divided by ten times the value of lagged total assets. Momentum is the 12-month return starting from month t-13 to month t-2. Volatility equals the standard deviation of monthly returns over the six-month period prior to the earnings announcement. Volume equals one-thousandth of the average monthly volume divided by shares outstanding over the six-month period prior to the earnings announcement. Dispersion is the standard deviation of sell-side analyst forecasts for the quarter. *Top Decile Extrapolated Return*is a dummy variable that equals one if the firm is in the top decile of the extrapolated return measure. The extrapolated return measure is equal to $\sum_{j=1}^{8} \frac{1}{j} R_{t-j}$, where R_{t-j} is the earnings announcement return from period $t-j$.

Table 5: Returns Before, During, and After Earnings Announcements

	(1)	(2)	(3)	(4)	(5)
	Previous 5-day Return	Previous 30-day Return	Earnings Announc. Return	Forward 5-day Return	Forward 30-day Return
Extrapolated Return	0.0346***	0.0013	-0.0011	-0.0313***	-0.0656***
	(3.29)	(0.08)	(-0.10)	(-3.04)	(-3.06)
Size	1.96	2.07	-7.93***	-0.09	-1.41*
	(0.78)	(0.21)	(-3.09)	(-0.03)	(-1.91)
Book-to-Market	-2.52	0.46	0.14	1.40	1.07
	(-0.61)	(0.03)	(0.03)	(0.29)	(0.07)
Surprise	1.56***	6.00***	15.7***	1.83***	2.34***
	(7.96)	(10.17)	(34.99)	(8.27)	(4.64)
Lagged Surprise	-0.28	-2.36**	-3.58***	-0.58**	-0.44
	(-1.13)	(-2.55)	(-12.67)	(-2.39)	(-0.61)
Earnings Growth	0.93**	6.03***	3.71***	2.86***	6.04***
	(2.38)	(5.06)	(5.77)	(4.51)	(5.66)
Accruals	-0.09	0.02	-1.17***	-0.66***	-1.92***
	(-1.22)	(0.08)	(-10.02)	(-7.13)	(-8.86)
Momentum	-0.047*	0.008	-0.020	-0.051**	-0.084
	(-1.96)	(0.09)	(-1.00)	(-2.47)	(-1.58)
Volatility	0.024*	0.145**	0.014	-0.011	0.029
	(1.66)	(2.21)	(1.51)	(-0.88)	(0.93)
Volume	0.53*	-2.66***	-1.07***	0.09	0.10
	(1.78)	(-3.38)	(-3.14)	(0.30)	(0.14)
Dispersion	-0.009**	-0.025*	0.035***	0.017***	0.013
	(-2.53)	(-1.92)	(5.07)	(4.35)	(1.28)
PEAD Rank	0.13	1.09***	-0.24	-0.22**	-0.73***
	(1.20)	(3.00)	(-1.64)	(-2.13)	(-3.22)
Loss Firm	-0.0016*	-0.0130***	-0.0135***	-0.0048***	-0.0091***
	(-1.86)	(-5.28)	(-14.49)	(-5.20)	(-3.47)
Seasonal Return	0.19	1.90*	1.93***	-0.17	-3.25***
	(0.39)	(1.77)	(2.87)	(-0.39)	(-3.35)
Constant	-0.15	-0.15	0.99***	0.45*	2.77***
	(-0.60)	(-0.12)	(4.40)	(1.67)	(4.21)
Observations	119,372	119,372	119,372	119,354	118,825
Industry-Quarter FE	YES	YES	YES	YES	YES

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The constant coefficient is multiplied by 100. The controls are as follows: Surprise and Lagged Surprise are the sell-side earnings announcement surprises from the current and previous quarters, respectively, where surprise is measured as the median difference between actual EPS and the analyst forecast divided by 10000. PEAD Rank is the decile that the firm's earnings surprise was in during the previous calendar quarter divided by 1000. Loss Firm is an indicator that equals 1 if earnings before extraordinary items is less than zero in the current quarter. Seasonal Return is the average of the firm's earnings-announcement return 4 quarters ago and 8 quarters ago divided by 100. Size is the logarithm of the market value of equity as of the beginning of the quarter divided by 10000. Book to Market is the logarithm of the beginning-of-the-quarter book-to-market ratio divided by 10000. Earnings growth is the growth in earnings year over year divided by 1000. Accruals is equal to the quantity of earnings before extraordinary items minus the cash flow divided by ten times the value of lagged total assets. Momentum is the 12-month return starting from month t-13 to month t-2. Volatility equals the standard deviation of monthly returns over the six-month period prior to the earnings announcement. Volume equals one-thousandth of the average monthly volume divided by shares outstanding over the six-month period prior to the earnings announcement. Dispersion is the standard deviation of sell-side analyst forecasts for the quarter. The extrapolated return measure is equal to $\sum_{j=1}^{8} \frac{1}{j} R_{t-j}$, where R_{t-j} is the earnings announcement return from period $t-j$.

Table 6: Returns Before, During, and After Earnings Announcements

	(1)	(2)	(3)	(4)	(5)
	Previous 5-day Return	Previous 30-day Return	Earnings Announc. Return	Forward 5-day Return	Forward 30-day Return
Top Decile Extrapolated Return	0.27***	0.17	-0.02	-0.23***	-0.47***
	(3.16)	(0.73)	(-0.25)	(-3.01)	(-2.50)
Size	1.59	1.39	-7.41***	-0.55	-1.48*
	(0.66)	(0.14)	(-3.07)	(-0.19)	(-1.98)
Book-to-Market	-1.44	5.86	0.91	0.20	-1.76
	(-0.34)	(0.44)	(0.20)	(0.04)	(-0.12)
Surprise	1.66***	6.34***	15.8***	1.87***	2.62***
	(7.91)	(10.26)	(35.81)	(8.94)	(5.21)
Lagged Surprise	-0.43*	-2.91***	-3.69***	-0.70**	-0.45
	(-1.69)	(-2.98)	(-14.10)	(-2.83)	(-0.65)
Earnings Growth	1.07***	6.57***	3.64***	2.51***	5.71***
	(2.82)	(5.54)	(6.04)	(4.10)	(5.58)
Accruals	-0.09	-0.04	-1.09***	-0.63***	-1.82***
	(-1.31)	(-0.23)	(-9.40)	(-7.15)	(-8.18)
Momentum	-0.038*	0.007	-0.019	-0.058***	-0.093*
	(-1.68)	(0.08)	(-1.01)	(-2.95)	(-1.86)
Volatility	0.023*	0.152**	0.014	-0.01	0.03
	(1.58)	(2.24)	(1.56)	(-0.80)	(1.02)
Volume	0.50*	-2.58***	-0.94***	0.18	0.38
	(1.69)	(-3.24)	(-2.78)	(0.60)	(0.53)
Dispersion	-0.008*	-0.034***	0.035***	0.018***	0.017
	(-1.96)	(-2.69)	(5.33)	(4.66)	(1.61)
PEAD Rank	0.24**	1.24***	-0.18	-0.23**	-0.88***
	(2.23)	(3.39)	(-1.26)	(-2.25)	(-3.97)
Loss Firm	-0.0020**	-0.0143***	-0.0132***	-0.0048***	-0.0092***
	(-2.51)	(-6.12)	(-14.49)	(-5.35)	(-3.61)
Seasonal Return	0.39	1.47	1.84***	-0.40	-3.20***
	(0.90)	(1.41)	(2.70)	(-0.91)	(-3.56)
Constant	-0.17	-0.20	0.90***	0.47*	2.83***
	(-0.70)	(-0.17)	(4.23)	(1.81)	(4.27)
Observations	132,533	132,533	132,533	132,514	131,944
Industry-Quarter FE	YES	YES	YES	YES	YES

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The constant coefficient is multiplied by 100. The controls are as follows: Surprise and Lagged Surprise are the sell-side earnings announcement surprises from the current and previous quarters, respectively, where surprise is measured as the median difference between actual EPS and the analyst forecast divided by 10000. PEAD Rank is the decile that the firm's earnings surprise was in during the previous calendar quarter divided by 1000. Loss Firm is an indicator that equals 1 if earnings before extraordinary items is less than zero in the current quarter. Seasonal Return is the average of the firm's earnings-announcement return 4 quarters ago and 8 quarters ago divided by 100. Size is the logarithm of the market value of equity as of the beginning of the quarter divided by 10000. Book to Market is the logarithm of the beginning-of-the-quarter book-to-market ratio divided by 10000. Earnings growth is the growth in earnings year over year divided by 1000. Accruals is equal to the quantity of earnings before extraordinary items minus the cash flow divided by ten times the value of lagged total assets. Momentum is the 12-month return starting from month t-13 to month t-2. Volatility equals the standard deviation of monthly returns over the six-month period prior to the earnings announcement. Volume equals one-thousandth of the average monthly volume divided by shares outstanding over the six-month period prior to the earnings announcement. Dispersion is the standard deviation of sell-side analyst forecasts for the quarter. *Top Decile Extrapolated Return*is a variable that equals one-hundredth if the firm is in the top decile of the extrapolated return measure. The extrapolated return measure is equal to $\sum_{j=1}^{8} \frac{1}{j} R_{t-j}$, where R_{t-j} is the earnings announcement return from period $t-j$.