The Post-Earnings Announcement Drift: An Anomalous Anomaly *

Jonathan N. Katz† Mathew D. McCubbins‡ Jeff L. McMullin§

November 3, 2018

Abstract

Prior literature uses portfolio analysis to document anomalies of the efficient market hypothesis (EMH). We argue portfolio analysis introduces an aggregation bias and clouds inferences about firm-level stock price behavior. We demonstrate this point in the context of the Post-Earnings Announcement Drift (PEAD), a well-known anomaly of the EMH. After replicating the PEAD portfolio analysis, we disaggregate the PEAD portfolios and find anomalies within the PEAD anomaly. Results of our analysis raise concerns over whether the PEAD exists at disaggregated levels, firms' prices actually drift in the direction of earnings news, and whether the PEAD is an anomaly of the EMH or just an artifact of aggregation.

*This project has benefited from discussions with Daniel Beneish, Brian Miller, Brady Twedt, and Jim Wahlen as well as conference participants at the 2016 BYU Accounting Research Symposium.

†Division of the Humanities and Social Sciences, California Institute of Technology, jkatz@caltech.edu

‡Department of Political Science and School of Law, Duke University, mathew.mccubbins@duke.edu

§Kelley School of Business, Indiana University, jemcmull@indiana.edu.

1 Introduction

The Efficient Market Hypothesis (EMH) is one of most tested hypotheses in modern financial economic research. This hypothesis states that "security prices fully reflect all available information," (Fama 1970) and is rejected by identifying profitable portfolio trading strategies.[1] Documenting these anomalous strategies requires aggregation of returns within and across periods. We argue that this introduces an aggregation bias and clouds inference about the veracity of theory. By aggregation bias we mean that inferences made from aggregated data analysis are not representative of underlying firm-level responses. In this paper, we demonstrate that one of the most famous and persistent anomalies, the Post-Earnings Announcement Drift (PEAD), does not persist as portfolio returns are disaggregated.

The PEAD trading strategy prescribes buying (selling) portfolios of firms that release extremely positive (negative) earnings news (unexpected quarterly earnings surprise). The prior literature documents the existence of the PEAD by aggregating these portfolios' returns over many time periods and finds a monotonic relation between the earnings news decile and the subsequent change in the portfolio value. That is, the value of the positive (negative) news portfolio drifts up (down). Buying the positive news portfolio and selling short the negative portfolio earns annual returns between 10% and 25%. From this analysis the literature concludes that a firm's stock price gradually drifts in the direction of the firms' earnings surprise.

Prior research argues the predictable pattern of the PEAD is an anomaly of the EMH. Bernard and Thomas (1989) examine whether the drift patterns have a risk-based explanation consistent with the EMH, but do not find evidence consistent with such an explanation. Interpreting their findings, they explain, "investors...fail to recognize fully the implications of current earnings for future earnings," and conclude that earnings information is not quickly

[1] For reviews see Zacks (2011) and Richardson, Tuna, and Wysocki (2010).

incorporated into price, but rather gradually drifts into price.[2] Subsequent research offers explanations for the existence of the drift.[3] Additionally, studies investigating the price discovery process use the magnitude of the PEAD as a measure of price inefficiency.[4] Implicit in this research is the acceptance of the existence of the PEAD at the firm level. Despite the extensive literature documenting the existence of and explanations for the PEAD what remains unexplored is whether the drift exists when portfolio returns are disaggregated.

To understand whether the PEAD patterns documented by prior research are present at disaggreated levels, we increasingly disaggregate the top (Good News) and bottom (Bad News) unexpected earnings deciles. First, we examine the PEADs of the twenty percentiles in the Good News and Bad News deciles (i.e., 0^{th} to 9^{th} and 90^{th} to 99^{th} percentiles.) We then examine the PEAD by quarter. Finally, we examine the PEADs of individual firms in the Good News and Bad News deciles. While prior literature suggests a monotonic relation between earnings surprise and the size of the firm price drift, we find no simple monotonic relation even after only slight disaggregation. Our analysis reveals disaggregated PEADs are highly heterogeneous and highly variable over time. This heterogeneity grows and monotonicity fades the further we disaggregate. We find firm-level PEADs do not vary in a consistent way with unexpected earnings. Rather our evidence indicates subsequent price responses to earnings surprises are far more complex than the aggregated portfolio analysis suggests. Whereas the PEAD literature concludes underreaction theory explains why aggregated portfolio returns drift, our results demonstrate that the question of why earnings surprise portfolios drift is still open. We conclude that the PEAD does not exist

---

2 More specifically, investors do not understand what the information in a firm's current quarterly earnings surprise implies about the firm's earnings next quarter, especially when the current quarter surprise is extreme.

3 Some explanations include liquidity risk (Sadka 2006), arbitrage risk (Mendenhall 2004), trading costs (Bhushan 1994), and information uncertainty (Francis, LaFond, Olsson, and Schipper 2007). See Taylor (2011) for a review of the PEAD literature.

4 Specifically, these studies examining the processing of earnings news by correlating the magnitude of the firms' price drift with various factors. Some examples of these factors include, investor sophistication (Bartov, Radhakrishnan, and Krinsky 2000 and Hirshleifer, Myers, Myers, and Teoh 2008), attention (Hirshleifer, Lim, and Teoh 2009 and DellaVigna and Pollet 2009), and short selling (Boehmer and Wu 2013 and Fang, Huang, and Karpoff 2016).

when returns are dissaggreated and do not conclude, as many before us have that PEAD, is an anomaly of the EMH, but rather an artifact of aggregation.

Other areas of research within accounting and finance find differences across disaggregated and aggregated analysis. For example, prior research examining the contemporaneous earnings-return relation documents a positive relation at the firm level and a negative relation at the aggregated level (Ball and Brown (1968), Kothari, Lewellen, and Warner (2006), and Sadka and Sadka (2009)). Research examining the relation between analyst recommendations and excess returns documents a positive relation at the aggregate level (Howe, Unlu, and Yan (2009)), but no relation at the firm level (Bradshaw (2004)). However, from our reading, this prior research does not note the potential problems of aggregation bias in the analysis of aggregate data.

In Section 2, we review relevant prior literature. Section 3 discusses aggregation bias in the context of portfolio analysis. The section following discusses models and methods we use to examine the PEAD. Section 5 and Section 6 discuss data and results, respectively. We then conclude in Section 7.

2 Related Literature

The theory of efficient markets proposed in Fama (1965) and Fama (1970) argues actions taken by market participants lead prices of assets to fully reflect all information relevant for pricing the assets. An implication of this theory is that market participants should not earn excess, risk-adjusted returns using readily available information to guide their investment decisions as the information is already impounded into price. Examining this theory, empirical researchers test the Efficient Market Hypothesis (EMH) by examining the profitability of trading strategies based on public information.

Although early empirical evidence provides support of the EMH (Fama, Fisher, Jensen, and Roll 1969), more recently researchers document empirical results anomalous to the theory of market efficiency.[5] These researchers document significantly positive risk-adjusted returns to hedge portfolio trading strategies. De Bondt and Thaler (1987) document an overreaction to past performance where selling past winners and buying past losers earns abnormal returns. Lakonishok, Shleifer, and Vishny (1994) document positive returns to a trading strategy that buys value stocks and sells glamor stocks. Sloan (1996) documents an anomaly where firms with low accruals outperform firms with high accruals. Piotroski (2000) and Mohanram (2005) both develop indexes based on common financial statement variables and show these indexes can be used to construct profitable trading strategies.

Although this empirical evidence is often interpreted to contradict the theory of market efficiency, these anomalous results are often dismissed on account of measuring risk incorrectly. For example, Fama and French (1993), Fama and French (1995), and Fama (1997) dismiss the size and book-to-market anomalies by creating market-wide measures of two risk factors. Additionally, these findings are also dismissed by noting that enough analysis of the data will yield such results Fama (1998). Despite these counter-arguments, Fama (1998) notes that one anomaly remains above suspicion, the post-earnings announcement drift (PEAD).

Initially discovered by Ball and Brown (1968) and later examined in greater detail by Bernard and Thomas (1989) and Bernard and Thomas (1990), the PEAD is a trading strategy based on the news in firms' quarterly earning announcements. In this strategy unexpected earnings or earnings surprise (the difference between actual earnings and expected earnings) is taken as the news signal. Investors implementing this trading strategy buy firms with extremely good news (high unexpected earnings) and sell firms with extremely bad news (low unexpected earnings). The finance and accounting literature is replete with

[5] This evidence is grouped as the behavioral finance literature with the work primarily following from Shiller (1981), the seminal paper in this field. Shleifer (2000) reviews behavioral finance's theoretical and empirical arguments. Recently, Lo (2017) puts forward the Adaptive Markets Hypothesis in an attempt to combine both the EMH and behavioral finance theories.

papers that document the post earnings announcement drift.[6] Despite variation in research design choices, earnings surprise measurement, and risk adjustment methods this literature consistently documents positive average returns to this trading strategy.[7]

Anomaly papers typically offer an under-reaction or over-reaction theory as an alternative to the efficient markets theory. In describing why the PEAD exisits, Bernard and Thomas (1989) explain that investors fail "to recognize fully the implications of current earnings for future earnings". In other words, investors underreact to the information in current earnings that informs predictions of future earnings. This theory describes how individual investors examine news about a firm when making investment decisions.[8] Interestingly, to provide empirical support for these individual-level theories, researchers focus on aggregated portfolio analysis. We argue using portfolio analysis to test firm-level theory offered to explain the drift may lead to incorrect ecological inferences.

We are not the first to raise concerns about making inferences about firm-level theory by documenting the profitability of a portfolio trading strategy. Kraft, Leone, and Wasley (2006, p. 299, their emphasis) "highlight and quantify the impact of robustness tests on *causal* inferences drawn from trading strategies based on accrual-related information". They find that a simple robustness test of truncating the top and bottom of the sample based on returns at 1% reveals no evidence of investors fixation on earnings, the theory offered by this literature to explain the accrual anomaly. From this evidence, Kraft, Leone, and Wasley (2006) argue researchers should perform robustness tests of portfolio analysis when

[6] For brevity, we do not cite all of these papers. Rather, the interested reader should turn to the recent survey, Taylor (2011).

[7] Some research design choices that vary within the PEAD literature include holding period and rebalancing frequency. Earnings surprise is measured by some papers as the difference between prior year's same quarter's earnings and actual earnings, while other papers use some form of a consensus analyst earnings forecast as the expected earnings measure. To adjust returns for risk, some papers use a market model, while others adjust returns using a reference portfolio, still others use a version of the Fama-French multiple factor model.

[8] Barberis, Shleifer, and Vishny (1998) and Daniel, Hirshleifer, and Subrahmanyam (1997) are two papers that attempt to build theories to replace the efficient markets theory. However, as Fama (1998) explains, these "behavioral models work well on the anomalies they are designed to explain."

making causal inferences. In contrast, we argue the validity of firm-level theories should not be tested by estimating the average return to portfolio strategies.

Prior research raise concerns about making inferences from aggregated statistics (e.g. averages) about individual firm's market responses.[9] Bamber, Christensen, and Gaver (2000) revisit Beaver (1968), a seminal paper about market reactions to earnings announcements. In describing how subsequent research interpreted Beaver (1968)'s findings, Bamber, Christensen, and Gaver (2000) (p. 104, their emphasis) explain "*a focus on mean effects obscures the fact that most individual earnings announcements are not associated with unusual price reactions*". We likewise argue that a focus on the aggregated portfolio returns obscures the fact that practically all firm-specific drifts do not follow the portfolio drift path.

Our findings also contribute to our understanding of prior literature that examines the implementability of the anomalies documented in academic research (see Hirshleifer, Teoh, and Yu (2011), Drechsler and Drechsler (2014), and Beneish, Lee, and Nichols (2015)). Generally, this literature finds that many of the documented anomalies are difficult to implement. Finding an abnormal return to a trading strategy using *ex post* data is one thing; however, actually implementing the strategy using real investments requires overcoming an expanded set of constraints, such as short selling and liquidity constraints. Thus, the papers examining implementability of trading strategies, and the results of our study indicate that our empirical and theoretical understanding of the anomalies may need to be enhanced. This improved understanding would allow for improvements in theories that explain the market's response to news and identification of profitable trading strategies.

[9] Studies outside the accounting and finance literature also document problems with making inferences from averages (i.e., ecological inferences). Schuessler (1999) defines ecological inference as "the process of using aggregate data to draw conclusions about individual-level behavior." The ecological inference problem was initially demonstrated by Robinson (1950) who documented that the correlation between literacy and race falls from 0.946 to 0.773 and then to 0.203 as the data are disaggregated from the nine geographic regions to the 50 states and then to the individual level. Schuessler (1999) explains, "The fundamental difficulty with such inferences is that many different possible relationships at the individual level can generate the same observation at the aggregate level."

3 Aggregation Bias

Aggregation bias is the difference between effects estimated with aggregate data and effects estimated with individual-level data (Stoker 2009). This differences increases as the heterogeneity of the aggregated individuals grows. Stoker (2009) notes, “Aggregates reflect a smear of individual responses,” and “without careful attention, the smear is unpredictable and uninterpretable.” Luloff and Greenwood (1980) explore the effect of aggregation and document aggregation can result in increases in, decreases in, or no effect on the coefficient of determination (R^2) as well as variation in the sign, magnitude, and statistical significance of coefficients. However, if this aggregation bias is insignificant, results will not vary at different aggregation levels, thus allowing analysis at the aggregate level to inform individual level behavior. However, if aggregation bias is present but painlessly assumed away, this costly assumption results in biased and distorted inferences of the veracity of theory.

Using portfolio analysis to test individual level theories may be sensitive to an aggregation bias. While we understand that traders form portfolios when investing, researchers often use portfolio analysis to examine theories of effect of *individual* investors behavior on *single* firm's security price. By performing portfolio analysis, researchers may conclude strong correlations at the portfolio level confirm individual level theories, when in fact the correlation at the individual level may be much lower. Stated differently, by aggregating firms returns to the portfolio level, researchers may incorrectly conclude that the correlations at the aggregate level persist at disaggreated levels.

Other fields of social science have recognized the problem of drawing inferences about individuals from aggregate data and labeled it the “ecological fallacy.” In these fields, unlike when examining the stock market, disaggregated data is often unavailable. As a result, researchers in these other fields have developed “ecological inference” as a tool for overcoming the lack of data at the individual level. Schuessler (1999) explains that, “Ecological inference is the process of drawing conclusions about individual-level behavior from aggregate-level data".[10] This type of inference is typically attempted when researchers only have aggregate-level data, but are interested in individual-level phenomenon. For example, when researchers only have precinct level voting results, but are interested in the impact of an individual's race on her voting behavior, they may turn to ecological inference to inform such questions. However, the conditions under which valid ecological inferences are possible are quite narrow and typically require strong assumptions.

The problem of aggregation bias was not broadly recognized until Robinson (1950). Using the correlation between race and illiteracy to demonstrate his point, Robinson (1950) notes while the conditions are theoretically possible where the correlation at aggregated levels could be used as a substitute for an individual-level correlation, "the conditions under which this can happen are far removed from those ordinarily encountered in data." He goes on to argue researchers should never use aggregate data to make inferences at the individual level. King (1997) explains, "[Robinson (1950)] remains one of the most influential works in social science methodology," and greatly impacted future research.[11]

The concerns raised about making ecological inferences apply to portfolio analysis. By focusing on the aggregated portfolio returns, researchers may conclude the results from the aggregated analysis reflect disaggregated returns. However, such conclusions may not be accurate. Although the portfolio-level estimates may appear to confirm investor-level theories, such portfolio-level estimates may result from an infinite number of combinations of individual-level returns that may not reflect the individual-level theory.

[10] At least since Ogburn and Goltra (1919) political science researchers have noted the ecological inference problem. In their article, Ogburn and Goltra (1919) attempt to infer the voting behavior of women from precinct level data while noting, "the method of voting makes it impossible to count women's votes, one wonders if there is not some indirect method of solving the problem." These types of "indirect methods" to make inferences(i.e., ecological inferences) were used in some of the earliest statistical works (Graunt (1662), Petty (1690), and Petty (1691)).

[11] The response to Robinson (1950) by researchers was two fold. First, work using aggregate statistics declined being replaced by other methods, such as survey analysis (see Achen and Shively (1995)). In fact, many important research questions received little attention due to researchers avoiding the use of aggregate data (King (1990)). Second, researchers went to work either explaining away the problem of ecological inference or proposing solutions to the problem. King (1997) describes a solution to the ecological inference problem by outlining the conditions required to allow a researcher to make ecological inferences. These conditions require assumptions that are difficult to satisfy.

As an example, consider a portfolio of ten firms with an average return of 5%. An infinite number of combinations of ten firm-specific returns exist that average to a portfolio return of 5%. A researcher focusing only on the portfolio's average return would make the same inference from each of the infinite combinations of firm specific returns, regardless of the variability of the firm-specific returns within each combination.[12] While it is possible that firm-specific returns are consistent with the conclusions made from the aggregated analysis (e.g., all ten firms experience a 5% return); it may also be the case that results at the aggregated level are not representative of firm-level returns. This is an example of the potential problems of making inferences from aggregated data (Schuessler 1999). In the context of the PEAD, without disaggregating the returns of the Good News and Bad News portfolios, researchers may incorrectly infer that patterns documented at aggregated levels persist at disaggregated levels.

We argue an examination of the firm-level returns is required to draw tight conclusions about firm-level theories and whether the empirical patterns documented at the aggregate level are representative of returns at the firm level. To highlight the importance of this disaggregted analysis we extend the example from above. Using the following four cases of the ten-firm portfolio described above, we exemplify the infinitude of possible combinations of individual returns that, on average, yield a 5% return:

- Case 1: 10 firms' return = 5%
- Case 2: 9 firms' return = 0%, 1 firm's return = 50%
- Case 3: 5 firms' return = -5%, 5 firms' return = 15%
- Case 4: Value weighted Portfolio
  - 9 small ($10M market cap each) firms' return = 0%

[12] We note that requiring statistical significance of the portfolio return excludes a number of these possibilities. Nonetheless, this subset remains infinite.

– 1 large ($45B market cap each) firms' return = 5.01%

While in each of these cases only observing the portfolio level average would lead researchers to identical conclusions, the inference one would draw about firm-level and investor-level behavior after analyzing the firm-level returns of these four examples is vastly different.[13] In Case 1, researchers would conclude there is a homogeneous effect (i.e., firms' responses to the treatment are an identical 5% increase). Under Case 2, a researcher would note that generally the news had no effect, except for one firm, in which case the effect was extremely large. An examination of the firm-level returns would uncover the outlier firm. Case 3 exemplifies when the reaction to treatment is very heterogeneous. That is, some firms are positively impacted while other firms are negatively impacted by the treatment. In cases like this, inferring an average treatment effect of 5% from portfolio analysis can easily be seen as making inferences not supported by the data. Case 4 demonstrates the impact weighting returns by market capitalization can have on portfolio analysis. In this case, only the large firm responded to the treatment, but portfolio analysis would conclude that all firms large and small responded with a return of 5%. Thus, the portfolio average is representative of the individual firms' responses in only Case 1.

The problems of relying on the average of a portfolio to test firm-level theory can also be seen through a simple example. Suppose a medical study examines the impact of a new drug on the health of the study's participants. For simplicity, assume the average improvement in heath for participants that take the new drug is 5%. Similar to Case 1 above, if all individuals' health level increases by 5%, the conclusion that the drug improves health is warranted. However, if individuals' responses to the drug are akin to Case 3, the drug only improves the health of half of the participants and while it decreases the health of the other

[13] Of course, researchers would consider statistical significance of these averages which may lead to differing conclusions among these four stylized examples. Nonetheless, similar cases could be found where average returns are identical and statistical significance results in the same inference, but the underlying firms' returns differ greatly.

half. It is clear to see the problems of doctors making conclusions about the impact of the drug on health from the average response alone.

By examining the individual firms returns researchers avoid making ecological inferences. Interestingly however, the literature tends to not examine the firm-specific returns in detail. This is in spite of MacKinlay (1997) who argues that after estimating the average return, researchers should calculate the percent positive and percent negative returns to firms in the portfolio to help understand the extent to which the effect is homogeneous in direction. This type of analysis, if performed, is rarely reported.

Despite having access to daily firm-level returns and even data capturing the individual trades and quotes made in the market, the typical analysis of returns data is aggregated to a portfolio level and estimated over monthly intervals. This is unlike many political science settings where individual-level voting behavior is not observable, thus requiring ecological inference or some other technique. Thus, these daily firm-level data make possible a more careful examination of returns at disaggregated levels. We argue that rather than continuing to test individual-level theory by performing portfolio analysis, future research should focus on developing new and applying existing methodologies to exploit the rich set of disaggregated stock market data that exists.

4 Models and Methods

The typical research design used to examine the EMH and theories of anomalies is the long-run event study methodology.[14] This methodology was developed by Fama, Fisher, Jensen, and Roll (1969) to examine the market response to the announcement of stock splits. To perform a long-run event study, researchers first identify an event common to many firms. These "event firms" are grouped together into a portfolio. Next, researchers calculate the

[14] This long-run event study methodology is explained in greater detail in Kothari and Warner (2006). Barber and Lyon (1997) and Kothari and Warner (1997) examine the power and specification of the test statistics that are commonly used.

risk-adjusted return to those firms experiencing the event. A modification of this method is the hedge trading strategy where the firms from two extremes of the variable of interest form two distinct portfolios. Firms in one extreme are purchased and those in the other extreme are sold short. The combination of these two portfolios is often referred to as a zero-investment portfolio. The primary estimand of these methods is the average return to these portfolios. Finding a non-zero, average return to the event-firm portfolio or the zero-investment portfolio is taken as evidence either confirming or rejecting theory. In the case of the EMH, abnormally positive average returns to such trading strategies are taken as evidence that contradicts the EMH.

In using this design, researchers make assumptions about the exchangeability of the firm-year observations, the similarity of the news signals, and time. In essence, researchers assume that firm-quarters are exchangeable, information signals are identical, and time is ignorable. Violations of these assumptions lead to variance in firms' returns. Consequently, when firm returns are aggregated to the portfolio level to test hypotheses and theory, whether the portfolio returns are due to the theory being tested or violations of these assumptions is uncertain.[15] Violations of these assumptions decrease the plausibility that researchers can appropriately use ecological inference to test individual-level theory.

Although long-run event study methodology is widely accepted in tests of the EMH and used extensively in the accounting and finance literature, these assumptions are rarely mentioned or examined. In a similar way to how the literature views test of market efficiency as joint tests of market efficiency and the expected return model, we argue that these assumptions are part of the joint test. Thus, the long-run event study methodology is a joint test between the theory being examined, the risk-adjust method being used, and these

[15] The long run event methodology requires researchers to adjust returns for risk. Variations of two methods are typically used, the reference portfolio approach (Daniel, Grinblatt, Titman, and Wermers (1997)) or the multi-factor model approach (Fama and French (1993)). Under the reference portfolio approach, returns are adjusted by subtracting the return to a portfolio of relatively similar firm. The multi-factor approach adjusts returns for market-wide risk factors. Our points regarding the problem of making inferences about firm-level theory from portfolio-level analysis are valid regardless of the risk-adjustment methodology used to adjust returns.

assumptions. In the case of the PEAD, we offer empirical evidence that challenges these assumptions. We argue that failure to meet these assumptions calls into question whether this research design does test the EMH, whether the PEAD estimates are valid, and whether this research design allows researchers to make inference about individual-level theory (e.g., over reaction or under reaction to information explanations of anomalies).

We argue that when researchers posit individual-level theory, such theory should be tested at the individual level, rather than at the aggregated portfolio level, especially when individual data is readily accessible. This approach avoids the problems of ecological inference. Something as simple as examining the percent positive and percent negative of firms in the portfolio is a good first step in understanding whether the portfolio average effect can be used to infer the individual level responses.

The tendency of researchers to utilize portfolio analysis to understand the stock market is likely primarily due to the theoretical underpinnings of asset pricing and the benefits of diversification. Additionally, one might argue that aggregating returns to the portfolio level cancels noise in returns, thus allowing the portfolio return to represent the response to the earnings news. In this context noise in returns reflects a price changes that are not a result of the firm's earnings announcement. For example, this noise may be due to the release of other news by the firm itself, other firms in the market, or economy wide news as well as the effects of noise traders. However, achieving such noise cancellation via aggregations requires strong assumptions about this noise. These assumptions include that the noise in each firm's return must occur completely at random and must not be correlated with the noise in other firms' returns. That is, observing one stock's return must not improve our expectations about another stock's return.

The evidence in the prior literature questions the tenability of assumptions required for aggregation to cancel out price movements unrelated to earnings news. For example, a set of studies document the comovement of stock prices (Roll 1988; Morck, Yeung, and Yu 2000; Barberis, Shleifer, and Wurgler 2005; Wahal and Yavuz 2013; Chan, Hameed, and Kang 2013; and Honghui, Singal, and Whitelaw 2016). Also, many studies document return correlation across firms that differ on size (Lo and MacKinlay 1990; Hou 2007), analyst following (Brennan, Jegadeesh, and Swaminathan 1993), institutional ownership (Badrinath, Kale, and Noe 1995), volume (Chordia and Swaminathan 2000), and business complexity (Cohen and Lou 2012). More directly related to the earnings announcement setting, Ramnath (2002) documents information transfers from firms that announce earnings early to firms that announce later. Thus, the argument that aggregation removes “noise” in returns is suspect.

Since modern portfolio theory was introduced Markowitz (1952) and since the long-window event study methodology was created Fama, Fisher, Jensen, and Roll (1969), statistical techniques have expanded to allow for better examination of large, complex datasets. These new tools will likely allow researchers to uncover new insights and bolster the credibility of firm-level inferences.

5 Data

In this section we discuss the data collection and variable measurement. The PEAD is well studied and replicated since originally discovered by Ball and Brown (1968). There are often minor variations across these replications in how the earnings surprise is defined, how returns are adjusted for risk, and how long the trading strategy is held. Although these variations produce different estimates of the PEAD, they are unlikely to impact the conclusions we draw from our empirical analysis. For simplicity, we therefore follow a typical method to estimate the PEAD.

Our data come from three separate datasets. Specifically, we use data from Thomson Reuters' Institutional Brokers' Estimate System (IBES), the Center for Research in Security Prices' U.S. Stock databases (CRSP), and Standard & Poor's Compustat North America Fundamentals dataset (Compustat). These datasets provide analyst estimates, security prices, and company characteristics, respectively. The intersection of these datasets allow for a sample period from the first quarter of 1985 to the second quarter of 2014 and a sample size of 243,826 firm quarters. Descriptive statistics of the following variables, except for returns, are in Table 1.

To estimate PEAD, we first measure *Earnings Surprise*. This measure captures the amount of new information in reported earnings and is measured as the difference between actual earnings and expected earnings. While expected earnings is the level of earnings the market expects the company to report, actual earnings is the earnings figure reported by the company. Thus, earnings surprise captures the amount of earnings reported by company that the market did not expect. If this surprise is positive (actual earnings is greater than expected earnings), the surprise is considered good news. If it is negative, it is considered bad news.

Researchers calculate expected earnings in various ways. For example, Foster, Olsen, and Shevlin (1984) and Bernard and Thomas (1989) estimate expected earnings as a seasonal random walk (i.e., earnings from the prior year's same quarter are used for estimate of expected quarterly earnings); and Doyle, Lundholm, and Soliman (2006) and Livnat and Mendenhall (2006) use equity analysts estimates from IBES. Brandt, Kishore, Santa-Clara, and Venkatachalam (2008) avoid measuring expected earnings directly by measuring *Earnings Surprise* as the announcement day stock returns. We follow a common method and use analysts' estimates from IBES to measure the market expectation of earnings. Specifically, to measure *Expected Earnings* we first identify all analyst earnings forecasts in the 90 days prior to the earnings announcement. If an analyst updates her earnings forecasts during this 90 day period, we exclude the early earnings forecasts and only use the analyst's forecast closest to the actual earnings announcement date. From these remaining forecasts, we calculate *Expected Earnings* as the median analyst forecast.

Actual earnings is measured by prior studies using either the Compustat or the IBES dataset as both datasets have a measure of actual earnings. While Livnat and Mendenhall (2006) discuss reasons why actual earnings in these two datasets might differ, and while

Ljungqvist, Malloy, and Marston (2009) document systematic adjustments made to actual earnings in IBES, both measures have their trade-offs. (See Bradshaw and Sloan (2002) for a discussion of these issues.) Noting the presence of these tradeoffs, we use the actual earnings variable in the IBES dataset to measure *Actual Earnings*.[16] As mentioned above, although these measurement variations are the focus of important papers in the PEAD literature, these measurement choices are unlikely to impact the inferences we make from the analysis.[17]

To calculate the unexpected earnings level or *Earnings Surprise*, we subtract *Actual Earnings* from *Expected Earnings*. Following the prior literature, we standardize this measure of earnings surprise by dividing by the company's stock price at the end of the calendar year. As does the prior literature, we refer to this measure of standardized unexpected earnings as *SUE*.

$$SUE = \frac{ActualEarnings - ExpectedEarnings}{Price} \tag{1}$$

Estimating the returns to the PEAD trading strategy requires assigning firms with similar earnings surprises into *SUE* decile portfolios. Since ranking all firms based on the current quarter's *SUE* distribution injects a look-a-head bias into this trading strategy, prior research typically identifies *SUE* decile cutoffs based on the distribution of *SUE* in the prior quarter. We follow this design and calculate decile cutoffs (10^{th} percentile, 20^{th} percentile, etc.) based on the prior quarter's distribution of *SUE*. Firm-quarters where the *SUE* measure is above the 90^{th} percentile cutoff are part of the "Good News" portfolio. Firm-quarters where the *SUE* measure is below the 10^{th} percentile cutoff are part of the "Bad News" portfolio. The PEAD trading strategy buys firms in the "Good News" portfolio and sells short the firms in the "Bad News" portfolio. The PEAD estimate is the returns to this trading strategy. Although the prior literature examines holding period ranging from a month to three years,

[16] We adjust the forecasts for stock splits using the adjustment factor variables in the CRSP dataset. This puts actual and estimated earnings on the same basis (i.e., same number of shares).

[17] Livnat and Mendenhall (2006) compare measuring expected earnings from analyst forecasts of earnings and time series forecasts and argue that the PEADs estimated using each measure are distinct.

the typical holding period examined is usually over the next quarter, or three months. Thus, we calculate the buy and hold return for each firm over the 13 weeks after the earnings announcement.

$$RetBH_{it} = \prod_{t=1}^{T}[1 + R_{it}] \tag{2}$$

To adjust each firm's return for risk, we subtract the return to a reference portfolio of firms of similar size and similar book-to-market ratio R_{rp}. Specifically, following Daniel, Grinblatt, Titman, and Wermers (1997) we sort all firm-quarters into size quintiles and book-to-market quintiles. The reference portfolio for a firm is defined as all firms in the same book-to-market to market quintile.

$$AbnRet_{it} = RetBH_{it} - \prod_{t=1}^{T}[1 + R_{rp}] \tag{3}$$

We use data from Compustat to measure firm characteristics. To measure firm size *Market Value of Equity* we multiple the stock price at the end of the fiscal year (Compustat item prcc_f) by the number of common shares outstanding (Compustat item csho). We measure the Book-to-Market ratio as a firm's book value of common equity divided by its market value of common equity. Book value of common equity is Common/Ordinary Equity (Compustat item ceq) plus Deferred Taxes and Investment Tax Credit (Compustat item txditc) less preferred stockholders equity.[18]

[18] We define preferred stockholders equity as the first non-missing value of Compustat items pstkrv, pstkl, or pstk in this order. If all are missing, we set preferred stockholders equity to zero.

6 Results

6.1 PEAD Replication

To begin our analysis, we follow prior research and estimate the post-earnings announcement drift. Using standardized unexpected earnings *SUE*, we sort firms into *SUE* decile portfolios using cutoffs based on the prior quarter's distribution of *SUE*. We create the variable *SUE_Decile* to capture which decile portfolio the firm is assigned. For example, if a firm's value of *SUE* for the quarter was in between the prior quarter's 10^{th} percentile and 20^{th} percentile of *SUE* the firm would be assigned to *SUE_Decile* 2. We identify firms with "Good News" and firms with "Bad News" as firms in the decile ten and decile zero, respectively.[19]

Figure 1 depicts the returns for all ten *SUE* deciles over the 13 weeks following the earnings announcement. For each firm in these decile portfolios, we calculate the size and book-to-market reference portfolio adjusted buy-and-hold cumulative return from one day following the announcement of earnings to the end of each of the next 13 weeks. Using these returns, we calculate the 13 average cumulative abnormal return estimates of each decile portfolio and draw a line representing the decile portfolios' drifts. This figure is of the same spirit of the graphs produced by Ball and Brown (1968) and Bernard and Thomas (1989), some of the most well know figures produced by accounting academic researchers in the past 50 years.

Table 2 shows the mean and standard deviation of returns to the ten decile portfolios the over three months following the earnings announcement. The returns to these portfolios generally show the monotonically increasing pattern going from the lowest SUE decile to the highest SUE decile. Table 3 shows the t-statistics and corresponding p-values testing whether the average returns in the ten portfolios are significantly different from zero. The PEAD trading strategy meant to exploit this predictable pattern of returns requires buying

[19] Specifically, to determine if a firm's SUE is good or bad news we define cuttoffs based on the prior quarter's distribution of *SUE*. If a firm's *SUE* is greater (less) than the 90th (10th) percentile of prior quarter's *SUE* distribution, then the firm is included in the good (bad) news portfolio.

"Good News" firms and selling "Bad News" firms. Table 2 reports this hedge trading strategy results in a risk adjusted return of 5.1% over three months following the earnings announcement, respectively. This translates into an annual return of over 20% and is inline with prior research that documents the annual return to the PEAD strategy being approximately between 8.76% (Sadka (2006)) and 43.08% (Battalio and Mendenhall (2007)).

The predictable pattern of portfolio drifts depicted in Figure 1 and shown in Table 2 and the abnormal returns to the PEAD trading strategy are inconsistent with the theory of market efficiency. Bernard and Thomas (1989) carefully examine the PEAD to determine if these drifts can be explained by a mismeasurement of risk, excessive transaction costs, or by investors' underreaction to information. Analyzing the "Good News" and "Bad News" portfolios, they produce evidence the pattern is consistent with the investor underreaction theory. This underreaction theory argues that individual investors fail to immediately and completely respond to the news in a single firm's earnings surprise and that the full pricing impact of the earnings surprise is gradually incorporated into price. Bernard and Thomas (1989) explain that "some investors may fail to form an unbiased expectation of future earnings immediately upon revelation of current earnings," and that the full price reaction to the earnings news is gradually incorporated in the months following the earnings announcement. The figures produced by Ball and Brown (1968) and Bernard and Thomas (1989) and our Figure 1 show *portfolio* drifts that seem consistent with this underreaction theory. However, whether this pattern persists as we disaggregate the returns cannot be determined from this analysis. Additionally, although these aggregated portfolio patterns are consistent with the underreaction theory, whether these patterns are observed at the individual firm stock price drift level remains to be tested.

6.2 PEAD Disaggregation

In this section, we disaggregate the PEAD at increasing levels of disaggregation. We first form smaller portfolios (percentiles, rather than deciles). Then, we dissaggregate by time.

Finally, we disaggregate the PEAD to the firm level. This disaggregation reveals anomalies within the PEAD anomally. Primarily, we find evidence that questions the conclusions from the aggregated PEAD analysis that the PEAD is an anomaly of the EMH and that investors underreact to the news in earnings.

6.2.1 PEAD Percentile Portfolios

Disaggregting the PEAD into smaller portfolios allows us to examine the impact of a slight change to how PEAD is traditionally calculated. We disaggregate the PEAD into smaller portfolios by assigning each firm-quarter into a PEAD *percentile*. For example, if a firm's value of *SUE* for the quarter was in between the prior quarter's 15^{th} percentile and 16^{th} percentile of *SUE* the firm would be assigned to *SUE_Percentile* 15. In Figure 2, we plot the PEAD for the percentile portfolios in the Good News Decile (percentiles 90 through 99) and Bad News Decile (percentiles 0 through 9) portfolios. At this level of disaggregation, we find the monotonic return drift patterns traditionally associated with the PEAD do not persist. For example, we find the *SUE_Percentile* portolio with the most negative news (portfolio 0), experiences the least negative drift of all percentile portfolios in the Bad News Decile portfolio. According this analysis, a refined PEAD trading strategy would be to buy firms in *SUE_Percentile* 99 and sell firms in *SUE_Percentile* 3. This strategy would return around 10% per quarter. However, within the underraction theory, we know of no reason to predict ex-ante that *SUE_Percentile* 3 should result in the most negative drift.

6.2.2 Disaggregating PEAD by Time

Next, we disaggregate the PEAD by time. In Figure 1, each drift line is the average of 118 portfolio drifts since our data consist of 118 quarters. We disaggregate these average drifts by plotting Good News and Bad News Decile portfolio drifts for each of the 118 quarters. Panel A and Panel B depict the quarterly drifts for the Good News Decile and Bad News Decile portfolios, respectively. These graphs reveal significant heterogeneity around the average drifts. Specifically, for the Good News drifts we find that 16.1% of drifts (or, 19 quarters) actually experience negative returns. While the average return to the Good News portfolio over the 118 quarters is 3.3%, the standard deviation is 3.8%. During one quarter, a Good News portfolio experienced a -8% return. Turning to the Bad News portfolios, we find 28.0% of drifts (or, 33 quarters) experience positive drift, with one Bad-News decile portfolio experiencing a return of 16.4%. The average return to the Bad News Portfolio is -1.9% and the standard deviation is 4.5%. This heterogeneity demonstrates that portfolios do not tend to drift closely along the average portfolio drift lines depicted in Figure 1.

We also disaggregate the hedge trading strategy by time and examine each quarter, rather than all 118 quarters jointly. We plot in Figure 4 the average return as well as the confidence interval for each of the 118 quarters in our sample. Although in 95% of the quarters the difference between the Good-News and Bad-News portfolios is positive, our analysis reveals these portfolios contain firms returns of significant heterogeneity and the average return is significantly different from zero in less than half of the quarters. Specifically, we calculate the average of the firm return for each firm in this trading strategy by quarter and test whether this average is significantly different from zero. Given that we perform multiple t-tests, we adjust the typical level of significance (p-value = 5%) using the Bonferroni correction. This correction reduces false discovery risk. Since we perform 118 t-tests, the corrected threshold is 0.04% (5%/118). We find that in only 13.6% of quarters (16 quarters) this average is significantly different from zero.[20] Thus, by disaggregating by time and examining the statistical significance, we document that the PEAD is present in only a small number of quarters.

[20] This percentage falls to 5.9% (7 quarters) if we use the more conservative statistically significance threshold 1%, corrected to 0.0085% (1%/118).

6.2.3 Firm-Level PEAD Analysis

In our next analysis, we disaggregate PEAD returns to the firm-quarter level. At this level of disaggregation we can directly examine whether firm returns tend to follow the patterns observed at highly aggregated levels and predicted by theory. To begin, we create simulated firm-level drift patterns. Specifically, we simulate firm specific drifts by adding some noise to the Good News and Bad News portfolio drifts patterns in Figure 1. We generate 500 simulated firm specific drifts based on the "Good News" and "Bad News" portfolios' drifts. The average drift path of these *simulated* firm-level drifts is the average of the *actual* "Good News" and "Bad News" portfolio drifts. We depict these simulated firm-specific drifts and the actual "Good News" and "Bad News" portfolio drifts in Figure 5.

We next turn to analyzing actual firm-level returns. Finding that actual firm-level drifts are generally similar to the simulated firm-level drift lines in Figure 5 would be evidence supporting the individual-level underreaction theory posited as an explanation for the PEAD anomaly. Moreover, this would lend credibility to the conclusions about firm-level returns drawn from aggregated portfolio analysis. However, as mentioned above, the actual *portfolio* drifts can result from an infinitude of possible firm-specific drifts. Unlike other settings where ecological inferences rely on aggregate measures since aggregate data is all that is available, data to measure firm-specific drifts is readily available. Thus, examining firm-specific drifts allows us to assess the extent to which the aggregate outcomes (portfolio drifts) are good estimates of the individual outcomes (firm-specific drifts). In other words, comparing actual firm-specific drifts to portfolio drifts helps determine the validity of ecological inferences made about individual firms when doing portfolio analysis and allows us to gauge whether it is reasonable to use portfolio level estimates to make inferences about individual level theory.

We measure firm-specific returns from two days following the announcement of earnings to the end of each of the next 13 weeks. To facilitate depicting these drifts we randomly select 200 firm-specific drifts from both the "Good News" and "Bad News" portfolios. We plot these firm specific drifts in Figure 6. Figure 6 show that firm drifts rarely, if ever, follow the well-known drift patterns associated with the PEAD anomaly. Given that many of the firm-specific drifts go beyond the range of the Y-axis in Figure 6, we increase the range of the Y-axis by a factor of ten and plot the data in Figure 6 to create Figure 7.

Given the erratic firm-specific drifts in Figures 6 and 7, it is difficult to determine the distribution of the firm-specific buy-and-hold returns from one day after the announcement to 13 weeks later. To better see the total return over 13 weeks for firms in "Good News" and "Bad News" deciles, we randomly select and plot 800 firm specific drifts from each extreme decile and draw a line from 0% on two days following the earnings announcement to the firm's buy-and-hold abnormal return 13 weeks later. This removes the variance in the returns across the 13 weeks and allows for a better assessment of the distribution of returns after 13 weeks. Figure 8 plots the firm specific drifts for 800 randomly selected firm-quarters in the "Good News" ("Bad News") decile. These plots further demonstrate that firm-specific drifts only rarely follow the portfolio drift lines depicted in the well-known PEAD graphs of Ball and Brown (1968) and Bernard and Thomas (1989).

MacKinlay (1997) suggests examining the firm-specific returns of firms in the portfolio. Specifically, this approach centers on comparing the percent of observations with positive returns to the percent of observations with negative returns. In Table 3 we report the percent of 13-week returns that are positive as well as the percent of 13-week returns that are negative for each PEAD decile. In the "Good News" portfolio, 51.8% of observations have positive returns and 48.2% of observations have negative returns. In the "Bad News" portfolio, 41.7% of observations have positive returns and 58.3% of observations have negative returns. Put simply, the well-known portfolio drifts in Figure 1, are not present at the firm return level and raises concern about making inferences of the exisitence of the PEAD at the firm-level. Additionally, the underreaction theory incorrectly describes close to half of the observations. This highlights the potential problems of using portfolio analysis to test firm-level theory offered to explain the PEAD. That is, despite finding the *average* returns to the "Good News" and “Bad News” portfolios are consistent with the direction of the theory, an examination of firm-level data (the level the theory is meant to explain) does not support similar conclusions.

Our next analysis is motivated by the prior research’s conclusion that prices gradually “drift” in the direction of the earnings surprise. It is also motivated by the definition of drift, “a continuous movement from one place to another”. If firm returns are drifting, we expect ”Good News“ (“Bad News”) portfolio firm experience consistently positive (negative) returns over the following 13 weeks. To assess whether firms returns are indeed “drifting,” we disaggregate the 3-month return into 13 weekly returns. If firms’ stock prices are drifting, we should observe that firms in the “Good News” (“Bad News”) decile experience 13 (0) positive weeks. For each firm, we calculate the percentage of the 13-weeks that experience a positive stock price return.[21] In Table 4, we report that only 17.3% of “Good News” decile firms have 13 weeks of positive returns. For the “Bad News” decile, 21.0% of firms had 13 weeks with negative returns. This means that close to 83% of firms in the “Good News” portfolio and 79% of firms in the “Bad News” portfolio do not experience the gradual stock price drift as depicted in Figure 1. Moreover, we find close to 15% of “Good News” firms had no weeks with positive returns and over 11% of “Bad News” firms had 13 weeks with only positive returns. This analysis provides evidence the stock price drifts do not consistently move in the direction depicted in Figure 1 and predicted by underreaction theory.

6.3 Correlation between SUE and Returns

To better examine the association between SUE and post-earnings announcement returns we calculate the correlation between these two pivotal variables to the PEAD literature. Table 5 reports the correlation between measures of earnings surprise (SUE, $\text{SUE}_{\text{Decile}}$, and $\text{SUE}_{\text{Percentile}}$) and 3-Months risk-adjusted returns. Prior research’s empirical evidence and the underreaction theory predicts a positive association between earnings surprise and returns. Table 5 reports the correlation between the average SUE and average return for two levels of

[21] We examine returns at a weekly time frame as a smaller time frame (e.g., daily) may capture noise.

portfolio aggregation, SUE deciles and SUE percentiles. The correlation of the 10 values of the mean SUE within each SUE decile and average returns for the 10 SUE decile portfolios is 0.693. This tight of a correlation is in evident in Figure 1 and prior literature's PEAD graphs. However, the correlation of the 100 average SUE and average returns for the 10 values of mean SUE within each percentiles is 0.320. This initial disaggregation reduces the correlation by 53.8%; nonetheless, the correlation remains moderately positive. This is consistent with observing some consistency between Figure 1 and Figure 2. Attempting to make inferences about underreaction theory from these correlations may lead one to infer support for the existence of PEAD as well as the underreaction theory. However, disaggregating the data to the firm-level and examining the correlation between SUE and returns indicates a near zero correlation, 0.0073. The similarity to Robinson (1950) finding that correlation decreases when the data is disaggregated is striking.

7 Conclusion

While the EMH is one of the most tested hypotheses in economic research, various researchers document anomalous empirical results. These anomalies are primarily documented by showing profitable portfolio trading strategies. Researchers typically offer under- or over-reaction theories to explain the empirical findings the theory of market efficiency does not predict. However, the tests of these alternative theories are typically ecological in nature, in that researchers make conclusions about firm-specific theory from returns aggregated to portfolios. We examine the ecological inference made by the PEAD literature. We demonstrate that these ecological inferences are problematic by carefully examining the firm-specific PEADs. This examination reveals the pattern of returns suggested by portfolio analysis is not found at the firm-specific drift level.

We argue that portfolio analysis research design techniques pioneered by (Fama, Fisher, Jensen, and Roll (1969)) to examine market efficiency, should not be used to examine theories explaining how individuals respond to firm-specific information. Unlike other settings, such as voting data where voter-specific data is unavailable, the data required to examine investors' responses to firm-level information is available. Since the introduction of modern portfolio theory Markowitz (1952) and since the creation of the long-window event study methodology Fama, Fisher, Jensen, and Roll (1969), statistical techniques have expanded to allow for better examination of large, complex datasets, like the CRSP and TAQ datasets. Given the data of individual firm stock prices and individual investor trades and quotes are readily accessible, exploring how these advances can be applied to these datasets is a reasonable trajectory for future research to facilitate better tests of individual-level theory.

Examination and development of individual-level theory is paramount to developing investment strategies. Indeed, in surveying empirical research covering accounting anomalies and fundamental analysis Richardson, Tuna, and Wysocki (2010) state, "empirical research is (or should be) informed by theory, because the interpretation of empirical analysis is impossible without theoretical guidance." Thus, by moving away from portfolio analysis toward techniques that better understand investor-level responses to information, we enhance our understanding of how information flows through markets and impacts equity prices. This enhanced understanding should allow for improved evaluation of theory and could help guide investment decisions.

References

Achen, C. H. and W. P. Shively (1995, December). *Cross-Level Inference*. The University of Chicago Press.

Badrinath, S. G., J. R. Kale, and T. H. Noe (1995, January). Of Shepherds, Sheep, and the Cross-autocorrelations in Equity Returns. *The Review of Financial Studies 8*(2), 401–430.

Ball, R. and P. Brown (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research 6*(2), 159–178.

Bamber, L. S., T. E. Christensen, and K. M. Gaver (2000, January). Do we really 'know' what we think we know? A case study of seminal research and its subsequent overgeneralization. *Accounting Organizations and Socieity 52*, 103–129.

Barber, B. M. and J. D. Lyon (1997). Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. *Journal of Financial Economics 43*(3), 341–372.

Barberis, N., A. Shleifer, and R. W. Vishny (1998, August). A model of investor sentiment. *Journal of Financial Economics 49*(3), 307–343.

Barberis, N., A. Shleifer, and J. Wurgler (2005, February). Comovement. *Journal of Financial Economics 75*(2), 283–317.

Bartov, E., S. Radhakrishnan, and I. Krinsky (2000). Investor sophistication and patterns in stock returns after earnings announcements. *The Accounting Review*.

Battalio, R. H. and R. R. Mendenhall (2007, February). Post-Earnings Announcement Drift: Intra-day Timing and Liquidity Costs. *Working Paper*, 1–40.

Beaver, W. H. (1968). The information content of annual earnings announcements. *Journal of Accounting Research 6*, 67–92.

Beneish, M. D., C. M. C. Lee, and D. C. Nichols (2015, November). Journal of Accounting and Economics. *Journal of Accounting and Economics 60*(2-3), 33–57.

Bernard, V. L. and J. K. Thomas (1989). Post-earnings-announcement drift: delayed price response or risk premium? *Journal of Accounting Research 27*, 1–36.

Bernard, V. L. and J. K. Thomas (1990, December). 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.

Bhushan, R. (1994). An informational efficiency perspective on the post-earnings announcement drift. *Journal of Accounting and Economics 18*(1), 45–65.

Boehmer, E. and J. J. Wu (2013, January). Short Selling and the Price Discovery Process. *The Review of Financial Studies 26*(2), 287–322.

Bradshaw, M. T. (2004, January). How do analysts use their earnings forecasts in generating stock recommendations? *The Accounting Review 79*(1), 25–50.

Bradshaw, M. T. and R. G. Sloan (2002, March). GAAP versus The Street: An Empirical Assessment of Two Alternative Definitions of Earnings. *Journal of Accounting Research 40*(1), 41–66.

Brandt, M. W., R. Kishore, P. Santa-Clara, and M. Venkatachalam (2008, January). Earnings announcements are full of surprises. *Working Paper*, 1–37.

Brennan, M. J., N. Jegadeesh, and B. Swaminathan (1993, January). Investment Analysis and the Adjustment of Stock Prices to Common Information. *The Review of Financial Studies 6*(4), 799–824.

Chan, K., A. Hameed, and W. Kang (2013, August). Stock price synchronicity and liquidity. *Journal of Financial Markets 16*(3), 414–438.

Chordia, T. and B. Swaminathan (2000, April). Trading Volume and Cross-Autocorrelations in Stock Returns. *The Journal of Finance 55*(2), 913–935.

Cohen, L. and D. Lou (2012, May). Complicated firms. *Journal of Financial Economics 104*(2), 383–400.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers (1997, July). Measuring Mutual Fund Performance with Characteristic-Based Benchmarks. *The Journal of Finance 52*(3), 1035–1058.

Daniel, K., D. A. Hirshleifer, and A. Subrahmanyam (1997, February). A theory of overconfidence, self-attribution, and security market under- and over-reactions.

De Bondt, W. F. M. and R. H. Thaler (1987, July). Further Evidence on Investor Overreaction and Stock Market Seasonality. *The Journal of Finance 42*(3), 557–581.

DellaVigna, S. and J. M. Pollet (2009, April). Investor Inattention and Friday Earnings AnnouncementsInvestor Inattention and Friday Earnings Announcements. *The Journal of Finance 64*(2), 709–749.

Doyle, J. T., R. J. Lundholm, and M. T. Soliman (2006, December). The Extreme Future Stock Returns Following I/B/E/S Earnings Surprises. *Journal of Accounting Research 44*(5), 849–887.

Drechsler, I. and Q. F. Drechsler (2014, December). The Shorting Premium and Asset Pricing Anomalies. *Working Paper*, 1–60.

Fama, E. F. (1965). The behavior of stock market movements. *The Journal of Business 38*(1), 34–105.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance 25*(2), 383–417.

Fama, E. F. (1997). Industry costs of equity. *Journal of Financial Economics 43*(2), 153–193.

Fama, E. F. (1998, January). Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics 49*(3), 283–306.

Fama, E. F., L. Fisher, M. Jensen, and R. Roll (1969). The adjustment of stock prices to new information. *International Economic Review 10*(1), 1–21.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*(1), 3–56.

Fama, E. F. and K. R. French (1995, March). Size and Book-to-Market Factors in Earnings and Returns. *The Journal of Finance 50*(1), 131–155.

Fang, V. W., A. H. Huang, and J. M. Karpoff (2016, May). Short Selling and Earnings Management: A Controlled Experiment. *The Journal of Finance 71*(3), 1251–1294.

Foster, G., C. Olsen, and T. Shevlin (1984, October). Earnings Releases, Anomalies, and the Behavior of Security Returns. *The Accounting Review 59*(4), 574–603.

Francis, J., R. LaFond, P. Olsson, and K. Schipper (2007, April). Information Uncertainty and Post-Earnings-Announcement-Drift. *Journal of Business Finance & Accounting 34*(3-4), 403–433.

Graunt, J. (1662, December). Natural and Political Observations Mentioned in a Following Index, and Made Upon the Bills of Mortality. London: John Martyn and James Allestry.

Hirshleifer, D. A., S. S. Lim, and S. H. Teoh (2009, August). Driven to Distraction: Extraneous Events adn Underreaction to Earnings News. *The Journal of Finance LXIV*(5), 2289–2325.

Hirshleifer, D. A., J. N. Myers, L. A. Myers, and S. H. Teoh (2008, November). Do Individual Investors Cause Post-Earnings Announcement Drift? Direct Evidence from Personal Trades. *The Accounting Review 83*(6), 1521–1550.

Hirshleifer, D. A., S. H. Teoh, and J. J. Yu (2011, June). Short Arbitrage, Return Asymmetry, and the Accrual Anomaly. *The Review of Financial Studies 24*(7), 2429–2461.

Honghui, C., V. Singal, and R. F. Whitelaw (2016, August). Comovement revisited. *Journal of Financial Economics 121*(3), 624–644.

Hou, K. (2007, June). Industry Information Diffusion and the Lead-lag Effect in Stock Returns. *The Review of Financial Studies 20*(4), 1113–1138.

Howe, J. S., E. Unlu, and X. S. Yan (2009, June). The Predictive Content of Aggregate Analyst Recommendations. *Journal of Accounting Research 47*(3), 799–821.

King, G. (1990, May). Electoral Responsiveness and Partisan Bias in Multiparty Democracies. *Legislative Studies Quarterly 15*(2), 159–181.

King, G. (1997, March). *A Solution to the Ecological Inference Problem*. Princeton University Press.

Kothari, S. P., J. Lewellen, and J. Warner (2006). Stock returns, aggregate earnings surprises, and behavioral finance. *Journal of Financial Economics 79*(3), 537–568.

Kothari, S. P. and J. B. Warner (1997, March). Measuring long-horizon security price performance. *Journal of Financial Economics 43*(3), 301–339.

Kothari, S. P. and J. B. Warner (2006, September). Econometrics of Event Studies. In B. E. Eckbo (Ed.), *Handbook of Corporate Finance Empirical Corporate Finance*, pp. 1–35. Elsevier-North-Holland.

Kraft, A., A. J. Leone, and C. E. Wasley (2006, February). An Analysis of the Theories and Explanations Offered for the Mispricing of Accruals and Accrual Components. *Journal of Accounting Research 44*(2), 297–339.

Lakonishok, J., A. Shleifer, and R. W. Vishny (1994, December). Contrarian Investment, Extrapolation, and Risk. *The Journal of Finance 49*(5), 1541–1578.

Lee, K. C., M. H. Pesaran, and R. G. Pierse (1990, January). Testing for Aggregation Bias in Linear Models. *100*(400), 137–150.

Livnat, J. and R. R. Mendenhall (2006, March). Comparing the Post-Earnings Announcement Drift for Surprises Calculated from Analyst and Time Series Forecasts. *Journal of Accounting Research 44*(1), 177–205.

Ljungqvist, A., C. Malloy, and F. Marston (2009, July). Rewriting History. *The Journal of Finance 64*(4), 1935–1960.

Lo, A. W. (2017, May). *Adaptive Markets: Financial evolution at the speed of thought*. Princeton, NJ: Princeton University Press.

Lo, A. W. and A. C. MacKinlay (1990, January). When are Contrarian Profits Due to Stock Market Overreaction? *The Review of Financial Studies 3*(2), 175–205.

Luloff, A. E. and P. H. Greenwood (1980, December). *Definitions of community an illustration of aggregation bias*. New Hampshire Agricultural Experiment Station.

MacKinlay, A. C. (1997). Event studies in economics and finance. *Journal of Economic Literature 35*(1), 13–39.

Markowitz, H. (1952, March). Portfolio Selection. *The Journal of Finance 7*(1), 77–91.

Mendenhall, R. R. (2004, October). Arbitrage Risk and Post-Earnings-Announcement Drift. *The Journal of Business 77*(4), 875–894.

Mohanram, P. S. (2005, September). Separating Winners from Losers among Low Book-to-Market Stocks using Financial Statement Analysis. *Review of Accounting Studies 10*(2), 133–170.

Morck, R., B. Yeung, and W. Yu (2000, January). The information content of stock markets: why do emerging markets have synchronous stock price movements? *Journal of Financial Economics 58*(1-2), 215–260.

Ogburn, W. F. and I. Goltra (1919). How women vote. *Political Science Quarterly*.

Petty, S. W. (1690, December). *Political Arithmetick*. London: Robert Clavel and Hen. Mortlock.

Petty, S. W. (1691, December). *The Political Anatomy of Ireland*. London: D. Brown and W. Rogers.

Piotroski, J. D. (2000, December). Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers. *Journal of Accounting Research 38*(Supplement: Studies on Accounting Information and the Economics of the Firm), 1–41.

Ramnath, S. (2002, December). Investor and Analyst Reactions to Earnings Announcements of Related Firms: An Empirical Analysis. *Journal of Accounting Research 40*(5), 1351–1376.

Richardson, S., İ. Tuna, and P. Wysocki (2010, December). Accounting anomalies and fundamental analysis: A review of recent research advances. *Journal of Accounting and Economics 50*(2-3), 410–454.

Robinson, W. S. (1950, June). Ecological Correlations and the Behavior of Individuals. *American Sociological Review 15*(3), 351–357.

Roll, R. (1988, July). R^2. *The Journal of Finance 43*(3), 541–566.

Sadka, G. and R. Sadka (2009, October). Predictability and the earnings–returns relation. *Journal of Financial Economics 94*(1), 87–106.

Sadka, R. (2006, May). Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk. *Journal of Financial Economics 80*(2), 309–349.

Schuessler, A. A. (1999). Ecological inference. *Proceedings of the National Academy of Sciences of the United States of America 96*(19), 10578–10581.

Shiller, R. J. (1981, June). Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? *The American Economic Review 71*(3), 421–436.

Shleifer, A. (2000, April). *Inefficient Markets: An Introduction to Behavioral Finance.* Claredon Lectures in Economics. Oxford University Press.

Sloan, R. G. (1996, July). Do Stock Prices Fully Reflect Information in Accruals and Cash Flows About Future Earnings? *The Accounting Review 71*(3), 289–315.

Stoker, T. M. (2009, November). Aggregation (econometrics). In S. N. Durlauf and L. E. Blume (Eds.), *Macroeconometrics and Time Series Analysis*, pp. 1–13. Palgrave Macmillan.

Taylor, D. J. (2011, October). Post-Earnings Announcement Drift and Related Anomalies. In *Handbook of Equity Market Anomalies: Translating Market Inefficiencies into Effective Investment Strategies*, pp. 91–115. Wiley.

Wahal, S. and M. D. Yavuz (2013, January). Style investing, comovement and return predictability. *Journal of Financial Economics 107*(1), 136–154.

Zacks, L. (2011, October). *The Handbook of Equity Market Anomalies: Translating Market Inefficiencies into Effective Investment Strategies.* Wiley Finance Series. Wiley.

Figure 1 *Post-Earnings Announcement Drift*: This figure depicts the returns to ten portfolios (Standardized Unexpected Earnings (SUE) deciles) formed each quarter for 118 quarters (1984 Q1 to 2014 Q2) based on the prior quarter's distribution of SUE. The returns start 2 days after the earnings announcement and continue for 13 weeks.

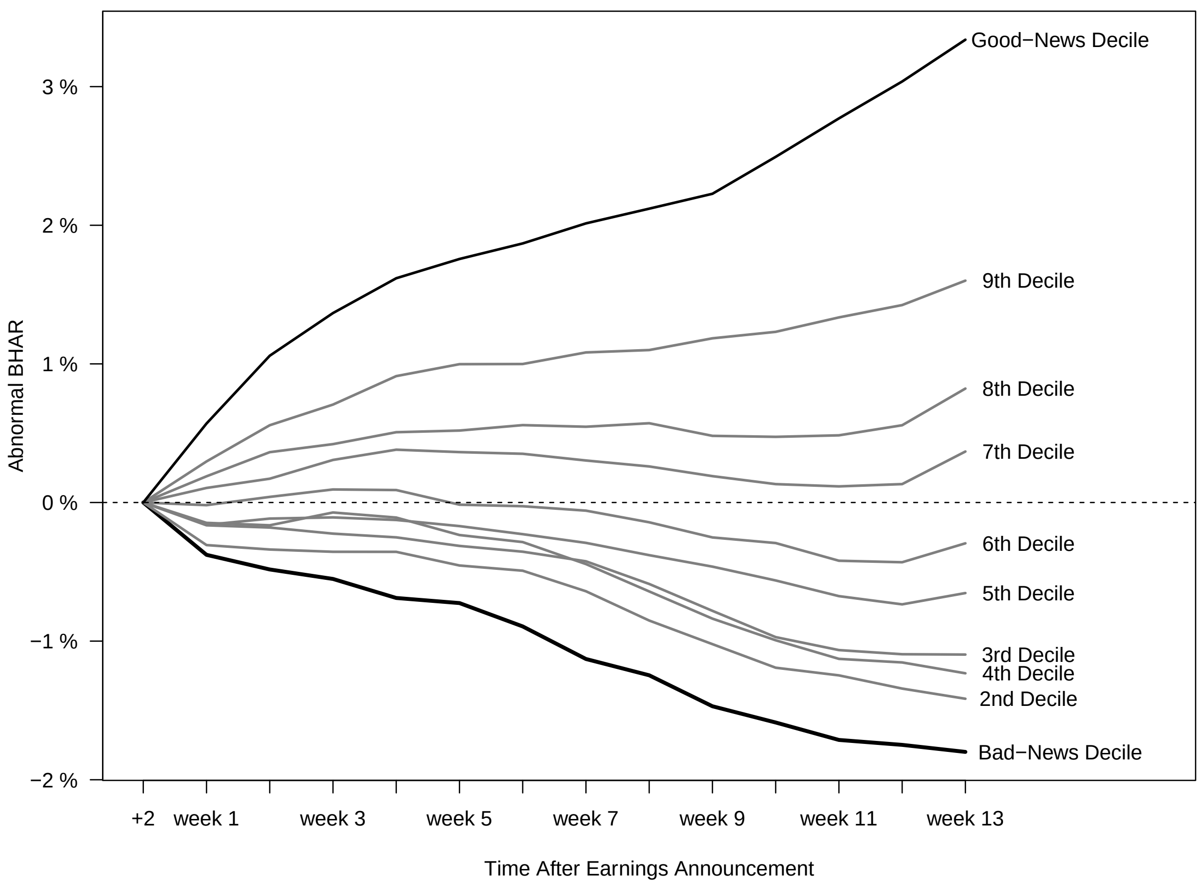


Figure 2 *Percentile Drifts in Extreme Deciles*: This figure depicts the returns to the firms in ten percentiles (90-99) within the Good News portfolio and the ten percentiles (0-9) within the Bad News portfolio. The Good News and Bad News portfolios drifts for the entire sample period are depicted in black in panels (a) and (b), respectively.

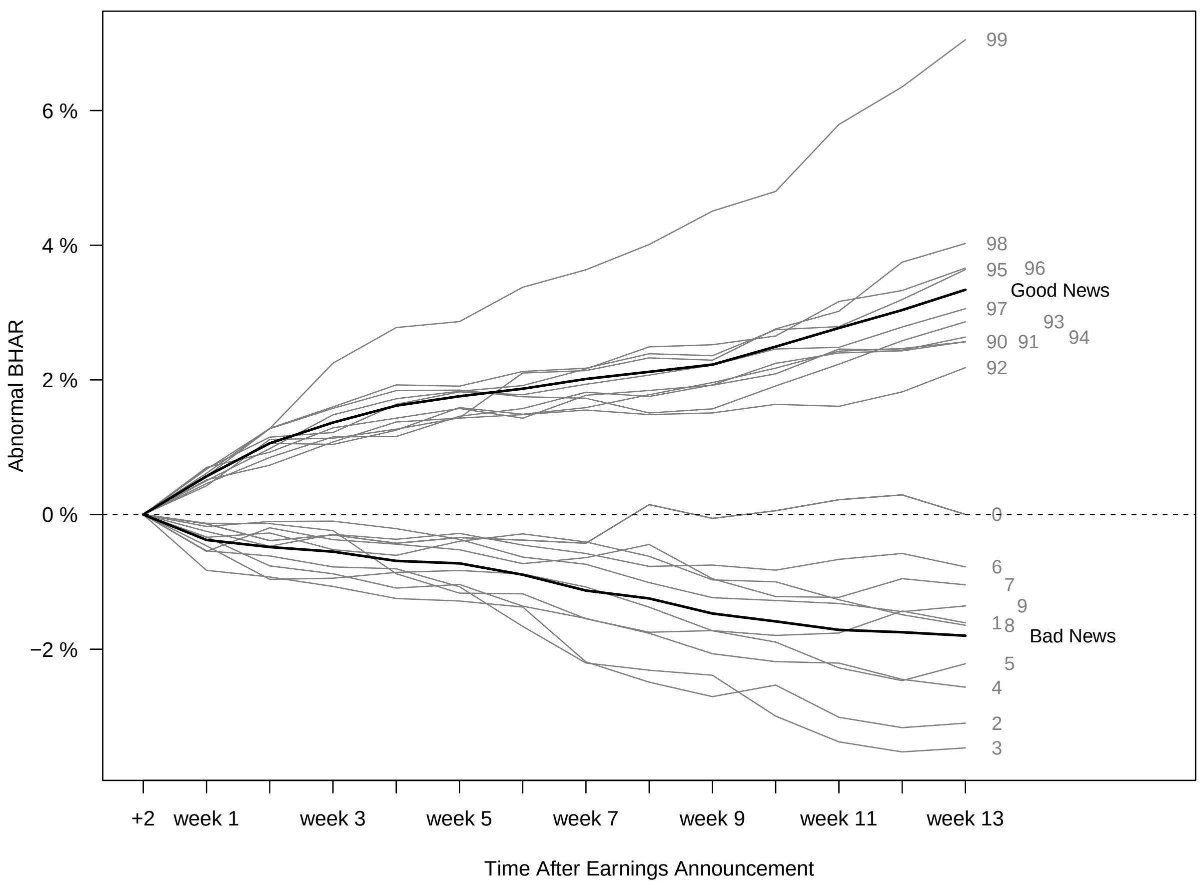
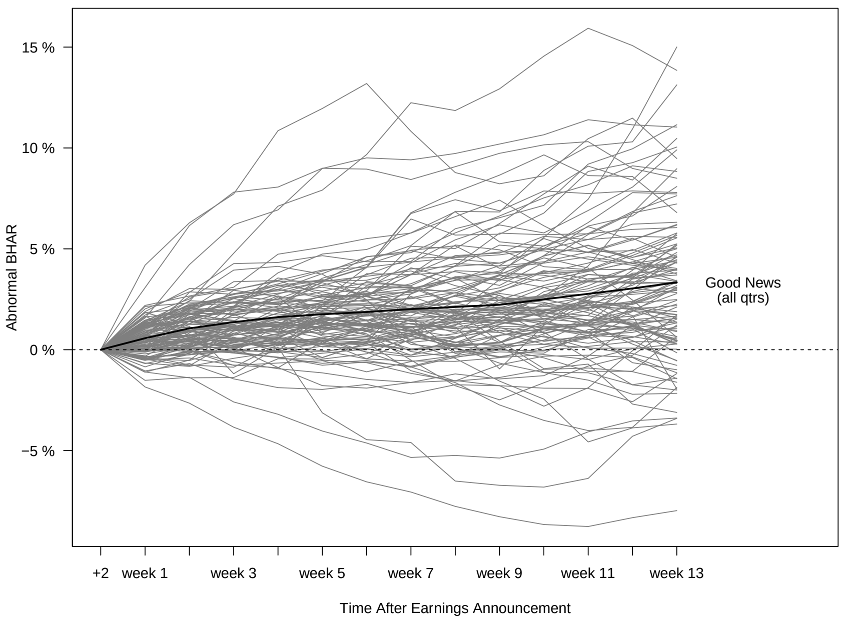
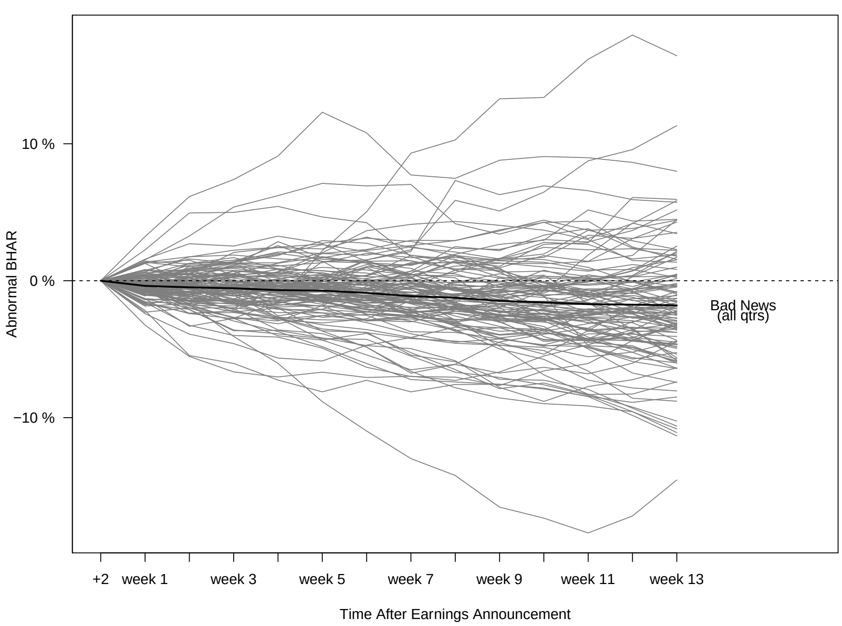


Figure 3 *Quarterly Post-Earnings Announcement Drifts*: This figure depicts in gray the drift paths of (a) Good News and (b) Bad News portfolios for each quarter from 1984 Q1 to 2014 Q2. The Good News and Bad News portfolios drifts for the entire sample period are depicted in black in panels (a) and (b), respectively. Returns are adjusted using size and book-to-market quintile reference portfolios.



(a) Good News



(b) Bad News

Figure 4 *Post-Earnings Announcement Drift Hedge Trading Strategy by Quarter*: This figure depicts the return to the PEAD hedge trading strategy of buying firms in the Good News decile and shorting firms in the Bad News decile for each quarter from 1984 Q1 to 2014 Q4. The return of to this trading strategy is depicted by a black circle. The confidence interval (+/- two standard deviations) around each quarterly returns is depicted by a black line. Returns are adjusted using size and book-to-market quintile reference portfolios.

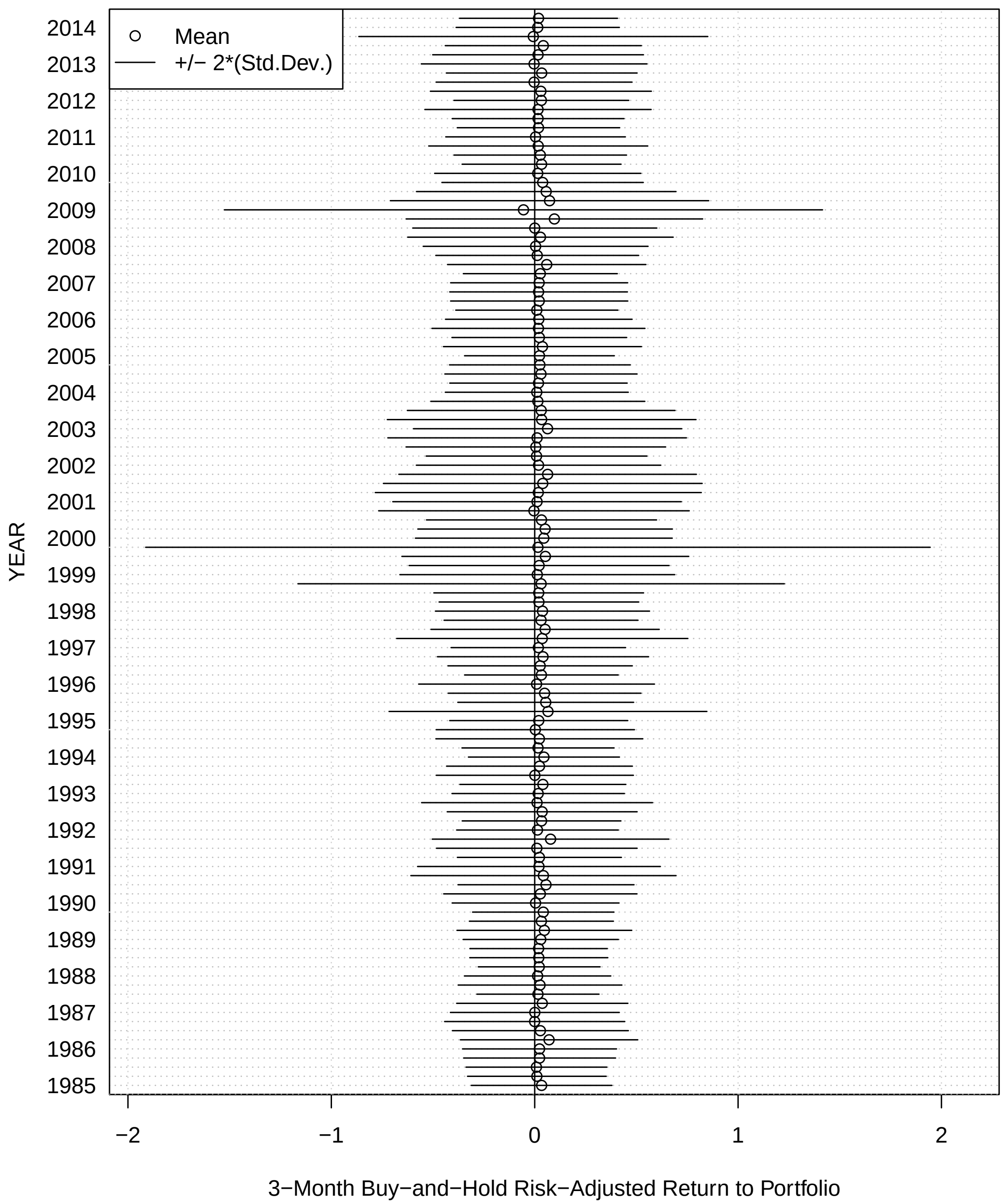
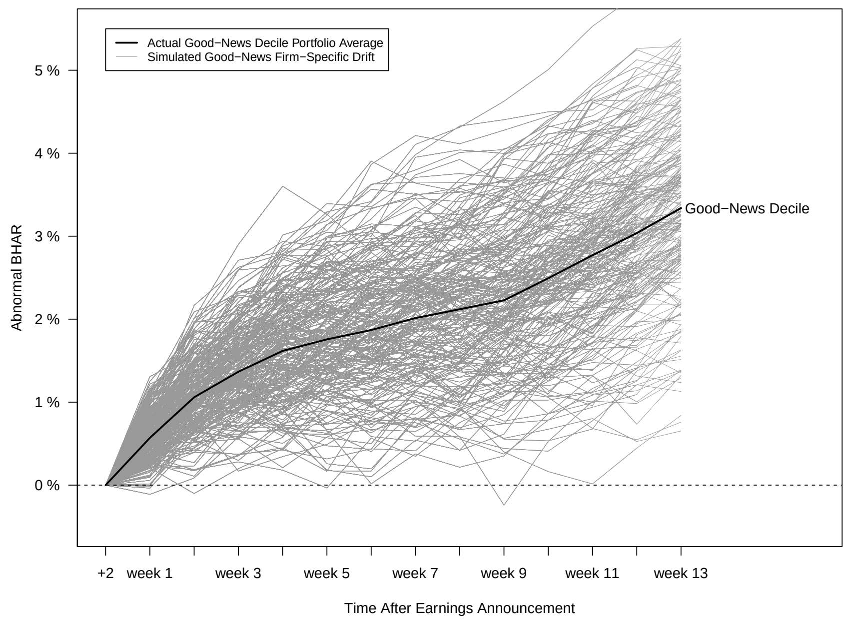
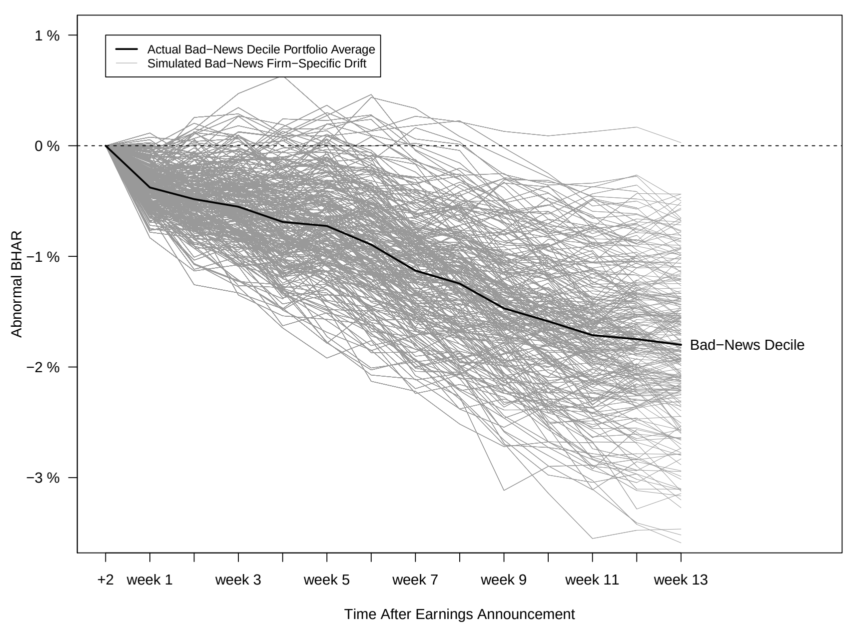


Figure 5 *Simulated Firm-Specific Post-Earnings Announcement Drifts*: This figure depicts in gray simulated drift paths of (a) Good News and (b) Bad News portfolios based on expected drift paths predicted by underreaction theory. For each path starting two days after the earnings announcement, the simulated returns for each of the 13 weeks following the earnings announcement are depicted. The average Good News and Bad News portfolios drifts for the entire sample period are depicted in black in panels (a) and (b), respectively.

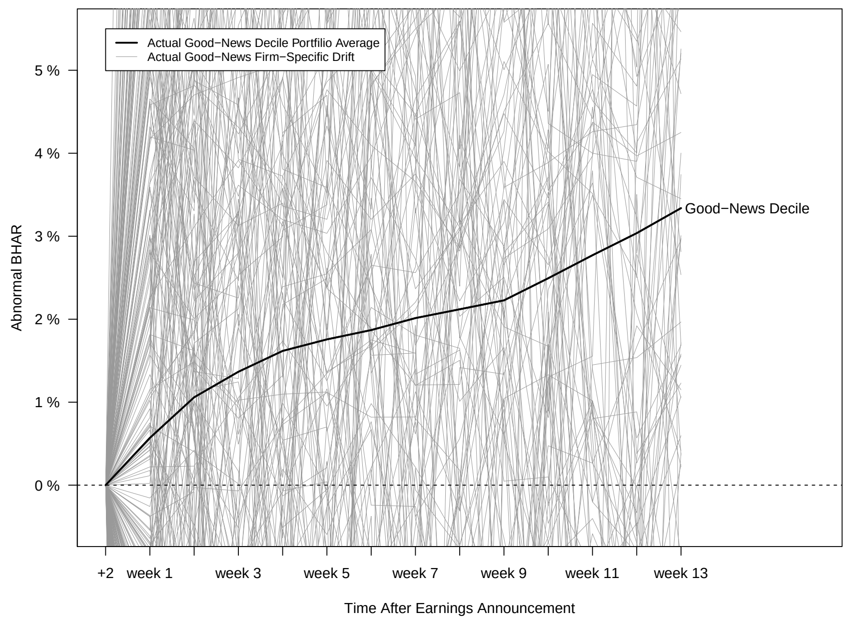


(a) Good News

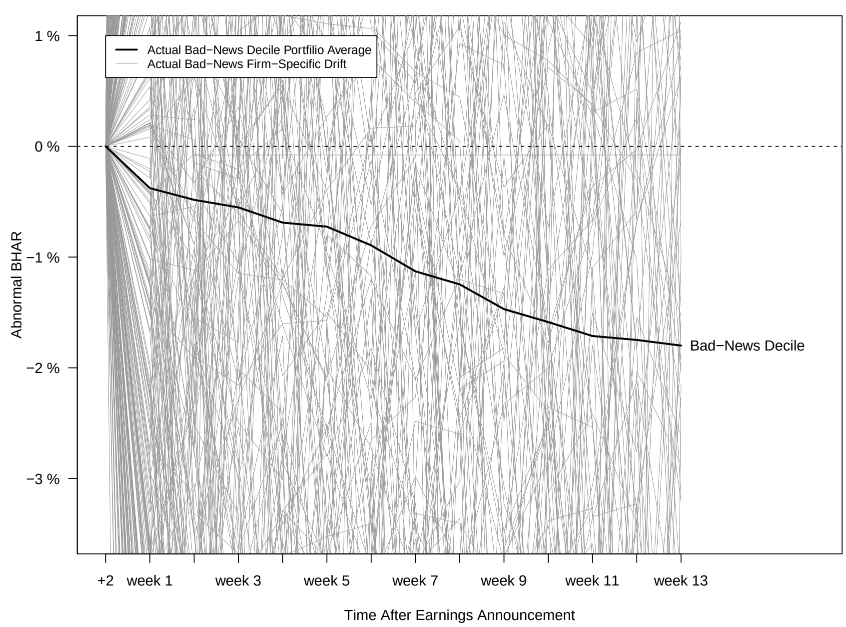


(b) Bad News

Figure 6 *Actual Firm-Specific Post-Earnings Announcement Drifts*: This figure depicts in gray actual firm-specific drift paths of (a) a random sample of 250 firms in the Good News portfolio and (b) a random sample of 250 firms in the Bad News portfolios based. For each path starting two days after the earnings announcement, the returns for each of the 13 weeks following the earnings announcement are depicted. The Good News and Bad News portfolios drifts for the entire sample period are depicted in black in panels (a) and (b), respectively. Returns are adjusted using size and book-to-market quintile reference portfolios.

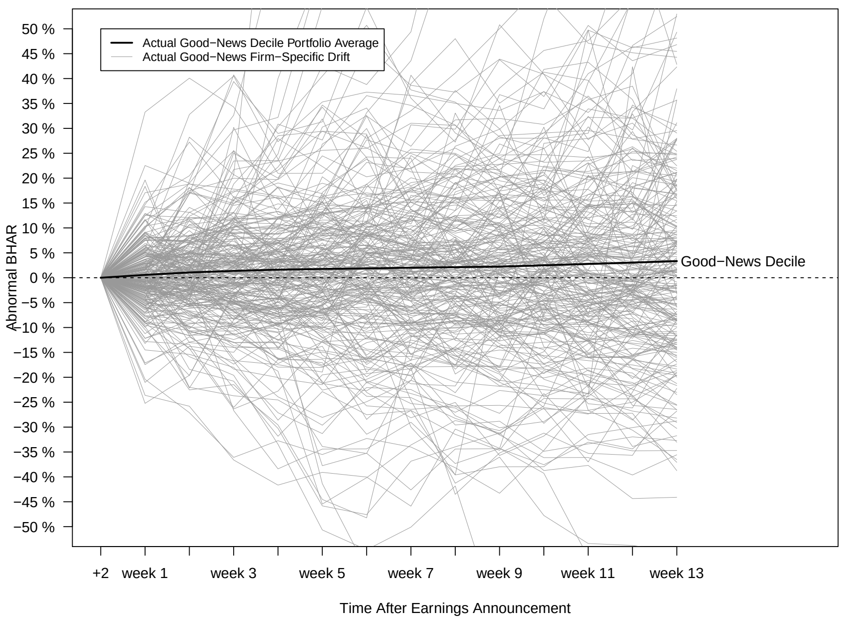


(a) Good News

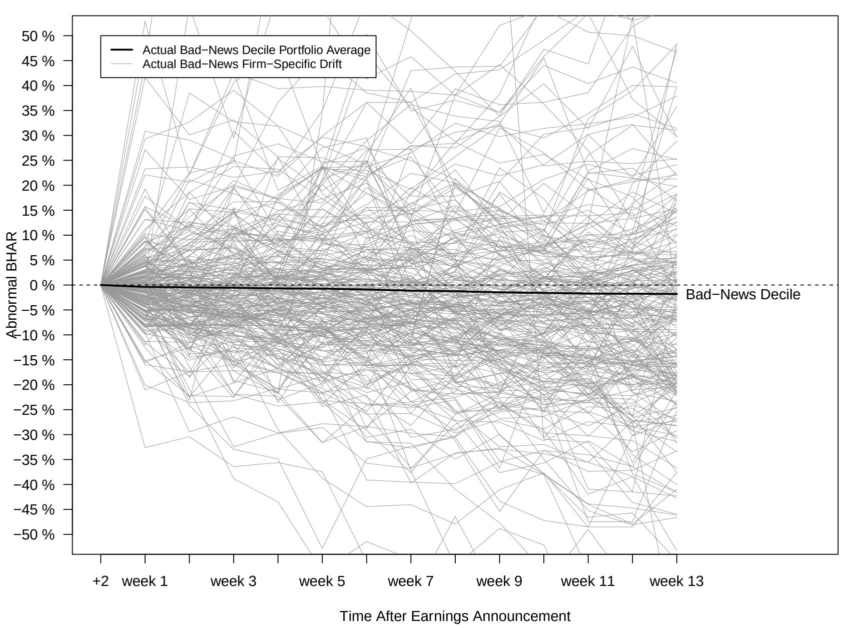


(b) Bad News

Figure 7 *Actual Firm-Specific Post-Earnings Announcement Drifts (Increased Y-Axis Range)*: This figure depicts in gray actual firm-specific drift paths of (a) a random sample of 250 firms in the Good News portfolio and (b) a random sample of 250 firms in the Bad News portfolios based. For each path starting two days after the earnings announcement, the returns for each of the 13 weeks following the earnings announcement are depicted. The Y-Axis range is -50% to +50% The Good News and Bad News portfolios drifts for the entire sample period are depicted in black in panels (a) and (b), respectively. Returns are adjusted using size and book-to-market quintile reference portfolios.

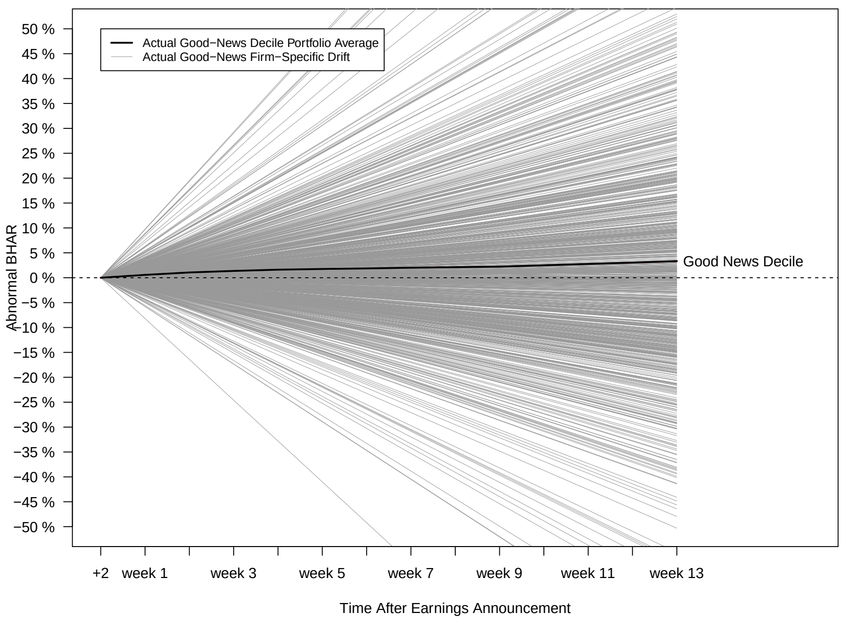


(a) Good News

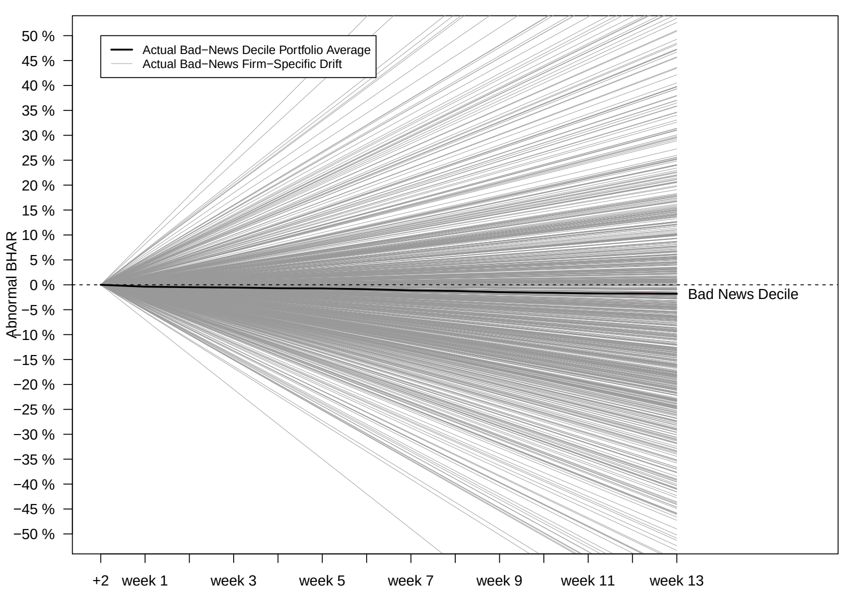


(b) Bad News

Figure 8 *Actual Firm-Specific PEADS (Variation from t to t+13 removed)* This figure depicts in gray actual firm-specific drift paths of (a) a random sample of 800 firms in the Good News portfolio and (b) a random sample of 800 firms in the Bad News portfolios based. For each path we plot a line from zero at two days after the earnings announcement to the risk-adjusted buy and hold 13 weeks after the earnings announcement. The Y-Axis range is -50% to +50%. The Good News and Bad News portfolios drifts for the entire sample period are depicted in black in panels (a) and (b), respectively. Returns are adjusted using size and book-to-market quintile reference portfolios.



(a) Good News



(b) Bad News

Table 1 Descriptive Statistics - Full Sample

	Mean	Std. Dev.	P25	Median	P75
SUE	-0.002	0.067	-0.001	0.000	0.002
SUE_Decile	4.521	2.778	2.000	5.000	7.000
Market Value of Equity	4015.157	16764.022	192.084	611.703	2119.154
Book-to-Market	0.592	0.436	0.300	0.500	0.774

TABLE NOTE: This table presents descriptives statistics for the entire sample of firm-quarters. Variables are defined in the Appendix A. N=250,186.

Table 2 Three-Month Returns to SUE-Decile Portfolios and PEAD Hedge Trading Strategy

	N Firms Rets.	Mean Ret.	Std. Dev.	T-Stat.	P-Val.
Good-News Portfolio	22156	0.033	0.299	16.617	0.000
SUE Decile 9	24823	0.016	0.227	11.129	0.000
SUE Decile 8	25651	0.008	0.209	6.285	0.000
SUE Decile 7	26215	0.004	0.193	3.085	0.002
SUE Decile 6	26810	-0.003	0.186	-2.592	0.010
SUE Decile 5	24804	-0.006	0.179	-5.732	0.000
SUE Decile 4	30534	-0.012	0.211	-10.192	0.000
SUE Decile 3	23637	-0.011	0.185	-9.127	0.000
SUE Decile 2	24124	-0.014	0.220	-9.985	0.000
Bad-News Portfolio	21432	-0.018	0.309	-8.531	0.000
Hedge Portfolio (Good - Bad)	43588	0.051	0.305	-17.639	0.000

TABLE NOTE: This table reports the adjusted returns, T-statistics and the corresponding P-Values testing whether three- month portfolio returns are significantly different from zero for the 10 PEAD portfolios. This table also reports the return to the hedge trading strategy of buying the Good-News Portfolio (SUE decile = 10) and selling the Bad-News Portfolio (SUE decile = 1). These results are labled Hedge Portfolio. Three-month returns are measured from 2 days to 92 days after the earnings announcement. Portfolios are identified by SUE and cuttoffs deteremined in prior quarter. Returns are adjusted using size and book-to-market quintile reference portfolios.

Table 3 Analsysis of Firm-Level Three-Month Returns Following an Earning Announcement

	N	% of 3 Month Firm Drifts with a Positive Return	% of 3 Month Firm Drifts with a Negative Return
Entire Sample	250186	47.38	52.62
Good-News	22156	51.75	48.25
SUE Decile 8	24823	50.72	49.28
SUE Decile 7	25651	50.18	49.82
SUE Decile 6	26215	49.62	50.38
SUE Decile 5	26810	47.70	52.30
SUE Decile 4	24804	47.44	52.56
SUE Decile 3	30534	45.33	54.67
SUE Decile 2	23637	45.38	54.62
SUE Decile 1	24124	43.73	56.27
Bad-News	21432	41.68	58.32

TABLE NOTE: This table reports the percentage of firm-quarters with postive returns and percentage of firm-quarters with negative returns for the whole sample and within each SUE Decile. Three-month returns are measured from 2 days to 92 days after the earnings announcement. Portfolios are identified by the observation's value of SUE and SUE decile cuttoffs deteremined using the prior quarter's distribution of SUE. Returns are adjusted using size and book-to-market quintile reference portfolios.

Table 4 Analysis of the 13 Weekly Returns Following an Earning Announcement

	Mean(% Positive Weeks)	% of Firm Drifts with No Weeks with Positive Return	% of Firm Drifts with All Weeks with Positive Return
Entire Sample	47.98	17.16	14.25
Good-News	51.81	14.89	17.32
SUE Decile 8	51.29	15.25	16.68
SUE Decile 7	50.39	15.66	16.07
SUE Decile 6	49.82	15.71	14.99
SUE Decile 5	48.45	16.27	13.98
SUE Decile 4	47.76	16.97	13.88
SUE Decile 3	46.43	18.13	13.02
SUE Decile 2	46.07	18.21	12.93
SUE Decile 1	44.73	19.91	12.29
Bad-News	42.74	21.04	11.40

TABLE NOTE: This table reports the mean of the percentage of positive weeks for each firm drift (N Weeks with Positive Return/13) within each SUE Decile. Portfolios are identified by the observation's value of SUE and SUE decile cuttoffs deteremined using the prior quarter's distribution of SUE. Returns are adjusted using size and book-to-market quintile reference portfolios.

Table 5 Correlation between SUE and Returns at Various Aggregation Levels

	N	corr(Average SUE, 3-Month Return)
Decile	10	0.6927
Percentile	100	0.3200
Quarterly Deciles	1180	0.2452
Quarterly Percentile	10825	0.0429
Firm-Level (no aggregation)	250186	0.0073

TABLE NOTE: This table presents the correlation between standardized unexpected earnings (SUE) and returns at the following aggregation levels: decile, percentile, quarterly decile, quarterly percentile, firm-level (no aggregation). SUE is measured as the difference between actual and expected earnings earnings divided by price. Three-month returns are measured from 2 days to 92 days after the earnings announcement.