Anomalies and News[Ψ]

Joseph Engelberg

R. David McLean

and

Jeffrey Pontiff

July 15, 2015

Abstract

Using a sample of 97 stock return anomalies documented in published studies, we find that anomaly returns are 7 times higher on earnings announcement days and 2 times higher on corporate news days. The effects are similar on both the long and short sides, and they survive adjustments for risk exposure and data mining. We also find that anomaly signals predict analyst forecast errors of earnings announcements. Taken together, our results support the view that anomaly returns are the result of mispricing, which is at least partially corrected upon news arrival.

Keywords: News, cross-sectional return predictability, earnings announcements, market efficiency.

JEL Code: G00, G14, L3, C1.

[Ψ] Engelberg (jengelberg@ucsd.edu) is at UCSD, McLean (rdmclean@ualberta.ca) is at the University of Alberta, and Pontiff (pontiff@bc.edu) is at Boston College. McLean is grateful to the Kipnes Chair for financial support. We thank Mark Bradshaw, Juhani Linnainmaa, and Philip Strahan for helpful comments, and seminar participants at Auburn, Boston College, UC Irvine, Michigan State, Cornell, DePaul, UC Riverside, University of Utah, Arizona State, ESCP, the Tinbergen Institute, and the University of Washington.

Academic research shows that a large number of characteristics can predict the cross-section of stock returns. This research goes back to at least Blume and Husick (1973), yet 42 years later academics still disagree on what causes return predictability (see the 2013 Nobel Prize).

There are three popular explanations for cross-sectional predictability. First, predictability could be the result of cross-sectional differences in risk, reflected in discount rates (see Fama (1991, 1998)). In this framework, cross-sectional return predictability is expected because return differences simply reflect ex-ante differences in discount rates that were used to value the stocks. There are no surprises here: what happens with returns ex-post was *expected* by rational investors ex-ante (e.g., Fama and French (1992, 1996)).[1]

The second explanation comes from behavioral finance, and argues that return predictability reflects mispricing (e.g., Barberis and Thaler (2003)). For example, investors may systematically have biased expectations of cash flows, and the anomaly variables are correlated with these mistakes among the cross-section of stocks. When new information arrives, investors update their beliefs, which corrects prices and creates the return-predictability. This theory of biased expectations has been used to explain predictability resulting from price-to earnings ratios (Basu, 1977), long-term reversal (Debondt and Thaler (1985, 1987)), and the value-growth anomaly (Lakonishok, Shleifer and Vishny (1994), La Porta et al. (1997)).

[1] In an efficient market the relation between discount rates and average returns is virtually mechanical. To see this consider a simple pricing model that says: P = D / (k - g). This equation can be rewritten so that: k = D/P + g. This just says that discount rates = returns, and that returns = dividend yield + capital gains. So if rational investors discount with k, then average returns will equal k, and firms with higher k will have higher average returns.

A third explanation for return predictability is data mining. As Fama (1998) points out, academics have likely tested thousands of variables, so it is not surprising to find that some of them predict returns in-sample, even if in reality none of them do. Recognition of a "multiple testing bias" in all types of empirical research dates at least back to Bonferroni (1935) and is stressed more recently in the finance literature by Harvey, Lin, and Zhu (2014).

To differentiate between these three views we compare predictability on days where firm-specific information is publicly released to days where we do not observe news. Most studies concerned with cross-sectional return predictability do not ask whether predictability is associated with earnings releases, and no studies (to our knowledge) examine where predictability is associated with news in general. The study of such effects is essential to our understanding of the source of the predictability, as the market's ability to process news is the backbone of the efficient market hypothesis.

We start with a sample of 97 anomalies studied in McLean and Pontiff (2015). Each of these 97 predictors has been shown to predict the cross-section of stock returns in a published academic study. We compare the average anomaly returns *on* versus *off* days with firm-specific news. If anomaly returns are expected, as in the risk-based theory of returns, then intertemporal differences in anomaly portfolios returns only occur due to intertemporal differences in priced risk. In contrast, if anomaly returns are the result of expectational errors, then anomaly portfolios should perform better on days when new information is released, as the new information causes investors to update their beliefs in a predictable way.

Using 489,996 earnings announcements and 6,223,007 Dow Jones news items during the period 1979-2013, we find support for the idea that anomalies are the result of mispricing. We find that anomaly portfolios have higher returns on news days compared to non-news days. Comparisons of anomaly returns on and off earnings announcement days shows that anomaly returns are 7 times higher on earnings announcement days and 2 times higher on corporate news days. This effect is robust across all types of anomalies, although we find interesting cross-sectional differences among the different anomaly types.

Market-based anomalies (e.g., momentum and idiosyncratic risk), which are constructed with exchange data, have the largest increase in predictability on non-earnings news days and the smallest increase on earnings days. Fundamental anomalies (e.g., accruals to assets and debt to equity), which are based purely on accounting data, have the smallest increases on non-earnings news days and the largest increase on earnings days. Valuation anomalies, which combine exchange data and market data, and event anomalies (e.g., share issues and changes in analysts' recommendations), exhibit increases that are in between those of market and fundamental anomalies on both types of days.

Like most tests of market efficiency, our tests are tests of the joint hypothesis of market efficiency and the correct asset-pricing model. However, it's important to note what a rational expectations asset-pricing model would need in order for it to explain our results. When we examine both the long and short side of anomaly portfolios separately, we find that returns are 5.5 times *higher* on earnings day for long-side stocks and 10 times *lower* for short-side stocks. If returns reflect risk as in a rational expectations asset-pricing model, the model would require some stocks to be 5.5 times riskier on their earnings announcement day and other stocks to be 10 times less risky on their earnings announcement day. This seems implausible.

Nevertheless, we consider two factor-based risk-explanations, where expected returns are generated either from exposure to the market portfolio or exposure to a factor that is based on an aggregate anomaly portfolio. Neither of these specifications changes our results. In fact, specifications that include these controls produce higher anomaly alphas on earnings and news days.

To further address the joint hypothesis problem we examine the expectations of an important group of market participants: sell-side analysts. This is a more direct test of the biased expectations hypothesis. If analysts have biased expectations regarding anomaly stocks then their forecasts should be too optimistic for stocks on the short side of anomaly portfolios and too pessimistic for stocks on the long side of anomaly portfolios. With one exception, this is precisely what we find; for stocks in the long leg of anomaly portfolios, analysts' forecasts are too low, and for stocks in the short leg, analysts' forecasts are too high. The one exception is that analysts tend to overestimate earnings for stocks that valuation ratios (such as price-to-earnings and book-to-market) suggest are undervalued. Yet for the other 7 of our 8 categories (4 anomaly categories, each with a long and short leg), our forecast error results mirror our stock return results, in that analyst forecast are too low for anomaly buys, and too high for anomaly sells.

In our final set of tests we provide evidence that data mining cannot fully explain why anomaly returns are higher on earnings announcement days and news days. Stocks with high (low) ex-post returns over a given period are more likely to have high (low) returns on news days because news days have more variance than non-news days. Because a data miner selects stocks based on ex-post performance, data mining implies that we would expect to find higher anomaly returns on news days compared to non-news days. We introduce a novel data-mining test which creates pseudo anomaly portfolios that have the same return properties as anomaly stocks and demonstrates that a pure data-mining explanation is not the source of our findings.

Our paper builds on previous studies, which show for a specific anomaly that returns are higher on earnings announcement days (e.g., Bernard and Thomas (1992), Chopra, Lakonishok and Ritter (1992), La Porta et al. (1994), Sloan (1996), and Jegadeesh and Titman (1993)). Our findings also build on Edelen, Ince, and Kadlec (forthcoming), who show that institutions tend to trade against 12 popular anomalies, and that such trading activities portend higher anomaly returns on earnings announcement days.

Our paper differs from the previous literature in several ways. First, we investigate not only earnings announcement days but also consider almost 6 million news days that do not coincide with earnings announcements. We also use a broad set of 97 anomalies which gives us more statistical power than previous studies and allows us to draw novel comparisons between categories of anomalies. Our paper is also the first to relate a broad set of anomalies to analyst forecast errors, which complements our news and earnings announcement findings. Finally, we introduce the first news day data-mining test that allows us to rule out the possibility that our results are driven by data mining.

1. Sample and Data

We begin our sample with 97 cross-sectional "predictors" studied in McLean and Pontiff (2015). These predictors are drawn from 80 studies published in peer-reviewed finance, accounting, and economics journals. Each of the predictor variables is shown to predict the cross-section of stock returns. All of the variables can be constructed with data from CRSP, Compustat, or IBES.

To create the predictor portfolios, stocks are sorted each month on each of the predictor characteristics. We define the extreme quintiles as the long and short side of each predictor strategy. 16 of our 97 predictors are indicator variables (e.g, credit rating downgrades). For these cases, there is only a long or short side, based on the binary value of the indicator. We remake the predictor portfolios each month. As in McLean and Pontiff (2015), the sample selection for each predictor follows the original study. So if a study only uses NYSE firms, then we only create that predictor variable for NYSE firms.

We obtain earnings announcement dates from the Compustat quarterly database. Compustat reports the earnings announcement day, but not the time. Many firms report earnings after the market closes. In these cases, the information will be reflected in the stock return on the following day (CRSP returns are from close to close). We therefore examine the firm's trading volume scaled by market trading volume for the day before, the day of, and the day after the reported earnings announcement date. We define the day with the highest volume as the earnings announcement day.

We obtain news stories dates from the Dow Jones news archive. Dow Jones reports both the date and time of its news stories. This archive contains all news stories from Dow Jones newswire and all Wall Street Journal stories for the period 1979-2013. These news data are also used in Tetlock (2010, 2011) and Engelberg, Reed, and Ringgenberg (2012). We merge this data with daily stock return data, so that we can test whether predictor portfolio returns are higher on information days as compared to off information days.

For consistency, we conduct all of our tests during the period 1979-2013, which is the period that we have news data for. We also exclude stocks with prices under $5. These low-priced stocks are excluded from many of the predictor portfolios to begin with, and low-priced stocks are less likely to have news or earnings announcement data.

1.1. Sample Descriptive Statistics

Table 1 provides some descriptive statistics for our sample, which consists of 40,165,651 firm-day observations for the period 1979-2013. Each observation is in the CRSP daily return database with reported stock returns and a stock price greater than $5. Among these observations, 16% have Dow Jones news stories, while 1.2% have earnings announcements reported in Compustat.

There is overlap between the news days and the earnings announcement days. Of the 489,966 earnings announcement days, 256,745, or 52%, are also Dow

Jones news days. This is however, a small percentage of the total news days. The total number of news days is 6,453,258 so only 4% of these are also earnings announcements that are reported in Compustat. It could be that Dow Jones stories cover a significant number of earnings announcements not covered in Compustat, so 4% is a lower bound on the percentage of news stories that likely reflect earnings announcements. Table 2 provides descriptive descriptions of the portfolio variables.

3. Main Results

3.1 Predictor Portfolio Returns On and Off Information Days

In this section of the paper we report our main findings. In our first set of tests we estimate the following regression equation:

$$R_{i,t} = \alpha_t + \beta_1 Net_{i,t} + \beta_2 Net_{i,t} \times Eday_{i,t} + \beta_3 Net_{i,t} \times Nday_{i,t} + \beta_4 Eday_{i,t} + \beta_5 Nday_{i,t} + \sum_{i=1}^{10} \gamma_i Lag\ Return_{t-i} + \sum_{i=1}^{10} \delta_i Lag\ Return^2{}_{t-i} + \sum_{i=1}^{10} \rho_i Volume_{t-i} + \varepsilon_{i,t}$$

The regression includes day fixed effects (α_t). In the above equation $R_{i,t}$ is the daily return of stock *i* on day *t* in percent (returns are multiplied by 100). *Net* is our aggregate anomaly variable; it is the difference between the number of long-side anomaly portfolios a firm is in minus the number of short side anomaly portfolios the firm is in. The anomaly portfolios are measured at the beginning of each month, and returns are measured on each day throughout the month. Thus, although news such as earnings announcements may affect future values of *Net* for a given stock, the value of *Net* that we use in our regressions remains the same throughout a month. We describe *Net* in more detail below.

The variables *Eday and Nday* are dummy variables equal to 1 on earnings and news days for firm *i* and zero otherwise. An information day can be an earnings announcement day, a Dow Jones News day, or both. Our hypotheses are tested with the interaction term, i.e., are anomaly returns higher on information days? We include lagged return, volatility (return squared) and volume as controls. For brevity we do not report these coefficients. We also report specifications without the controls and the results do not change.

The variables *Long* and *Short* are counts of the respective long and short portfolios that a particular stock belongs in for a given month. The average stock is in 8.61 long portfolios and 9.23 short portfolios. If the portfolios were solely based on 97 random quintiles groupings, we would expect long and short to equal to 19.4 (97 x 0.20). Our counts are much lower, since some characteristics are indicator variables, and thus, lack either a long or short side, and, following the original study, some characteristics are only constructed for a subset of stocks (for example, NYSE stocks). For characteristics that are subset based, stocks that fall out of the subset are not assigned to a long or short side.

The variable *Net* is an aggregate anomaly variable. For each firm-month observation, we sum up the number of long side (*Long*) and short side (*Short*) anomaly portfolios that the observation belongs to. *Net* is the difference between *Long* and *Short*: *Net* = *Long* – *Short*. Summary statistics for *Net, Long,* and *Short* are provided in Table 1. The mean value for *Net* is -0.61, the maximum is 32, and the minimum is -36.

With respect to the above regression equation, market efficiency (in the absence of data mining and changes in risk exposure) suggests that the interactions terms should be zero, i.e., predictor returns should not be any stronger on information days as compared to other days. This is because in the rational expectations framework return-predictability is explained by ex-ante differences in discount rates, which should not change in a predictable manner on information days.

In contrast, the biased expectations framework suggests that coefficient for the interaction between *Net* and the earnings and news day dummies should be positive, or that anomaly returns should be greater when new information is released. This is because in the biased expectations framework return-predictability is the result of ex-post releases of information that cause investors to update their expectations, which were systematically biased ex-ante.

Panel A of Table 3 reports the regression results. We estimate regressions with and without the lagged controls. The slope coefficients are virtually identical, but the t-statistics tend to increase with observation controls, demonstrating that the controls reduce residual error. In order to economize on tables, we focus on results from the regressions that include controls.

The coefficients are reported in percent. In Panel A we define the information day as a 1-day window, while in Panel B we use a 3-day window, i.e., days t-1, t, and t+1. In the first regression in Panel A the *Net* coefficient is 0.003, while the *Net x*

Earnings Announcement interaction coefficient is 0.020. Taken together, the coefficients show that for a *Net* value that is higher by 10 (about 1 ½ standard deviations) expected returns are higher by 3 basis points on non-earnings announcement days, and by an additional 20 basis points, or 667% more, on earnings announcement days. The *Net x News Day* interaction coefficient is 0.003, showing that anomaly returns are higher by 100% on news days that are not also earnings announcement days, which is also a sizeable effect. Taken together, the coefficients show that anomaly returns are higher by 0.023, or 767%, on earnings days that are also Dow Jones news days. All of the coefficients are significant at the 1% level.

In the second regression reported in Panel A we replace the day fixed effect with a day-information event fixed effect. That is, we have a unique slope for each day and each firm-day that is either a news day or an earnings announcement day. In this regression the comparison is therefore between two firms that both have a news story or earnings announcement on the same day, but have different values of *Net*. The coefficients in this regression are very similar to those in the first regression. The Net coefficient is still 0.003, while the earnings day and news day interactions are 0.018 and 0.004 respectively.

In the next few regressions we dig deeper into the idea that systematic risk can explain anomaly returns. To this end we add either an anomaly factor or a market portfolio factor to our regressions, and test whether our inferences change. The anomaly factor is the long-short portfolio return for a portfolio that is long in the top 20% percentile of *Net* and short in the bottom 20% percentile of *Net*. We add the returns of this anomaly factor (*Factor*) and an interaction between *Factor* and an information day dummy to the regression specification.

The coefficients on *Factor* or *Market*, and the interactions with *Factor* or *Market*, jointly estimate beta coefficients on *Factor* or *Market*. For example, for the specifications that include information day interactions with *Market*, and double interactions with *Market* and *Net*, the slope on the *Market* estimates the factor beta for zero net stocks on non-information days. The coefficient on *Market* interacted with information day dummy tells us how much the typical beta increases on information days. The *Net* interactions with *Market* allows beta to be a linear function of *Net*.

In regression 3 we see that including *Factor* has virtually no impact on the either the *Net* coefficient or the coefficient for the interaction between *Net* and earnings days. The *Net* coefficient is still 0.003, and the *Net* earnings day interaction is 0.020, which is the same result that we report in regression 1. The Net news day interaction is 0.002, similar to the value of 0.003 estimated in regression 1. Thus, controlling for beta exposure to the market or an anomaly factor has little impact on the cross-sectional return variation that *Net* explains.

The coefficient for *Factor* is -0.931 and statistically significant. Hence, when *Factor* has high returns, expected stock returns are lower. This finding says that when anomalies do well, average stock returns tend to be lower. Note that anomaly portfolios are equally long and short, so there is no reason to expect this coefficient to be positive. This result poses a challenge to risk based models of anomaly returns. If anomalies represent compensation for some source of risk, then stocks ought to have higher expected returns when anomalies do well, yet we find the opposite.

The coefficient for the interaction between *Factor* and the earnings announcement day dummy is 0.030 and not significant. The interaction between *Factor* and the news day dummy is 0.461 and significant. Hence, if a stock has an news announcement its exposure to *Factor* becomes more negative. Note that this is the opposite result that we find with *Net*; if a stock has a higher value of *Net*, its expected return is higher, and this effect is greater on news days. The results with *Factor* therefore contradict the idea that the *Net* results are explained by covariance with some underlying risk; including *Factor* does not affect the *Net* coefficient, and the *Factor* and *Net* coefficients are of the opposite sign.

In regression 4 we estimate a specification in which we replace *Factor* with the market portfolio, which is the return of the CRSP value-weighted index minus the risk free rate. This later specification tells us whether controlling for market risk changes our inferences. The coefficients for *Net* and the *Net* earnings day and news day interactions are still 0.003, 0.020, and 0.002 respectively, so the results for *Net* cannot be explained by market risk. The coefficient for the market portfolio is 0.737, whereas the market portfolio earnings day and news day interactions are 0.033 and 0.319 respectively. These results make sense; the average beta is close to one, and beta increases earnings announcement and news days.

The fifth regression is like the fourth, but it includes an interaction between *Net* and the market portfolio, and a three-way interaction between *Net*, the market portfolio, and the earnings announcement and news dummies. The interaction between *Net* and the market portfolio is negative and significant; stocks with higher values of *Net* have lower covariance with the market portfolio. The results show that for every unit increase in *Net*, market beta falls by -0.023. Moreover, this effect is greater on earnings announcement days; for every unit increase in *Net*, market beta falls by an additional -0.003 on earnings announcement days, although the coefficient is insignificant. For every unit increase in *Net*, market beta increases by 0.004 on news days. These findings are difficult to reconcile with risk-based explanations for anomalies, as they show that stocks with higher anomaly exposure have less market risk, which increases only slightly on news days, and declines on earnings days.

The results in Panel B, which studies returns news and earnings announcement returns over 3-day windows, are similar. The information day coefficients are smaller as compared to Panel A, which is to be expected because Panel B uses a 3-day window. Yet there are still significantly higher returns on information days, and these effects and are unchanged in the presence of various fixed effects and controls for market factor and market risk.

The coefficients reported in both panels document substantial positive returns on both earnings days and news-not-earnings days. The earnings day result is consistent with Franzini and Lamont (2006). We do not know of previous research that has documented our news-not-earnings day finding—news days are associated with positive stock price reactions.

3.2. Estimating Separate Long and Short Anomaly Effects

In Table 4 we removes the *Net* variable from the regressions and replace it with *Long* and *Short*, which as we explain above are the sums of the number of long side and short side anomaly portfolios that the stock belongs. Using *Long* and *Short* separately allows us to examine whether the effects of information are different for the long and short sides of anomalies.

The first three regressions in Table 4 use the 1-day announcement window. We use the lagged controls in all of the regressions reported in Table 4 along with day fixed effects, and interactions between net and the daily market returns, and two 3-way interactions among *Net*, the daily market return, and either a news day or earnings day dummy to control for market risk. These interactions are also included in the fifth regressions in Panels A and B of Table 3.

In the first regression of Table 4 the information day variable is an earnings announcement. In this regression the *Long* coefficient is 0.004, while the *Long x Earnings Announcement* interaction coefficient is 0.022, showing an increase in expected returns of slightly more than 550% on earnings announcement days. The news day interaction is 0.001 and not significant. Hence, on the long side, the effects are largely from earnings announcements.

The effects on the short side are even stronger. The *Short* coefficient is -0.002, while the *Short x Earnings Announcement* interaction coefficient is -0.022, showing that the incremental impact of short anomalies on earnings announcement days is 11 fold that of a typical day. The news day interaction is -0.066 and highly significant.

Various authors [for example, Miller (1977)] argue that if short-selling imposes extra-costs on short sellers, overvaluation situations will be more frequent than undervaluation situations. On the surface, the symmetry of the long and short interactions run counter to such an argument. The overall effect is that on earnings days that are also news days the overall short coefficient is -0.002 + -0.020+ -0.006 = -0.028, whereas the overall long coefficient is 0.004 + 0.022 + 0.001 = 0.027. One reason is that short-specific costs are holding costs, which are proportional to holding period length [Pontiff, 1996]. In this case, we expect the incremental costs of shorting around news days and earnings days to be minor.

In column 2 we replace the 1-day window with a 3-day window for the news and earnings announcements. The results are similar. The magnitudes are smaller, which is to be expected with the longer window, however the signs and significance of the coefficients are unchanged.

Taken together, the results in Tables 3 and 4 are consistent with the idea that mispricing, and specifically, biased expectations play an important role in explaining cross-sectional return predictability. The long side of anomaly strategies tend to do especially well on days when new information is released, whereas stocks on the short side have especially low returns on days when information is released. Hence, investors seem to be expecting too much from the short side firms, and too little form the long side firms.

The results here are very different than what we should observe in an efficient market where investors have rational expectations. In the rational expectations world, cross-sectional differences in stock returns are explained by cross-sectional differences in expected returns. New information is random, since the release of this new information should not have a predictable impact on returns. Instead, Tables 3 and 4 shows that the effect of new information on prices is predicted ex-ante. These results are also summarized in Figure 1.

3.3 Do the Effects vary Across Predictor Types?

In this section of the paper we ask whether the type of information used to create the predictor affects the results in the previous section. McLean and Pontiff (2014) categorize predictors into four different types: (i) Event; (ii) Market; (iii) Valuation; and (iv) Fundamentals. The categorization is based on the information needed to construct the predictor.

Event predictors are based on events within the firm, external events that affect the firm, and changes in firm-performance. Examples of event predictors include share issues, changes in financial analyst recommendations, and unexpected increases in R&D spending. Market predictors are predictors that can be constructed using only financial data, such as volume, prices, returns and shares outstanding. Momentum, long-term reversal, and market value of equity are included in our sample of market predictors.

Valuation predictors are ratios, where one of the numbers reflects a market value and the other reflects fundamentals. Examples of valuation predictors include sales-to-price and market-to-book. Finally, fundamental predictors are those that are constructed with financial statement data and nothing else. Leverage, taxes, and accruals are fundamental predictors.

We construct the same *Net* variable as before, only we sum up the portfolio memberships within each of the four groups. As in the previous tables, the regressions include time fixed effects, the lagged control variables used in the previous tables, controls for the market factor interacted with information day, and standard errors clustered on time.

We report the results from these tests in Table 5. Panel A reports the results from the regression, while Panel B reports the results from linear restriction tests that compare the effects among the four anomaly types.

The regression in Panel A shows that all four of the anomaly types have significantly higher returns on earnings announcement days. Hence, the results in the previous tables are not driven by a few anomalies or just one type of anomaly; instead the effects are common across all types of anomalies. With respect to news days, 3 of the 4 anomaly types have positive and significant interactions. Fundamental anomalies have a negative and significant interaction. The coefficient for fundamental anomalies is 0.001, whereas the news day interaction is -0.004. Taken together, the two coefficients show that fundamental anomalies tend to have negative returns on news days, in stark contrast to the other anomaly types. The earnings day interaction for fundamental anomalies is 0.020, showing that on earnings days fundamental anomalies have positive and significant alphas.

Panel B tests whether the interactions vary across the predictor types. One salient result is that market predictors, which are based solely prices, returns, variance of returns, and trading volume, have the lowest earnings day effects but the highest news day effects. Valuation predictors, which are based on ratios of price to fundamentals, have the highest earnings day effects, although the difference relative to fundamental predictors is not statistically significant.

Focusing on the relative, as opposed to absolute, change in the net interaction coefficient and the net coefficient, produces a starker comparison. Characteristics based on accounting data produce more extreme long-short returns on earnings days, characteristics based on market data have more extreme returns on non-earnings news days, and characteristics which mix this data fall in the middle. The highest relative earnings day change is 20 times for fundamentals, and the lowest is 3 times for market anomalies. The highest relative non-earnings news day change is over 4 times for market anomalies, and lowest is negative 4 times for fundamentals.

3.4. What Portion of Abnormal Returns are Earned on Information Days?

In this section of the paper we decompose each predictor's return into returns earned on information days and returns earned on non-information days. This decomposition allows us to place a lower bound on the importance of information releases. As we explain before, this is a lower bound because it is well-documented that earnings announcements are persistent and produce drifts in stock returns, and because there can be information about the firm that is released but not covered by Dow Jones

To conduct this exercise we do the following. For each firm-day observation, we first measure the firm's abnormal return as the firm's return minus the value-weighted market return on the same day. Then, for each predictor portfolio, we sum up all of the abnormal returns on information days and on non-information days separately. We also count the number of information firm-days and the number of non-information firm-days in each predictor portfolio. This exercise allows us to say what percentage of a predictor's returns are earned on information days and what percentage of a predictor's returns are earned on non-information days. Here we define an information day as the 3-day window around either an earnings announcement or news story.

As an example, consider a predictor that over our sample period has 1,000 firm-day observations in total. Assume that 300 of these are information days. Assume that the abnormal firm-day returns in total sum to 5,000 basis points; 3,000 of which are earned on information days and 2,000 of which are earned on non-information days. This allows us to state that for this predictor information days account for 30% of the total days, and 60% of the total returns. We conduct this exercise of each of the predictor portfolios in our sample and report the averages in Table 6.

We report results for the full 97-predictor sample, and for the four predictor types. With respect to the full 97-predictor sample, we see that information days account for 34.5% of the firm-days on the long side, and 80.1% of the returns. The results are similar on the short side. Information days account for 34.6% of the firm-days, and 84.8% of the returns. These results are consistent with the previous tables.

The results are robust across the four different predictor types. Among the predictor types, the results are strongest for the market predictors. Within this group of predictors on the short side information days account for 33.6% of the firm day returns, and 107.7% of the market returns. The price, bid ask spreads, volume, and Amihud illiquidity measure predictors drive this effect, as the long side returns for these predictors are almost entirely explained by returns on information days. This effect is not as salient in Tables 3 and 4, which reports results from tests that use an aggregate predictor variable that mutes the effect of any single predictor.

Taken in their entirety, the results in Table 6 reinforce the idea that biased expectations play a pivotal in explaining cross-sectional return predictability. Returns on information days are approximately 2 to 3 times more important in explaining predictor portfolio returns as compared to returns on non-information days.

3.5. Analysts Forecast Errors

In this section of the paper we ask whether our anomaly variables predict analyst forecast errors. The results thus far suggest that cross-sectional return predictability is the result if biased expectations. It seems that investors' expectations are too negative (positive) for stocks on the long (short) side of predictor portfolios. When new information is released, investors update their beliefs, resulting in high (low) returns for stocks on the (long) short side of predictor portfolios. If biased expectations do explain these effects, then we might also find that analysts' forecasts are too low (high) for stocks on the long (short) side of the predictor portfolios. We report tests of this hypothesis in Table 6.

Our analyst forecast error variable is from IBES. It is the difference between a stock's last reported median sell-side forecast and the actual reported earnings, divided by the closing price in the previous month. We have data from IBES for the period 1983 through 2014. The biased expectations framework predicts that this variable will be negative for the long side stocks (forecast too low) and positive for the short side stocks (forecast too high). We merge the forecast data with our anomaly data, and test whether anomaly portfolio membership can predict forecast error.

We control for the number of analysts making forecasts, whether there is only a single forecast, and the standard deviation of the forecast. If there is only a single forecast we set the standard deviation of the forecast equal to zero. We also include time fixed effects and cluster our standard errors on time. We do not include firm-level controls because the firm level variables that we would include are also predictors (e.g., size, price, book-to-market).

We report the results from these tests in Table 7. The first regression reports the findings for the full 97-anomaly samples. The regression coefficients show that analyst forecasts are too high for stocks in the short side of anomaly portfolios, and too low for stocks in the long side of anomaly portfolios. Both of these effects are statistically significant. These results share similarities with Edlen, Ince, and Kadlec's (forthcoming) who show that earnings day anomaly returns are more pronounced when institutional investors are underinvested in the high return leg and overinvested in the low return leg.

The effects are economically significant too. Our forecast error variable has a mean value of 0.107 (not in tables). Table 2 shows that *Long* and *Short* have standard deviations of 5.02 and 5.94. Combining these statistics with the coefficients in Table 7, we see that a one standard deviation increase in *Long* results in a -0.041 decrease in expected forecast error, whereas a one standard deviation increase in *Short* leads to a 0.080 increase in expected forecast error.

Table 7 also reports the effects across the 4 anomaly groups. We see that in all four groups the *Short* variable is positive and significant, showing that analysts' expectations are too pessimistic for firms in all types of short side anomaly portfolios. With respect to the *Long* variable, it is negative and significant for three of the anomaly groups, but positive and significant for the valuation anomaly group. Hence, for the long side of valuation anomalies earnings tend to be lower than what analysts expect. As we explain earlier, valuation anomalies include variables are ratios of price to some accounting variable, e.g., sales-to-price, earnings-to-price, etc.

Taken in their entirety, the results in Table 7 largely agree with the results in the other tables. Investors and analysts seem to be too pessimistic (optimistic) about stocks in the long (short) side of anomaly portfolios. This bias is revealed in stock returns when firms announce earnings and other news, and in analysts' forecast errors.

3.6. Can Data-Mining Explain Cross-Sectional Return Predictability?

As we explain in the Introduction, Fama (1998) and Harvey, Lin, and Zhu (2014) stress that data-mining could explain at good deal of cross-sectional return predictability. We conjecture that an anomaly that earns high returns due to luck should have lower information day effects as compared to an anomaly that is driven by the biased expectations of investors. Our tests are complicated, since returns are more volatile on information days, so even an anomaly that is the result of data-mining might do especially well on information days. In our sample, the typical earnings day has a return standard deviation that is 108% greater than a non-news day, and the typical non-earnings news day has a return deviation that is 30% greater than non-news days. Our conjecture is that if an anomaly's returns are the result of biased expectations, then the anomaly should have a greater information effect than an anomaly that reflects pure data-mining.

To conduct our data-mining test we first create a *Net Portfolio* variable that is equal to 1 if the stock is in the top quintile of a sort based on Net, -1 if the stock is in the bottom quintile, and zero otherwise. Then, for each stock in month *t* with a *Net Portfolio* value of 1 we find a stock with the same return (or as close as possible) in month *t*, which does not have a *Net Portfolio* value of 1. Similarly, for each stock in month *t* with a *Net Portfolio* value of -1, we find a stock with same the same return in month *t* that does not have a *Net Portfolio* value of -1. We do this for every stock in every month with a *Net Portfolio* value of either 1 or -1, thereby creating a *Pseudo Net Portfolio* variable.

As an example, assume that GE has a *Net Portfolio* value of 1 in June 1988, and that GE had a return of 1.5% for that month. Apple also had a return of 1.5% for June, but did not have a *Net Portfolio* value of 1. Apple could then be used in the *Pseudo Net Portfolio* for June 1988.

Table 9 reports the results for our pseudo tests. We report two different specifications. The first regression compares the effects of earnings days and news days between the real and pseudo *Net* portfolios. Regression2 excludes the worst matches from both the real and pseudo *Net* portfolios. To evaluate the closeness of the match we measure the difference in monthly returns (real anomaly stock's return – matched stock's return) between the real and matched firms. We then exclude the top and bottom 1% of the sample based on the return matches.

The first result of interest in Table 9 is that as expected the pseudo portfolio has positive and significant information day interactions. As we explain above, returns are more volatile on information days, so a strategy that generates returns by luck would almost have to perform well on these days too in order to have high returns.

Table 9 shows that it is also the case that in every specification the interaction terms for the real portfolios are greater. Regression 2 is perhaps our most reliable specification, as it excludes the bad matches. In regression 2, the earnings day interaction is 0.141 for the pseudo portfolio, whereas the interaction for the real portfolio is 0.212, or 50% higher. The bottom row reports a test of whether this difference is statistically significant, and we find that this is the case. In regression 2 the news day interaction is 0.006 for the pseudo portfolio, and 0.029 for the real anomaly portfolio, or 383% higher. This difference is also statistically significant.

Taken together, the results in Table 9 show that predictor portfolio returns have stronger information day effects as compared to what one would expect if return predictability were entirely the outcome of data-mining.

Conclusions

Evidence of cross-sectional return-predictability goes back at least 41 years to Blume and Husick (1972), yet to this day academics disagree about the cause. In this paper we compare return predictability on news and non-news days, and provide evidence that is consistent with return predictability being caused by mispricing, and in particular, mispricing based on biased expectations. Our findings are consistent with investors have overly optimistic expectations about the cash flows of some firms, and overly pessimistic expectations about the cash flows of other firms. Our results suggest that investors are surprised by news. When new information is released investors revise their biased beliefs, which in turn, causes prices to change, which in turn, causes the observed return predictability. Evidence from sell-side equity earnings forecasts dovetail with the stock return evidence: analysts overestimate the earnings for firms on the short-side of anomaly portfolios and underestimate earnings for firms on the long-side.

References

Baker, Malcolm & Litov, Lubomir & Wachter, Jessica A. & Wurgler, Jeffrey, 2010. "Can Mutual Fund Managers Pick Stocks? Evidence from Their Trades Prior to Earnings Announcements," Journal of Financial and Quantitative Analysis 45, 1111-1131.

Barberis, N. and R. Thaler (2003), "A Survey of Behavioral Finance," in George Constantinides, Milton Harris, Rene Stulz eds., Handbook of the Economics of Finance, North-Holland.

Basu, S., 1977, "Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Markets Hypothesis", Journal of Finance 32, 663-682.

Bernard, V.L., Thomas, J.K., 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. Journal of Accounting and Economics 13, 305-340.

Bernard, V. and J. Thomas, 1989, Post-earnings-announcement drift: Delayed price response or risk premium? Journal of Accounting Research, Suppl. 27, l-36.

Blume, Marshal E. and Frank, Husic, 1973, "Price, beta, and exchange listing," Journal of Finance 28, 283-299.

Bonferroni, C. E. 1935, "Il calcolo delle assicurazioni su gruppi di teste." In Studi in Onore del Professore Salvatore Ortu Carboni. Rome: Italy, pp. 13-60.

Chopra, N., J. Lakonishok, and J. Ritter, 1992, Measuring abnormal performance: Do stocks overreact?, Journal of Financial Economics 31, 235-268.

De Bondt, Werner and Richard Thaler, 1985, " Does the Stock Market Overreact?," Journal of Finance 40, 793-805.

De Bondt, Werner and Richard Thaler, 1987. "Further evidence of stock marker overreaction and seasonality", Journal of Finance 42, 557-81.

De Long, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1990, "Noise trader risk in financial markets," Journal of Political Economy 98, 703–738.

Duan, Ying, Gang Hu, and R. David McLean, 2009, "When is Stock-Picking Likely to be Successful? Evidence from Mutual Funds," Financial Analysts Journal 65, 55-65.

Duan, Ying, Gang Hu, and R. David McLean, 2009, "Costly Arbitrage and Idiosyncratic Risk: Evidence from Short Sellers," Journal of Financial Intermediation 19, 564-579.

Engelberg, Joey, Adam Reed, and Matt Ringgenberg, 2012, "How are shorts informed?: Short sellers, news, and information processing," Journal of Financial Economics 105, 260-278.

Edelen, Roger, Ozgur Ince, and Gregoy Kadlec, forthcoming,"Institutional Investors and Stock Return Anomalies," Journal of Financial Economics.

Fama, Eugene F., 1991, "Efficient capital markets: II," Journal of Finance 46, 1575-1617.

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

Fama, E., and K. R. French, 1996, Multifactor Explanations of Asset Pricing Anomalies, Journal of Finance 51, 55–84.

Fama, E. 1998, Market Efficiency, Long-term Returns and Behavioral Finance, Journal of Financial Economics 49, 283-306.

Foster, G., C. Olsen, and T. Shevlin. 1984, Earnings Releases, Anomalies, and the Behavior of Security Returns, The Accounting Review, 574-603.

Frazzini Andrea and Owen Lamont, 2006, The earnings announcement premium and trading volume, NBER Working Paper.

Harvey, Campbell R., Yan Liu, and Heqing Zhu, 2014, "… and the cross-section of expected returns," unpublished working paper, Duke University

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

Lakonishok J., A. Shleifer, and R. Vishny, 1994, Contrarian Investment, Extrapolation, and Risk, Journal of Finance 49, 1541-1578.

La Porta, R., J. Lakonishok, A. Shleifer, and R. Vishny. 1997. Good news for value stocks: further evidence on market efficiency. Journal of Finance 52, 859-874.

McLean, R. David and Jeffrey Pontiff, 2014, "Does academic research destroy stock return predictability?," *Journal of Finance*, Forthcoming.

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

Pontiff, Jeffrey, 1996, "Costly arbitrage: Evidence from closed-end funds," *Quarterly Journal of Economics* 111, 1135-1151.

Pontiff, Jeffrey, 2006, "Costly arbitrage and the myth of idiosyncratic risk," *Journal of Accounting and Economics* 42, 35-52.

Shleifer, Andrei, and Robert W. Vishny, 1992, "Liquidation Values and Debt Capacity: A Market Equilibrium Approach," Journal of Finance 47, 1343-1366.

Figure 1: Anomaly Returns around Earnings Announcement Days

This table reports the coefficients from regressions of daily returns on the aggregate anomaly variables *Long* and *Short*, dummies for 3-day windows around earnings announcements, interactions between *Long* and *Short* and the 3-day window dummies, and day fixed effects. *Long* and *Short* are defined in Table 2. The Figure plots the sum of the coefficients for the interactions and the coefficients for *Long and Short*, i.e, we plot the overall effect of *Long* and *Short* for each of the seven different 3-day windows.

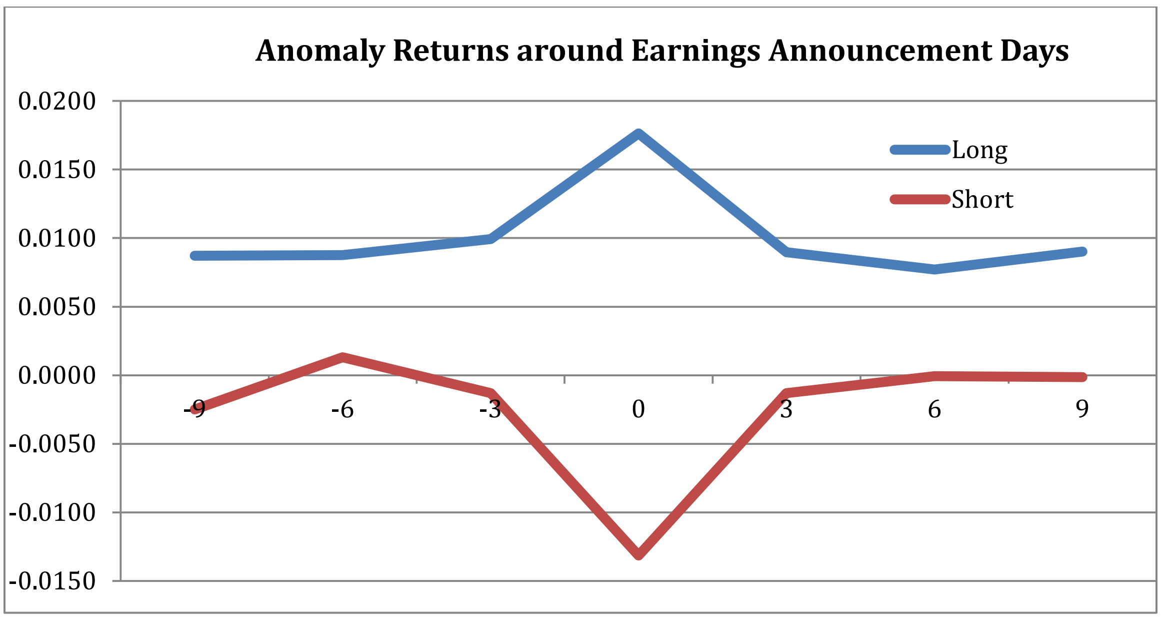

Table 1: Earnings Announcement and News Data

This table describes our sample in terms of earnings announcements and news releases. The unit of observation is at the firm-day level. To be included in our sample a stock must have return data reported in both the CRSP monthly and daily stock returns databases, and have a stock price that is at least $5. We obtain earnings announcement dates from the Compustat quarterly database, and news announcements from the Dow Jones news archive. We define an earnings day or news day as the day of an earnings announcement or news release. If the announcement is made after hours then the following day is the event day. The sample period is from 1979-2013.

Number of Firm-Day Returns

	News Day		Total
Earnings Day	No	Yes	
No	33,510,434	6,223,007	39,733,441
Yes	256,745	230,251	486,996
Total	33,767,179	6,453,258	40,220,437

Percentage of Firm-Day Returns

	News Day		Total
Earnings Day	No	Yes	
No	0.833	0.155	0.988
Yes	0.006	0.006	0.012
Total	0.840	0.160	1.000

Table 2: Descriptive Statistics for the Portfolio Variables

This table provides descriptive statistics for the predictor variables. We use the 97 cross-sectional predictors studied in McLean and Pontiff (2015). Each month, stocks are sorted on each predictor characteristic (e.g., size, book-to-market, accruals, etc.). We use the extreme quintiles to define long and short side of each predictor strategy. 16 of our 97 predictors are indicator variables (e.g, credit rating downgrades). For these predictors, there is only a long or short side, based on the binary value of the indicator. We remake the predictor portfolios each month. For each firm-day observation, we sum up the number of long side and short side predictor portfolios that the firm belongs to; this creates the variables *Long* and *Short*. The variable *Net* is equal to *Long* – *Short*.

Aggregate Anomaly Variables

Variable	**Observations**	**Mean**	**Std. Dev.**	**Min**	**Max**
Long	40,220,437	8.61	5.07	0	37
Short	40,220,437	9.23	5.94	0	44
Net	40,220,437	-0.61	6.12	-36	32

Table 3: Predictor Returns on Information Days vs. Off Information Days

This table reports results from a regression of daily returns on time fixed effects, the *Net* predictor variable, an information day dummy variable, interactions between the *Net* and the information day variables, and control variables (coefficients unreported). The control variables include lagged values for each of the past 10 days for stock returns, stock returns squared, and trading volume. To create the *Net* predictor variable we use the 97 cross-sectional predictors studied in McLean and Pontiff (2015). For each stock-month observation, we sum up the number of long side and short side predictor portfolios that the stock belongs to, thereby creating *Long* and *Short*. *Net* is equal to *Long* minus *Short*. We then merge this monthly dataset with daily stock return data from CRSP and with daily indicators for earnings announcement days and Dow Jones News stories, which we refer to as information days. We define an earnings day or news day as the 1-day or 3-day window around an earnings announcement or news release, i.e., days t-1, t, and t+1. *Factor* is the returns of a portfolio that is long the stocks in the highest quintile of *Net* and short the stocks in the lowest quintile of *Net*. *Market Portfolio* is the return of the CRSP value-weighted portfolio. The sample period is from 1979-2013. Standard errors are clustered on time.

Table 3: (Continued)

Panel A: 1-day Window					
Net	0.003	0.003	0.003	0.003	0.004
	(6.97)***	(6.89)***	(6.28)***	(6.69)***	(13.22)***
*Net * Eday*	0.020	0.018	0.020	0.020	0.021
	(12.11)***	(10.60)***	(12.29)***	(12.24)***	(13.01)***
*Net * Nday*	0.003	0.004	0.002	0.002	0.002
	(5.77)***	(5.05)***	(3.33)***	(4.26)***	(4.02)***
Eday	0.202		0.199	0.207	0.207
	(19.33)***		(12.14)***	(17.74)***	(17.76)***
Nday	0.150		0.118	0.106	0.106
	(23.35)***		(7.53)***	(17.55)***	(17.78)***
Factor			-0.931		
			(38.02)***		
*Factor * Eday*			0.030		
			(0.61)		
*Factor * Nday*			-0.461		
			(11.20)***		
Market				0.737	0.726
				(114.86)***	(112.69)***
*Market * Eday*				0.033	0.030
				(2.04)**	(1.88)*
*Market * Nday*				0.319	0.309
				(34.38)***	(31.44)***
*Net * Market*					-0.023
					(31.99)***
*Net * Mrkt. * Eday*					-0.003
					(1.45)
*Net * Mrkt. * Nday*					0.004
					(5.20)***
Observations	39,860,610	39,860,610	39,860,610	39,860,610	39,860,610
Fixed Effects	Day	Day * Event	None	None	None

Table 3 (Continued)

Panel B: 3-day Window					
Net	0.003	0.003	0.003	0.003	0.004
	(6.89)***	(6.95)***	(6.38)***	(6.73)***	(13.22)***
*Net * Eday*	0.010	0.010	0.011	0.011	0.011
	(11.58)***	(11.87)***	(12.23)***	(11.93)***	(13.19)***
*Net * Nday*	0.002	0.002	0.001	0.001	0.001
	(4.23)***	(4.74)***	(1.47)	(2.66)***	(2.10)**
Eday	0.082		0.069	0.083	0.084
	(15.23)***		(5.46)***	(11.73)***	(11.73)***
Nday	0.102		0.080	0.065	0.065
	(19.36)***		(6.46)***	(12.59)***	(12.64)***
Factor			-0.886		
			(36.64)***		
*Factor * Eday*			0.083		
			(1.93)*		
*Factor * Nday*			-0.414		
			(11.81)***		
Market				0.705	0.696
				(105.28)***	(104.63)***
*Market * Eday*				-0.004	-0.005
				(0.33)	(0.44)
*Market * Nday*				0.293	0.282
				(33.31)***	(32.69)***
*Net * Market*					-0.023
					(30.28)***
*Net * Mrkt. * Eday*					0.000
					(0.38)
*Net * Mrkt. * Nday*					0.003
					(3.99)***
Observations	39,860,610	39,860,610	39,860,610	39,860,610	39,860,610
Fixed Effects	Day	Day * Event	Day	Day	Day

Table 4: Long and Short Predictor Returns on Information Days vs. Off Information Days

This table reports results from a regression of daily returns on time fixed effects, the *Long and Short* predictor variables, an information day dummy variable, interactions between *Long and Short* and the information day variables, and control variables (coefficients unreported). The controls include lagged values for each of the past 10 days for stock returns, stock returns squared, and trading volume. We also include as controls for market risk interactions between the information day dummies and the daily return of the market portfolio (info day x market), and this variable interacted with *Net* (info day x market x *Net*). To create the *Long and Short* predictor variable we use the 97 cross-sectional predictors studied in McLean and Pontiff (2015). For each stock-month observation, we sum up the number of long side and short side predictor portfolios that the stock belongs to, thereby creating *Long* and *Short*. We then merge this monthly dataset with daily stock return data from CRSP and with daily indicators for earnings announcement days and Dow Jones News stories, which we refer to as information days. We define an earnings day or news day as the 1-day or 3-day window around an earnings announcement or news release, i.e., days t-1, t, and t+1. The sample period is from 1979-2013. Standard errors are clustered on time.

Table 4 (Continued)

	1-day Window	3-Day Window
Long	0.004	0.004
	(10.84)***	(11.47)***
Short	-0.002	-0.002
	(4.24)***	(3.89)***
*Long * Eday*	0.022	0.010
	(9.89)***	(9.08)***
*Short * Eday*	-0.020	-0.011
	(10.49)***	(11.33)***
*Long * Nday*	0.001	0.001
	(0.88)	(1.52)
*Short * Nday*	-0.006	-0.004
	(9.39)***	(7.28)***
Nday	0.194	0.118
	(17.87)***	(13.64)***
Eday	0.182	0.093
	(6.67)***	(6.57)***
Observations	39,860,610	39,860,610
Day Fixed Effects?	Yes	Yes
Market Risk Controls?	Yes	Yes

Table 5: The Effect of Information Across Predictor Types

This table tests whether the effect of information on predictor returns varies across different types of predictors. To conduct this exercise we split our predictors into the four groups created in McLean and Pontiff (2015): (i) Event; (ii) Market; (iii) Valuation; and (iv) Fundamentals. Event predictors are those based on corporate events or changes in performance. Examples of event predictors are share issues, changes in financial analyst recommendations, and unexpected increases in R&D spending. Market predictors are predictors that can be constructed using only financial data, such as volume, prices, returns and shares outstanding. Momentum, long-term reversal, and market value of equity (size) are included in our sample of market predictors. Valuation predictors are ratios, where one of the numbers reflects a market value and the other reflects fundamentals. Examples of valuation predictors include sales-to-price and market-to-book. Fundamental predictors are those that are constructed with financial statement data and nothing else. Leverage, taxes, and accruals are fundamental predictors. The regressions include time fixed effects and controls for lagged values for each of the past 10 days for stock returns, stock returns squared, and trading volume (coefficients unreported). We also include as controls for market risk interactions between the information day dummies and the daily return of the market portfolio (info day x market), and this variable interacted with *Net* (info day x market x *Net*). Standard errors are clustered on time. Panel B reports the results of linear restriction tests that ask whether the various coefficients are equal or different.

Table 5: (Continued)

Panel A: Regression Results

Market	0.003
	(4.42)**
*Market * Eday*	0.010
	(2.88)**
*Market * Nday*	0.013
	(9.92)**
Valuation	0.004
	(4.95)**
*Valuation * Eday*	0.034
	(8.35)**
*Valuation * Nday*	0.005
	(4.51)**
Fundamental	0.001*
	(1.67)
*Fundamental * Eday*	0.020
	(4.88)**
*Fundamental * Nday*	-0.004
	(3.77)**
Event	0.003
	(7.10)**
*Event * Eday*	0.022
	(6.21)**
*Event * Nday*	0.002
	(2.58)**
Eday	0.191
	(18.41)**
Nday	0.147
	(31.52)**
Observations	39,860,610
Day Fixed Effects?	Yes
Market Risk Controls?	Yes

Table 5: (Continued)

Panel B: Linear Restriction Tests

Earnings Day Tests	Difference	p-value
Market – Valuation = 0	-0.024	0.000
Market – Fundamental = 0	-0.010	0.050
Market – Event = 0	-0.012	0.023
Valuation – Fundamental = 0	0.014	0.819
Valuation – Event = 0	0.012	0.000
Fundamental – Event = 0	-0.002	0.023

News Day Tests	Difference	p-value
Market – Valuation = 0	0.008	0.000
Market – Fundamental = 0	0.017	0.000
Market – Event = 0	0.011	0.000
Valuation – Fundamental = 0	0.009	0.000
Valuation – Event = 0	0.003	0.021
Fundamental – Event = 0	-0.006	0.000

Table 6: The Relative Importance of Information Days

In this Table we document the relative importance of information days in explaining predictor returns. For each firm-day observation, we first measure the firm's abnormal return as the firm's return minus the value-weighted market return on the same day. Then, for each predictor portfolio, we sum up all of the abnormal returns on information days and on non-information days separately. We also count the number of days that are information days and the number non-information days for each predictor portfolio. This exercise allows us to say what percentage of a predictor's days are information days and what percentage of the predictor's returns is from information days. We conduct this exercise of each of the predictor portfolios in our sample and report the average. We define an information day as the 3-day window around an earnings announcement or news release, i.e., days t-1, t, and t+1. The sample period is from 1979-2013.

Table 6: (Continued)

Long Side	**Full Sample**	**Market**	**Valuation**	**Fundamental**	**Event**
Percentage of Days	0.345	0.319	0.326	0.358	0.367
Percentage of Returns	0.801	0.959	0.863	0.741	0.683

Short Side	**Full Sample**	**Market**	**Valuation**	**Fundamental**	**Event**
Percentage of Days	0.346	0.336	0.345	0.367	0.338
Percentage of Returns	0.848	1.077	0.747	0.766	0.766

Table 7: Analysts' Forecast Errors

In this table we test whether predictors are related to analysts' forecast errors. The dependent variable is analysts' forecast error, which is measured as the median earnings forecast minus the actual reported earnings, scaled by last month's closing stock price. We use the median quarterly forecast from the latest IBES statistical period, or the last date that IBES computed its summary statistics for the firms' earnings forecasts. *Number of Estimates* is the number of analysts issuing forecasts. *Single Forecast* is a dummy equal to 1 if only one analyst makes a forecast for the firm, and zero otherwise. *Dispersion* is the standard deviation of the forecasts. We set dispersion equal to zero if *Single Forecast* is equal to 1. The variables *Long* and *Short* and the different anomaly samples are defined in the previous tables. The regressions include time fixed effects. Standard errors are clustered on time.

	Full Anomalies Sample	Market	Valuation	Fundamental	Event
Long	-0.008	-0.017	0.020	-0.009	-0.018
	(14.79)***	(11.71)***	(7.58)***	(7.10)***	(14.21)***
Short	0.014	0.025	0.022	0.028	0.025
	(22.91)***	(14.18)***	(13.82)***	(15.23)***	(20.31)***
Number of Estimates	-0.010	-0.010	-0.008	-0.006	-0.007
	(20.15)***	(19.04)***	(18.30)***	(17.04)***	(18.68)***
Single Forecast	0.133	0.126	0.112	0.118	0.128
	(17.20)***	(16.17)***	(14.94)***	(15.50)***	(16.51)***
Dispersion	0.000	0.000	0.000	0.000	0.000
	(3.46)***	(3.45)***	(3.50)***	(3.50)***	(3.46)***
Intercept	0.063	0.109	0.066	0.083	0.096
	(9.86)***	(20.92)***	(14.25)***	(17.63)***	(19.51)***
Month Fixed Effects?	Yes	Yes	Yes	Yes	Yes
Observations	294,535	294,535	294,535	294,535	294,535

Table 8: Real Anomalies vs. Pseudo anomalies

In this Table we compare the effects of information releases on real anomaly portfolios vs. pseudo anomaly portfolios. We first create a real anomaly portfolio variable, *Net Portfolio*, which is based on *Net*. *Net Portfolio* is equal to 1 if the stock is in the highest quintile of *Net*, -1 if the stock is in the lowest *Net* quintile, and zero otherwise. To create the pseudo variable we find stocks that are not in the highest (lowest) *Net* portfolio, but have the same return as the stocks in the highest (lowest) *Net* portfolio. As an example, assume GE and DELL both have a 1% return in June. GE is in in the long (high) *Net* portfolio in June, but DELL is not. DELL can therefore be in the pseudo long *Net* portfolio for June. We repeat this procedure for every stock in the long and short *Net* portfolios for every month in our sample. The bottom row of the table reports tests of whether the information day effects are greater for the real *Net Portfolio* as compared to the *Pseudo Net Portfolio*. The regressions in the third and fourth columns exclude the worst matches, which are the 1st and 99th percentiles for differences in returns between the actual and matched firms. The regressions include time fixed effects and controls for lagged values for each of the past 10 days for stock returns, stock returns squared, and trading volume coefficients unreported). The standard errors are clustered on time.

	Full Sample	Excluding Bad Matches
Net Portfolio	0.032	0.032
	(9.68)***	(9.59)***
Pseudo Net Portfolio	0.037	0.034
	(25.62)***	(24.06)***
*Net Port. * Eday*	0.223	0.212
	(13.91)***	(13.38)***
*Pseudo Net Port * Eday*	0.156	0.141
	(9.22)***	(8.92)***
*Net Port * Nday*	0.032	0.029
	(6.16)***	(5.61)***
*Pseudo Net Port * Nday*	0.016	0.006
	(4.76)***	(1.94)*
Eday	0.192	0.182
	(18.28)***	(17.80)***
Nday	0.148	0.140
	(22.56)***	(21.72)***
Observations	39,861,109	39,713,540
Day Fixed Effects?	Yes	Yes
*Net * Eday = Pseudo * Eday*	0.000	0.000
*Net * Nday = Pseudo * Nday*	0.000	0.000