Anomalies and News

Authors: Joseph Engelberg , R. David McLean and Jeffrey Pontiff

Summary

The paper "Anomalies and News" examines 97 stock return anomalies, finding that their returns increase significantly on corporate news days (50% higher) and earnings announcement days (six times higher). These effects are consistent across anomaly types and remain decently robust after controlling for macroeconomic news, day-of-the-week effects, and extreme returns. Long-side anomalies see elevated returns, while short-side anomalies show declines on news days. The paper links these patterns to biased investor expectations, evidenced by analyst forecast errors, rather than risk or data mining. Out-of-sample anomalies display similar results, further supporting the mispricing explanation. While previous studies have suggested dynamic risk models help explain return predictability, they lack sufficient evidence to explain anomalies' heightened returns during information events. This paper is motivated to address these gaps and advance the understanding of anomaly behavior in the presence of firm-specific news. The study's goal is to distinguish among risk, mispricing, and data-mining explanations for anomaly returns. The authors provide evidence that biased expectations primarily drive anomaly returns, with mispricing corrected during news events. Their findings challenge traditional risk-based views and offer insights into the role of investor behavior in shaping market inefficiencies.

General Comments

The idea behind the paper is compelling, as it addresses an important question about why anomalies behave distinctively during news events or earnings announcements. The analysis provides valuable insights into a probable cause—biased expectations—but does not advance our understanding of why anomalies exist in the first place, limiting its contribution to the broader anomaly literature. The approach is well-executed and feasible for achieving the stated goals. The authors' methodological rigor and robust

controls, including macroeconomic and day-specific effects, fixed effects regressions, and out-of-sample testing, add credibility to the findings. However, their approach is not robust enough which weakens the overall makes the impact of their analysis.

Major Concerns

1) Robustness Issues

The paper's robustness is a significant concern, as it limits its analysis to NYSE firms and excludes stocks priced under $5. While this methodological choice aims to replicate the study by Mclean and Pontiff (2016), it inadvertently reduces the generalizability of the findings and increases the risk of data mining bias. Clipping or removing outliers might make more sense. Including firms from other exchanges, such as NASDAQ, and incorporating data from international markets would enhance the applicability of the results and reduce potential biases associated with limiting the sample. Moreover, analyzing data from other exchanges and countries would serve two important purposes: testing whether the anomalies still hold out-of-sample and determining if the news day and earnings day effects are consistent across different markets.

The authors also mention on page 11 that “omission and inclusion of correlated variables may cause changes in statistical significance” but they don’t consider conducting more factor-based tests to validate the analysis. For example, the authors mention on page 5 “ We show that this intuition … during the same month” which implies there might be short term momentum driving those anomaly returns. So, the authors could improve their robustness by introducing a short-term time-series momentum factor into their regressions and seeing if their earnings and news announcement results are still statistically significant.

Overall, expanding the dataset, testing across other markets and incorporating additional factors would make the analysis more rigorous and provide stronger evidence for the conclusions.

2) Standard errors seem understated

Another concern is the potential understatement of standard errors in the analysis. The authors use standard regression models and cluster standard errors by time, which is reasonable since Fama-MacBeth regressions might not yield significantly different results given that daily returns are unlikely to be independent. While this approach is suitable for controlling time-based dependencies, it fails to account for potential correlations within anomaly types or across firms. Specifically, anomalies within the same category (e.g., Event anomalies) are likely to be correlated. For example, within Event anomalies, variables such as share issues, changes in financial analyst recommendations, and unexpected increases in R&D spending could exhibit interdependence. Ignoring these correlations can lead to understated standard errors, inflating the statistical significance of the results and potentially yielding false positives.

To address this issue, the authors should consider clustering standard errors by firm or anomaly type in addition to time. This approach would account for intra-category correlations and firm-specific effects, providing a more accurate measure of statistical significance.

3) Results are not Convincing enough

The results presented in the paper are not entirely convincing due to several limitations and inconsistencies. On page 12, the authors report that expected returns are higher by an additional 21.64 basis points on earnings announcement days, which, while statistically significant, are objectively very small in percentage terms. When combined with concerns about potentially understated standard errors, as discussed earlier, these small differences raise doubts about whether they are meaningfully or reliably

different from zero. It is plausible that these small results are driven by measurement errors, further reducing confidence in their significance. Additionally, inconsistencies emerge in the analysis. For example, on page 16, one out of four short-side anomaly portfolios (market anomalies) is an exception for negative earnings day interactions, while fundamental anomalies are exceptions for short-side interactions on news days. Furthermore, on the long side, valuation and event interactions on news days are found to be insignificant. On page 32, another inconsistency arises, as the news day interaction is reported to be insignificant for large stocks but positive and significant for small stocks.

These exceptions and insignificant findings weaken the overall narrative of robust and pervasive effects. They could stem from a lack of robustness in the analysis, possible data overfitting, measurement error, or the restricted dataset focusing solely on NYSE firms with share prices greater than $5.

Overall, the combination of negligible effect sizes, inconsistencies across anomaly types and firm sizes, and potential methodological weaknesses undermines the persuasiveness of the results. The authors should address these concerns by expanding their dataset, strengthening their robustness checks, and reassessing the statistical significance of their findings to provide more compelling evidence.

Minor Concerns

1) Reorganize Results Presentation

The results section would benefit from clearer organization. For instance, separating findings by anomaly type (event, market, valuation, fundamental) into distinct subsections would improve readability and help highlight key differences.

2) Creating a methodology section

Adding to the Sample and Data section, it would be good to have a methodology section which outlines how the tests were conducted and an explicit list of different types of regressions performed (eg – one regression is for all anomalies, another one for individual anomaly types, one for large cap stocks, etc).

3) Add Robustness Checks

Include robustness tests that analyze the anomalies across different datasets, such as firms listed on NASDAQ or international exchanges. This would strengthen the generalizability of the findings and reduce concerns about data mining.

4) Reorganizing result tables

It is very inconvenient when the author mentions a table and we have to scroll all the way down to the end of the paper to look at the table results. It would help to have the tables where the authors discuss the results.

Unanswered Questions

Question 1

On Page 10, the authors mention that their counts for average stock being in a long or short portfolio is less as compared to 97 random quantile groupings because 16 of the variables they consider are indicator variables lacking a long and short side (example- credit rating downgrade). It would be helpful to know why they chose to exclude those variables when sorting portfolios. For example, if there is a rating downgrade for a stock, it might make sense to short it. Although the indicator variable lacks a short side, it can be inferred if It is negative or positive for a stock.

Question 2 - Helpful to explore in future research

Can you make money from the findings? What would the anomaly returns look like on news or earning events after taking into account transaction costs? What about after taking into account taxes for short term buying and selling?

Question 3 - Helpful to explore in future research

What is the effect on the anomaly returns when stocks for an anomaly do not have any news events, but other stocks to which they are highly correlated to have. For example, Delta and American airlines might be highly correlated stocks and news or earnings release comes out for American airlines. So, both stocks would move (and this usually happens with those stocks in practice). If both stocks are a part of different anomaly, would that affect the anomaly returns?

Directions for Future Research

Future research should take two key directions. First, it should focus on identifying why anomalies exist, ensuring they are not driven by data-mining bias, and enhancing our understanding of pricing and returns. Second, researchers should investigate the root causes of investor optimism and pessimism about cash flows, as highlighted in the conclusion of the paper. This includes exploring whether these biases are consistent across specific categories of stocks, industries, or other patterns. If a consistent pattern is found, it may suggest that abnormal earnings-day returns are not the cause of anomalies but stem from other underlying factors. If no pattern emerges, it strengthens the argument that anomaly returns may result from data mining. These directions would provide critical insights into the origins and drivers of anomalies.