HEC MONTRÉAL

Assessing the Credibility of Anomalies in Asset Pricing Research

par

Yilin Zhang

Vincent Grégoire
HEC Montréal
Directeur de recherche

Charles Martineau
University of Toronto
Codirecteur de recherche

Sciences de la gestion
Spécialisation Applied Financial Economics

Mémoire présenté en vue de l'obtention
du grade de maîtrise ès sciences
(M. Sc.)

December 2022

Résumé

Un grand nombre d'anomalies ont été découvertes par les chercheurs au cours des dernières décennies. Les méthodologies pour identifier les anomalies ont évolué au fil des années grâce aux innovations en matière de recherche et de technologie de négociation, ce qui suscite des inquiétudes quant à la validité des anomalies au fil du temps. Ce projet contribue à la littérature récente visant à réévaluer les anomalies établies, en répliquant 299 anomalies provenant de 145 publications et en réévaluant leur validité à l'aide des tests hors échantillon, des ajustements pour les firmes à petite capitalisation boursières (microcaps) et en utilisant un seuil de signification novateur. Les anomalies reproduites montrent une forte cohérence avec Chen & Zimmermann (2021) et les études originales. La régression des statistiques t répliquées dans la présente étude sur les statistiques t de Chen & Zimmermann (2021) produit une pente de 1 et un R carré de 99%. Une régression similaire sur les statistiques t originales donne une pente de 0,87 et un R carré de 80%. En examinant les périodes hors échantillon, on constate que les t-statistiques post-publication diminuent de 68,84% par rapport aux statistiques t pré-publication. La suppression des microcaps réduit les statistiques t des anomalies à la fois pour les échantillons pré-publication (-27,92% de 3,82 à 2,75) et hors échantillon (par exemple, post-publication : -35,83% de 1,23 à 0,79). De plus, après adoption d'un seuil de statistique t plus élevé de 3,0 proposé par Harvey et al. (2016), 57,81% des prédicteurs de rendement des actions en sections transversales restent significatifs dans l'échantillon pré-publication, tandis que seulement 9,38% des prédicteurs survivent après publication. Ainsi, les méthodologies actuelles pour identifier les anomalies devraient faire l'objet d'une analyse critique.

Abstract

A large number of anomalies have been discovered by researchers throughout the past decades. The methodologies for identifying anomalies have evolved through years of innovations in research and trading technology, which sparks concerns in the validity of the anomalies over time. As an extension to previous work in re-evaluating the established anomalies, the present study replicated 299 anomalies from 145 publications and reassessed their validity via out-of-sample tests, microcaps adjustments, and a novel significance threshold. My reproduced anomalies demonstrate high consistencies with Chen & Zimmermann (2021) and the original studies. The regression of the replicated t-statistics in the present study on Chen & Zimmermann's (2021) t-statistics produces a slope of 1 and an R squared of 99%. A similar regression on the original t-statistics yields a slope of 0.87 and an R squared of 80%. By examining out-of-sample periods, post-publication t-statistics are found to diminish by 68.84% compared to pre-publication. The removal of microcaps pulls down the t-statistics of anomalies for both in-sample (-27.92% from 3.82 to 2.75) and out-of-sample (e.g. post-publication: -35.83% from 1.23 to 0.79). Furthermore, after adopting a higher t-statistic threshold of 3.0 proposed by Harvey et al. (2016), 57.81% of the cross-sectional stock return predictors remain significant in-sample, while only 9.38% of the predictors survive post-publication. Thus, the current methodologies for identifying anomalies should be subject to scrutiny.

Acknowledgements

I would like to express my sincere gratitude to my supervisors Dr. Vincent Grégoire and Dr. Charles Martineau for their kindness, encouragement and generosity over the entire journey of this thesis. I am extremely grateful for their constructive feedback and guidance, which allows me to have a wonderful and unique experience. In addition, I am thankful for my family and my dog for their emotional and financial supports. Special thanks to my boyfriend for his comprehensive and wholehearted support.

Contents

List of Figures

List of Tables

List of Acronyms

AMEX American Stock Exchange

capex Capital Expenditure

CAPM Capital Asset Pricing Model

CRSP Center for Research in Security Prices

EPS Earnings per share

FRED Federal Reserve Economic Data

IBES Institutional Brokers' Estimate System

ISSM Institute for the Study of Security Markets

NASDAQ National Association of Securities Dealers Automated Quotations

NYSE New York Stock Exchange

SIC Standard Industrial Classification

TAQ Trades and Quotes

WRDS Wharton Research Data Services

1 Introduction

Replication crisis has been a challenge in scientific research across multiple disciplines over the past decade. Generally, this phenomenon could be attributed to less rigorous methods and false discoveries driven by research biases. To address these problems in empirical finance literature, the present study extends the efforts of Chen & Zimmermann (2021) in open collaboration by replicating and assessing 299 cross-sectional anomalies using Python. Anomaly performance is further evaluated with multiple testing, including out-of-sample tests, microcaps adjustments, and imposing a novel t-statistic threshold.

The present study is inspired by Chen & Zimmermann (2021), in which nearly 100% of the predictability results in the literature are successfully reproduced using the same methods and samples as the original studies. The aims of the present study are fourfold: (1) to reproduce cross-sectional anomalies from Chen & Zimmermann's (2021) study; (2) to reassess their predictability out-of-sample (McLean & Pontiff, 2016); (3) to reassess their in- and out-of-sample predictability with microcaps excluded (Hou et al., 2020), and (4) to reassess the predictability using Harvey et al.'s (2016) t-statistic threshold of 3.0, which is higher than the standard 1.96 in the finance literature. The novelty of the present work is (1) to reproduce a large number of published anomalies using an open-source programming language, Python, and (2) to combine evaluation methods established in the literature to re-evaluate the overall predictability performance of anomalies. Based on previous research, the performance of reproduced anomalies is expected to decline both out-of-sample (McLean & Pontiff, 2016; Chen & Zimmermann, 2021; Chordia et al., 2020; Linnainmaa & Roberts, 2018; Chordia et al., 2014) and with microcaps excluded (Hou et al., 2020; Fama & French, 2008; Novy-Marx & Velikov, 2016). Additionally, the proportion of significant anomalies is expected to decrease after adopting Harvey et al.'s (2016) higher t-statistic threshold.

To achieve a comprehensive reproduction, the detailed methodologies used by Chen & Zimmermann (2021) have been adopted to the greatest extent possible. In total, 299 published cross-sectional anomalies selected from 145 studies are reproduced. The benchmark are the long-short portfolios constructed using the original studies' in-sample period, quantile sort, stock-weighting, as well as filters on share type, price, and exchange. To evaluate the replication results, a regression of the reproduced t-statistics in the present study on Chen & Zimmermann's (2021) t-statistics is performed, which produces a slope of 1 and an R squared of 99%. A similar regression on the original t-statistics yields a slope of 0.87 and an R squared of 80%. Thus, the replication is successful.

To investigate the return predictability out-of-sample, each anomaly in the present study is examined across three eras in addition to the original in-sample periods: (1) full sample, (2) pre-publication, and (3) post-publication. Specifically, the full sample period, which includes all data available for download, is used to examine the overall robustness of anomalies. The pre- and post- publications are subsets of the full sample that contain the sample periods before and after the original publication year, reflecting the predictability performance before and after the discovery of anomalies, respectively. The present study finds that compared to in-sample, the average full-sample portfolio return declines by 26.46%. Such decline is mainly driven by the post-publication decay. Specifically, the average post-publication return and t-statistic diminish by 52.97% and 67.92% compared to in-sample. In addition, the same metrics decrease by 44.76% and 68.84%, compared to pre-publication. These results confirm the previous finding of performance degradation out-of-sample, especially post-publication (McLean & Pontiff, 2016; Linnainmaa & Roberts, 2018; Chen & Zimmermann, 2021).

To address the issues related to microcaps and assess the return predictability of anomalies in a more realistic trading scenario, all microcaps are excluded with New York Stock Exchange

(NYSE) breakpoints at portfolio construction. This filter allocates a fair proportion of small and big stocks into extreme quantiles while alleviating concerns regarding transaction costs and generalizability (Novy-Marx & Velikov, 2016; Hou et al., 2020). Consistent with Fama & French (2008) and Hou et al. (2020), the performance of anomalies is found to deteriorate after controlling for microcaps in the present study. Specifically, the average in-sample return and t-statistic decrease about 25.14% and 27.92%, respectively.

To examine the effects of Harvey et al.'s (2016) higher t-statistic threshold of 3.0 on the survival rate and performance of predictors, their returns and t-statistics under different conditions of sample periods and microcaps presence are computed. Predictors are assigned into two groups based on the reproduced in-sample t-statistics: (1) above 1.96 and below 3.0, (2) above 3.0. Within each group, predictor performance is assessed in portfolios containing all stocks and without microcaps, using in-sample and post-publication samples. Independent two sample t-tests are also performed to assess the true difference of post-publication performance between two groups of predictors. It is shown that only five predictors remain robust post-publication at the novel threshold of 3.0, regardless of microcaps adjustment. Specifically, the five predictors are: EPS forecast revision (Hawkins et al., 1984), earnings announcement return (Chan et al., 1996), predicted dividend yield next month (Litzenberger & Ramaswamy, 1979), down forecast EPS (Barber et al., 2001), and put volatility minus call volatility (Yan, 2011). The proportion of robust anomalies also decreases drastically after imposing the higher t-statistic threshold. The majority of the t-tests produces non-significant results, suggesting that imposing the higher t-statistic threshold in-sample has little effects on post-publication performance.

1.1 Literature Review

Many claimed research findings across multiple disciplines are hard to reproduce. Ioannidis (2005) presents a theoretical model to demonstrate that for the majority of disciplines, over 50% of the published research findings are likely false. Prinz et al. (2011) report that around 65% of the results from 67 published pharmaceutical research studies cannot be validated by other scientists. Similarly, Begley & Ellis (2012) show that out of 53 landmark studies in preclinical cancer research, only 11% of them can be reproduced. The Open Science Collaboration (2015) replicates 100 psychological research studies published in top journals, however, only 36% of the replications produce significant results.

The low reproducibility is also evident in economic research. In a replication study involving 67 publications from 13 economics journals, less than 50% of the results are successfully reproduced (Chang & Li, 2015). Similarly, for experimental economics, Camerer et al. (2016) only successfully replicate 61% of the findings out of 18 studies. Generally, this could be attributed to less rigorous methods and false discoveries driven by research biases. Empirical results are vulnerable to even the slightest change in model parameters (Leamer, 1983). In addition, researchers often fail to account for different implementations of analytical functions and algorithms across software packages, such as the nonlinear maximization routines in the publications of *American Economic Review* (McCullough & Vinod, 2003).

In terms of biases, both confirmation bias and conflicting incentives play a role in the poor reproducibility. To measure the confirmation bias, Fanelli (2010) randomly selects 2434 publications across all disciplines. It is found that Economics and Business demonstrate a probability of reporting supportive results five times higher than physical sciences, reflecting a greater confirmation bias. With regards to incentives, vague and close-source methodologies are often the basis of realizing profits in finance. When asymmetric information creates profit opportunity for market participants, incentives are not always aligned with transparency. In addition, publication is closely related to researchers' professional success (i.e. reputation, research funding). Further, journals strive to publish novel and high-impact studies, rather than making a balanced coverage for the field by including replication studies and low-impact studies with less significant results (Harvey, 2017). The publication biases along with the material reward and prestige associated with publication create a highly competitive research environment, which leads to implicit conflicts of interest. Specifically, researchers driven by such incentives are inclined to not only employ methods that produce significant and publishable results, but also selectively report novel and positive findings (Nosek et al., 2012; Harvey et al., 2016; Harvey, 2017). As a consequence, the odds of supportive results are likely inflated in publications (John et al., 2012; Simmons et al., 2016). According to a *Nature*'s survey of over 1500 researchers from diverse scientific disciplines, pressure to publish and selective reporting are believed to be the leading causes of irreproducibility (Baker, 2016). Such phenomenon could lead to false positive discoveries and cast doubts in the results of finance literature.

One prominent example would be cross-sectional stock return predictors, also known as anomalies. Over the last 40 years, hundreds of anomalies have been uncovered by researchers in both academia and industry at an impressive rate. According to Harvey et al. (2016), the true discovery rate has likely decreased as the anomalies with sound theoretical support had mostly been discovered; thus, many of the recently published factors in financial economics are likely false. Later, Harvey (2017) confirms that the incentives for publishing the most significant results put an emphasis on publishing new anomalies rather than verifying the existing ones, which exacerbates the concerns for p-hacking in empirical finance. As summarized by Cochrane (2011), it is a zoo of new factors and a systemic consolidation is needed.

Currently, extensive efforts have been put into replications, especially meta analyses, to navigate the zoo and re-evaluate the anomalies (Harvey et al., 2016; Linnainmaa & Roberts, 2018; Chordia et al., 2020; Chen & Zimmermann, 2021).

Multiple meta-analysis studies have revealed that many published anomalies are not reproducible. The irreproducibility in anomaly literature takes two basic forms: either (1) the main results cannot be replicated with the same or very similar data and methodologies, or (2) the successfully replicated in-sample results are attributed to data mining, and the significance fails to persist in other samples.

Hou et al. (2020) attempt to reproduce 447 published anomalies in literature by applying the same methods on the same samples but fail to achieve the original results in approximately 50% of the anomalies. For the reproduced anomalies that are found significant, however, their returns and t-statistics are often much lower in magnitudes than those reported in original publications. Furthermore, Hou et al. (2020) use NYSE breakpoints and value-weights to control for microcaps and evaluate anomaly performance in a more realistic trading scenario, which reveals that only 36% of the anomalies are significant. Linnainmaa & Roberts (2018) replicate 36 accounting-based anomalies and examine their in-sample, pre-sample, and post-sample performance, which shows that most of these anomalies are likely a product of data mining. In addition, Chordia et al. (2020) estimate that 45.3% of strategies are falsely accepted under conventional thresholds in classical hypothesis testing, because finance researchers fail to account for multiple hypothesis testing. However, in cases where multiple testing is applied, false discoveries are still present. For instance, Harvey et al. (2016) replicate 296 significant anomalies from top journals, among which 27% to 53% are considered as false discoveries after adjusting for multiple testing (Bonferroni, Holm, and BHY). Thus, false discoveries is one of the major contributing factors to the poor reproducibility rate in anomaly literature, which warrants the replication of published anomalies for cross-validation.

When the robust in-sample performance of an anomaly is replicated, the out-of-sample performance may fail to persist nevertheless. Previous studies on anomalies have shown contradictory findings on the performance of out-of-sample return predictability. For example, Jegadeesh & Titman (2001) show that the momentum effects continued to persist after their initial discovery in 1993. Whereas, Schwert (2007) demonstrates that since the initial publication of the value and size effects, the index funds based on these strategies fail to generate abnormal returns. Therefore, to fully understand the out-of-sample performance, meta-analysis studies including McLean & Pontiff (2016) and Chen & Zimmermann (2021) are conducted, which show a consistent post-publication decay in the performance of selected anomalies.

The identification and the performance of anomalies could also be affected by the selection of stocks. In the literature, two frequently used methods for identifying anomalies are: (1) sorts of returns on anomaly variables, and (2) cross-sectional regressions of returns on anomaly variables (Fama & MacBeth, 1973). However, stocks with a market capitalization less than the 20th percentile of the NYSE distribution, namely microcaps, are influential in both approaches (Fama & French, 2008). Although microcaps only account for 3% of the total market capitalization in NYSE-AMEX-NASDAQ universe, they make up about 60% of the total number of stocks (Fama & French, 2008). This suggests that microcaps are economically trivial but could make up about 60% of the stocks in each portfolio if it is sampled randomly. In addition, microcaps have the highest average cross-section standard deviations of the anomaly variables, compared to small or big stocks. Therefore, microcaps are more likely to end up in long-short portfolios constructed using sorts of all stocks and account for more than 60% of the stocks. In other words, for studies that employ sorts, the return of equal weighted long-short portfolios containing all stocks could be dominated by microcaps, resulting in a biased assessment of anomaly robustness. As for studies that employ Fama & MacBeth's (1973) cross-sectional linear regressions, returns are more sensitive to outliers such as microcaps than equal-weight in sorts (Fama & French, 2008; Hou et al., 2020). From a practical standpoint, if the abnormal returns of an anomaly are primarily driven by microcaps, the strategy is less likely to be profitable due to the high trading costs associated with microcaps. Therefore, it is important to determine whether the anomaly returns are skewed by microcaps to comprehensively assess the performance.

To improve the reproducibility of anomalies, researchers in the field have advocated to employ stricter statistical criteria for anomaly identification. Harvey et al. (2016) propose a higher t-statistic threshold of 3.0 for the current asset pricing research as opposed to the conventional threshold of 1.96. The asset pricing literature is heavily based on the same data from the CRSP database, which has led to overfitting. Furthermore, because truly robust anomalies with sound theoretical support have mostly been uncovered already, the true discovery rate has likely decreased (Harvey et al., 2016). As the technology advances, data mining has become more accessible. Thus, the anomalies derived from such approach are likely false discoveries. Therefore, it is crucial to examine the proportion of anomalies that survives Harvey et al.'s (2016) higher t-statistic threshold.

This paper is organized as follows. Section 2 describes the methodologies adopted in the present study, including anomaly categorization, anomaly generation, and portfolio construction. Section 3 details replication results and the main findings of the predictability assessments under different conditions. Section 4 concludes.

2 Methods

An anomaly in asset pricing research refers to a deviation from the expected return that is predicted by standard financial models, such as the Capital Asset Pricing Model (CAPM). It is a pattern in the financial market data that cannot be explained by traditional risk and return models and contradicts the efficient market hypothesis. These anomalies are seen as investment opportunities for generating abnormal returns. The present meta-analysis reproduces and evaluates 299 published anomalies in finance and accounting literature using Python. This work is largely based on a previous meta-analysis study on 319 anomalies using Stata, SAS, and R by Chen & Zimmermann (2021).

The data are mostly acquired from public databases including CRSP, Compustat, 13F, IBES, etc. Although a complete replication covering all 319 anomalies in Chen & Zimmermann's (2021) study is intended, 20 anomalies could not be reconstructed due to difficulties with code compilation and data download, which leads to minor deviations from Chen & Zimmermann's (2021) results. Specifically, due to time constraints, a Stata package *ffind* by Judson Caskey, that creates Fama and French industry classifications based on given SIC codes, is not compiled. In addition, the only piece of data that is not downloaded is the bid-ask spreads from Institute for the Study of Security Markets (ISSM) and Trades and Quotes (TAQ) database, resulting in several missing anomalies. Consequently, in the present study, 299 anomalies are recreated with Python, out of the 319 generated by Chen & Zimmermann (2021).

2.1 Predictability Categorization

The anomalies selected from 145 studies provide a comprehensive coverage of anomalies at the firm level in the literature. Based on the statistical significance reported in the original publications, anomalies are categorized into two predictability groups: (1) “predictors” and (2) “placebos”. The predictor category contains anomalies with clear predictive power in the original publications (e.g. long-short t-statistic ≥ 1.96). The placebo category includes anomalies with t-statistics below 1.96, with implied predictability, and without empirical evidence on predictability. For consistency, the predictability categorization adopted the same criteria as Chen & Zimmermann (2021). The initial four categories imposed by Chen & Zimmermann (2021) are now merged into two for simplicity: “clear predictors” and “likely predictors” are classified as “predictors”, whereas “not-predictors” and “indirect signal” are classified as “placebos” (Table 1).

The predictor category contains anomalies with t-statistics above the threshold of either 1.96 in long-short portfolios or 2 in regressions in the original studies. Exemplary anomalies that fit the former criterion include change in capex (Anderson & Garcia-Feijoo, 2006), return skewness (Bali et al., 2014), and net debt financing (Bradshaw et al., 2006). Anomalies such as illiquidity (Amihud, 2002), share volume (Datar et al., 1998), and change in financial liability (Richardson et al., 2005) fit the latter.

In contrast, anomalies with ambiguous predictability evidence are classified as placebos, which have t-statistics below 1.96 in the original studies, including long-short portfolios and regressions (e.g., accounting component of price (Callen et al., 2013), change in short-term investment (Richardson et al., 2005)). However, due to the varying methodologies and report metrics used in 145 original studies, the categorization could not follow an arbitrary t-statistic rule. Therefore, anomalies with implied predictability are also categorized as placebos in the present study. These anomalies are mostly derived from other established anomalies, which contain predictability-related information, but their link to return prediction is weak and usually not statistically significant. For example, Soliman’s (2008) change in noncurrent operating assets and change in noncurrent operation liability are used as ingredients to construct predictors, while themselves have little predictability. In addition, the placebo category includes other anomalies that were not empirically examined for return predictability in the original publications, such as earnings predictability, earnings smoothness, and earnings timeliness (Francis et al., 2005).

2.2 Generating Anomalies

The detailed methods used by Chen & Zimmermann (2021) have been adopted to the greatest extent possible to achieve a comprehensive reproduction. This leads to slight differences in t-statistics from some original studies because of the standardization enforced by Chen & Zimmermann (2021) in certain procedures. For example, three- and four-month lags are employed to indicate accounting data availability in past studies by Xie (2001) and Piotroski (2000) respectively. However, in the present study, the lag of annual accounting data is set at six months. In these cases, the increased lag diminishes the time effectiveness in the input information, which might result in lower and less time-relevant returns. Moreover, t-statistics are computed with raw long-short portfolio returns instead of the factors-adjusted returns in the present study, consistent with Chen & Zimmermann (2021). This might lead to lower average returns and t-statistics, compared to original studies. For example, Rosenberg et al. (1985) construct a set of indexes to remove the exposure to a variety of risk factors, including size and leverage, and achieve a high level of statistical significance (t-statistic = 5.7) for the book-to-market anomaly.

Raw data are automatically downloaded from Wharton Research Data Services (WRDS) and Federal Reserve Economic Data (FRED) in March 2021, preprocessed, and then stored for subsequent analyses. As an initial part of the preprocessing pipeline, a standard data lag is imposed across signals to indicate the availability of accounting data. Annual accounting data are lagged for six months, and quarterly accounting data are lagged for one quarter for most of the signals, as per Chen & Zimmermann's (2021) standardization approach. In some circumstances, the earning reporting date is used as data availability indicator for quarterly data instead, in order to reflect the methods used in the original studies more precisely. The remaining preprocessing procedures are largely consistent with the original studies. All anomalies are computed and generated on a monthly basis.

A couple of reproduced anomalies slightly deviate from both the original publications and Chen & Zimmermann (2021), due to different function implementations in time-series analysis between Stata and Python. Unlike the time gaps in daily signals due to weekends and holidays, the few ones in the monthly signals are likely caused by missing data. In Stata, once a time series has been defined by *tsset* command, observations are indexed by their associated date. In analyses that involve moving windows, such as rolling regression and moving average, the Stata window takes missing time points into account if the time series is not continuous. However, in Python, the time gaps are ignored when computing rolling statistics or rolling regressions, which results in using data that exceed the defined timeframe. As a consequence, the signals generated have less time-relevant information and potentially less predictive power. Anomalies that affected by this issue include liquidity beta (Pástor & Stambaugh, 2003), citations to R&D expenses (Hirshleifer et al., 2013), and dividend omission (Michaely et al., 1995).

Moreover, the approach to reproduce predicted analyst forecast error (Frankel & Lee, 1998) in the present study is slightly different from Chen & Zimmermann (2021), but it is consistent with the original study. Specifically, Chen & Zimmermann (2021) use the *relrank* function in Stata to generate the relative rankings of the variable's values in the distribution of themselves. However, Frankel & Lee's (1998) approach to express the variables of interest in terms of their percentile ranks is better implemented in Python.

2.3 Portfolio Construction and Predictability Assessments

The long-short portfolios are built following the signal construction. As suggested in the original studies, many signals only have predictive power in certain subsets of the data. Thus, the security's share type, price, exchange, among other factors, are filtered to construct portfolios.

Other parameters of portfolios collected from the original studies, such as quantile sort and stock-weighting, are also strictly followed when generating portfolios for each anomaly. If these parameters are not specified in the original papers, a default equal weighted quintile sort is applied. In accordance with Chen & Zimmermann's (2021) methodology, the reproduced long-short portfolios are rebalanced on a monthly basis to maintain the choice of weighting scheme (equal- or value-weighting) in the original studies. This monthly rebalancing deviates somewhat from a couple of the original publications where different rebalancing frequencies were imposed, such as weekly (Hou, 2007; Johnson & So, 2012).

The long-short portfolios in the present study are constructed from signal sorts without adjusting for factor exposures, consistent with Chen & Zimmermann (2021). For anomalies that predict negative returns, the signs of signals are flipped when constructing portfolios to ensure a consistently positive sign in mean returns. In these anomalies, such as idiosyncratic volatility (Ang et al., 2006), credit rating downgrade (Dichev & Piotroski, 2001), and size (Banz, 1981), a larger magnitude implies a lower mean return.

Consistent with Chen & Zimmermann (2021), the null hypothesis is that the average monthly return of an anomaly's long-short portfolio is zero. The predictability performance of an anomaly is first assessed with in-sample test, where the monthly portfolio returns and corresponding t-statistics are computed with the sample periods that used to identify anomaly in the original study. Specifically, the t-statistic of an anomaly is calculated by dividing the difference between its mean portfolio return and zero with its standard error. The standard error is obtained by dividing the standard deviation with the square root of the sample size. In subsequent analysis, the sample periods are replaced with pre-publication samples, post-publication samples and full samples in order to assess the post-publication effect and overall performance of each anomaly.

As an extension from previous work, the novel null hypothesis tested in the present study is that the average monthly return of an anomaly's long-short portfolio is zero if microcaps are excluded from investment universe. Microcaps are typically associated with high trading costs; thus, if the abnormal returns of an anomaly are primarily driven by microcaps, the strategy is difficult to exploit in practice. By removing microcaps, the role of frictions in anomalies' performance is examined. The return predictability of anomalies is further assessed in a realistic trading scenario. In this case, the long-short portfolio for each anomaly is constructed using stocks with market equity above the NYSE 20th percentile breakpoints across the sample periods. The monthly portfolio returns and t-statistics are computed with pre-publication samples, post-publication samples and full samples.

3 Empirical Results

This section shows the main findings. The regression of the replicated t-statistics in the present study on Chen & Zimmermann's (2021) t-statistics produces a slope of 1 and an R squared of 99%. A similar regression on the original long-short t-statistics yields a slope of 0.87 and an R squared of 80%. Post-publication t-statistics are diminished by 68.84% compared to pre-publication. With microcaps removed from the investment universe, the average in-sample t-statistic drops by 27.92% from 3.82 to 2.75. After raising the t-statistic threshold from 1.96 to 3.0 (Harvey et al., 2016), 57.81% of the cross-sectional stock return predictors remain significant in-sample, while only 9.38% of the predictors survive the novel threshold post-publication.

3.1 Replication Performance

The reproduction success is assessed using t-statistics. If the average return of long-short portfolio is significant at 5% level (with a t-statistic above 1.96), the corresponding anomaly is considered as a reproduction success. Using original in-sample periods, the reproduced t-statistics and the average monthly portfolio returns for each anomaly are presented in Table 2 and 3. Chen & Zimmermann's (2021) results are listed alongside for comparison. The two tables display predictors and placebos, respectively, which provide a summary to the dataset. Anomalies are sorted by author names. The acronyms used in the present study are also provided.

Figure 1 illustrates the overall reproduction performance for all anomalies. As shown, 87.5% of the predictors have t-statistics over 1.96. The reproduced t-statistics for predictors are averaged around 3.82, and many t-statistics are much larger. However, more than half of the placebos fail to achieve statistical significance in the present study. Specifically, 66.67% of the placebos have t-statistics below the 1.96 cutoff. Both the t-statistics and mean returns reproduced in the present study echo Chen & Zimmermann's (2021) results.

Figure 2 provides a more qualitative assessment of the reproduction success among predictors by regressing the reproduced t-statistics in the present study against those in Chen & Zimmermann's (2021) work. Each marker represents a long-short portfolio t-statistic for one anomaly using the original in-sample periods. The regression yields a coefficient of 1 and an R squared of 99%, indicating closely matching results. Thus, the replication is successful. The remaining deviations may stem from inconsistencies in algorithm implementations and difficulties in code compilation.

Figure 3 compares the reproduced in-sample t-statistics of predictors to the t-statistics from the original publications. To ensure comparability, only anomalies with predictability evidence based on long-short portfolios in the original publications are included. The excluded anomalies were empirically examined either using regressions or in event studies. The predictor markers centre around the 45-degree dotted line. The regression of reproduced t-statistics on original long-short t-statistics produces a slope of 0.87 and an R squared of 80%, indicating that the reproductions do not deviate far from the original studies. Because many t-statistics in the original studies are computed with factor-adjusted returns, the observed variances might be attributed to the fact that the reproduced t-statistics use the gross long-short portfolio returns without factor adjustments. The variances may also be explained by the discrepancies between the methods in certain original studies and the standardization procedures enforced in the present study.

3.2 Out-of-Sample Tests

Certain original literature uses data from as early as 1920s, while over 96% of the literature does not capture the data from the past decade. These in-sample data are stale and fail to account for the recent development in trading technology. Lack of the latest data raises concerns about whether these strategies are still valid. Thus, the original in-sample periods are replaced with full samples, which include all data available for download, to examine the overall performance of 192 predictors. The average length of original in-sample and full-sample is 328 months and 782 months, respectively.

Notably, predictors' portfolio returns with full-sample experienced a noticeable drop compared to that with in-sample (Figure 5). The average full-sample return (0.49% per month) is 26.46% lower. 14 previously significant in-sample predictors become insignificant at the 5%

level in the full sample. Out of 192 predictors, the number of significant predictors declines by 6% compared to that with in-sample.

To further investigate this decline in predictors' performance, the full sample is split into a pre-publication and a post-publication sample based on the publication year of each original study. The average span for pre- and post-publication is 573 months and 209 months, respectively. The individual performance of 192 predictors under these three sample periods are reported in Table 4. Figure 5 illustrates the mean returns and t-statistics of predictors under different conditions of sample periods and microcaps presence. Each dot represents a long-short portfolio of an anomaly. The middle line inside the box indicates the median. The upper and lower limits of the box indicate the first and third quartiles, respectively.

As shown, the pre-publication sample demonstrates 14.9% lower average return compared to in-sample, although the t-statistics between the two are nearly identical. In contrast, post-publication sample exhibits a far greater drop in predictors' performance (mean return: -52.97%, t-statistic: -67.92%). Therefore, the performance drop observed in full-sample could be largely attributed to post-publication effects. Because the pre-publication sample contains data with wider time range, it is chosen as the baseline to further assess the post-publication effects.

As shown in the top panel of Figure 5, post-publication returns are smaller than pre-publication returns on average. Compared to average pre-publication return (0.56% per month), the average post-publication return (0.31% per month) decays 44.76%. The average t-statistics exhibit a similar trend, as shown in the bottom panel. Out of 192 predictors, 153 (80.21%) predictors are significant at the 5% level in the pre-publication samples. When the post-publication samples are employed, 105 significant pre-publication predictors become insignificant. As a result, only 48 predictors (25%) remain significant post publication. The average post-publication t-statistic is 1.23, which is 68.84% lower than the average pre-publication t-statistic of 3.93.

The large post-publication decay suggests that the return predictability of anomalies in literature is likely attributed to mispricing, consistent with McLean & Pontiff (2016), Yan & Zheng (2017) and Jacobs & Müller (2020). If market participants become aware of the mispricing through research publications, they will adapt promptly to trade against it, leading to diminishing or vanishing returns after the anomaly publication (Timmermann & Granger, 2004). The poor out-of-sample performance could also be explained by data mining, in which the in-sample performance tend to stand out compared to other sample periods. To the contrary, the return predictability is likely to persist if it stems from a true risk factor, regardless of the anomaly publication and the choice of sample periods (Cochrane, 1999).

3.3 Controlling for Microcaps

Microcaps are typically associated with high trading costs; thus, anomalies inflated by them are difficult to exploit in practice. The NYSE 20th percentile breakpoints are used to construct new long-short portfolios of each anomaly in order to control for microcaps and further assess the return predictability of anomalies in a more realistic trading scenario.

With original in-sample period, when microcaps are excluded from the investment universe, there is an evident decrease in both the average returns and t-statistics (Figure 4). The top panel compares the average in-sample returns including microcaps versus excluding them within each predictability category. The bottom panel shows the effects of microcaps in average t-stats. The error bars indicate the standard error.

With microcaps removed, long-short portfolio returns of approximately 77% anomalies are found to decrease. The average in-sample returns are 0.49% per month for predictors and 0.15% per month for placebos, which are 25.14% and 34.8% lower than the average returns before microcaps adjustment, respectively. Similarly, after microcaps adjustment, the in-sample t-statistics of 76% of the anomalies are also found to decrease. For predictors, the average t-statistic drops by 27.92% from 3.82 to 2.75 in-sample. For placebos, the average t-statistic shows a 26.89% decline from 1.19 to 0.87. Out of 299 anomalies, excluding microcaps decreases the number of significant anomalies from 203 (68%) to 152 (51%). These results are consistent with Chen and Velikov's (2021) finding that controlling for trading costs lowers the in-sample performance by about 1/3.

A similar trend has been found outside the original studies' sample periods. As shown in Table 5, controlling for microcaps leads to a drastic and consistent decay in the predictors' performance throughout all sample periods of interest. Using full samples, the average monthly return and t-statistic lower by 36.13% and 28.58%, respectively, after controlling for the microcaps. Using pre-publication samples, the average monthly return and t-statistic shrink 33.29% and 27.28%, respectively, after controlling for the microcaps. Among all sample periods of interest, post-publication suffers the largest decay (average monthly return: -37.22%, t-statistic: -35.83%). Therefore, it is crucial to determine whether the anomaly returns are inflated by microcaps to accurately assess the anomaly performance.

Furthermore, the post-publication decays become greater with microcaps adjustment, compared to without. After controlling for microcaps, the average full-sample return (0.31% per month) and the average full-sample t-statistic (2.81) decrease 17.32% and 1.76% from pre-publication, respectively. When limiting the sample periods to post-publication, the decays are even larger. The average post-publication return (0.19% per month) and t-statistic (0.79) decline relative to the pre-publication by 48.02% and 72.51%, respectively. With microcaps removed, only 25 out of 192 (13.02%) predictors remain significant at the 5% level, using the conventional t-statistic threshold of 1.96.

To illustrate the evolving significance of individual anomalies over time, the moving t-statistic of portfolio returns for a representative anomaly, earnings surprise (Foster et al., 1984), is presented (Figure 6). Specifically, the past 36-month observations are used to compute the t-statistic at each time point. The post-publication effect and microcaps adjustment are found to largely diminish the t-statistic of the anomaly, consistent with the trend of average t-statistics shown in the bottom panel of Figure 5.

3.4 Higher t-Statistic Threshold

With significant developments in research innovations and trading technologies over the past decades, there is a growing concern regarding data mining in asset pricing research. Therefore, Harvey et al. (2016) argue that the conventional t-statistic threshold of 1.96 for statistical significance is obsolete and propose a higher t-statistic threshold of 3.0 to account for the extensive data mining in the field.

Panel A of Table 6 indicates the percentages of predictors that survive the conventional threshold of 1.96 and the Harvey et al.'s (2016) threshold of 3.0 across sample periods. A consistent and drastic decrease in survival rate is found across sample periods and stock sets. As shown, when long-short portfolios are constructed with all stocks, 87.5% of the predictors have in-sample t-statistics over 1.96. Once the threshold is raised to 3.0, only 57.81% of the predictors remain significant in-sample. Post-publication suffers the largest decrease with only 9.38% of the predictors survive the threshold of 3.0. A similar trend is found in portfolios using all-but-microcap stocks. The in-sample survival rate drops from 63.02% to 38.54%. Less than 3% of the predictors survive the threshold of 3.0 post-publication. This suggests that only a small portion of the established anomalies pass the novel threshold and remain robust over decades of research innovations. Thus, current methodologies for identifying anomalies should be critically examined.

Figure 7 illustrates the post-publication performance of predictors that are significant at both the conventional t-statistic threshold of 1.96 and the novel threshold of 3.0. Specifically, it shows how many predictors with t-statistics above 3.0 in-sample remain robust at the same statistical significance level post-publication. Predictors are assigned into two groups based on the reproduced in-sample t-statistics: (1) above 1.96 and below 3.0, (2) above 3.0. Within each group, predictor performance is assessed in long-short portfolios containing all stocks and without microcaps, using in-sample and post-publication samples. The descriptive statistics for predictor performance are displayed in Panel B of Table 6.

A noticeable decline in t-statistics is found for predictors post-publication, regardless of the presence of microcaps. When portfolios are constructed using all stocks, the predictors with in-sample t-statistics below 3.0 yield average t-statistics of 2.52 and 1.04 for in-sample and post-publication, respectively. The ones with in-sample t-statistics above 3.0 produce average t-statistics of 5.07 and 1.58 for in-sample and post-publication, respectively. The performance drop is larger for the predictors with in-sample t-statistics above 3.0 (-68.78%), compared to the less significant ones (-58.57%). This result echos previous findings that the post-publication decay is larger for predictors with higher in-sample t-statistics (McLean & Pontiff, 2016; Chen & Zimmermann, 2021). Despite imposing the higher t-statistic threshold in-sample, the post-publication performance of predictors deteriorate drastically, with only a few predictors remain significant. Similar results are reported for the portfolios constructed without microcaps, in which even fewer predictors survive (Figure 7, bottom panel).

Four independent two sample t-tests are performed to assess the true difference of post-publication performance between two groups of predictors (below and above the t-statistic threshold of 3.0). Specifically, the null hypothesis is that the true difference in the post-publication means is equal to zero for (1) portfolio returns and (2) t-statistics with all stocks, as well as (3) portfolio returns and (4) t-statistics with no microcaps. The post-publication means of t-statistics for portfolios consisting of all stocks are found to be significantly different between the two groups ($p < 0.05$). The remaining t-tests produce non-significant results, suggesting zero difference in the post-publication means between the two groups. This is consistent with the findings in Figure 7 that imposing a higher t-statistic threshold in-sample makes little difference in post-publication performance.

The low survival rate post-publication corroborated Harvey et al.'s (2016) argument that the t-statistic threshold of 3.0 is still too low, even for anomalies published in renowned journals. The asset pricing literature, including the present study, has been heavily based on the same data from the CRSP database, resulting in overfitting. Generally, models that suffer from overfitting tend to perform badly out-of-sample. This explains the poor out-of-sample performance of reproduced anomalies in the present study. Furthermore, truly robust anomalies with strong theoretical support have mostly been uncovered in the early days of asset pricing research. As the technology develops over the past decades, data mining in the field has become more accessible. Thus, the anomalies derived from such approach could be subjected to both type I and type II errors, contributing to the falsely high discovery rate of anomalies.

4 Conclusion

Hundreds of anomalies have been identified continuously by researchers since the last century. The growing numbers of anomaly discoveries have exacerbated the concerns in their validity. Extensive efforts have been put into meta analyses to re-evaluate the anomalies. The present study successfully replicated 299 anomalies (192 predictors, 107 placebos) from 145 publications at a high consistency with Chen & Zimmermann (2021) and the original studies. For replication, 87.5% of the reproduced predictors are significant in-sample under the conventional t-statistic threshold of 1.96. Through out-of-sample tests, post-publication performance decay is evident, with 25% of the predictors remain significant. It is also confirmed that the absence of microcaps pulls down the performance of anomalies across all sample periods. Specifically, only 13.02% of the predictors survive post-publication after accounting for microcaps.

Furthermore, after adopting the higher t-statistic threshold of 3.0 proposed by Harvey et al. (2016), 57.81% of the reproduced predictors are significant in-sample, while 9.38% are significant post-publication. Excluding microcaps leads to only 2.6% of the predictors surpassing the novel threshold.

In summary, after imposing the t-statistic threshold of 3.0, 18 predictors stay significant post-publication for portfolios containing all stocks, while only 5 remain robust for portfolios with microcaps removed. Specifically, the 5 predictors are: EPS forecast revision (Hawkins et al., 1984), earnings announcement return (Chan et al., 1996), predicted dividend yield next month (Litzenberger & Ramaswamy, 1979), down forecast EPS (Barber et al., 2001), and put volatility minus call volatility (Yan, 2011). The persistent return predictability of these anomalies implies a strong connection with true risk factors. Thus, empirical finance research on anomalies should be built on the economic foundation of first principles.

Bibliography

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, 5(1), 31–56.

Anderson, C. W. & Garcia-Feijoo, L. (2006). Empirical evidence on capital investment, growth options, and security returns. *The Journal of Finance*, 61(1), 171–194.

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1), 259–299.

Baker, M. (2016). Reproducibility crisis. *Nature*, 533(26), 353–66.

Bali, T., Engle, R., & Murray, S. (2014). Empirical asset pricing: The cross section of stock returns: An overview. *Wiley StatsRef: Statistics Reference Online*, (pp. 1–8).

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1), 3–18.

Barber, B., Lehavy, R., McNichols, M., & Trueman, B. (2001). Can investors profit from the prophets? security analyst recommendations and stock returns. *The Journal of Finance*, 56(2), 531–563.

Begley, C. G. & Ellis, L. M. (2012). Raise standards for preclinical cancer research. *Nature*, 483(7391), 531–533.

Bradshaw, M. T., Richardson, S. A., & Sloan, R. G. (2006). The relation between corporate financing activities, analysts' forecasts and stock returns. *Journal of Accounting and Economics*, 42(1-2), 53–85.

Callen, J. L., Khan, M., & Lu, H. (2013). Accounting quality, stock price delay, and future stock returns. *Contemporary Accounting Research*, 30(1), 269–295.

Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., et al. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, 351(6280), 1433–1436.

Chan, L. K., Jegadeesh, N., & Lakonishok, J. (1996). Momentum strategies. *The Journal of Finance*, 51(5), 1681–1713.

Chang, A. C. & Li, P. (2015). Is economics research replicable? sixty published papers from thirteen journals say'usually not'.

Chen, A. Y. & Zimmermann, T. (2021). Open source cross-sectional asset pricing. *Critical Finance Review, Forthcoming*.

Chordia, T., Goyal, A., & Saretto, A. (2020). Anomalies and false rejections. *The Review of Financial Studies*, 33(5), 2134–2179.

Chordia, T., Subrahmanyam, A., & Tong, Q. (2014). Have capital market anomalies attenuated in the recent era of high liquidity and trading activity? *Journal of Accounting and Economics*, 58(1), 41–58.

Cochrane, J. H. (1999). Portfolio advice for a multifactor world.

Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of Finance*, 66(4), 1047–1108.

Datar, V. T., Naik, N. Y., & Radcliffe, R. (1998). Liquidity and stock returns: An alternative test. *Journal of Financial Markets*, 1(2), 203–219.

Dichev, I. D. & Piotroski, J. D. (2001). The long-run stock returns following bond ratings changes. *The Journal of Finance*, 56(1), 173–203.

Fama, E. F. & French, K. R. (2008). Dissecting anomalies. *The Journal of Finance*, 63(4), 1653–1678.

Fama, E. F. & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3), 607–636.

Fanelli, D. (2010). "positive" results increase down the hierarchy of the sciences. *PloS One*, 5(4), e10068.

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

Francis, J., LaFond, R., Olsson, P., & Schipper, K. (2005). The market pricing of accruals quality. *Journal of Accounting and Economics*, 39(2), 295–327.

Frankel, R. & Lee, C. M. (1998). Accounting valuation, market expectation, and cross-sectional stock returns. *Journal of Accounting and Economics*, 25(3), 283–319.

Harvey, C. R. (2017). Presidential address: The scientific outlook in financial economics. *The Journal of Finance*, 72(4), 1399–1440.

Harvey, C. R., Liu, Y., & Zhu, H. (2016). . . . and the cross-section of expected returns. *The Review of Financial Studies*, 29(1), 5–68.

Hawkins, E. H., Chamberlin, S. C., & Daniel, W. E. (1984). Earnings expectations and security prices. *Financial Analysts Journal*, 40(5), 24–38.

Hirshleifer, D., Hsu, P.-H., & Li, D. (2013). Innovative efficiency and stock returns. *Journal of Financial Economics*, 107(3), 632–654.

Hou, K. (2007). Industry information diffusion and the lead-lag effect in stock returns. *The Review of Financial Studies*, 20(4), 1113–1138.

Hou, K., Xue, C., & Zhang, L. (2020). Replicating anomalies. *The Review of Financial Studies*, 33(5), 2019–2133.

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS medicine*, 2(8), e124.

Jacobs, H. & Müller, S. (2020). Anomalies across the globe: Once public, no longer existent? *Journal of Financial Economics*, 135(1), 213–230.

Jegadeesh, N. & Titman, S. (2001). Profitability of momentum strategies: An evaluation of alternative explanations. *The Journal of Finance*, 56(2), 699–720.

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, 23(5), 524–532.

Johnson, T. L. & So, E. C. (2012). The option to stock volume ratio and future returns. *Journal of Financial Economics*, 106(2), 262–286.

Leamer, E. E. (1983). Let's take the con out of econometrics. *American Economic Review*, 73(1), 31–43.

Linnainmaa, J. T. & Roberts, M. R. (2018). The history of the cross-section of stock returns. *The Review of Financial Studies*, 31(7), 2606–2649.

Litzenberger, R. H. & Ramaswamy, K. (1979). The effect of personal taxes and dividends on capital asset prices: Theory and empirical evidence. *Journal of Financial Economics*, 7(2), 163–195.

McCullough, B. D. & Vinod, H. D. (2003). Verifying the solution from a nonlinear solver: A case study. *American Economic Review*, 93(3), 873–892.

McLean, R. D. & Pontiff, J. (2016). Does academic research destroy stock return predictability? *The Journal of Finance*, 71(1), 5–32.

Michaely, R., Thaler, R. H., & Womack, K. L. (1995). Price reactions to dividend initiations and omissions: Overreaction or drift? *the Journal of Finance*, 50(2), 573–608.

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia: Ii. restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science*, 7(6), 615–631.

Novy-Marx, R. & Velikov, M. (2016). A taxonomy of anomalies and their trading costs. *The Review of Financial Studies*, 29(1), 104–147.

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716.

Pástor, L. & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, 111(3), 642–685.

Piotroski, J. D. (2000). Value investing: The use of historical financial statement information to separate winners from losers. *Journal of Accounting Research*, (pp. 1–41).

Prinz, F., Schlange, T., & Asadullah, K. (2011). Believe it or not: how much can we rely on published data on potential drug targets? *Nature Reviews Drug Discovery*, 10(9), 712–712.

Richardson, S. A., Sloan, R. G., Soliman, M. T., & Tuna, I. (2005). Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics*, 39(3), 437–485.

Rosenberg, B., Reid, K., & Lanstein, R. (1985). Persuasive evidence of market inefficiency. *The Journal of Portfolio Management*, 11(3), 9–16.

Schwert, G. W. (2007). Chapter 15 anomalies and market efficiency. *Financial Markets and Asset Pricing*, (pp. 939–974).

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2016). False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant.

Soliman, M. T. (2008). The use of dupont analysis by market participants. *The Accounting Review*, 83(3), 823–853.

Timmermann, A. & Granger, C. W. (2004). Efficient market hypothesis and forecasting. *International Journal of Forecasting*, 20(1), 15–27.

Xie, H. (2001). The mispricing of abnormal accruals. *The Accounting Review*, 76(3), 357–373.

Yan, S. (2011). Jump risk, stock returns, and slope of implied volatility smile. *Journal of Financial Economics*, 99(1), 216–233.

Yan, X. S. & Zheng, L. (2017). Fundamental analysis and the cross-section of stock returns: A data-mining approach. *The Review of Financial Studies*, 30(4), 1382–1423.

Figure 1: **Reproduction Performance for All Anomalies**

This figure shows the reproduced in-sample t-statistics for all anomalies in categories. The categorization is based on the t-statistics reported in the original publications: the predictor category contains anomalies with clear predictive power (e.g. long-short t-statistic ≥ 1.96), whereas the placebo category includes anomalies (1) with t-statistics below 1.96, (2) with implied predictability, and (3) without empirical evidence on predictability. The null hypothesis is that the average portfolio return of an anomaly is zero. Each marker represents a long-short portfolio for an anomaly reconstructed based on the original study. The vertical dotted line indicates the t-statistics threshold of 1.96. This figure illustrates the reproduction performance of anomalies in each predictability category: the majority of the predictors surpasses the t-statistics threshold of 1.96 and remain significant, while more than half of the placebos are found to be not significant.

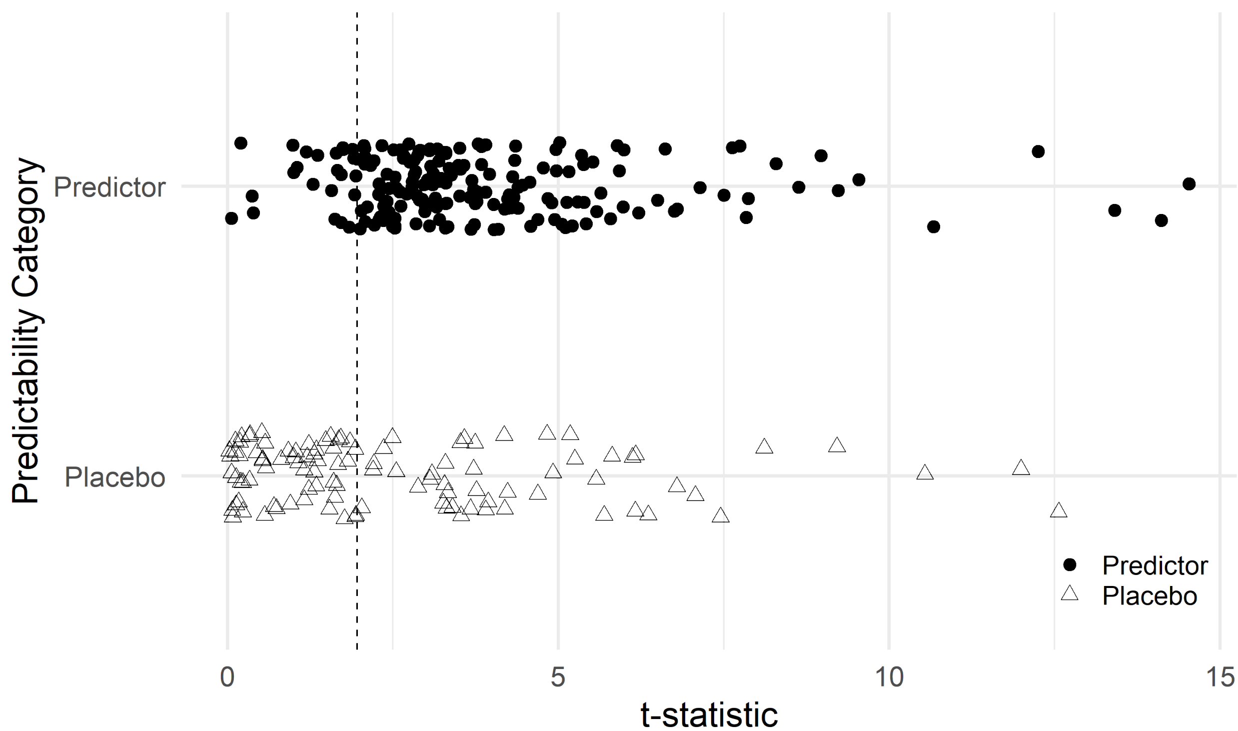


Figure 2: **Comparison of t-Statistics for Predictors: My Replication versus Chen & Zimmermann (2021)**

This figure compares the reproduced t-statistics in the present study to the ones in Chen & Zimmermann's (2021) study. Each marker represents a long-short portfolio t-statistic for one predictor using the original in-sample periods. Clear and likely predictors defined by Chen & Zimmermann (2021) are indicated with solid circle and hollow triangle markers in this figure, respectively. Clear predictors have long-short t-statistics over 2.5 in the original studies, whereas likely predictors have long-short t-statistics around 2 but above 1.96 in the original studies. The dotted line is a 45-degree line, indicating an ideal situation where my reproduced t-statistics are equal to the ones in Chen & Zimmermann (2021). The solid line is the ordinary least squares (OLS) regression fit, which indicates the regression of the reproduced t-statistics against those in Chen & Zimmermann (2021). Axes are in log-scale for visualization purposes. The regression yields a coefficient of 1 and an R squared of 99%, indicating closely matching results.

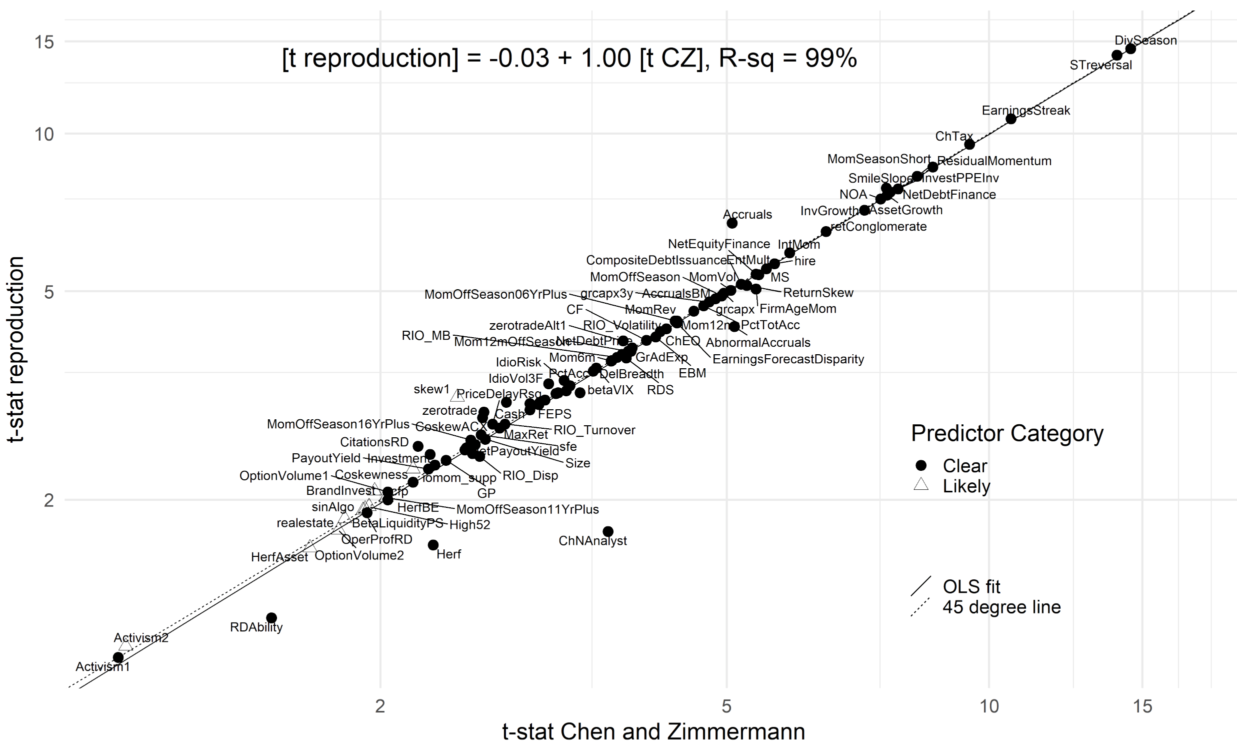


Figure 3: **Comparison of t-Statistics for Predictors: My Replication versus Original Studies**

This figure compares the reproduced t-statistics in the present study to the long-short t-statistics in the original studies. Each marker represents a long-short portfolio t-statistic for one predictor using the original in-sample periods. Clear and likely predictors defined by Chen & Zimmermann (2021) are indicated with solid circle and hollow triangle markers in this figure, respectively. Clear predictors have long-short t-statistics over 2.5 in the original studies, whereas likely predictors have long-short t-statistics around 2 but above 1.96 in the original studies. The dotted line is a 45-degree line, indicating an ideal situation where my reproduced t-statistics are equal to the ones in the original studies. The solid line is the ordinary least squares (OLS) regression fit, which indicates the regression of the reproduced t-statistics in the present study against those in the original study. Axes are in log-scale for visualization purposes. The regression yields a coefficient of 0.87 and an R squared of 80%, indicating small deviations from the original studies.

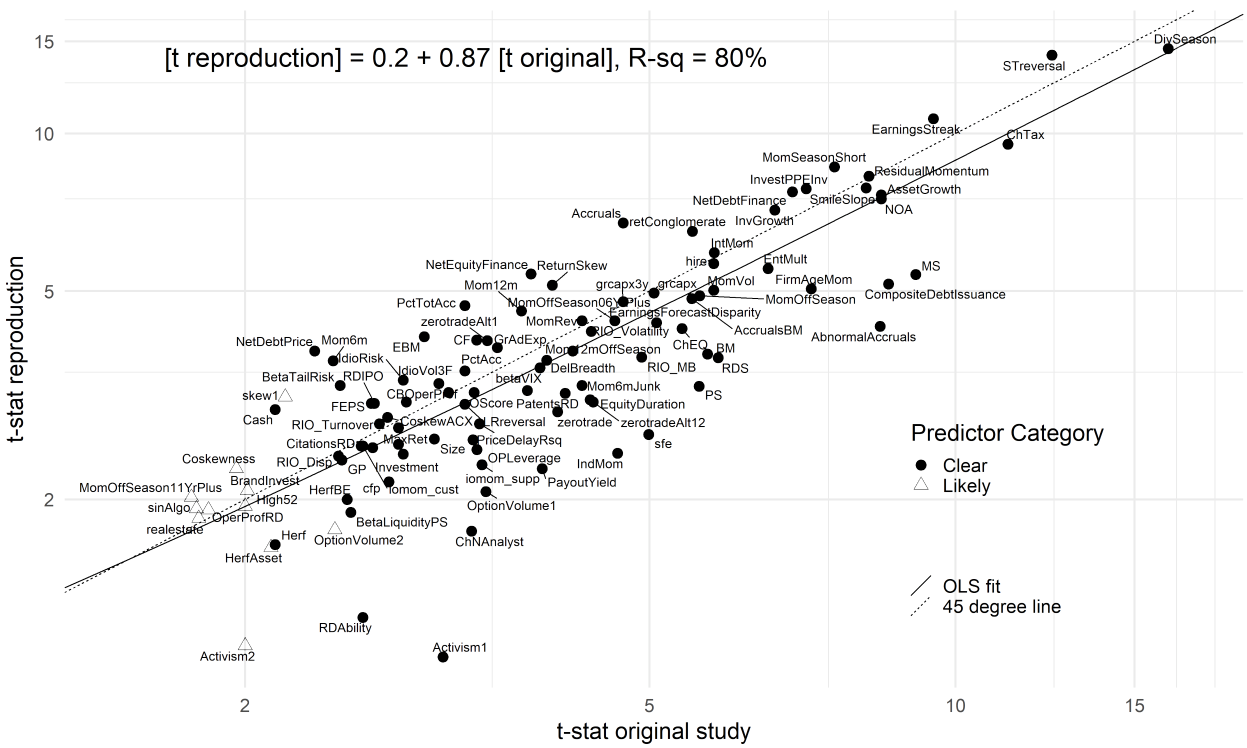


Figure 4: **Anomaly Performance with Microcaps Adjustment**

This figure illustrates the in-sample mean returns and t-statistics of all anomalies. The top panel compares the average in-sample returns including microcaps versus excluding them within each predictability category. The bottom panel shows the effects of microcaps in average t-statistics. It is found that the absence of microcaps pulls down the performance of anomalies. The error bars indicate the standard error.

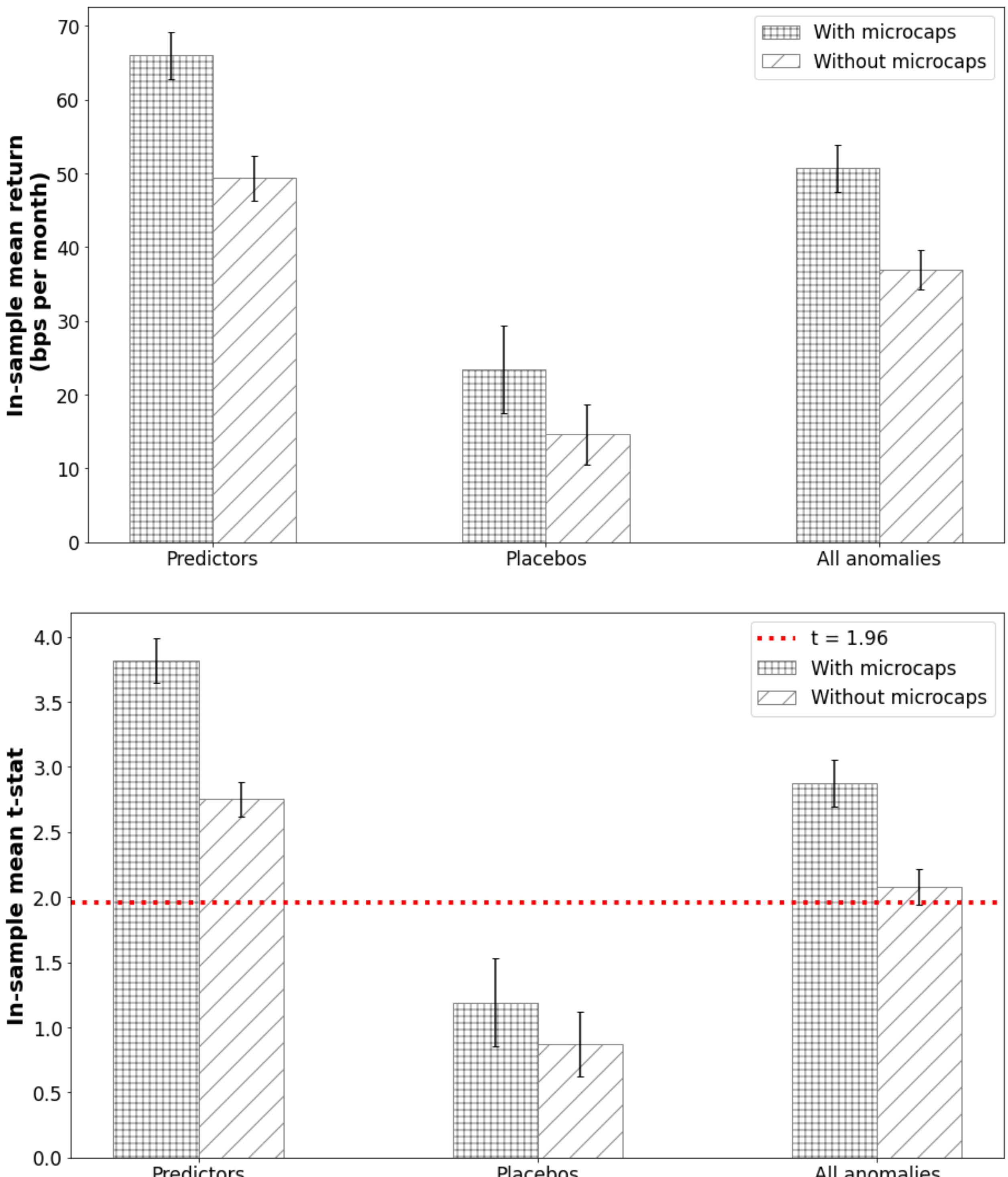
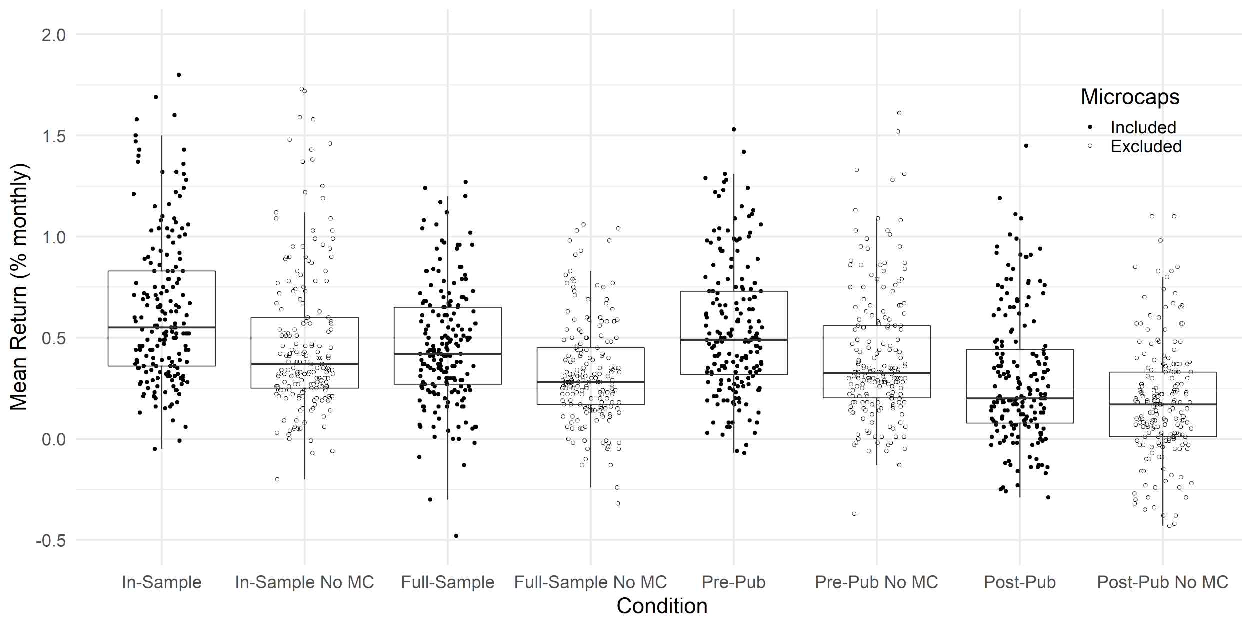


Figure 5: **Predictor Performance by Sample Periods**

This figure illustrates the mean returns and t-statistics of 192 predictors under different conditions of sample periods and microcaps presence. The top panel compares the average returns including microcaps versus excluding them across sample periods. The bottom panel shows similar effects of microcaps in average t-statistics. Each dot represents a long-short portfolio of a predictor. The middle line inside the box indicates the median. The upper and lower limits of the box indicate the first and third quartiles, respectively.



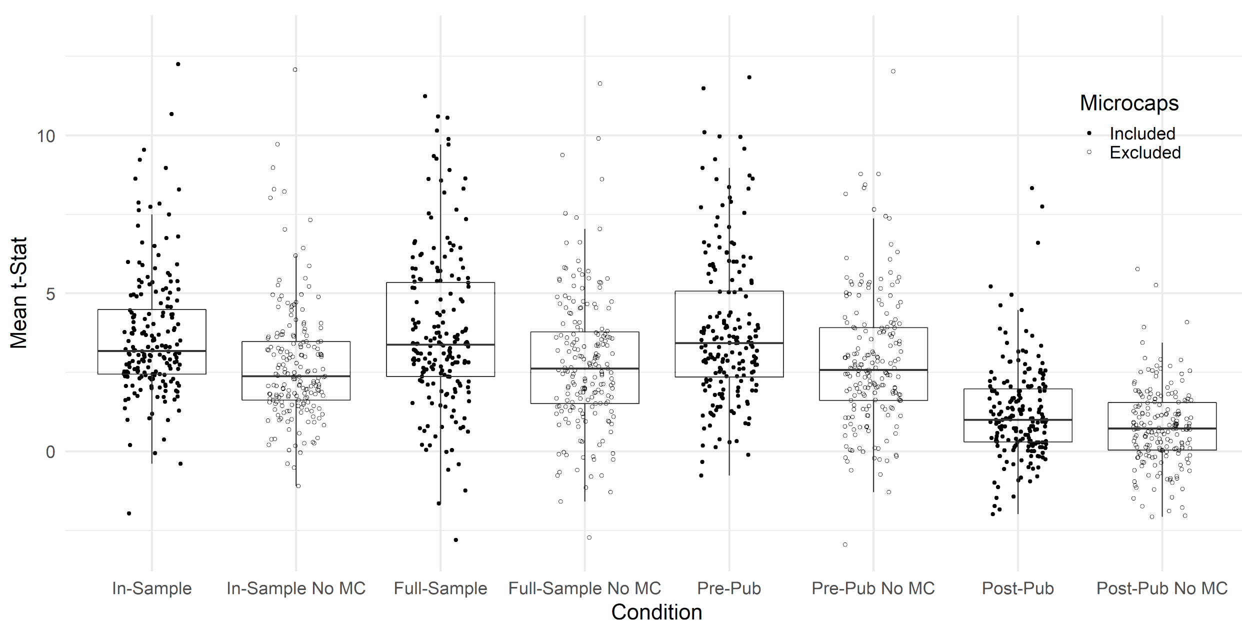


Figure 6: **Moving t-Statistic for a Representative Anomaly**

This figure shows the evolving significance of an individual anomaly, earnings surprise (Foster et al., 1984), over time. The past 36-month observations are used to compute the t-statistic at each time point. The solid black line represents the moving t-statistic of the portfolio returns with all stocks. The dotted blue line represents the moving t-statistic of the portfolio returns with microcaps removed. The vertical grey line indicates the original publication year of the anomaly.

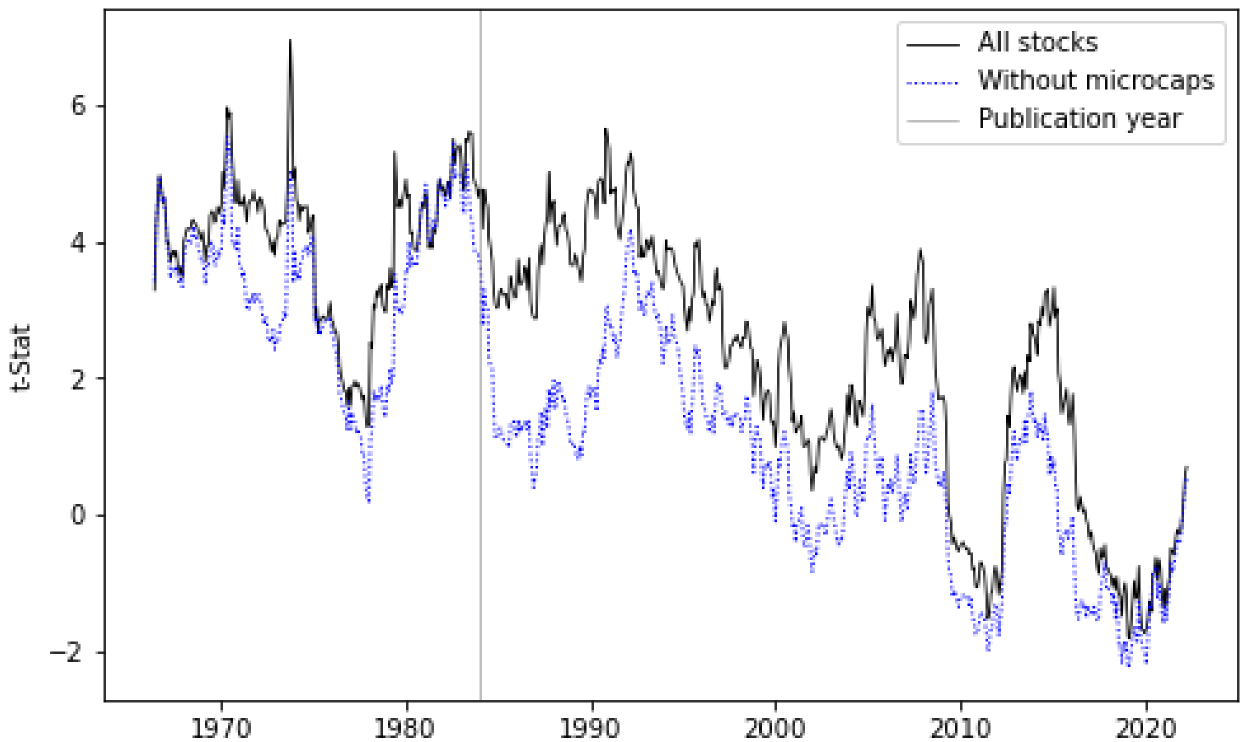


Figure 7: **Predictor Performance at the t-Statistic Threshold of 3.0**

This figure illustrates the t-statistics of 192 predictors under different conditions of sample periods and microcaps presence. Specifically, the figure shows how many predictors with t-statistics above 3.0 in-sample remain robust at the same statistical significance level post-publication. Predictors are assigned into two groups based on the reproduced in-sample t-statistics: (1) above 1.96 and below 3.0, (2) above 3.0. The top panel shows the t-statistics of portfolios constructed using all stocks, while the bottom panel shows the ones with microcaps removed. Both panels demonstrate that only a few predictors remain robust post-publication at the novel threshold of 3.0. Each dot represents a long-short portfolio of a predictor. The middle line inside the box indicates the median. The upper and lower limits of the box indicate the first and third quartiles, respectively.

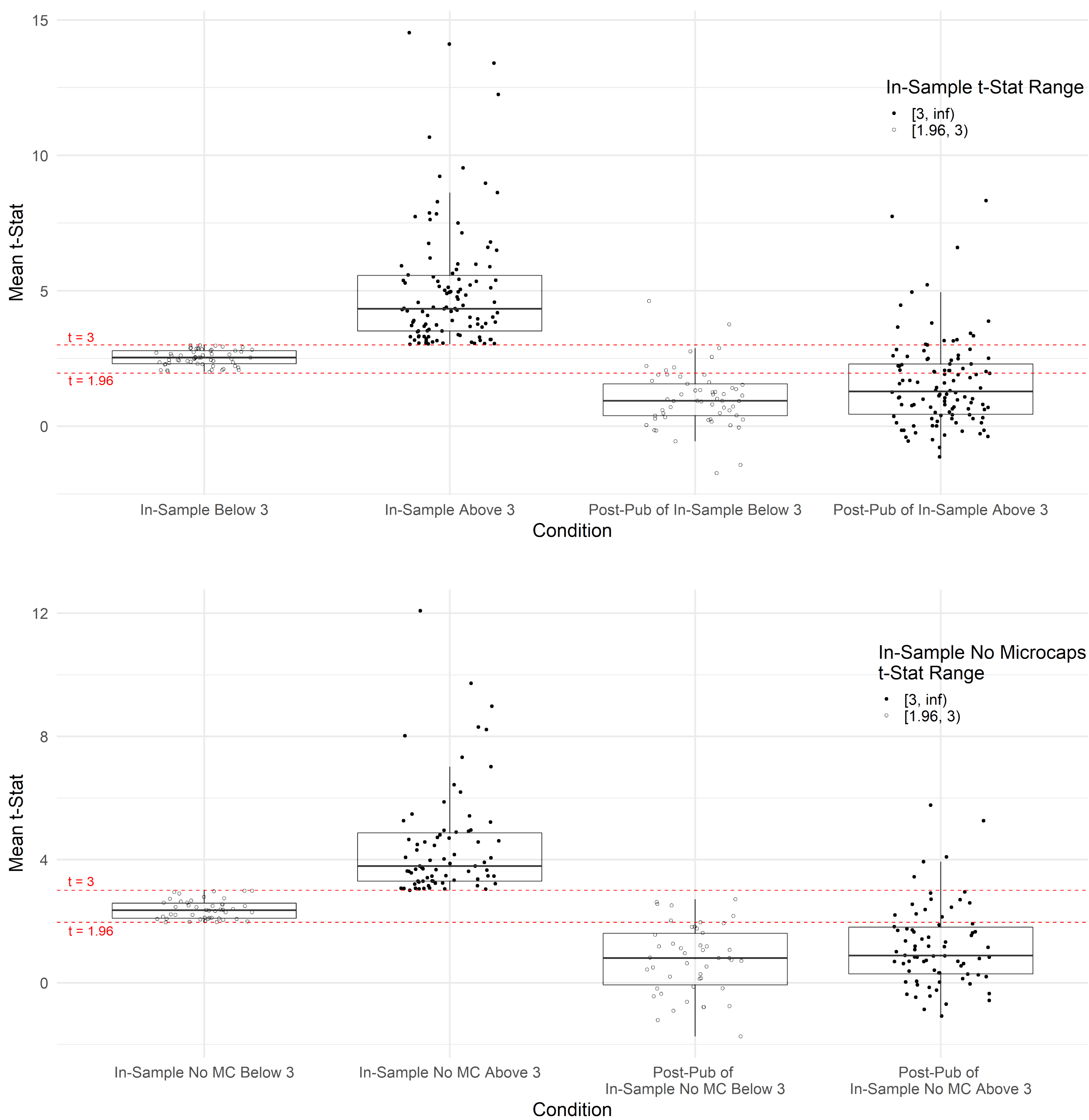


Table 1: **Category Assignment.**

This table shows the predictability categorization in the present study, which adopts the same criteria as Chen & Zimmermann's (2021) study. The categorization is based on the t-statistics reported in the original publications. The predictor category contains anomalies with clear predictive power (e.g. long-short t-statistic ≥ 1.96). The placebo category includes anomalies (1) with t-statistics below 1.96, (2) with implied predictability, and (3) without empirical evidence on predictability.

		My Replication	**CZ**
Predictor	Clear Predictor	150	161
	Likely Predictor	42	44
	Total	192	205
Placebo	Not-Predictor	12	14
	Indirect Signals	95	100
	Total	107	114
Total Signals		299	319

Table 2: **In-sample Performance of Individual Predictors.**

This table reports the average return (% monthly) and t-statistic in the reproduced long-short portfolio for each predictor, using the original in-sample periods. Chen & Zimmermann's (2021) results are listed alongside for comparison. Predictors are sorted by author names. The acronyms used in the present study are also provided. The reproduced mean returns and t-statistics are nearly the same as Chen & Zimmermann's (2021) results.

Original Study	Predictor	Acronym	Sample	Reproduction		CZ's Results	
				Mean Ret	t-Stat	Mean Ret	t-Stat
Abarbanell and Bushee (1998)	Change in capital inv (ind adj)	ChInvIA	1974-1988	0.49	5.42	0.50	5.50
Abarbanell and Bushee (1998)	Sales growth over inventory growth	GrSaleToGrInv	1974-1988	0.31	3.31	0.31	3.30
Abarbanell and Bushee (1998)	Sales growth over overhead growth	GrSaleToGrOverhead	1974-1988	-0.05	-0.39	-0.06	-0.44
Ali, Hwang, and Trombley (2003)	Idiosyncratic risk (AHT)	IdioVolAHT	1976-1997	0.89	2.53	0.89	2.53
Alwathainani (2009)	Earnings consistency	EarningsConsistency	1971-2002	0.21	2.51	0.21	2.51
Amihud (2002)	Amihud's illiquidity	Illiquidity	1964-1997	0.57	3.51	0.57	3.51
Anderson and Garcia-Feijoo (2006)	Change in capex (two years)	grcapx	1976-1999	0.50	4.96	0.50	4.96
Anderson and Garcia-Feijoo (2006)	Change in capex (three years)	grcapx3y	1976-1999	0.59	4.77	0.59	4.77
Ang et al. (2006)	Systematic volatility	betaVIX	1986-2000	1.08	3.57	1.07	3.54
Ang et al. (2006)	Idiosyncratic risk	IdioRisk	1963-2000	1.04	3.38	0.99	3.25
Ang et al. (2006)	Idiosyncratic risk (3 factor)	IdioVol3F	1963-2000	1.03	3.33	0.96	3.12
Ang, Chen and Xing (2006)	Coskewness using daily returns	CoskewACX	1963-2001	0.31	2.87	0.29	2.62
Asquith Pathak and Ritter (2005)	Inst own among high short interest	IO_ShortInterest	1980-2002	2.24	3.07	2.22	3.04
Avramov et al (2007)	Junk Stock Momentum	Mom6mJunk	1985-2003	1.58	3.30	1.58	3.30
Baik and Ahn (2007)	Change in order backlog	OrderBacklogChg	1971-1999	0.38	2.49	0.38	2.49
Balakrishnan, Bartov and Faurel (2010)	Return on assets (qtrly)	roaq	1976-2005	1.69	5.89	1.69	5.90
Bali, Cakici, and Whitelaw (2011)	Maximum return over month	MaxRet	1962-2005	0.89	2.74	0.89	2.74
Bali, Engle and Murray (2015)	Return skewness	ReturnSkew	1963-2012	0.42	5.13	0.41	5.27
Ball et al. (2016)	Cash-based operating profitability	CBOperProf	1963-2014	0.46	3.20	0.46	3.20
Ball et al. (2016)	Operating profitability R&D adjusted	OperProfRD	1963-2014	0.33	1.91	0.33	1.91
Banz (1981)	Size	Size	1926-1975	0.50	2.61	0.50	2.64
Barbee, Mukherji and Raines (1996)	Sales-to-price	SP	1979-1991	0.71	2.84	0.71	2.86
Barber et al. (2001)	Consensus Recommendation	ConsRecomm	1985-1997	0.53	1.36	0.53	1.35
Barber et al. (2001)	Down forecast EPS	DownRecomm	1985-1997	0.63	5.58	0.63	5.54
Barber et al. (2001)	Up Forecast	UpRecomm	1985-1997	0.60	4.57	0.61	4.62
Barry and Brown (1984)	Firm age based on CRSP	FirmAge	1931-1980	-0.01	-0.06	-0.01	-0.06
Barth and Hutton (2004)	Change in Forecast and Accrual	ChForecastAccrual	1981-1996	0.35	4.26	0.36	4.28
Bartov and Kim (2004)	Book-to-market and accruals	AccrualsBM	1980-1998	1.43	4.84	1.44	4.85
Basu (1977)	Earnings-to-Price Ratio	EP	1957-1971	0.39	2.21	0.39	2.21
Bazdresch, Belo and Lin (2014)	Employment growth	hire	1965-2010	0.51	5.64	0.51	5.67

Continued on next page

Table 2: (continued)

Original Study	Predictor	Acronym	Sample	Reproduction		CZ's Results	
				Mean Ret	t-Stat	Mean Ret	t-Stat
Belo and Lin (2012)	Inventory Growth	InvGrowth	1965-2009	0.86	7.14	0.87	7.19
Belo, Lin and Vitorino (2014)	Brand capital investment	BrandInvest	1975-2010	0.60	2.08	0.56	1.97
Bhandari (1988)	Market leverage	Leverage	1952-1981	0.36	2.64	0.36	2.64
Blitz, Huij and Martens (2011)	Momentum based on FF3 residuals	ResidualMomentum	1930-2009	0.94	8.29	0.95	8.27
Blume and Husic (1973)	Price	Price	1932-1971	1.43	3.06	1.42	3.07
Boudoukh et al. (2007)	Net Payout Yield	NetPayoutYield	1984-2003	0.87	2.55	0.87	2.57
Boudoukh et al. (2007)	Payout Yield	PayoutYield	1984-2003	0.44	2.29	0.43	2.27
Bradshaw, Richardson, Sloan (2006)	Net debt financing	NetDebtFinance	1971-2000	0.75	7.74	0.75	7.70
Bradshaw, Richardson, Sloan (2006)	Net equity financing	NetEquityFinance	1971-2000	1.06	5.39	1.06	5.40
Brennan, Chordia, Subra (1998)	Past trading volume	DolVol	1966-1995	0.75	2.71	0.75	2.72
Cen, Wei, and Zhang (2006)	Analyst earnings per share	FEPS	1983-2002	1.47	3.05	1.46	3.04
Chan and Ko (2006)	Momentum and LT Reversal	MomRev	1965-2001	1.20	4.39	1.19	4.38
Chan, Jegadeesh and Lakonishok (1996)	Earnings announcement return	AnnouncementReturn	1977-1992	1.21	13.41	1.20	13.34
Chan, Jegadeesh and Lakonishok (1996)	Earnings forecast revisions	REV6	1977-1992	1.28	5.35	1.29	5.44
Chan, Lakonishok and Sougiannis (2001)	Advertising Expense	AdExp	1975-1996	0.65	3.16	0.65	3.15
Chan, Lakonishok and Sougiannis (2001)	R&D over market cap	RD	1975-1995	1.00	5.79	1.01	5.82
Chandrashekar and Rao (2009)	Cash Productivity	CashProd	1963-2003	0.52	3.18	0.56	3.40
Chen, Hong and Stein (2002)	Breadth of ownership	DelBreadth	1979-1998	0.69	3.69	0.69	3.69
Chordia, Subra, Anshuman (2001)	Share turnover volatility	std_turn	1966-1995	0.77	3.30	0.80	3.42
Chordia, Subra, Anshuman (2001)	Volume Variance	VolSD	1966-1995	0.38	2.82	0.38	2.82
Cohen and Frazzini (2008)	Customer momentum	CustomerMomentum	1980-2004	1.15	3.23	1.16	3.27
Cohen and Lou (2012)	Conglomerate return	retConglomerate	1977-2009	1.32	6.50	1.32	6.50
Cohen, Diether and Malloy (2013)	R&D ability	RDAbility	1980-2009	0.22	1.19	0.27	1.50
Cooper, Gulen and Schill (2008)	Asset growth	AssetGrowth	1968-2003	1.50	7.63	1.50	7.64
Cremers and Nair (2005)	Takeover vulnerability	Activism1	1990-2001	0.24	1.00	0.24	1.00
Cremers and Nair (2005)	Active shareholders	Activism2	1990-2001	0.44	1.05	0.43	1.02
Cusatis, Miles and Woolridge (1993)	Spinoffs	Spinoff	1965-1988	0.40	2.22	0.40	2.22
Da and Warachka (2011)	Long-vs-short EPS forecasts	EarningsForecastDisparity	1983-2006	0.66	4.35	0.66	4.38
Daniel and Titman (2006)	Composite equity issuance	CompEquIss	1968-2003	0.27	2.41	2.27	2.41
Daniel and Titman (2006)	Intangible return using BM	IntanBM	1968-2003	0.39	2.28	0.40	2.29
Daniel and Titman (2006)	Intangible return using CFtoP	IntanCFP	1968-2003	0.40	2.31	0.40	2.32
Daniel and Titman (2006)	Intangible return using EP	IntanEP	1968-2003	0.33	2.44	0.34	2.46
Daniel and Titman (2006)	Intangible return using Sale2P	IntanSP	1968-2003	0.53	2.41	0.53	2.42
Daniel and Titman (2006)	Share issuance (5 year)	ShareIss5Y	1968-2003	0.52	4.31	0.52	4.32
Datar, Naik and Radcliffe (1998)	Share Volume	ShareVol	1962-1991	0.93	3.77	0.91	3.87
De Bondt and Thaler (1985)	Long-run reversal	LRreversal	1929-1982	0.79	3.04	0.79	3.04
De Bondt and Thaler (1985)	Medium-run reversal	MRreversal	1933-1980	0.39	2.08	0.39	2.08
Dechow et al. (2001)	Short Interest	ShortInterest	1976-1993	0.83	5.29	0.83	5.30

Continued on next page

Table 2: (continued)

Original Study	Predictor	Acronym	Sample	Reproduction		CZ's Results	
				Mean Ret	t-Stat	Mean Ret	t-Stat
Dechow, Sloan and Soliman (2004)	Equity Duration	EquityDuration	1962-1998	0.56	3.10	0.56	3.09
Desai, Rajgopal, Venkatachalam (2004)	Operating Cash flows to price	cfp	1973-1997	0.36	2.16	0.36	2.18
Dharan and Ikenberry (1995)	Exchange Switch	ExchSwitch	1962-1990	0.45	2.89	0.45	2.89
Dichev (1998)	O Score	OScore	1981-1995	0.79	3.20	1.01	3.39
Dichev and Piotroski (2001)	Credit Rating Downgrade	CredRatDG	1986-1998	0.72	2.91	0.73	2.92
Diether, Malloy and Scherbina (2002)	EPS Forecast Dispersion	ForecastDispersion	1976-2000	0.65	3.07	0.65	3.05
Doyle, Lundholm and Soliman (2003)	Excluded Expenses	ExclExp	1988-1999	0.25	2.85	0.27	3.02
Eberhart, Maxwell and Siddique (2004)	Unexpected R&D increase	SurpriseRD	1974-2001	0.29	2.98	0.29	3.00
Elgers, Lo and Pfeiffer (2001)	Earnings Forecast to price	sfe	1982-1998	0.83	2.66	0.81	2.61
Fairfield, Whisenant and Yohn (2003)	Growth in long term operating assets	GrLTNOA	1964-1993	0.37	3.72	0.37	3.73
Fama and French (1992)	Total assets to market	AM	1963-1990	0.63	3.51	0.63	3.50
Fama and French (1992)	Book to market using December ME	BMdec	1963-1990	0.97	5.35	0.98	5.37
Fama and French (1992)	Book leverage (annual)	BookLeverage	1963-1990	0.28	3.30	0.28	3.30
Fama and French (2006)	operating profits / book equity	OperProf	1977-2003	0.72	3.02	0.72	3.00
Fama and MacBeth (1973)	CAPM beta	Beta	1929-1968	0.66	1.72	0.66	1.72
Foster, Olsen and Shevlin (1984)	Earnings Surprise	EarningsSurprise	1974-1981	1.16	4.94	1.16	4.94
Frankel and Lee (1998)	Analyst Value	AnalystValue	1975-1993	0.26	1.72	0.26	1.73
Frankel and Lee (1998)	Analyst Optimism	AOP	1975-1993	0.37	2.07	0.36	2.01
Frankel and Lee (1998)	Predicted Analyst forecast error	PredictedFE	1979-1993	0.31	0.99	0.30	0.96
Franzoni and Marin (2006)	Pension Funding Status	FR	1980-2002	0.09	0.37	0.31	1.74
Frazzini and Pedersen (2014)	Frazzini-Pedersen Beta	BetaFP	1929-2012	0.06	0.20	0.03	0.08
George and Hwang (2004)	52 week high	High52	1963-2001	0.51	1.94	0.51	1.94
Gompers, Ishii and Metrick (2003)	Governance Index	Governance	1990-1999	0.52	2.11	0.52	2.11
Gou, Lev and Shi (2006)	IPO and no R&D spending	RDIPO	1980-1995	1.00	3.05	0.97	2.97
Grinblatt and Moskowitz (1999)	Industry Momentum	IndMom	1963-1995	0.26	2.45	0.27	2.55
Hafzalla, Lundholm, Van Winkle (2011)	Percent Operating Accruals	PctAcc	1989-2008	0.46	3.52	0.46	3.51
Hafzalla, Lundholm, Van Winkle (2011)	Percent Total Accruals	PctTotAcc	1989-2008	0.49	4.69	0.50	4.70
Hahn and Lee (2009)	Tangibility	tang	1973-2001	0.71	3.66	0.71	3.67
Hartzmark and Salomon (2013)	Dividend seasonality	DivSeason	1927-2011	0.33	14.53	0.33	14.54
Harvey and Siddique (2000)	Coskewness	Coskewness	1964-1993	0.28	2.29	0.27	2.18
Haugen and Baker (1996)	net income / book equity	RoE	1979-1993	0.32	2.82	0.32	2.82
Haugen and Baker (1996)	Cash-flow to price variance	VarCF	1979-1993	-0.57	-1.96	-0.56	-1.91
Haugen and Baker (1996)	Volume to market equity	VolMkt	1979-1993	0.44	1.57	0.45	1.59
Haugen and Baker (1996)	Volume Trend	VolumeTrend	1979-1993	0.54	2.90	0.54	2.93
Hawkins, Chamberlin, Daniel (1984)	EPS forecast revision	AnalystRevision	1975-1980	0.91	5.11	0.91	5.12
Heston and Sadka (2008)	Momentum without the seasonal part	Mom12mOffSeason	1965-2002	1.22	3.84	1.23	3.85
Heston and Sadka (2008)	Off season long-term reversal	MomOffSeason	1965-2002	1.31	4.90	1.31	4.93
Heston and Sadka (2008)	Off season reversal years 6 to 10	MomOffSeason06YrPlus	1965-2002	0.59	4.39	0.59	4.36

Continued on next page

Table 2: (continued)

Original Study	Predictor	Acronym	Sample	Reproduction		CZ's Results	
				Mean Ret	t-Stat	Mean Ret	t-Stat
Heston and Sadka (2008)	Off season reversal years 11 to 15	MomOffSeason11YrPlus	1965-2002	0.24	2.02	0.24	2.03
Heston and Sadka (2008)	Off season reversal years 16 to 20	MomOffSeason16YrPlus	1965-2002	0.36	2.60	0.35	2.54
Heston and Sadka (2008)	Return seasonality last year	MomSeasonShort	1965-2002	1.36	8.63	1.36	8.62
Hirschleifer, Hsu and Li (2013)	Citations to RD expenses	CitationsRD	1982-2008	0.23	2.53	0.21	2.21
Hirschleifer, Hsu and Li (2013)	Patents to RD expenses	PatentsRD	1982-2008	0.30	3.19	0.30	3.18
Hirshleifer et al. (2004)	Net Operating Assets	NOA	1964-2002	1.07	7.50	1.07	7.51
Hirshleifer, Hou, Teoh, Zhang (2004)	change in net operating assets	dNoa	1964-2002	1.04	9.23	1.05	9.25
Hong and Kacperczyk (2009)	Sin Stock (selection criteria)	sinAlgo	1926-2006	0.21	1.92	0.21	1.92
Hou and Moskowitz (2005)	Price delay r square	PriceDelayRsq	1964-2001	0.50	2.79	0.48	2.69
Hou and Moskowitz (2005)	Price delay coeff	PriceDelaySlope	1964-2001	0.20	2.36	0.17	2.01
Hou and Moskowitz (2005)	Price delay SE adjusted	PriceDelayTstat	1964-2001	0.13	1.29	0.15	1.53
Hou and Robinson (2006)	Industry concentration (sales)	Herf	1963-2001	0.15	1.64	0.21	2.30
Hou and Robinson (2006)	Industry concentration (assets)	HerfAsset	1963-2001	0.18	1.62	0.18	1.66
Hou and Robinson (2006)	Industry concentration (equity)	HerfBE	1963-2001	0.22	2.00	0.22	2.04
Ikenberry, Lakonishok, Vermaelen (1995)	Share repurchases	ShareRepurchase	1980-1990	0.32	3.96	0.32	4.01
Jegadeesh (1990)	Short term reversal	STreversal	1934-1987	2.95	14.11	2.94	14.02
Jegadeesh and Livnat (2006)	Revenue Surprise	RevenueSurprise	1987-2003	0.75	5.99	0.75	5.99
Jegadeesh and Titman (1993)	Momentum (12 month)	Mom12m	1964-1989	1.37	4.58	1.37	4.58
Jegadeesh and Titman (1993)	Momentum (6 month)	Mom6m	1964-1989	1.04	3.68	1.04	3.68
Jegadeesh et al. (2004)	Change in recommendation	ChangeInRecommendation	1985-1998	1.03	6.61	1.04	6.65
Johnson and So (2012)	Option to stock volume	OptionVolume1	1996-2010	0.68	2.07	0.68	2.04
Johnson and So (2012)	Option volume to average	OptionVolume2	1996-2010	0.52	1.75	0.53	1.79
Kelly and Jiang (2014)	Tail risk beta	BetaTailRisk	1963-2010	0.46	3.30	0.46	3.30
La Porta (1996)	Long-term EPS forecast	fgr5yrLag	1983-1990	0.83	2.06	0.83	2.06
Lakonishok, Shleifer, Vishny (1994)	Cash flow to market	CF	1968-1990	0.83	4.03	0.83	4.04
Lakonishok, Shleifer, Vishny (1994)	Revenue Growth Rank	MeanRankRevGrowth	1968-1990	0.55	3.90	0.55	3.94
Landsman et al. (2011)	Real dirty surplus	RDS	1976-2003	0.49	3.73	0.49	3.83
Lee and Swaminathan (2000)	Momentum in high volume stocks	MomVol	1965-1995	1.58	5.02	1.59	5.05
Lev and Nissim (2004)	Taxable income to income	Tax	1973-2000	0.44	3.49	0.45	3.52
Litzenberger and Ramaswamy (1979)	Predicted div yield next month	DivYieldST	1936-1977	0.41	4.23	0.41	4.23
Liu (2006)	Days with zero trades	zerotrade	1960-2003	0.55	2.94	0.49	2.63
Liu (2006)	Days with zero trades	zerotradeAlt1	1960-2003	0.67	4.02	0.64	3.80
Liu (2006)	Days with zero trades	zerotradeAlt12	1960-2003	0.44	3.07	0.40	2.79
Lockwood and Prombutr (2010)	Growth in book equity	ChEQ	1964-2007	0.56	4.24	0.56	4.26
Loh and Warachka (2012)	Earnings surprise streak	EarningsStreak	1987-2009	1.10	10.67	1.09	10.60
Loh and Warachka (2012)	Earnings streak length	NumEarnIncrease	1987-2009	0.52	6.80	0.52	6.75
Lou (2014)	Growth in advertising expenses	GrAdExp	1974-2010	0.44	3.90	0.44	3.89
Loughran and Wellman (2011)	Enterprise Multiple	EntMult	1963-2009	0.85	5.52	0.85	5.55

Continued on next page

Table 2: (continued)

Original Study	Predictor	Acronym	Sample	Reproduction		CZ's Results	
				Mean Ret	t-Stat	Mean Ret	t-Stat
Lyandres, Sun and Zhang (2008)	Composite debt issuance	CompositeDebtIssuance	1970-2005	0.31	5.16	0.31	5.19
Lyandres, Sun and Zhang (2008)	change in ppe and inv/assets	InvestPPEInv	1970-2005	0.79	7.84	0.80	7.86
Menzly and Ozbas (2010)	Customers momentum	iomom_cust	1986-2005	0.71	2.53	0.71	2.53
Menzly and Ozbas (2010)	Suppliers momentum	iomom_supp	1986-2005	0.61	2.33	0.60	2.31
Michaely, Thaler and Womack (1995)	Dividend Initiation	DivInit	1964-1988	0.58	5.21	0.58	5.65
Michaely, Thaler and Womack (1995)	Dividend Omission	DivOmit	1964-1988	0.52	3.11	0.51	3.06
Mohanram (2005)	Mohanram G-score	MS	1978-2001	1.32	5.38	1.34	5.44
Nagel (2005)	Inst Own and Forecast Dispersion	RIO_Disp	1980-2003	0.57	2.42	0.62	2.60
Nagel (2005)	Inst Own and Market to Book	RIO_MB	1980-2003	0.90	3.74	0.90	3.74
Nagel (2005)	Inst Own and Turnover	RIO_Turnover	1980-2003	0.65	2.79	0.65	2.78
Nagel (2005)	Inst Own and Idio Vol	RIO_Volatility	1980-2003	1.01	4.19	1.01	4.19
Novy-Marx (2011)	Operating leverage	OPLeverage	1963-2008	0.35	2.49	0.35	2.50
Novy-Marx (2012)	Intermediate Momentum	IntMom	1927-2010	1.24	5.92	1.24	5.90
Novy-Marx (2013)	gross profits / total assets	GP	1963-2010	0.30	2.38	0.30	2.38
Palazzo (2012)	Cash to assets	Cash	1972-2009	0.70	2.97	0.70	2.97
Pastor and Stambaugh (2003)	Pastor-Stambaugh liquidity beta	BetaLiquidityPS	1968-1999	0.34	1.89	0.35	1.93
Penman, Richardson and Tuna (2007)	Leverage component of BM	BPEBM	1963-2001	0.23	2.86	0.22	2.83
Penman, Richardson and Tuna (2007)	Enterprise component of BM	EBM	1963-2001	0.30	4.09	0.31	4.14
Penman, Richardson and Tuna (2007)	Net debt to price	NetDebtPrice	1963-2001	0.55	3.84	0.55	3.88
Piotroski (2000)	Piotroski F-score	PS	1976-1996	0.92	3.29	0.92	3.29
Pontiff and Woodgate (2008)	Share issuance (1 year)	ShareIss1Y	1970-2003	0.62	4.97	0.62	4.97
Prakash and Sinha (2013)	Deferred Revenue	DelDRC	2002-2007	0.71	1.66	0.71	1.66
Rajgopal, Shevlin, Venkatachalam (2003)	Order backlog	OrderBacklog	1981-1999	0.49	3.35	0.51	3.43
Richardson et al. (2005)	Change in current operating assets	DelCOA	1962-2001	0.53	5.98	0.54	6.01
Richardson et al. (2005)	Change in current operating liabilities	DelCOL	1962-2001	0.35	4.33	0.35	4.35
Richardson et al. (2005)	Change in equity to assets	DelEqu	1963-2001	0.46	3.17	0.47	3.18
Richardson et al. (2005)	Change in financial liabilities	DelFINL	1962-2001	0.73	12.25	0.73	12.23
Richardson et al. (2005)	Change in long-term investment	DelLTI	1962-2001	0.16	2.53	0.17	2.55
Richardson et al. (2005)	Change in net financial assets	DelNetFin	1962-2001	0.55	8.97	0.55	9.00
Richardson et al. (2005)	Total accruals	TotalAccruals	1962-2001	0.28	2.62	0.28	2.63
Ritter (1991)	IPO and age	AgeIPO	1981-1984	1.40	2.67	1.41	2.68
Ritter (1991)	Initial Public Offerings	IndIPO	1975-1987	0.66	2.36	0.66	2.36
Rosenberg, Reid, and Lanstein (1985)	Book to market using most recent ME	BM	1973-1984	1.60	3.79	1.60	3.79
Scherbina (2008)	Decline in Analyst Coverage	ChNAnalyst	1982-2005	0.65	1.74	1.09	3.65
Sloan (1996)	Accruals	Accruals	1962-1991	0.69	6.75	0.56	5.07
Soliman (2008)	Change in Asset Turnover	ChAssetTurnover	1984-2002	0.29	3.77	0.29	3.77
Soliman (2008)	Change in Net Noncurrent Op Assets	ChNNCOA	1984-2002	0.36	4.46	0.35	4.43
Soliman (2008)	Change in Net Working Capital	ChNWC	1984-2002	0.16	2.75	0.16	2.83

Continued on next page

Table 2: (continued)

Original Study	**Predictor**	**Acronym**	**Sample**	**Reproduction**		**CZ's Results**	
				Mean Ret	t-Stat	Mean Ret	t-Stat
Spiess and Affleck-Graves (1999)	Debt Issuance	DebtIssuance	1975-1989	0.17	2.99	0.17	2.98
Thomas and Zhang (2002)	Inventory Growth	ChInv	1970-1997	0.77	6.21	0.77	6.24
Thomas and Zhang (2011)	Change in Taxes	ChTax	1977-2006	1.09	9.54	1.09	9.50
Titman, Wei and Xie (2004)	Investment to revenue	Investment	1973-1996	0.27	2.44	0.25	2.28
Tuzel (2010)	Real estate holdings	realestate	1971-2005	0.29	1.84	0.28	1.82
Valta (2016)	Convertible debt indicator	ConvDebt	1985-2012	0.38	4.34	0.38	4.34
Xie (2001)	Abnormal Accruals	AbnormalAccruals	1971-1992	0.47	4.28	0.54	5.10
Xing, Zhang and Zhao (2010)	Volatility smirk near the money	skew1	1996-2005	0.63	3.14	0.55	2.45
Yan (2011)	Put volatility minus call volatility	SmileSlope	1996-2005	1.80	7.87	1.78	7.62
Zhang (2006)	Firm Age - Momentum	FirmAgeMom	1983-2001	2.09	5.05	2.29	5.40

Table 3: **In-sample Performance of Individual Placebos.**

This table reports the average return (% monthly) and t-statistic in the reproduced long-short portfolio for each placebo, using the original in-sample periods. Chen & Zimmermann's (2021) results are listed alongside for comparison. Placebos are sorted by author names. The acronyms used in the present study are also provided.

Original Study	Predictor	Acronym	Sample	Reproduction		CZ's Results	
				Mean Ret	t-Stat	Mean Ret	t-Stat
Abarbanell and Bushee (1998)	Effective Tax Rate	ETR	1974-1988	0.01	0.13	0.01	0.18
Abarbanell and Bushee (1998)	Gross margin growth to sales growth	GrGMToGrSales	1974-1988	0.36	3.29	0.37	3.30
Abarbanell and Bushee (1998)	Laborforce efficiency	LaborforceEfficiency	1974-1988	-0.07	-0.74	-0.07	-0.74
Abarbanell and Bushee (1998)	Change in gross margin vs sales	pchgm_pchsale	1974-1988	0.42	3.72	0.42	3.76
Acharya and Pedersen (2005)	Illiquidity-illiquidity beta (beta2i)	betaCC	1964-1999	0.26	1.62	0.33	1.87
Acharya and Pedersen (2005)	Illiquidity-market return beta (beta4i)	betaCR	1964-1999	-0.02	-0.23	-0.09	-0.98
Acharya and Pedersen (2005)	Net liquidity beta (betanet,p)	betaNet	1964-1999	0.29	1.67	0.35	1.97
Acharya and Pedersen (2005)	Return-market illiquidity beta	betaRC	1964-1999	0.10	0.53	0.06	0.30
Acharya and Pedersen (2005)	Return-market return illiquidity beta	betaRR	1964-1999	-0.03	-0.13	-0.03	-0.14
Adrian, Etula and Muir (2014)	Broker-Dealer Leverage Beta	BetaBDLeverage	1973-2009	0.40	2.20	0.41	2.22
Anderson and Garcia-Feijoo (2006)	Investment growth (1 year)	grcapx1y	1964-2003	-0.28	-3.74	-0.28	-3.74
Anderson, Ghysels, and Juergens (2005)	Long-term forecast dispersion	ForecastDispersionLT	1991-1997	-0.01	-0.04	0.00	0.00
Ang et al. (2006)	Idiosyncratic risk (CAPM)	IdioVolCAPM	1963-2000	-0.33	-1.07	-0.31	-1.01
Ang et al. (2006)	Idiosyncratic risk (q factor)	IdioVolQF	1967-2000	-0.39	-1.16	-0.39	-1.16
Ang, Chen and Xing (2006)	Downside beta	DownsideBeta	1963-2001	0.05	0.23	0.07	0.31
Balakrishnan, Bartov and Faurel (2010)	Change in Return on assets	ChangeRoA	1976-2005	1.32	12.56	1.32	12.59
Balakrishnan, Bartov and Faurel (2010)	Change in Return on equity	ChangeRoE	1976-2005	1.08	10.54	1.08	10.53
Bali, Engle and Murray (2015)	Idiosyncratic skewness (Q model)	ReturnSkewQF	1967-2012	-0.26	-4.18	-0.26	-4.29
Ball et al. (2016)	Cash-based oper prof lagged assets	CBOperProfLagAT	1963-2014	0.46	3.33	0.46	3.33
Ball et al. (2016)	Cash-based oper prof lagged assets qtrly	CBOperProfLagAT_q	1963-2014	0.86	6.17	0.86	6.19
Ball et al. (2016)	Oper prof R&D adj lagged assets	OperProfRDLagAT	1963-2014	0.05	0.33	0.05	0.33
Ball et al. (2016)	Oper prof R&D adj lagged assets (qtrly)	OperProfRDLagAT_q	1963-2014	1.17	5.57	1.17	5.56
Barbee, Mukherji and Raines (1996)	Sales-to-price quarterly	SP_q	1979-1991	1.18	4.92	1.18	4.93
Basu (1977)	Earnings-to-Price Ratio	EPq	1963-1971	1.31	6.79	1.31	6.79
Belo, Lin and Vitorino (2014)	Brand capital to assets	BrandCapital	1975-2010	0.22	1.16	0.24	1.25
Bhandari (1988)	Market leverage quarterly	Leverage_q	1952-1981	0.26	0.81	0.26	0.79
Blitz, Huij and Martens (2011)	6 month residual momentum	ResidualMomentum6m	1930-2009	0.40	4.19	0.39	4.08
Boudoukh et al. (2007)	Net Payout Yield quarterly	NetPayoutYield_q	1984-2003	0.75	2.03	0.76	2.04
Boudoukh et al. (2007)	Payout Yield quarterly	PayoutYield_q	1984-2003	0.71	6.12	0.72	6.16
Brown and Rowe (2007)	Return on invested capital	roic	1970-2005	0.08	0.34	0.08	0.33
Callen, Khan and Lu (2013)	Accounting component of price delay	DelayAcct	1981-2006	-0.10	-0.57	-0.16	-0.83
Callen, Khan and Lu (2013)	Non-accounting component of price delay	DelayNonAcct	1981-2006	0.27	1.68	0.27	1.70

Continued on next page

Table 3: (continued)

Original Study	Predictor	Acronym	Sample	Reproduction		CZ's Results	
				Mean Ret	t-Stat	Mean Ret	t-Stat
Campbell, Hilscher and Szilagyi (2008)	Failure probability	FailureProbability	1981-2003	0.40	0.92	0.40	0.91
Campbell, Hilscher and Szilagyi (2008)	Failure probability	FailureProbabilityJune	1981-2003	0.04	0.08	0.03	0.07
Chan, Lakonishok and Sougiannis (2001)	R&D over market cap quarterly	RD_q	1975-1995	1.87	5.18	1.89	5.23
Chan, Lakonishok and Sougiannis (2001)	R&D to sales	rd_sale	1975-1995	0.13	0.56	0.12	0.53
Chan, Lakonishok and Sougiannis (2001)	R&D to sales	rd_sale_q	1975-1995	0.71	1.48	0.71	1.48
Cooper, Gulen and Schill (2008)	Asset growth quarterly	AssetGrowth_q	1968-2003	-0.94	-4.83	-0.94	-4.84
Desai, Rajgopal, Venkatachalam (2004)	Operating Cash flows to price quarterly	cfpq	1973-1997	1.07	8.11	1.07	8.12
Dichev (1998)	O Score quarterly	OScore_q	1981-1995	-1.11	-3.06	-1.10	-3.02
Dichev (1998)	Altman Z-Score	ZScore	1981-1995	-0.35	-1.20	-0.35	-1.20
Dichev (1998)	Altman Z-Score quarterly	ZScore_q	1981-1995	-0.13	-0.44	-0.13	-0.47
Dimson (1979)	Dimson Beta	BetaDimson	1955-1974	-0.30	-1.93	-0.23	-1.53
Elgers, Lo and Pfeiffer (2001)	Number of analysts	nanalyst	1982-1998	0.19	1.03	0.19	1.02
Fama and French (1992)	Total assets to market (quarterly)	AMq	1975-1990	0.78	3.31	0.78	3.31
Fama and French (1992)	Book leverage (quarterly)	BookLeverageQuarterly	1973-1990	-0.23	-1.60	-0.23	-1.59
Fama and French (2006)	operating profits / book equity	OperProfLag	1977-2003	0.40	1.82	0.40	1.81
Fama and French (2006)	operating profits / book equity	OperProfLag_q	1977-2003	1.02	3.40	1.02	3.40
Fama and MacBeth (1973)	CAPM beta squred	BetaSquared	1929-1968	-0.66	-1.71	-0.66	-1.71
Francis, LaFond, Olsson, Schipper (2004)	Earnings conservatism	EarningsConservatism	1975-2001	0.00	-0.07	0.00	-0.01
Francis, LaFond, Olsson, Schipper (2004)	Earnings persistence	EarningsPersistence	1975-2001	-0.21	-1.60	-0.21	-1.59
Francis, LaFond, Olsson, Schipper (2004)	Earnings Predictability	EarningsPredictability	1975-2001	-0.60	-3.52	-0.60	-3.49
Francis, LaFond, Olsson, Schipper (2004)	Earnings Smoothness	EarningsSmoothness	1975-2001	0.02	0.12	0.02	0.11
Francis, LaFond, Olsson, Schipper (2004)	Earnings timeliness	EarningsTimeliness	1975-2001	-0.01	-0.17	-0.01	-0.21
Francis, LaFond, Olsson, Schipper (2004)	Value relevance of earnings	EarningsValueRelevance	1975-2001	-0.02	-0.33	-0.02	-0.32
Francis, LaFond, Olsson, Schipper (2004)	RoA volatility	roavol	1975-2001	-0.07	-0.19	-0.07	-0.19
Francis, LaFond, Olsson, Schipper (2005)	Accrual Quality	AccrualQuality	1971-2002	0.15	0.58	0.16	0.62
Frankel and Lee (1998)	Intrinsic or historical value	IntrinsicValue	1975-1993	0.49	2.49	0.48	2.45
Hafzalla, Lundholm, Van Winkle (2011)	Percent Abnormal Accruals	AbnormalAccrualsPercent	1989-2008	-0.28	-3.90	-0.29	-4.09
Hahn and Lee (2009)	Tangibility quarterly	tang_q	1973-2001	0.84	5.81	0.85	5.84
Haugen and Baker (1996)	Capital turnover	CapTurnover	1979-1993	0.25	1.23	0.25	1.23
Haugen and Baker (1996)	Capital turnover (quarterly)	CapTurnover_q	1979-1993	0.85	4.69	0.85	4.68
Holthausen and Larcker (1992)	Depreciation to PPE	depr	1978-1988	0.27	1.00	0.28	1.02
Holthausen and Larcker (1992)	Change in depreciation to PPE	pchdepr	1978-1988	0.17	1.65	0.18	1.66
La Porta (1996)	Long-term EPS forecast (Monthly)	fgr5yrNoLag	1983-1990	-0.64	-1.56	-0.66	-1.60
Lakonishok, Shleifer, Vishny (1994)	Annual sales growth	sgr	1968-1990	-0.60	-4.23	-0.60	-4.25
Lakonishok, Shleifer, Vishny (1994)	Annual sales growth quarterly	sgr_q	1968-1990	0.61	3.53	0.60	3.51
Lamont, Polk and Saa-Requejo (2001)	Kaplan Zingales index	KZ	1968-1997	0.08	0.53	0.08	0.53
Lamont, Polk and Saa-Requejo (2001)	Kaplan Zingales index quarterly	KZ_q	1968-1997	-1.48	-9.21	-1.49	-9.23
Lev and Nissim (2004)	Taxable income to income (qtrly)	Tax_q	1973-2000	0.02	0.20	0.03	0.23

Continued on next page

Table 3: (continued)

Original Study	Predictor	Acronym	Sample	Reproduction Mean Ret	Reproduction t-Stat	CZ's Results Mean Ret	CZ's Results t-Stat
Loughran and Wellman (2011)	Enterprise Multiple quarterly	EntMult_q	1963-2009	-1.59	-11.99	-1.59	-11.96
Naranjo, Nimalendran, Ryngaert (1998)	Dividend yield for small stocks	DivYield	1963-1994	0.37	1.23	0.34	1.11
Naranjo, Nimalendran, Ryngaert (1998)	Last year's dividends over price	DivYieldAnn	1963-1994	0.01	0.12	0.01	0.11
Novy-Marx (2011)	Operating leverage (qtrly)	OPLeverage_q	1963-2008	0.39	2.36	0.39	2.37
Novy-Marx (2013)	gross profits / total assets	GPlag	1963-2010	0.20	1.85	0.20	1.85
Novy-Marx (2013)	gross profits / total assets	GPlag_q	1963-2010	0.88	6.16	0.88	6.17
Ortiz-Molina and Phillips (2014)	Asset liquidity over book assets	AssetLiquidityBook	1984-2006	0.35	1.36	0.35	1.37
Ortiz-Molina and Phillips (2014)	Asset liquidity over book (qtrly)	AssetLiquidityBookQuart	1984-2006	0.31	0.95	0.31	0.96
Ortiz-Molina and Phillips (2014)	Asset liquidity over market	AssetLiquidityMarket	1984-2006	1.41	7.45	1.41	7.46
Ortiz-Molina and Phillips (2014)	Asset liquidity over market (qtrly)	AssetLiquidityMarketQuart	1984-2006	1.31	6.36	1.31	6.35
Ou and Penman (1989)	CF to debt	cashdebt	1973-1983	-0.06	-0.21	-0.06	-0.21
Ou and Penman (1989)	Current Ratio	currat	1973-1983	0.29	1.94	0.29	1.93
Ou and Penman (1989)	Change in Current Ratio	pchcurrat	1973-1983	0.21	2.55	0.21	2.54
Ou and Penman (1989)	Change in quick ratio	pchquick	1973-1983	0.32	3.28	0.32	3.27
Ou and Penman (1989)	Change in sales to inventory	pchsaleinv	1973-1983	0.49	5.25	0.49	5.24
Ou and Penman (1989)	Quick ratio	quick	1973-1983	0.29	1.77	0.30	1.77
Ou and Penman (1989)	Sales to cash ratio	salecash	1973-1983	0.22	1.31	0.22	1.31
Ou and Penman (1989)	Sales to inventory	saleinv	1973-1983	0.03	0.18	0.03	0.18
Ou and Penman (1989)	Sales to receivables	salerec	1973-1983	0.29	1.54	0.29	1.53
Penman, Richardson and Tuna (2007)	Enterprise component of BM	EBM_q	1963-2001	0.81	7.07	0.81	7.09
Penman, Richardson and Tuna (2007)	Net debt to price	NetDebtPrice_q	1963-2001	-0.73	-3.76	-0.73	-3.78
Piotroski (2000)	Piotroski F-score	PS_q	1976-1996	1.00	3.94	1.39	6.35
Richardson et al. (2005)	Change in short-term investment	DelSTI	1962-2001	-0.03	-0.52	-0.04	-0.53
Rosenberg, Reid, and Lanstein (1985)	Book to market (quarterly)	BMq	1973-1984	1.65	3.67	1.65	3.67
Soliman (2008)	Asset Turnover	AssetTurnover	1984-2002	0.40	2.21	0.40	2.23
Soliman (2008)	Asset Turnover	AssetTurnover_q	1984-2002	0.59	3.09	0.59	3.10
Soliman (2008)	Change in Noncurrent Operating Assets	ChNCOA	1984-2002	-1.02	-5.69	-1.02	-5.68
Soliman (2008)	Change in Noncurrent Operating Liab	ChNCOL	1984-2002	-0.54	-3.58	-0.54	-3.60
Soliman (2008)	Change in Profit Margin	ChPM	1984-2002	0.11	1.34	0.11	1.35
Soliman (2008)	Profit Margin	PM	1984-2002	0.52	1.93	0.52	1.94
Soliman (2008)	Profit Margin	PM_q	1984-2002	1.29	2.88	1.29	2.88
Soliman (2008)	Return on Net Operating Assets	RetNOA	1984-2002	0.01	0.06	0.01	0.05
Soliman (2008)	Return on Net Operating Assets	RetNOA_q	1984-2002	1.26	3.26	1.26	3.25
Valta (2016)	Secured debt	secured	1985-2012	0.00	-0.03	0.00	-0.04
Valta (2016)	Secured debt indicator	securedind	1985-2012	-0.06	-0.70	-0.06	-0.69
Whited and Wu (2006)	Whited-Wu index	WW	1975-2001	0.37	1.34	0.38	1.34
Whited and Wu (2006)	Whited-Wu index	WW_Q	1975-2001	0.50	1.31	0.50	1.31

Table 4: **Predictor Performance Across Sample Periods.**

This table presents the average return (% monthly) and t-statistic in the reproduced long-short portfolio for each predictor, with full sample, pre-publication and post-publication, respectively. The full sample includes all data available for download, the sample periods before the original publication year are pre-publication, the ones after are post-publication. Predictors are sorted by author names in Table 2.

		Full Sample				**Pre-Publication**				**Post-Publication**			
		All Stocks		**No Microcaps**		**All Stocks**		**No Microcaps**		**All Stocks**		**No Microcaps**	
Acronym	**Sample**	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat
ChInvIA (1998)	1952-2022	0.33	5.81	0.23	4.75	0.37	6.57	0.26	5.29	0.24	1.91	0.16	1.54
GrSaleToGrInv (1998)	1952-2022	0.24	5.17	0.19	4.08	0.36	6.45	0.3	5.37	0.01	0.13	-0.03	-0.37
GrSaleToGrOverhead (1998)	1952-2022	-0.02	-0.41	-0.04	-0.66	0.08	1.21	0.04	0.65	-0.23	-1.99	-0.19	-1.78
IdioVolAHT (2003)	1927-2021	0.16	0.77	-0.05	-0.26	0.24	1.05	0.03	0.16	-0.25	-0.56	-0.42	-1.19
EarningsConsistency (2009)	1953-2022	0.25	3.62	0.22	3.09	0.27	3.64	0.26	3.15	0.18	0.93	0.06	0.43
Illiquidity (2002)	1927-2022	0.34	3.33	0.29	3.32	0.45	3.64	0.35	3.4	-0.1	-0.79	0.03	0.26
grcapx (2006)	1953-2022	0.35	5.81	0.28	4.66	0.41	6.16	0.3	4.65	0.14	1.01	0.18	1.32
grcapx3y (2006)	1954-2022	0.39	5.56	0.31	5.55	0.48	6.28	0.34	5.51	0.06	0.4	0.23	1.7
IdioRisk (2006)	1926-2021	0.55	2.82	0.3	1.8	0.64	3	0.33	1.86	0	0.01	0.05	0.13
IdioVol3F (2006)	1926-2021	0.47	2.36	0.28	1.68	0.54	2.49	0.3	1.61	0.01	0.02	0.19	0.48
betaVIX (2006)	1986-2021	0.63	3.4	0.47	2.62	0.89	3.57	0.62	2.63	0.22	0.8	0.21	0.8
CoskewACX (2005)	1963-2021	0.36	3.4	0.29	2.95	0.36	3.55	0.31	3.42	0.36	1.18	0.21	0.74
IO_ShortInterest (2007)	1979-2022	3.32	3.75	2.43	2.85	2.24	3.29	1.61	2.36	5.07	2.49	3.76	1.93
Mom6mJunk (2007)	1978-2017	0.98	3.35	0.98	3.35	1.31	3.91	1.31	3.91	-0.09	-0.15	-0.09	-0.15
OrderBacklogChg (2007)	1972-2022	0.35	2.59	0.14	0.9	0.34	2.23	0.22	1.12	0.37	1.33	-0.05	-0.2
roaq (2010)	1966-2022	1.2	5.44	0.59	3.09	1.27	5.04	0.66	3.09	0.95	2.05	0.3	0.73
MaxRet (2011)	1926-2022	0.63	2.93	0.44	2.3	0.69	2.97	0.49	2.38	0.13	0.24	0.02	0.04
ReturnSkew(2015)	1926-2022	0.49	7.65	0.28	5.34	0.52	7.9	0.3	5.5	0.08	0.31	0.02	0.1
CBOperProf (2016)	1962-2022	0.5	3.58	0.45	3.37	0.49	3.43	0.43	3.16	0.69	1.07	0.72	1.18
OperProfRD (2016)	1963-2022	0.41	2.47	0.28	1.82	0.36	2.12	0.21	1.34	0.9	1.38	0.98	1.59
Size (1981)	1926-2022	0.37	3.11	0.14	1.87	0.53	3.09	0.26	2.48	0.13	0.91	-0.03	-0.33
SP (1996)	1951-2022	0.78	5.7	0.54	3.76	0.74	5.4	0.48	4.04	0.84	2.88	0.65	1.9
ConsRecomm (2001)	1993-2022	0.5	2.49	0.49	2.45	1.29	2.6	1.28	2.56	0.18	0.92	0.18	0.9
DownRecomm (2001)	1993-2022	0.37	6.21	0.29	5.35	0.75	5.84	0.55	4.24	0.22	3.43	0.18	3.44
UpRecomm (2001)	1993-2022	0.34	5.45	0.25	4.67	0.77	7.15	0.59	5.66	0.17	2.3	0.12	1.92
FirmAge (1984)	1928-2022	-0.09	-1.24	-0.05	-0.8	-0.06	-0.77	-0.02	-0.26	-0.12	-0.99	-0.1	-0.86
ChForecastAccrual (2004)	1976-2021	0.22	4.07	0.17	3.19	0.3	4.25	0.29	4.36	0.1	1.12	-0.02	-0.24
AccrualsBM (2004)	1964-2022	1.27	6.52	0.69	2.92	1.42	5.89	0.88	3.05	0.94	2.84	0.24	0.61
EP (1977)	1951-2022	0.31	3.17	0.31	3.22	0.48	3.13	0.46	2.85	0.2	1.62	0.23	1.84

Continued on next page

Table 4: (continued)

Acronym	Sample	Full Sample				Pre-Publication				Post-Publication			
		All Stocks		No Microcaps		All Stocks		No Microcaps		All Stocks		No Microcaps	
		Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat
hire (2014)	1965-2022	0.47	5.78	0.25	3.02	0.5	5.88	0.25	2.96	0.28	1.05	0.24	0.82
InvGrowth (2012)	1952-2021	0.66	7.35	0.39	4.66	0.73	7.55	0.45	4.93	0.19	0.79	-0.01	-0.04
BrandInvest (2014)	1965-2022	0.41	1.82	-0.01	-0.07	0.43	1.76	-0.06	-0.32	0.26	0.48	0.3	0.72
Leverage (1988)	1951-2022	0.41	2.89	0.34	2.23	0.49	4.06	0.32	2.43	0.31	1.17	0.36	1.27
ResidualMomentum (2011)	1930-2022	0.85	8.19	0.77	7.04	0.93	8.36	0.87	7.37	0.22	0.79	0.01	0.05
Price (1973)	1926-2022	0.76	3	0.76	3	1.03	2.53	1.03	2.53	0.48	1.62	0.48	1.62
NetPayoutYield (2007)	1964-2022	0.83	4.8	0.59	3.83	0.73	3.93	0.55	2.97	1.11	2.76	0.74	2.62
PayoutYield (2007)	1964-2022	0.26	2.14	0.23	1.71	0.27	2.02	0.23	1.44	0.21	0.8	0.22	0.96
NetDebtFinance (2006)	1972-2022	0.67	9.26	0.32	4.55	0.79	8.97	0.43	5.02	0.38	3.15	0.09	0.69
NetEquityFinance (2006)	1972-2022	0.94	5.45	0.75	5.45	1.02	5.05	0.79	4.59	0.76	2.3	0.65	2.95
DolVol (1998)	1926-2022	0.72	4.04	0.19	1.71	0.79	3.53	0.21	1.63	0.49	2.22	0.1	0.53
FEPS (2006)	1976-2021	0.79	3.09	0.64	2.45	0.98	2.94	0.83	2.44	0.41	1.08	0.23	0.63
MomRev (2006)	1929-2022	0.68	3.88	0.58	3	0.77	3.92	0.66	3.09	0.25	0.64	0.17	0.38
AnnouncementReturn (1996)	1971-2021	1.08	13.88	0.78	9.9	1.15	11.49	0.86	8.34	1.01	8.33	0.7	5.77
REV6 (1996)	1976-2021	0.65	2.82	0.28	1.43	1.28	6.3	0.95	4.94	0.14	0.36	-0.27	-0.86
AdExp (2001)	1955-2022	0.5	2.8	0.2	1.19	0.51	2.1	0.12	0.52	0.48	2.02	0.34	1.56
RD (2001)	1951-2022	0.85	5.21	0.5	3.62	1	5.62	0.63	3.9	0.46	1.31	0.19	0.71
CashProd (2009)	1951-2022	0.34	2.88	0.14	1.1	0.37	2.85	0.15	1.09	0.2	0.71	0.07	0.25
DelBreadth (2002)	1980-2022	0.58	3.03	0.58	3.03	0.86	2.86	0.86	2.86	0.26	1.17	0.26	1.17
VolSD (2001)	1928-2022	0.27	2.21	0.11	0.95	0.35	2.52	0.15	1.15	-0.01	-0.05	-0.04	-0.18
std_turn (2001)	1928-2022	0.46	2.68	0.56	1.41	0.62	3.65	0.95	1.86	-0.13	-0.25	-0.54	-1.15
CustomerMomentum (2008)	1977-2022	0.79	2.67	0.75	2.52	0.94	3.07	0.86	2.8	0.44	0.64	0.48	0.69
retConglomerate (2012)	1976-2019	1.17	6.76	1.06	5.74	1.22	6.52	1.09	5.45	0.78	1.79	0.83	1.82
RDAbility (2013)	1957-2022	0.06	0.47	0.05	0.33	0.06	0.38	0.05	0.32	0.12	0.32	0.03	0.07
AssetGrowth (2008)	1952-2022	0.94	7.53	0.58	5.82	0.98	7.1	0.58	5.28	0.75	2.62	0.54	2.45
Activism1 (2005)	1990-2007	0.14	0.77	0.16	0.84	0.14	0.72	0.16	0.79	0.15	0.32	0.17	0.38
Activism2 (2005)	1990-2007	0.59	1.79	0.5	1.52	0.62	1.79	0.56	1.6	0.16	0.16	-0.21	-0.2
Spinoff (1993)	1926-2022	0.27	2.23	0.26	2.13	0.33	2.2	0.33	2.09	0.11	0.59	0.11	0.59
EarningsForecastDisparity (2011)	1982-2021	0.44	3.63	0.32	2.53	0.56	3.91	0.4	2.68	0.1	0.42	0.07	0.32
CompEquIss (2006)	1931-2022	0.3	2.94	0.29	2.76	0.3	2.49	0.3	2.49	0.34	1.91	0.25	1.27
IntanBM (2006)	1966-2022	0.36	2.56	0.35	2.29	0.46	2.82	0.44	2.58	0.11	0.39	0.09	0.28
IntanCFP (2006)	1956-2022	0.34	2.81	0.28	2.16	0.34	2.62	0.28	2	0.32	1.13	0.27	0.88
IntanEP (2006)	1956-2022	0.3	3.09	0.28	2.64	0.33	3.06	0.3	2.55	0.21	0.93	0.22	0.9
IntanSP (2006)	1956-2022	0.44	2.89	0.27	2.34	0.47	2.83	0.28	2.38	0.33	0.94	0.24	0.77
ShareIss5Y (2006)	1931-2022	0.41	5.53	0.28	4.55	0.43	5.04	0.28	4.01	0.31	2.43	0.27	2.38
ShareVol (1998)	1926-2022	0.49	3.41	0.27	1.87	0.58	3.31	0.37	2.04	0.19	0.92	-0.03	-0.13
Lrreversal (1985)	1929-2022	0.68	3.76	0.33	2.49	0.7	2.81	0.33	1.77	0.65	2.58	0.33	1.87

Continued on next page

Table 4: (continued)

Acronym	Sample	Full Sample				Pre-Publication				Post-Publication			
		All Stocks		No Microcaps		All Stocks		No Microcaps		All Stocks		No Microcaps	
		Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat
Mrreversal (1985)	1927-2022	0.45	3.41	0.28	2.57	0.46	2.48	0.33	2.1	0.42	2.56	0.2	1.55
ShortInterest (2001)	1973-2022	0.84	6.25	0.5	3.87	0.73	5.33	0.55	3.77	0.99	3.81	0.42	1.81
EquityDuration (2004)	1963-2022	0.44	2.77	0.31	2.05	0.67	3.42	0.56	2.99	-0.14	-0.55	-0.29	-1.21
cfp (2004)	1964-2022	0.37	2.22	0.44	2.84	0.26	1.27	0.42	2.16	0.64	2.17	0.49	1.96
ExchSwitch (1995)	1926-2022	0.66	5.18	0.48	3.49	0.5	3.61	0.33	2.02	0.86	3.76	0.66	2.89
OScore 91998	1972-2022	0.7	3.46	0.59	2.44	0.64	3.27	0.79	2.94	0.77	2.06	0.37	0.87
CredRatDG (2001)	1979-2017	0.68	3.13	0.06	0.32	0.89	3.58	0.33	1.61	0.46	1.27	-0.24	-0.79
ForecastDispersion (2002)	1976-2021	0.43	2.25	0.24	1.29	0.67	2.73	0.44	1.8	0.08	0.28	-0.04	-0.15
ExclExp (2003)	1983-2022	0.1	0.92	0.16	2.01	0.17	0.86	0.29	2.34	0.03	0.33	0.01	0.1
SurpriseRD (2004)	1952-2022	0.06	1.05	0.05	0.78	0.08	1.13	0.06	0.77	0	0.04	0.02	0.2
sfe (2001)	1976-2022	0.56	2.08	0.67	2.57	0.52	1.41	0.73	1.91	0.61	1.56	0.6	1.75
GrLTNOA (2003)	1952-2022	0.19	3.09	0.13	1.99	0.26	3.46	0.2	2.41	-0.01	-0.15	-0.05	-0.43
AM (1992)	1951-2022	0.51	3.57	0.3	1.99	0.53	3.82	0.26	1.83	0.48	1.69	0.37	1.19
Bmdec (1992)	1952-2022	0.72	6.43	0.42	3.54	0.8	5.68	0.56	4.04	0.62	3.34	0.22	1.08
BookLeverage (1992)	1951-2022	0.08	0.73	-0.13	-1.15	0.09	1.22	-0.05	-0.6	0.07	0.28	-0.24	-0.99
OperProf (2006)	1963-2022	0.44	3.42	0.34	2.75	0.46	2.85	0.35	2.23	0.38	2.07	0.33	1.82
Beta (1973)	1928-2022	0.35	1.52	0.18	0.86	0.4	1.14	0.16	0.5	0.3	1	0.21	0.72
EarningsSurprise (1984)	1963-2022	0.73	9.88	0.44	5.79	1.23	9.24	0.99	7.44	0.45	5.22	0.13	1.42
AOP (1998)	1976-2022	0.16	1.45	0.11	0.96	0.35	2.27	0.18	1.06	-0.02	-0.15	0.05	0.29
AnalystValue (1998)	1976-2022	0.27	1.53	0.35	1.94	0.2	1.45	0.3	2.01	0.33	1.04	0.39	1.21
PredictedFE (1998)	1983-2022	0.01	0.05	-0.02	-0.09	0.21	0.81	0.18	0.69	-0.13	-0.51	-0.15	-0.6
FR (2006)	1987-2022	-0.3	-1.65	-0.24	-1.59	-0.07	-0.33	-0.13	-0.73	-0.61	-1.84	-0.38	-1.49
BetaFP (2014)	1929-2021	0.06	0.21	0.06	0.21	0.06	0.18	0	-0.01	0.1	0.11	0.8	0.86
High52 (2004)	1926-2022	0.04	0.19	0.27	1.67	0.03	0.13	0.35	1.89	0.07	0.22	-0.1	-0.35
Governance (2003)	1990-2007	-0.13	-0.58	-0.13	-0.59	-0.03	-0.11	-0.03	-0.12	-0.54	-1.73	-0.54	-1.74
RDIPO (2006)	1951-2022	0.73	3.41	0.35	1.57	0.74	2.66	0.33	1.13	0.69	2.21	0.4	1.15
IndMom (1999)	1926-2022	0.37	3.92	0.26	2.58	0.38	4.05	0.27	2.63	0.35	1.32	0.22	0.83
PctAcc (2011)	1964-2022	0.42	4.99	0.26	2.68	0.49	4.9	0.3	2.64	0.14	1.05	0.08	0.5
PctTotAcc (2011)	1988-2022	0.39	4.81	0.31	3.6	0.38	3.78	0.3	2.74	0.42	3.02	0.33	2.51
tang (tang)	1951-2022	0.32	2.82	0.14	1.2	0.42	3.55	0.17	1.33	-0.16	-0.5	0.01	0.03
DivSeason (2013)	1926-2021	0.3	14.46	0.25	11.64	0.33	14.64	0.27	12.03	0.04	0.87	-0.03	-0.57
Coskewness (2000)	1927-2022	0.06	0.73	0.05	0.61	0.03	0.32	0.01	0.1	0.17	0.99	0.19	1.18
RoE (1996)	1961-2022	0.26	2.6	0.22	2.18	0.18	1.81	0.14	1.33	0.37	1.9	0.35	1.73
VarCF (1996)	1953-2022	-0.48	-2.8	-0.32	-2.73	-0.59	-3.33	-0.37	-2.95	-0.29	-0.83	-0.22	-0.96
VolMkt (1996)	1927-2022	0.29	1.9	0.18	1.19	0.25	1.41	0.18	1.06	0.4	1.33	0.17	0.55
VolumeTrend (1996)	1928-2022	0.56	5.37	0.3	3.67	0.52	3.94	0.28	2.84	0.66	4.62	0.35	2.49
AnalystRevision (1984)	1976-2021	0.65	8.9	0.6	7.53	1.06	7.79	1.08	7.66	0.56	6.6	0.48	5.26

Continued on next page

Table 4: (continued)

Acronym	Sample	Full Sample				Pre-Publication				Post-Publication			
		All Stocks		No Microcaps		All Stocks		No Microcaps		All Stocks		No Microcaps	
		Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat
Mom12mOffSeason (2008)	1926-2021	0.89	3.94	0.91	4.55	0.97	4.13	1.01	4.92	0.3	0.42	0.19	0.28
MomOffSeason (2008)	1927-2022	1.04	5.97	0.69	4.89	1.11	5.93	0.75	5.15	0.65	1.32	0.33	0.7
MomOffSeason06YrPlus (2008)	1931-2021	0.67	6.14	0.37	4.36	0.63	5.36	0.34	3.81	0.91	3.16	0.57	2.14
MomOffSeason11YrPlus (2008)	1936-2022	0.29	2.99	0.2	2.75	0.31	2.95	0.21	2.78	0.18	0.73	0.15	0.64
MomOffSeason16YrPlus (2008)	1943-2022	0.28	3.19	0.15	1.85	0.32	3.35	0.18	2.09	0.09	0.4	-0.02	-0.12
MomSeasonShort (2008)	1927-2022	0.83	6.08	0.6	5.16	0.99	6.62	0.76	6.08	-0.13	-0.41	-0.35	-1.08
CitationsRD (2013)	1977-2022	0.18	2.37	0.17	2.11	0.17	2.06	0.16	1.79	0.25	1.17	0.24	1.12
NOA (2004)	1963-2022	0.81	6.59	0.53	4.09	1.09	7.72	0.75	4.72	0.12	0.51	0.01	0.03
dNoa (2004)	1962-2022	0.77	8.57	0.5	6.03	0.99	8.61	0.67	6.31	0.24	1.92	0.1	0.83
sinAlgo (2009)	1926-2021	0.24	2.03	0.12	1.02	0.23	1.9	0.1	0.78	0.32	0.8	0.23	0.72
PriceDelayRsq (2005)	1927-2021	0.54	3.51	-0.1	-0.77	0.62	3.52	-0.03	-0.2	0.12	0.47	-0.43	-2.07
PriceDelaySlope (2005)	1927-2021	0.2	2.52	-0.02	-0.26	0.24	2.64	0.01	0.11	0.05	0.25	-0.16	-0.88
PriceDelayTstat (2005)	1927-2021	0	-0.02	-0.05	-0.77	0.02	0.29	-0.02	-0.3	-0.11	-0.95	-0.18	-1.46
Herf (2006)	1951-2021	0.07	1.1	-0.02	-0.19	0.08	1.18	-0.02	-0.16	0.03	0.17	-0.01	-0.09
HerfAsset (2006)	1951-2022	0	0.04	-0.02	-0.17	0.08	0.89	0.06	0.56	-0.26	-1.47	-0.29	-1.89
HerfBE (2006)	1951-2022	0.05	0.62	-0.01	-0.12	0.13	1.5	0.08	0.76	-0.24	-1.43	-0.32	-2.04
ShareRepurchase (1995)	1951-2022	0.2	2.86	0.22	3.7	0.19	3.18	0.21	4.17	0.2	1.68	0.22	2.17
STreversal (1990)	1926-2022	2.76	13.68	1.04	6.62	3.39	14.93	1.52	8.78	1.45	3.66	0.06	0.2
RevenueSurprise (2006)	1963-2022	0.61	8.62	0.24	3.6	0.69	8.63	0.29	3.92	0.36	2.51	0.11	0.72
Mom12m (1993)	1927-2022	0.83	3.46	0.98	4.74	0.93	3.38	1.13	4.79	0.58	1.24	0.61	1.48
Mom6m (1993)	1926-2022	0.61	2.76	0.79	4.04	0.48	1.93	0.66	3.01	0.92	2.02	1.1	2.71
ChangeInRecommendation (2004)	1993-2022	0.57	7.4	0.38	5.48	1.03	8.73	0.75	6.55	0.27	2.83	0.14	1.72
OptionVolume1 (2012)	1996-2022	0.48	2.21	0.4	1.87	0.7	2.38	0.6	2.11	0.07	0.23	0.03	0.1
OptionVolume2 (2012)	1996-2022	0.29	1.59	0.25	2	0.52	1.95	0.44	2.42	-0.14	-0.92	-0.09	-0.76
BetaTailRisk (2014)	1932-2022	0.36	2.74	0.33	2.41	0.38	2.78	0.35	2.45	0.08	0.18	0.07	0.15
fgr5yrLag (1996)	1982-2022	0.19	0.8	0.22	0.96	0.41	1.4	0.36	1.14	0.06	0.17	0.15	0.46
CF (1994)	1951-2022	0.36	2.42	0.46	3.64	0.45	2.92	0.56	3.7	0.2	0.7	0.3	1.36
MeanRankRevGrowth (1994)	1957-2022	0.22	2.96	0.15	2.01	0.39	3.95	0.29	2.97	-0.03	-0.28	-0.05	-0.47
RDS (2011)	1973-2022	0.26	2.59	0.18	2.12	0.33	2.74	0.2	2.02	0.02	0.13	0.08	0.62
MomVol (2000)	1928-2022	1.06	4.16	0.93	3.7	1.13	4.3	1.03	3.93	0.79	1.18	0.57	0.88
Tax (2004)	1951-2022	0.36	5.34	0.33	4.53	0.35	4.38	0.33	3.81	0.38	3.19	0.31	2.56
DivYieldST (1979)	1926-2021	0.57	9.34	0.41	6.6	0.48	5.74	0.44	5.18	0.68	7.75	0.37	4.09
zerotrade (2006)	1926-2022	0.5	3.1	0.23	1.51	0.58	3.42	0.31	1.99	0.02	0.03	-0.3	-0.62
zerotradeAlt1 (2006)	1927-2022	0.54	3.45	0.24	1.61	0.61	3.81	0.3	1.97	0.06	0.13	-0.16	-0.36
zerotradeAlt12 (2006)	1927-2022	0.42	3.27	0.16	1.41	0.5	3.83	0.24	2.02	-0.14	-0.33	-0.34	-0.91
ChEQ (2010)	1962-2022	0.5	4.42	0.35	3.71	0.52	4	0.35	3.25	0.42	1.9	0.33	1.88
EarningsStreak (2012)	1984-2022	0.96	10.56	0.73	8.61	1.04	9.97	0.81	8.44	0.7	3.88	0.47	2.7

Continued on next page

Table 4: (continued)

Acronym	Sample	Full Sample				Pre-Publication				Post-Publication			
		All Stocks		No Microcaps		All Stocks		No Microcaps		All Stocks		No Microcaps	
		Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat
NumEarnIncrease (2012)	1964-2022	0.47	9.71	0.28	5.41	0.5	9.95	0.31	5.58	0.32	2.13	0.17	1.06
GrAdExp (2014)	1965-2022	0.27	1.7	0.32	1.87	0.29	1.68	0.36	1.88	0.07	0.28	0.03	0.13
EntMult (2011)	1951-2022	0.69	5.79	0.51	3.42	0.75	6	0.57	3.6	0.36	0.97	0.13	0.32
CompositeDebtIssuance (2008)	1956-2022	0.24	5.49	0.15	3.34	0.25	5.12	0.18	3.76	0.19	2	0	0.02
InvestPPEInv (2008)	1952-2022	0.54	8.64	0.32	5.24	0.6	8.31	0.37	5.35	0.3	2.5	0.1	0.79
iomom_cust (2010)	1986-2022	0.54	2.9	0.45	2.37	0.69	2.92	0.54	2.14	0.2	0.7	0.27	1.03
iomom_supp (2010)	1986-2022	0.52	2.64	0.42	1.97	0.51	2.27	0.36	1.41	0.54	1.37	0.54	1.42
DivInit (1995)	1926-2022	0.38	3.4	0.4	3.61	0.46	3.46	0.47	3.63	0.16	0.79	0.18	0.89
DivOmit (1995)	1926-2022	0.62	4.08	0.23	1.51	0.59	3.43	0.14	0.79	0.72	2.23	0.48	1.47
MS (2005)	1962-2022	0.96	4.72	0.6	3.93	1.2	4.16	0.61	2.92	0.53	2.27	0.57	2.91
RIO_Disp (2005)	1976-2021	0.61	3.4	0.54	2.97	0.72	2.89	0.63	2.48	0.42	1.83	0.38	1.66
RIO_MB (2005)	1962-2022	0.65	3.96	0.63	2.65	0.85	3.96	0.87	2.63	0.13	0.69	0.08	0.41
RIO_Turnover (2005)	1926-2022	0.3	2.37	0.48	2.13	0.28	1.93	0.49	1.85	0.42	1.67	0.4	1.62
RIO_Volatility (2005)	1926-2022	0.49	3.22	0.71	3.16	0.41	2.38	0.69	2.61	0.9	2.79	0.83	2.55
OPLeverage (2011)	1951-2022	0.38	3.59	0.3	3.35	0.41	3.6	0.28	2.85	0.18	0.68	0.41	2.02
IntMom (2012)	1927-2022	1.12	5.71	0.81	4.53	1.24	6.03	0.94	4.98	0.01	0.01	-0.38	-0.69
GP (2013)	1951-2022	0.39	3.68	0.39	3.81	0.33	3.16	0.33	3.05	0.79	1.9	0.85	2.71
Cash (2012)	1971-2021	0.66	3.2	0.33	1.41	0.64	2.87	0.22	0.86	0.72	1.41	0.85	1.66
BetaLiquidityPS (2003)	1966-2020	0.28	2.11	0.23	1.76	0.37	2.26	0.28	1.82	0.07	0.31	0.09	0.4
BPEBM (2007)	1962-2022	0.13	2.14	0.15	2.09	0.18	2.59	0.18	2.09	-0.02	-0.16	0.08	0.53
EBM (2007)	1962-2022	0.23	3.82	0.2	2.79	0.24	3.62	0.22	2.75	0.2	1.43	0.13	0.84
NetDebtPrice (2007)	1963-2022	0.46	3.32	0.11	0.94	0.55	3.91	0.22	1.77	0.19	0.52	-0.23	-0.79
PS (2000)	1972-2022	0.7	2.54	0.77	2.68	0.66	2.4	0.84	3.04	0.76	1.41	0.67	1.19
ShareIss1Y (2008)	1927-2022	0.44	6.44	0.3	4.92	0.42	6	0.28	4.38	0.61	2.47	0.44	2.24
DelDRC (2013)	2000-2022	0.58	1.41	0.09	0.31	0.85	1.3	0.02	0.05	0.14	0.68	0.2	0.92
OrderBacklog (2003)	1971-2022	0.06	0.65	0	-0.03	0.2	1.57	0.01	0.09	-0.17	-1.13	-0.03	-0.19
DelCOA (2005)	1952-2022	0.42	6.65	0.25	3.24	0.49	6.61	0.32	3.48	0.18	1.6	0.02	0.13
DelCOL (2005)	1952-2022	0.23	3.78	0.09	1.27	0.27	3.87	0.12	1.33	0.09	0.76	0.03	0.19
DelEqu (2005)	1962-2022	0.51	4.31	0.35	3.54	0.58	3.97	0.38	3.08	0.33	1.69	0.27	1.76
DelFINL (2005)	1952-2022	0.51	11.24	0.32	7.4	0.6	11.84	0.39	8.15	0.19	2.05	0.1	1.01
DelLTI (2005)	1953-2022	0.17	3.72	0.12	2.97	0.19	3.61	0.13	2.74	0.07	1.01	0.09	1.16
DelNetFin (2005)	1952-2022	0.33	6.59	0.22	4.85	0.4	7.41	0.27	5.29	0.07	0.62	0.06	0.62
TotalAccruals (2005)	1952-2022	0.3	3.55	0.26	4.38	0.36	3.74	0.28	4.08	0.1	0.58	0.19	1.65
AgeIPO (1991)	1980-2021	0.47	1.97	0.27	1.09	0.93	3.01	0.01	0.02	0.29	0.94	0.37	1.21
IndIPO (1991)	1926-2022	0.43	2.91	0.22	1.3	0.48	2.2	-0.06	-0.22	0.4	2.06	0.37	1.8
BM (1985)	1961-2022	0.96	5.4	0.28	1.63	0.77	3.03	0.31	1.35	1.09	4.47	0.26	1.07
ChNAnalyst (2008)	1976-2021	0.7	1.29	-5.45	-1.29	0.66	1.92	-5.45	-1.29	0.91	0.34		

Continued on next page

Table 4: (continued)

Acronym	Sample	Full Sample				Pre-Publication				Post-Publication			
		All Stocks		No Microcaps		All Stocks		No Microcaps		All Stocks		No Microcaps	
		Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat	Mean Ret	t-Stat
Accruals (1996)	1952-2022	0.41	6.37	0.39	5.59	0.55	6.78	0.47	5.37	0.17	1.57	0.26	2.2
ChAssetTurnover (2008)	1953-2022	0.16	3.86	0.16	3.44	0.21	4.4	0.2	3.73	-0.03	-0.38	-0.01	-0.07
ChNNCOA (2008)	1952-2022	0.23	5.39	0.22	5.19	0.28	6.13	0.25	5.27	-0.02	-0.19	0.07	0.83
ChNWC (2008)	1952-2022	0.16	4.48	0.13	3.57	0.17	4.34	0.14	3.55	0.09	1.25	0.06	0.8
DebtIssuance (1999)	1960-2022	0.26	4.75	0.09	1.41	0.35	6.29	0.15	2.3	0.16	1.53	0.01	0.11
ChInv (2002)	1952-2022	0.61	8.31	0.41	5.7	0.73	8.03	0.54	6.12	0.31	2.6	0.06	0.55
ChTax (2011)	1962-2022	0.97	10.15	0.58	5.47	1.1	10.1	0.65	5.37	0.33	1.95	0.22	1.19
Investment (2004)	1953-2022	0.16	1.71	0.11	1.26	0.19	1.81	0.17	1.64	0.05	0.28	-0.07	-0.44
realestate (2010)	1970-2022	0.23	1.88	0.17	1.55	0.24	1.73	0.18	1.39	0.16	0.73	0.14	0.69
ConvDebt (2016)	1951-2022	0.24	4.38	0.08	1.39	0.24	4.26	0.08	1.46	0.2	1	0	-0.02
AbnormalAccruals (2001)	1972-2022	0.16	1.87	0.28	3.38	0.29	2.85	0.4	3.91	-0.02	-0.15	0.09	0.7
skew1 (2010)	1996-2022	0.42	3.43	0.41	3.4	0.58	3.3	0.59	3.34	0.2	1.26	0.18	1.15
SmileSlope (2011)	1996-2022	1.24	10.6	1.03	9.38	1.53	9.58	1.33	8.78	0.78	4.95	0.57	3.93
FirmAgeMom (2006)	1926-2022	1.02	6.16	0.83	4.75	0.99	5.42	0.78	4.07	1.19	3	1.1	2.59

Table 5: **Effects of Microcaps and Publications on Predictors.**

This table illustrates the performance and its change in percentages associated with microcaps and publications across sample periods of predictors. Panel A presents the average returns (% monthly) and t-statistics of the long-short portfolios containing all stocks and with microcaps removed. A large and consistent performance drop is found in anomalies across sample periods after controlling for microcaps. Panel B shows the post-publication performance decay in the same metrics.

Panel A: Effects of Microcaps

	All Stocks		No Microcaps		Performance Change (%) Relative to All Stocks	
	Return	t-Stat	Return	t-Stat	Return	t-Stat
Full-Sample	0.49	3.93	0.31	2.81	-36.13	-28.58
In-Sample	0.66	3.82	0.49	2.75	-25.14	-27.92
Pre-Pub	0.56	3.93	0.37	2.86	-33.29	-27.28
Post-Pub	0.31	1.23	0.19	0.79	-37.22	-35.83

Panel B: Effects of Publications

	All Stocks		No Microcaps	
	Return	t-Stat	Return	t-Stat
Performance Change (%) Relative to In-Sample				
Post-Pub	-52.97	-67.92	-60.55	-71.44
Full-Sample	-26.46	3.00	-37.25	2.05
Performance Change (%) Relative to Pre-Publication				
Post-Pub	-44.76	-68.84	-48.02	-72.51
Full-Sample	-13.63	0.04	-17.32	-1.76

Table 6: **Survival Rate and Predictor Performance at a Higher t-Statistic Threshold.**

This table compares the survival rate and performance of predictors before and after imposing a higher t-statistic threshold of 3.0 (Harvey et al., 2016). Panel A reports the percentages of predictors that surpass the conventional t-statistic threshold of 1.96 and the novel threshold of 3.0 across sample periods. The proportion of robust anomalies largely decreases after imposing the higher t-statistic threshold. Panel B presents the descriptive statistics for the performance of predictors that are significant at the t-statistic thresholds of 1.96 and 3.0. Predictors are assigned into two groups based on the reproduced in-sample t-statistics: (1) above 1.96 and below 3.0, (2) above 3.0. Within each group, predictor performance is assessed in portfolios containing all stocks and without microcaps, using in-sample and post-publication samples. SD represents standard deviation. Range consists of minimum and maximum values.

Panel A: Survival Rate (%)

	$t \geq 1.96$		$t \geq 3.0$	
	All Stocks	No Microcaps	All Stocks	No Microcaps
Full-Sample	82.29	65.63	58.85	42.19
In-Sample	87.50	63.02	57.81	38.54
Pre-Pub	80.21	66.67	58.85	38.54
Post-Pub	25.00	13.02	9.38	2.60

Panel B: Predictor Performance

	$1.96 \leq t < 3$		$t \geq 3$	
	In-Sample	Post-Pub	In-Sample	Post-Pub
		All Stock		
Return				
Mean	0.47	0.26	0.84	0.38
SD	0.24	0.29	0.46	0.56
Range	[0.16, 1.40]	[-0.54, 1.11]	[0.28, 2.95]	[-0.17, 5.07]
t-Stat				
Mean	2.52	1.04	5.07	1.58
SD	0.29	1.08	2.33	1.62
Range	[2.00, 2.99]	[-1.73, 4.62]	[3.02, 14.53]	[-1.13, 8.33]
		No Microcaps		
Return				
Mean	0.43	0.23	0.73	0.24
SD	0.29	0.60	0.38	0.29
Range	[0.15, 1.73]	[-0.54, 3.76]	[0.21, 1.72]	[0.38, 1.10]
t-Stat				
Mean	2.38	0.75	4.47	1.19
SD	0.31	1.10	1.76	1.33
Range	[1.96, 2.99]	[-1.74, 2.71]	[3.00, 12.08]	[-1.08, 5.77]