Do Information Acquisition Costs Matter?
The Effect of SEC EDGAR on Stock Anomalies

Yong-Hyuck Kim†

First Draft: February 22, 2021
This Version: September 10, 2021

Abstract

I estimate the costs of information acquisition and the extent to which they explain stock anomaly returns. The SEC's staggered implementation of EDGAR from 1993 to 1996 greatly lowered the costs of acquiring accounting information. I study how this quasi-exogenous and staggered shock affects the profitability of 126 accounting and 108 non-accounting anomalies. The EDGAR introduction lowers the average alphas for the accounting anomalies by 4.0% per year, explaining more than half of the pre-EDGAR alphas. The attenuation is stronger for the accounting anomaly portfolios that require more up-to-date accounting information and those consisting of EDGAR filer stocks with less information available in the pre-EDGAR period. By contrast, alphas for the non-accounting anomalies remain unaffected. These results imply that the information acquisition costs, which are usually neglected, can be as important as the transaction or short sale costs.

Keywords: Information acquisition costs, stock anomalies, EDGAR, limits-to-arbitrage

---

† Michigan State University, Department of Finance. Email: kimyon32@broad.msu.edu
I would like to thank Zoran Ivkovich, Dmitriy Muravyev, Charles Hadlock, Andrei Simonov, Scott Weisbenner, and the participants in the MSU Finance Department Brown Bag Seminar for their valuable comments and discussion.

1. Introduction

Do costs of acquiring information matter? Many influential theories (e.g., Grossman and Stiglitz (1980), Verrecchia (1982)) argue that costly information acquisition affects investor decisions and market outcomes. In practice, a typical hedge fund spent over $1 million on data subscriptions in 2019 (Whyte (2020)), even though recent advances in technology made information cheaper to gather and disseminate. Yet, most prior studies assume that the information acquisition costs are negligible and instead focus on the transaction or short sale costs. In fact, hardly any studies attempt to undertake the empirically challenging task of quantifying the costs of acquiring information. This study fills the gap in the literature by presenting an estimate of information acquisition costs in the U.S. equity market from a clean-cut causal effect and showing that they can be as important as the transaction or short sale costs.

To estimate the costs of acquiring information, I examine the causal effect of costly information constraints on a comprehensive set of 234 anomaly portfolio returns by studying the SEC's staggered implementation of the Electronic Data Gathering, Analysis and Retrieval (EDGAR) system from February 1993 to May 1996. This exogenous shock substantially lowered the information acquisition costs to investors. Therefore, studying the staggered implementation of EDGAR allows me to establish causality, addressing the concern that the costs of acquiring information are endogenous (Grossman and Stiglitz (1980)). On the other hand, examining a comprehensive set of anomaly portfolio returns helps capture the full ramification of EDGAR's information acquisition cost-saving effect on the anomaly returns.[1] Prior literature documents that limits-to-arbitrage, such as noise trader risk, transaction costs, and short sale costs, partially explain anomaly returns, but hardly any study examines the effect of information acquisition costs *per se* on anomaly returns.[2] However, investors incur the costs of acquiring information even before they pay transaction or short sale costs because investors need first to identify which stocks to buy or

[1] Investors need to collect a complete set of financial information, on both accounting and non-accounting information, for the entire cross-section of stocks to identify which stocks to buy or sell in order to trade anomaly portfolios. Therefore, analyzing 234 anomaly portfolio returns allows to capture almost all the costs of acquiring public information that investors bear in the market.

[2] For example, De Long, Shleifer, Summers and Waldmann (1990) argue that noise trader risk can explain several asset pricing anomalies, Novy-Marx and Velikov (2015) study how trading costs affect the profitability of anomalies, and Chu, Hirshleifer, and Ma (2020) examine the causal effect of short sale constraints on 11 anomalies. Lam and Wei (2011) provide a thorough list of limits-to-arbitrage.

sell before they trade. I show that the information acquisition costs can be as important as the transaction or short sale costs.

EDGAR allows investors to access firms' *accounting* information for free, therefore lowers only the costs of acquiring accounting information. Accordingly, in a staggered difference-in-difference framework, I find that the Fama-French three-factor and five-factor alphas for the *accounting*-based anomaly portfolios decline on average by 32.6 basis points per month (4.0% per year) in response to the EDGAR introduction. The average Fama-French alphas of the accounting anomaly portfolios in the treatment group before the EDGAR implementation is 52.9 basis points per month. Therefore, 32.6 basis points drop accounts for nearly 61.6% of the pre-EDGAR alphas.[3] Figure 1 clearly shows how accounting anomalies sharply become much less profitable right after the EDGAR implementation. By contrast, EDGAR has not lowered other, *non*-accounting information costs. Accordingly, I find that the non-accounting anomaly alphas have not been affected by the EDGAR introduction.

My results have implications beyond the anomaly literature. The 4.0% per year attenuation translates to the costs of information acquisition that investors face in a market where accessing public information is not costless. EDGAR is free to use hence eliminates almost all the costs of acquiring accounting information for arbitrageurs. Therefore, that exact amount of information cost-saving effect allows an average of 4% higher profitability from trading against the accounting anomaly portfolios consisting of EDGAR filers. But these arbitrage opportunities are quickly exploited, and the prices adjust accordingly. Consequently, the profitability of the accounting anomaly portfolios attenuates by 4%. Therefore, the estimated 4% attenuation of profitability translates to the costs of acquiring accounting information that investors face in the absence of the

[3] Please see Figure 1 and Section 4.A for more detailed discussion. 32.6 basis-point is the average staggered difference-in-difference coefficient across all accounting-based anomaly portfolio specifications under analysis. 52.9 basis-point is the average monthly Fama-French alphas of treatment group (the accounting-based anomaly portfolios constructed with EDGAR filer stocks) in the pre-EDGAR period. 61.6% is given by 32.6 basis points divided by 52.9 basis points. This simple comparison is feasible because (1) the average alphas of the treatment group (52.9 basis points) and the control group (the accounting-based anomaly portfolios constructed with *non*-EDGAR filer stocks, 48.8 basis points) are very close to each other in the pre-EDGAR period, and (2) the average alphas of the control group remain almost unchanged even after the EDGAR is introduced (48.8 vs 51.1 basis points). Note that the accounting-based anomalies use accounting information to create signals, whereas the non-accounting anomalies rely on information other than accounting data to create signals.

EDGAR system. This paper is the first to document an unbiased estimation of investor information acquisition costs.

I examine three main hypotheses. My first hypothesis is that the profitability of the anomaly portfolios constructed from EDGAR filers becomes weaker relative to the profitability of the anomaly portfolios constructed from non-EDGAR filers. This attenuation should be observed only in accounting-based anomalies. The second hypothesis is that the profitability attenuation is more pronounced in the first month following the EDGAR implementation, coinciding with the period immediately following the implementation period (the period during which the novelty of EDGAR is at its peak). Lastly, my third hypothesis is that the accounting-based anomalies which rely on more recent information and the ones that have less information available before the EDGAR implementation weaken more. The intuition is that these types of anomalies attract more arbitrageurs upon EDGAR adoption because EDGAR's information acquisition cost-saving benefit is greater from trading these anomalies.

To examine the first hypothesis, I first construct a panel of long-short portfolio monthly returns for all anomalies and nine implementation stages. The anomaly portfolios consisting of EDGAR filers constitute the treatment group and those of non-EDGAR filers the control group in a series of staggered implementation phases. In other words, I create a panel data of anomaly returns by month by EDGAR phase for the treated and the controlled group. Note that I use a comprehensive set of anomalies documented in Chen and Zimmermann (2020), who replicate almost all the stock return anomalies that researchers have discovered to date. In addition, I show returns for four long-short portfolio specifications: equal-weighted and value-weighted returns for both decile and quintile portfolios. The presence of a series of treatment (EDGAR filer anomaly portfolios) and control (non-EDGAR filer anomaly portfolios) groups allows me to harness the staggered difference-in-difference panel regression framework. Later, the presence of non-accounting anomalies provides another reference group. Using the staggered difference-in-differences research design, I find that the anomaly portfolio profitability drops more for the treated portfolios (those constructed using EDGAR stocks) after the EDGAR adoption. I also find that this effect is limited to accounting-based anomalies. These findings support my first hypothesis.

For the second hypothesis, I introduce a separate indicator variable for the first month following each implementation phase to the baseline difference-in-difference specification discussed above. Repeating the difference-in-difference analysis after introducing the indicator variable for the first month of EDGAR implementation, I find that the accounting-based anomaly portfolio returns decline significantly more in the first month of EDGAR adoption than in the following months.

I examine the third hypothesis by introducing one measure of accounting anomaly portfolio's sensitivity to recent information and two measures of information availability before the EDGAR implementation. Using these measures, I test whether the accounting anomalies with higher sensitivity to recent information and less information availability in the pre-EDGAR period weaken more in response to EDGAR adoption. The first measure I use is the portfolio's turnover ratio, which measures the extent to which the anomaly portfolio replaces its stocks as it rebalances, to capture the anomaly's sensitivity to recent information. The intuition behind this is that if an anomaly portfolio is more sensitive to recent information, then the anomaly portfolio will be more active in consuming new information by replacing more stocks upon updating its signal. As a result, the anomaly will have a higher turnover ratio. Second, I use firm size and analyst coverage to measure the information availability prior to the EDGAR implementation. Therefore, my third hypothesis predicts that the attenuation should be more pronounced if an accounting-based anomaly portfolio has a higher turnover ratio or is constructed with smaller stocks or stocks with less analyst coverage in the pre-EDGAR period. And indeed, I find strong support for my third hypothesis as well.[4]

Finally, additional robustness tests further validate my findings. First, I show that the main results remain unchanged even after controlling for firm size, bid-ask spread, and industry effects. Second, I study how trading volume, liquidity, volatility, short interest, and institutional ownership respond to the EDGAR implementation to better understand why the anomaly profitability declines. Third, I show that if I split the accounting anomaly portfolios into the short and long legs,

[4] In other words, I find that the accounting-based anomalies with higher information turnover ratios in the pre-EDGAR period weaken more than those with lower information turnover ratios. Likewise, I document that the accounting-based anomalies with lower information availability weaken more than those with higher information availability in the pre-EDGAR period. Again, the rationale is that EDGAR's information cost-saving effect is larger for the accounting anomalies with higher sensitivity to recent information and less information availability in the pre-EDGAR period; therefore, these anomalies attract more arbitrageurs, resulting in a larger drop in alphas.

the alpha for the short leg drops the most. This attenuation of profitability in the short leg suggests that retail investors increased net buying of the EDGAR filer stocks as EDGAR rolled out and, as a result, squeezed the arbitrage profit on the short leg portfolio while prompting the institutional investors to cover their short position. Fourth, I address any potential selection bias issues regarding the first implementation phase. Lastly, I run the pre-trends and falsification tests.

This paper contributes to three strands of the literature. First, this paper contributes to the information costs and market outcomes literature. Merton (1987) and Shapiro (2002) study how costly information constraints compel investors to only trade the securities they possess adequate information and how the constraints affect the general equilibrium process and outcomes. Grossman and Stiglitz (1980) claim a perfect market efficiency is impossible to reach because information is costly to collect. Easley and O'Hara (2004) argue that private information presents a systematic risk because it cannot be diversified away, prompting uninformed investors to demand compensation for bearing it in equilibrium. Chen, Ma, Martin, and Michaely (2021) use the introduction of high-speed rail as an exogenous shock to the cost of acquiring information. The authors document that lower information acquisition costs increase information production. My paper contributes to this strand of literature by providing a clean-cut estimation of information acquisition costs in the context of the U.S. equity market.

This paper also contributes to the stock anomaly literature by identifying the causal effect of information constraints on anomaly returns. A few papers study how limits to arbitrage affect anomalies using exogenous shocks to address the endogeneity. For example, Chu, Hirshleifer, and Ma (2020) examine how shorting costs affect 11 anomalies using the Reg-SHO pilot program. However, these studies do not explore costly information constraints as limits to arbitrage on anomaly returns like my research does. One paper similar in spirit to my paper is McLean and Pontiff (2016). McLean and Pontiff (2016) document that the portfolio returns decline by 58% on average after publication. But my study focuses on information acquisition costs, whereas McLean and Pontiff (2016) focus on information processing costs (Blankespoor, deHaan, and Marinovic (2020)). I also confirm that the post-publication effect is not driving my results (Appendix Section B).

Finally, the last body of literature focuses on the effect of the EDGAR system on financial markets. Using the staggered EDGAR implementation, Gao and Huang (2020) explore whether corporate outsiders produce more information in response to the advances in information technology. Goldstein, Yang, and Zuo (2020) show that equity financing improves following the EDGAR adoption. Gomez (2020) studies the effect of EDGAR on the information asymmetry between managers and investors and among different investor groups. However, these studies do not investigate the anomaly portfolio profitability or attempt to estimate the costs of acquiring information. This paper is the first to study how the EDGAR implementation affects the anomaly portfolio returns and to estimate the costs of acquiring information investors bear in the absence of readily accessible accounting information.

2. Implementation of the EDGAR system

A. The Costs of Information Acquisition and EDGAR

The advances in information technology in the 1990s prompted the SEC to develop and introduce the EDGAR system. EDGAR allows the public firms to disclose their financial information electronically, and investors or any other information consumers to access filed corporate information instantaneously via the internet for free.

Prior to the EDGAR adoption, the costs of information acquisition were substantially larger. Investors were mostly limited to three options to access corporate filings. The first option was to physically visit one of the reference rooms in Washington DC, New York, or Chicago where the SEC kept the paper financial statements. The second option was to subscribe to the commercial data vendors' services such as the Compustat, the Value Line, the Mead Data Central's database (LEXIS/NEXIS), or the Dialog. Lastly, investors could ask the companies to mail their filing documents to them.

Anecdotal evidence confirms that the first option was costly. Investors had to be physically present in one of the SEC's reference rooms and make a painstaking effort to acquire information on the corporate filings. In some cases, investors could not even access the information they needed because some of the paper files in the SEC's reference rooms were lost.[5]

[5] A *Wall Street Journal* article reports that "…[n]owadays the SEC is being hit by a tidal wave of paper, receiving some 700,000 paper filings every year, amounting to about five million pieces of paper. Those documents are warehoused in the SEC's crowded public reference room, where investors, journalists and financial research organizations routinely comb through stacks of file folders in search of hot documents -- and don't always find them" (Block, 1991).

The second option was also costly because the data aggregators in the 1980s and early 1990s charged high fees to the customers. A petition filed to the SEC and the U.S. House of Representatives in 1992 documents complaints about these exorbitant fees. The partition demands free public access to corporate filings, showing that the Compustat CD-ROM database with historical filings for just 7,200 companies cost $18,000 (Love (1992)).[6] In addition, the petition reports that Mead Data Central was only available for a considerable fee that consisted of a $125 per month fixed fee, a $39 an hour connection fee, and a search fee ranging from $6 to $51 per search. The petition also reveals that Dialog, another financial data service provider at that time, charged $84 per hour on top of a $1 per page search fee. Chang, Ljungqvist, and Tseng (2020) argue that only large institutions realistically could afford online access. The authors also claim that retail investors and the small institutions most likely chose not to have access to the mandatory filings unless they were in the vicinity of the three SEC's reference rooms.[7]

In addition to the high fees, the Compustat suffered from production lag and inaccuracy, which also pushed up the costs of acquiring accurate financial information. D'Souza, Ramesh, and Shen (2010) find that the Compustat had an average dissemination lag of 24.69 weekdays prior to the EDGAR.[8] This lag dropped by almost 50% once EDGAR was adopted. Moreover, even if investors had subscribed to commercial data vendor services, there existed a significant mismatch between their databases. Kern and Morris (1994) compare two popular commercial databases at the time, the Value Line and the Compustat, and find material disagreements in the two datasets from 1985 to 1990. More importantly, they replicate Porcano's study (1986) using each database to show that empirical research could have different outcomes depending on the database the researcher used. Therefore, the costs of obtaining *accurate* financial information were still very high, even after paying the stiff fees that the commercial data vendors charged.

[6] According to Love (1992), the CD-ROM was called "COMPUSTAT PC Plus". A less expensive one, "COMPUSTAT Corporate Text", was for sale for $9,000. But "COMPUSTAT Corporate Text" had information for only 3,200 firms.

[7] The Compact Disclosure was another popular commercial database at the time. Richards (1988) documents that the Compact Disclosure had quarterly updated financial and management information on 10,150 public companies, and costs around $4,500 per year for commercial institutions. However, Richards (1988) notes that the Compact Disclosure's access software had technical issues retrieving the time-series data, and was missing information on brokerage houses, foreign companies, and microcap stocks with less than $5 million in assets.

[8] The authors compute dissemination lag as the number of weekdays between the SEC period report filing date and the Compustat *FINALQPRD* variable which represents the production date when a company's final quarterly financial data from period SEC filings first appeared in the Compustat.

Lastly, in principle, investors could receive the financial documents directly from the companies via mail. Besides the costs of a long wait, this was not a viable option for an investor who intended to perform cross-sectional firm characteristics analysis because such analyses require financial information on all the publicly listed companies.

EDGAR significantly lowers the costs of information acquisition by expediting electronic filing and the dissemination of electronic information via the internet. The SEC website points out that EDGAR "benefits investors, corporations, and the U.S. economy overall by increasing the efficiency, transparency, and fairness of the securities markets... Access to EDGAR's public database is *free*—allowing you to research, for example, a public company's financial information and operations by reviewing the filings the company makes with the SEC." EDGAR grants investors free and convenient access to the firms' financial information and thus substantially cuts the information acquisition costs. EDGAR's search function also lowered the costs of searching information by allowing the users to retrieve specific information in electronic documents (Gao and Huang, 2020).

Although this body of anecdotal strongly suggests that EDGAR substantially lowered the costs of information acquisition, it does not provide an accurate estimation of the extent to which the investor costs of information were lowered, a circumstance that I remedy in this paper. I study the *causal* effect of costly information constraints on anomalies using the staggered difference-in-difference framework to estimate unbiased costs of acquiring information in a market where public information is not accessible costlessly.

The Staggered Nature of the EDGAR Implementation

A feature of EDGAR implementation, of paramount importance to the empirical design in this study, is that the SEC adopted EDGAR following a phase-in schedule. The schedule spanned from April 1993 to May 1996, with 10 phases (from Group CF-01 to Group CF-10). Each public firm that required filing was assigned to one of the ten implementation groups. Each implementation group had a separate designated date as of which electronic filling was mandated (SEC Release No. 33-6977). Specifically, the companies in the first group, Group CF-01, were mandated to start uploading filings through EDGAR on April 26, 1993, and those in the last group, Group CF-10, on May 1, 1996. Table 1 shows the details of the implementation schedule.

The staggered nature of the EDGAR implementation helps to establish a strong causality and overcome the empirical challenges that previous studies faced by allowing me to employ the staggered difference-in-difference method.[9] The staggered difference-in-difference regression design provides a rich set of controlled groups for how the costs of acquiring information would have remained unchanged in the absence of EDGAR. This allows me to rule out other confounding factors and focus on the effect of information acquisition costs on the anomaly portfolio returns. Moreover, the staggered EDGAR implementation helps address reverse causality because the EDGAR implementation represents an exogenous shock that is independent of investors' actions. In addition, the parallel trends assumption is likely to hold in my difference-in-difference setting, given a strong indication of randomized assignment of EDGAR filers.[10] The average monthly alpha of the treatment group in the pre-EDGAR period (52.9 basis points) and that of the control group in the same pre-EDGAR period (48.8 basis points) are very similar in nature.[11] This similarity between the treatment and the control groups suggests that the assignment of firms to implementation phases is highly randomized and that selection bias is unlikely to be present in the sample.

3. Data and Methodology

A. The SEC EDGAR Implementation Data

To construct anomaly portfolios from EDGAR filers and, separately, from non-EDGAR filers, I first identify the date each company becomes an EDGAR filer by examining the SEC Release No. 33-6977 document. I also incorporate all the subsequent changes and corrections to the initial EDGAR phase-in list.[12] The SEC Release documents provide the list of company names and their Central Index Key (CIK) in each of the ten phases of EDGAR implementation. Using the company name and the CIK, I match each firm to their respective financial statement data from the

[9] For instance, many previous studies examine the effect of limits to arbitrage on the anomalies using *proxies* of limits to arbitrage. However, these studies can be misleading because the proxies may be capturing other confounding factors.
[10] A formal pre-trends test is presented in Section 5.E.
[11] 52.9 basis points is the average of 0.550 (the average of values in Column (4) Panel A of Table 6) and 0.510 (the average of values in Column (4) Panel B of Table 6). 48.8 basis points is the average of 0.532 (the average of values in Column (1) Panel A of Table 6) and 0.445 (the average of values in Column (1) Panel B of Table 6). Please see Section 4.A for detailed explanation.
[12] The subsequent changes and corrections to the initial EDGAR phase-in list reported in SEC Release No. 33-6977 can be found in the SEC Release documents No. 33-7063, No. 34-34097, No. 33-7156, No. 34-35572, No. 33-7258, No. 34-36737, No. 33-7215, and No. 34-36220.

Compustat, and stock price and return data from the Center for Research in Security Prices (CRSP). The last column of Table 1 shows the number of firms in each of the ten implementation phases that I successfully match to the two databases.

B. The Anomalies

I start by examining a total of 320 anomalies documented in Chen and Zimmermann (2020). Chen and Zimmermann (2020) replicate almost all the cross-sectional stock return signals that researchers have discovered.[13] By analyzing almost all the anomalies that the researchers have documented so far, I capture the full ramification of the information cost-saving effect of the EDGAR implementation on the anomalies' profitability. Following Chen and Zimmermann (2020). I construct anomaly predictors using data from the CRSP, the Compustat, the IBES dataset, the Option Metrics data, the SEC's Form 13Fs, the Federal Reserve Economic Data (FRED), and other datasets two authors used to create signals.[14]

All the quarterly versions of the anomalies documented in Chen and Zimmermann (2020) are included. Chen and Zimmermann (2020) provide the quarterly versions of the anomalies to encompass the work of Hou, Xue, and Zhang (2020) who modify the characteristics in the original paper such that quarterly accounting information, in lieu of annual information, is used to create signals. Likewise, in the spirit of Hou, Xue, and Zhang (2020), I replace the nine anomalies[15] that originally use annual versions of accounting variables with their quarterly versions according to the method described in Chen and Zimmermann (2020).[16]

---

[13] Specifically, Chen and Zimmermann (2020) documents all the anomalies in Hou, Xue, and Zhang (2020), 98% of the portfolios in McLean and Pontiff (2016), 90% of the characteristics from Green, Hand, and Zhang (2017), and 90% of the firm-level predictors in Harvey, Liu, and Zhu (2016). I would like to thank Andrew Chen and Tom Zimmerman for sharing the anomaly signal generating codes.

[14] Note that, in line with the standard practice, Chen and Zimmermann (2020) exclude firms with CRSP share codes greater than 11.

[15] The nine anomalies are: accruals, sales growth over inventory growth, sales growth over overhead growth, change in sales vs change in receivables, revenue growth rank, change in depreciation to gross PPE, change in gross margin versus sales, change in sales to inventory, net income/book equity.

[16] The nine anomalies were published prior to or during the EDGAR implementation period. This implies that (i) the investors were aware of these nine anomalies hence traded them more actively, and (ii) the investors would have capitalized more on timely information from EDGAR by using EDGAR's quarterly (rather than annual) information to update the signals. Therefore, the quarterly versions of these nine anomalies would better capture the arbitrage transactions. However, note that using the original annual versions of the nine anomalies does not significantly change the results. Table A.4 in the appendix shows the baseline difference-in-difference regression results when I strictly adhere to the set of accounting anomalies from Chen and Zimmerman (2020) (i.e., without replacing the annual versions of the nine accounting anomalies with their quarterly version). I assume 12 months of publication process following McLean and Pontiff (2016). The paper by Abarbanell and Bushee (1998) which documents three anomalies

Following Chen and Zimmermann (2020), I assume the standard one-quarter lag for quarterly accounting data availability. In other words, I assume that investors must wait for three months before they can start trading on the anomaly portfolios with EDGAR filer stocks using the previous quarter's financial information collected from EDGAR. Figure 2 exhibits an example of how the timeline of events is constructed for this study.

After generating the predictors, I exclude firms with a market capitalization below $50 million and with an end-of-month stock price lower than $5. Fama and French (2008) find that microcaps only average about 3% of the market capitalization of the NYSE-Amex-NASDAQ market, but account for about 60% of the total number of stocks. Moreover, these stocks have high transaction costs and inadequate liquidity. Therefore, they are unlikely to be exploited in the market (Hou, Xue, and Zhang (2020)). Applying the two stock-level filters mitigates the concern that the microcap returns might be driving my test results. (Green, Hand, and Zhang (2017), Hou, Xue, and Zhang (2020)). In addition, I also present the results using the *value*-weighted anomaly portfolio alphas since value-weighted portfolio formation avoids overweighting microcaps (Green, Hand, and Zhang (2017)).

I also address the issue of delisting return bias. Beaver, McNichols, and Price (2007) show that delisting returns can significantly affect anomaly portfolio returns, especially when they are omitted. Specifically, they show that when using monthly stock returns data from CRSP, ignoring or using replacement values for firms with missing delisting returns will not identify all the missing returns. In fact, Beaver, McNichols, and Price (2007) state that "… *daily* delisting returns are straightforward – they contain only the delisting return, or the return is given by using the last available price before delisting and the payment ultimately received by shareholders for the delisted security." Therefore, to address the issues related to the monthly delisting returns, I use the daily stock returns after adjusting for daily delisting returns, with the unknown daily delisting returns replaced as in Shumway (1997).[17] These corrected daily returns are then aggregated to

(sales growth over overhead growth, change in sales vs change in receivables, and change in gross margin versus sales predictors) first appeared on the SSRN in 1996. Therefore, sales growth over overhead growth, change in sales vs change in receivables, and change in gross margin versus sales anomalies are considered as published during the EDGAR implementation period.

[17] If neither the last return nor the delisting return is available and the deletion code is in the 500s—which includes 500 (reason unavailable), 520 (became traded over the counter), 551–573 and 580 (various reasons), 574 (bankruptcy), 580 (various reasons), and 584 (does not meet exchange financial guidelines)—the delisting return is assigned to be -30%. If the delisting code is not in the 500s, the last return is set to -1.0.

compute monthly returns. This allows me to address fully the issue of delisting return bias discussed in Beaver, McNichols, and Price (2007) and Shumway (1997).

I apply two additional filters at the anomaly return level: the negative alpha filter and the correlation filter. To apply the two filters, I first compute the Fama and French three-factor alphas[18] (Fama and French (1992, 1993)) and the pairwise return correlation of the decile equal-weighted anomaly portfolio returns over a ten-year period[19] prior to the EDGAR adoption (from October 1983 to September 1993).[20] Then, in the spirit of Green, Hand, and Zhang (2017), I exclude all the anomalies that have negative Fama and French three-factor alpha over the ten-year estimation period. The key rationale for the exclusion is that negative alphas indicate the strategies did not work at the time and, hence unlikely to have been traded by investors during the EDGAR implementation period. Finally, I apply the correlation filter by dropping one of two "twin" anomalies that have a pairwise return correlation above 0.9. This filter prevents an overestimation of the attenuation effect. Applying the two filters results in a set of 234 core anomaly portfolios. Table 2 and Table 3 present the most highly cited core anomalies that I study in this paper.[21]

Next, I compute the benchmark-adjusted anomaly monthly returns over January 1992 to December 1997 sample period for 234 core anomalies, after controlling for either Fama-French three or five factors, for both the equal-weighted and value-weighted decile and quintile portfolio returns. For brevity, Table 2 and Table 3 report the summary statistics of the Fama-French five-factor alphas for 20 accounting-based anomalies and the non-accounting-based anomalies with most citations, respectively. Table A.1 and Table A.2 of the appendix show the summary statistics for the Fama-French three-factor alphas.[22]

---

[18] I use the Fama and French *three*-factor alphas to filter out unprofitable anomalies because the Fama and French *three*-factor model was known to the investors at the time of EDGAR implementation whereas the Fama and French *five*-factor model was still not known to the public. Therefore, using the Fama and French *three*-factor alphas allows me to better capture the actual investors' trading activity prior to the introduction of EDGAR.

[19] In an untabulated test, I find that my results are essentially unchanged when I use shorter periods (3 years) for generating filters.

[20] The correlation and the alphas are estimated using the 10-year period prior to the first date which the investors start trading on the EDGAR filer stocks (i.e., October 1, 1993).

[21] Table IA.1 to Table IA.4 of the Internet Appendix report the full list and the summary statistics of all the 234 core anomalies that I investigate. The Internet Appendix is available upon request.

[22] Table IA.1, Table IA.2, Table IA.3 and Table IA.4 of the Internet Appendix shows the summary statistics for all the 234 anomalies. Note that I obtain the monthly Fama-French factor returns from the Ken French Data Library.

Finally, I construct a monthly panel consisting of the treated anomaly portfolio (treatment group) alphas and the controlled anomaly portfolio (control group) alphas for each anomaly for all implementation phases. In other words, I create a panel data of anomaly returns by month by EDGAR phase for the treated and the controlled anomaly portfolios. Given an implementation phase, the treated *stocks* consist of EDGAR filers assigned to the given implementation phase. On the other hand, the controlled *stocks* consist of the non-EDGAR filers (i.e., the firms still waiting their turn to start filing via EDGAR).[23] I construct treated anomaly portfolios using the treated stocks for each implementation phase. Likewise, I construct controlled anomaly portfolios using the controlled stocks for each implementation phase. After constructing the treated and controlled anomaly portfolios, I compute the risk-adjusted anomaly portfolio returns following the above methodology. Note that a control group cannot be formed for the last phase, Phase 10 (CF-10), because no non-EDGAR filler stocks are left in Phase 10. Consequently, the panel dataset does not include Phase 10.[24] Table A.3 shows the descriptive statistics for the treated and controlled anomaly portfolios.

4. Empirical Results

A. Main Staggered Difference-in-Difference Results

I first examine the effect of EDGAR implementation on the anomaly portfolio profitability using staggered difference-in-difference framework. EDGAR lowers investors' information acquisition costs by providing investors with free and instant access to SEC filings via the internet. This cost-saving effect attracts arbitrageurs to trade on EDGAR filer stocks. Accordingly, the profitability of anomaly portfolios constructed from EDGAR filer stocks weaken. However, only the

[23] For example, the treated stocks for Phase 5 consist of the stocks of the firms that are assigned to Phase 5. On the other hand, the control stocks for Phase 5 consist of the stocks that are still not an EDGAR filer as of August 1, 1994, namely the stocks of the firms that have been assigned to Phase 6, Phase 7, Phase 8, and Phase 9. Note that I exclude Phase 10. Also, note that the EDGAR-filers assigned to the previous phases (i.e., Phase 1, Phase 2, Phase 3, and Phase 4) are neither treated nor controlled stocks.

[24] Gao and Huang (2020) documents that the SEC assigned smaller firms to the last implementation phase (CF-10) to give them ample time to prepare for electronic filing. However, this introduces a potential selection bias to the assignment of treatment group, hence Gao and Huang (2020) repeat their regression estimations after dropping the last phase. Therefore, excluding the last phase also allows to address the concerns regarding possible non-random assignment of EDGAR filers. However, another way to address this concern is to explicitly control for firm size. So, in Section 5.A, I show the Fama-MacBeth regression results after controlling for firm size, bid-ask spread, and industry effects to fully address the selection bias issue.

accounting-based anomalies should show this attenuation of profitability because EDGAR only lowers the costs related to *accounting* information.

Using the monthly anomaly return panel data I described in the previous section, I estimate the following baseline staggered difference-in-difference regression equation:

$$\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post_{p,t} + \beta_3 * EDGAR_p * Post_{p,t} + \epsilon_{p,t}, \qquad (1)$$

where $\alpha_{p,t}$ is the Fama-French three or five-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with EDGAR filers (the treatment group) and equals zero if the portfolio is constructed from non-EDGAR filers (the control group); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first date[25] on which investors start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR (i.e., the effective date), zero if the month is before the effective date. I cluster standard error by anomaly and by month to address the potential correlation in errors (Petersen, 2009). Note that the indicator variable $Post_{p,t}$ is *not* subsumed by the monthly fixed effects γ_t because $Post_{p,t}$ is not fixed within every month level due to the staggered nature of the difference-in-difference framework. The difference-in-difference coefficient, β_3, is the coefficient of primary interest.

First, I perform a separate baseline regression analysis for the accounting-based anomalies and the non-accounting-based anomalies. Then, I run another regression to test the statistical difference between the two coefficients. Table 4 and Table A.5 report the baseline results for the accounting-based anomalies regression. Table 4 shows the results when I define the Fama-French five-factor alphas as the dependent variable, and Table A.5 in the appendix shows the results for the Fama-French three-factor alphas. In both cases, the difference-in-difference coefficients are negative, as well as statistically and economically significant. For example, the Fama-French *five*-factor alpha for the value-weighted decile portfolio (1-10 VW) in the last column of Table 4 decreases by 37.1 basis points per month, or 4.54% per year, on average.

[25] The first day of trading (i.e., the effective date in Table 1) is defined as the first date as of which investors start trading the EDGAR filer stocks that are assigned to a given implementation phase, using the latest financial information they retrieve from the EDGAR on those EDGAR filer stocks. Please see Table 1 for the list of all the effective dates. Figure 2 exhibits an exemplar timeline of events.

Moreover, the mean of the difference-in-difference coefficients across all eight portfolio specifications in Table 4 and A.5 is 32.6 basis points per month (4.0% per year). This number is pivotal because 32.6 basis points per month translate to the costs of information acquisition that investors incurred without EDGAR. EDGAR is free to use hence eliminates the costs of acquiring information for the arbitrageurs. Therefore, that exact amount of information cost-saving effect (the total arbitrage profit from the arbitrageurs' point of view) is quickly arbitraged away, resulting in a 32.6 basis points per month attenuation of profitability thereafter. Accordingly, the 32.6 basis points per month measure the precise amount of information acquisition costs investors faced in the absence of EDGAR: the costs of acquiring information in a market where public information is not accessible costlessly.

Table 5 displays the estimated baseline difference-in-difference regression coefficients for the *non*-accounting-based anomaly portfolios.[26] In line with my prediction, the difference-in-difference coefficients β_3 for the non-accounting anomaly portfolios are very close to zero and are uniformly insignificant. For example, the Fama-French five-factor alpha for the value-weighted decile portfolio (FF5 1-10 VW) in the last column of Table 5 shows the average of alphas declines by a mere 3.45 basis points per month, or 0.415% per year. These results are in stark contrast to the attenuation of profitability observed in the accounting-based anomalies. Again, this evidence confirms, as expected, that the attenuation of profitability is only found in the accounting anomalies and that EDGAR does not affect the profitability of the non-accounting anomalies.

Table 6 decomposes the estimated baseline difference-in-difference coefficients.[27] In other words, Table 6 shows the average Fama-French alphas before and after the EDGAR implementation for the anomaly portfolios constructed with either EDGAR filer stocks (treatment group) or non-EDGAR filer stocks (control group). Specifically, Panel A and Panel B of Table 6 show the decomposition of the results for the accounting-based anomaly portfolios. Columns (4) and (5) in Panel A and B confirm once again that the profitability of the treatment group (the accounting-based anomaly portfolios constructed with EDGAR filers) attenuates significantly in

[26] Table A.6 in the appendix shows the baseline regression results for the non-accounting-based anomaly portfolios with the Fama-French three-factor alphas as the dependent variable.
[27] The baseline difference-in-difference regression results are presented in Table 4, Table 5, Table A.5, and Table A.6

response to the EDGAR introduction (52.9 basis points versus 22.6 basis points).[28] In clear contrast, columns (1) and (2) in Panel A and B (the pre-versus-post changes for the control group) show that the average alphas of the control group remain almost unchanged even after EDGAR is introduced (48.8 basis points versus 51.1 basis points).[29] For example, the Fama-French five-factor alphas for the equal-weighted decile accounting-based portfolio (row "1-10 EW" of Panel B) in Panel B show that the profitability of the treatment group drops from 55.6 basis points per month to 22.3 basis points per month. In contrast, the profitability of the control group remains almost unchanged at close to 54 basis points throughout the sample period (55.2 basis points per month before EDGAR is introduced and 53.5 basis points per month after EDGAR is introduced).

In addition, comparing the average alphas of the treatment group in the pre-EDGAR period (52.9 basis points) to those of the control group in the same pre-EDGAR period (48.8 basis points) shows that the returns to the two groups are very similar.[30] This suggests that the parallel trends assumption is likely to hold in my difference-in-difference setting. Moreover, the selection bias is unlikely to be present in the sample because the similarity between the two groups indicates randomized assignment.

Given that the treatment and the control groups exhibit similar alphas in the pre-EDGAR period (52.9 versus 48.8 basis points) and that the control group's alphas remain unchanged after EDGAR implementation (48.8 versus 51.1 basis points), the estimated costs of acquiring information (i.e., the average of 32.6 basis points per month attenuation in the profitability of the accounting anomaly portfolios) explain approximately 61.6% of the 52.9 basis-point average Fama-French alphas in the pre-EDGAR period.[31]

[28] 52.9 basis points is the average of 0.550 (the average of values in Column (4) of Panel A) and 0.510 (the average of values in Column (4) of Panel B). 22.6 basis points is the average of 0.272 (the average of values in Column (5) of Panel A) and 0.180 (the average of values in Column (5) of Panel B).

[29] 48.8 basis points is the average of 0.532 (the average of values in Column (1) of Panel A) and 0.445 (the average of values in Column (1) of Panel B). 51.1 basis points is the average of 0.586 (the average of values in Column (2) of Panel A) and 0.436 (the average of values in Column (2) of Panel B).

[30] Again, 52.9 basis points is the average of 0.550 (the average of values in Column (4) of Panel A, Table 6) and 0.510 (the average of values in Column (4) of Panel B, Table 6). 48.8 basis points is the average of 0.532 (the average of values in Column (1) of Panel A, Table 6) and 0.445 (the average of values in Column (1) of Panel B, Table 6).

[31] 32.6 basis-point is the average staggered difference-in-difference coefficients across all accounting-based anomaly portfolio specifications under analysis (the average value of difference-in-difference coefficients in Column (7) of Table 6 Panel A and Table 6 Panel B). 52.9 basis-point is the average monthly Fama-French alphas of treatment group (the accounting-based anomaly portfolios constructed with EDGAR filer stocks) in the pre-EDGAR period (the average of values in column (4) of Table 6 Panel A and Table 6 Panel B). 61.6% is given by 32.6 basis points divided by 52.9 basis points. This simple comparison is feasible because (1) the average alphas of the treatment group (52.9

These findings show that the information acquisition costs can be as important as the transaction or short sale costs. Prior literature documents limits-to-arbitrage, such as noise trader risk, transaction costs, and short sale costs can explain anomaly returns to some extent. For example, Chu, Hirshleifer, and Ma (2020), document that relaxed short sale constraints reduce abnormal returns on 11 anomaly portfolios by 72 basis points per month.[32] Novy-Marx and Velikov (2015) show that the average trading costs range from 20 to 57 basis points per month for the mid-turnover anomalies. However, investors need to acquire information on stocks to identify which stocks to buy or sell even before they start trading. Therefore, investors incur the costs of acquiring information even before they pay transaction or short sale costs. I find that these costs of acquiring information can be as large as 32.6 basis points per month, explaining over 60% of the anomaly returns. These findings thus show that the information acquisition costs can be as important as other limits-to-arbitrage related costs that arbitrageurs face in a market with frictions.

Figure 1 summarizes the results in Table 6 that I discussed above. Figure 1 illustrates how the average pre-EDGAR versus post-EDGAR Fama-French alphas of the treatment and control groups for the accounting-based anomaly portfolios responded to the EDGAR implementation, clearly showing all the attributes of the average alphas that I discussed.

Panel C and D of Table 6 show that the average alphas for the *non-accounting*-based anomalies also remain mostly unchanged. Although there exists a slight upward trend for pre-versus-post alphas in both the treatment and the control groups, the rate of change (the slope) for the two groups is very similar. For example, in Panel D of Table 6, the average rate of pre-versus-post alpha change is +37% for non-EDGAR filers (i.e., Column (1) versus Column (2)), and +44% for EDGAR filers (i.e., Column (4) versus Column (5)). These results bolster the evidence that the parallel trends assumption is likely to hold in this difference-in-difference setting as well. One possible explanation of this observed increase in profitability is that fewer resources are left to work on arbitraging away the non-accounting-based anomalies once the arbitrageurs direct their

basis points) and the control group (the accounting-based anomaly portfolios constructed with *non*-EDGAR filer stocks, 48.8 basis points) are very close to each other in the pre-EDGAR period, and (2) the average alphas of the control group remain almost unchanged even after EDGAR is introduced (48.8 vs 51.1 basis points).

32 Note that Chu, Hirshleifer, and Ma (2020) analyze only 11 anomaly portfolios which fall short of 320 anomalies analyzed in this paper. Therefore, the estimated effect of 72 basis points is unlikely to be an accurate estimation of short sale costs. In fact, given that large institutions do not pay high short sale costs in practice, the estimated effect might be an overestimation.

limited resources and attention toward exploiting the accounting-based anomalies. However, the economic magnitude of the increase in profitability is too small to provide conclusive evidence of this explanation.

To formally test whether the difference-in-difference coefficients for the accounting-based anomaly portfolios and the non-accounting-based anomaly portfolios are statistically different, I run the following *triple* difference-in-difference regression:

$$\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_{p,\,t} + \beta_2 * ACC_{p,t} + \beta_3 * Post_{p,t} + \beta_4 * EDGAR_{p,t} * ACC_{p,t} + \beta_5 * ACC_{p,t} * Post_{p,t} + \beta_6 * Post_{p,t} * EDGAR_{p,t} + \beta_7 * EDGAR_{p,t} * ACC_{p,t} * Post_{p,t} + \epsilon_{p,t}, \quad (2)$$

where $ACC_{p,t}$ is an indicator variable that equals one if the anomaly portfolio is constructed using accounting variables and equals zero otherwise. The other variables are defined as in Equation (1). In this specification, the coefficient of primary interest is β_7. β_7 measures the differential attenuation effect between the accounting anomalies and the non-accounting anomalies.

Table 7 exhibits the estimates for Equation (2).[33] The results in Table 7 confirm that the difference in profitability attenuation between accounting and non-accounting anomalies is statistically significant. For instance, the Fama-French five-factor alpha for the value-weighted decile portfolio (1-10 VW) in the last column of Table 7 shows that the difference in the coefficients between the accounting-based anomaly portfolio and the non-accounting-based anomaly is -0.336, statistically significant. These results suggest that the EDGAR implementation resulted in attenuated profitability of the accounting-based anomaly portfolios, and the accounting-based anomaly portfolios alone because EDGAR reduces only the costs of acquiring accounting information for investors.

B. Profitability Attenuation Dynamics in the Post-EDGAR Period

This subsection examines whether the profitability attenuation is concentrated in the first month following an effective date[34] for a given implementation phase. The intuition behind this is that the arbitrageurs quickly exploit any profit opportunities involving higher average returns with no

[33] Please see Table A.7 in the appendix for the regression results using the Fama-French three-factor alphas as the dependent variable.
[34] An effective date is defined as the first date as of which investors start trading the EDGAR filer stocks that are assigned to a given implementation phase, using the latest financial information they retrieve from the EDGAR on those EDGAR filer stocks. Please see Table 1 for the list of all the effective dates.

extra exposure to risk. In addition, the arbitrageurs have the most advantage immediately after the information release.

To test this hypothesis, I first estimate the following regression equation:

$$\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post1_{p,t} + \beta_3 * Post2_{p,t} + \beta_5 * EDGAR_p * Post1_{p,t}$$
$$+\beta_6 * EDGAR_p * Post2_{p,t} + \epsilon_{p,t}, \quad (3)$$

where $Post1_{p,t}$ is an indicator variable that equals one if the month is the first month that investors start trading on EDGAR filer stocks for a given implementation phase and zero otherwise; and $Post2_{p,t}$ is an indicator variable that equals one if the month falls under the period after the first month investors start trading on EDGAR filer stocks for a given implementation phase and zero otherwise. Given the predictions above, the estimation of the two coefficients of central interest, β_5 and β_6, should be such that $\hat{\beta}_5 < 0, \hat{\beta}_6 < 0$, and $\hat{\beta}_5 < \hat{\beta}_6$.

The results in Table 8 and Table A.8 indeed confirm that $\hat{\beta}_5 < 0, \hat{\beta}_6 < 0$, and $\hat{\beta}_5 < \hat{\beta}_6$ in all portfolio specifications.[35] The difference between the estimated coefficients can be as large as 78.9 basis points[36] per month, which translates to 9.89% per annum. This again provides an estimate of the extent to which information constraints explain the anomaly portfolio returns as time progresses and shows how the arbitrage trade dynamics of EDGAR's cost-saving effect evolve in the post-EDGAR period.

C. Information Sensitivity and Profitability Attenuation of Anomaly Portfolios

In this section, I study how the cross-sectional variation in the sensitivity to recent information in the pre-EDGAR period affects the attenuation of the anomaly portfolio profitability following the EDGAR introduction. First, to capture how sensitive an anomaly portfolio is to recent information, I introduce the anomaly portfolio information turnover ratio measure. The information turnover ratio is defined as the total number of new incoming stocks divided by the total number of stocks in the existing portfolio when the anomaly portfolio updates its signal and rebalances its stocks.[37]

---

[35] Table A.8 in the appendix shows the results of the regression Equation (3) when Fama-French three-factor alphas are defined as the dependent variable.

[36] Please see column "FF3 Alpha 1-5 VW" (the second column) of Table A.8 in the appendix. 78.9 basis points is the difference between -1.104 and -0.315.

[37] I define portfolio rebalancing as a process that includes updating both the weights of the stocks (rebalancing in the traditional sense) and the stocks that constitute the portfolio (reconstruction).

Suppose an anomaly portfolio relies on recent information to a larger extent. In that case, the number of new stocks entering the portfolio upon rebalancing will be greater because the anomaly actively consumes new information to update its stock composition. On the other hand, if an anomaly portfolio is less sensitive to recent information, most of the stocks will remain in the portfolio when rebalanced and the number of new stocks that enter the portfolio will be smaller. Therefore, the information turnover ratio captures the sensitivity of a given anomaly portfolio to recent information.

In the context of the EDGAR implementation, the accounting-based anomalies with a higher information turnover ratio will experience greater attenuation of profitability. EDGAR expedites the dissemination of the latest accounting information, and the information cost-saving effect is larger for the accounting anomaly portfolio that leverages more on the latest accounting information. Naturally, a higher cost-reduction effect attracts more arbitrageurs and results in greater profitability attenuation. On the other hand, this implies that the attenuation effect should exist, but to a lesser degree for the accounting anomalies with low turnover ratio. In addition, the *non*-accounting anomalies should not show any attenuation effect.

To test these hypotheses, I first compute the pre-EDGAR information turnover ratio for 234 anomalies from October 1, 1983, to September 30, 1993, using all the listed stocks[38] in the sample period. After categorizing the anomalies into accounting-based and non-accounting-based anomalies, I sort the anomalies in each category based on their information turnover ratio rank percentile. The anomalies that exceed the 50th percentile are classified as the "High Turnover" anomalies, and the remaining anomalies are classified as the "Low Turnover." I also exclude three outlier anomalies from each turnover category in each portfolio specification which distort the results.[39] Finally, I re-estimate the baseline regression (Equation (1) in Section 4.A) separately for the "High Turnover" and the "Low Turnover" accounting-based anomalies, and then those for the non-accounting anomalies.[40]

[38] For consistency, I apply the stock-level filters discussed in Section 3.B. The results remain mostly unchanged even if I use a different time window, for example from January 1990 to December 1992 (as in Section 4.D), to compute the pre-EDGAR turnover ratio. Please see Table IA.5 of the Internet Appendix.

[39] For example, pension funding status (book value), pension funding status (market value), earnings predictability, quarterly Piotroski F-score, accounting component of price delay, O score anomalies are excluded in quintile portfolios.

[40] Note that both the accounting and the non-accounting anomaly portfolios are constructed with EDGAR filers (treated stocks) for each implementation phases since the EDGAR cost-saving effect is only expected from the

Table 9 compares the estimated difference-in-difference coefficients for the high turnover ratio versus the low turnover ratio accounting-based anomalies.[41] Table 9 also shows whether the two difference-in-difference coefficients are statistically different. The results strongly support the hypothesis that the *accounting*-based anomalies with higher information turnover ratios show greater attenuation of profitability.

Table 10 shows the results for the same tests, but for the *non*-accounting anomalies.[42] In stark contrast, all difference-in-difference coefficients do not exhibit higher attenuation of profitability for the anomaly portfolios with *higher* turnover ratios. In fact, most of the difference-in-difference coefficients are close to zero. However, Table 10 also shows that the difference-in-difference coefficients of the value-weighted portfolios with *low* information turnover ratio are negative, ranging from -0.142% (Table 10 column (4)) to -0.244% (Table A.10 column (2)). But the negative coefficients are limited to value-weighted portfolios. Therefore, a minor price spillover effect from a partial overlap of large stocks (which drive the value-weighted returns) between the accounting-based and non-accounting-based anomaly portfolios might be driving these results.[43]

Overall, these results provide a clearer picture of the extent and the channels through which EDGAR's information cost-saving effect transmits to the anomaly portfolio returns.

D. Pre-EDGAR Informational Availability and Accounting-based Anomaly Portfolios

In this subsection, I study how the cross-sectional variation in information availability during the pre-EDGAR period affects the accounting-based anomaly portfolios after the EDGAR adoption. Stocks differ in terms of information availability that investors can access. For example, investors

---

treatment group. Therefore, this limits the samples to the treated "High Turnover" anomaly portfolios and the treated "Low Turnover anomaly portfolios. The statistical difference between the difference-in-difference coefficients on the "High Turnover" anomaly portfolios and those on the "Low Turnover" is tested using the triple difference-in-difference Equation (2) introduced in Section 4.A, after replacing the $ACC_{p,t}$ term in Equation (2) with an indicator variable that identifies high versus low turnover anomaly.

[41] Table 9 shows the results for the Fama-French five-factor alphas, whereas Table A.9 shows the results for the Fama-French three-factor alphas.

[42] Table 10 shows the results for the Fama-French five-factor alphas, whereas Table A.10 shows the results for the Fama-French three-factor alphas.

[43] Note that these negative coefficients are not observed in the equal-weighted portfolios with the same low turnover ratio. Therefore, the results might be driven by a small group of large stocks that remain for relatively longer time in the long (short) leg of the non-accounting anomaly portfolio and experience price drop (rise) in response to the EDGAR implementation, but not for long time in the long (short) leg of the accounting anomaly portfolio.

typically have less information on the stocks (hereafter, the low information stocks) that are smaller. Investors also have less information on stocks that are covered by fewer, or if any, analysts. If an investor intends to trade on these stocks, the investor must bear higher costs of information to acquire the relevant information on these firms. Higher information costs would conceivably discourage traders from exploiting any arbitrage opportunities related to these types of low information stocks.

However, EDGAR makes accounting information available for all public companies for free, especially for those with low information availability for which investors had to pay a higher price to access the corporate filings. This implies that stocks with *lower* information availability are also the ones that would render investors a higher information cost-saving effect once they become EDGAR filers. As a result, these stocks attract more arbitrageurs who try to capitalize on their higher information cost-saving effect. Therefore, I expect the accounting anomaly portfolios consisting of lower information availability EDGAR filers stocks to weaken more when compared to those constructed with higher information availability EDGAR filer stocks.

To test this hypothesis, I use two variables to measure the level of information availability in the sample stocks: the number of analysts covering a given stock and the firm size (the market capitalization of equity). I first compute the average number of analyst coverage and the firm size for all companies from January 1990 to December 1992 and then sort the firms into their respective rank percentiles of the two measures.[44] I assign the above-median stocks (those in the top 50th percentile rank) to the high analyst coverage (large size) group and the below-median (those in the bottom 50th percentile rank) to the low analyst coverage (small size) group.

To exploit the stock-level variation, I proceed by estimating the Fama-MacBeth regression (1973) separately for the high and the low information availability groups of accounting-based anomaly portfolios. In the first pass, I estimate a cross-section regression for 126 accounting-based anomalies, for every EDGAR implementation phase, and for every month from January 1992 to December 1997 (i.e., estimate cross-sectional regressions by-anomaly, by-group, by-month) using the following equation:

$$R_{i,a,p,t+1} = \alpha + \beta * SignalPercentile_{i,a,p,t} + \epsilon_{i,a,p,t}, \quad (4)$$

[44] I obtain the number of analysts data from IBES (IBES dataset item NUMEST).

where i is an EDGAR-filer assigned to a given EDGAR implementation phase p for a given accounting-based anomaly a in the month t; $R_{i,a,p,t+1}$ is the simple monthly return of the following month for the given EDGAR-filer (stock) i; β is the factor loading; $SignalPercentile_{i,a,p,t}$ is the stock i's signal (characteristic) percentile within the EDGAR filers assigned to implementation phase p for anomaly a in month t. Normalizing the signal percentile allows me to standardize the units of signals and thus better compare the $\hat{\beta}$ across the anomalies. The first pass regression creates a panel of monthly beta estimates for each accounting-based anomaly portfolio consisting of EDGAR filer stocks in a given implementation phase. To obtain the time-series average of these beta estimates and to measure the effect of EDGAR implementation on all the $\hat{\beta}$, I estimate the following second-pass regression:

$$\widehat{\beta_{a,p,t}} = \gamma_t + \delta * Post_{a,p,t} + \epsilon_{a,t}, \qquad (5)$$

where a is the anomaly portfolio formed on EDGAR filer stocks in a given implementation phase p; $\widehat{\beta_{a,p,t}}$ are the monthly beta estimates from the first-pass; γ_t is the monthly time fixed effect; $Post_{a,p,t}$ is an indicator variable that equals one if the month is after the first date (the effective date for implementation phase p) on which investors start to trade on the anomaly a using latest information on new EDGAR filers that investors obtain from EDGAR and equals zero otherwise; and δ is the coefficient of primary interest. Again, the standard errors are clustered by month and by anomaly to address the potential correlation in errors.[45]

Table 11 displays the estimates of δ. These coefficients show a significant decline in return predictability in response to the EDGAR adoption for the accounting-based anomaly portfolios consisting of low information EDGAR filers. For example, the Fama MacBeth coefficients on the anomaly portfolios with low analyst coverage decline by 0.728% per month upon the EDGAR introduction. In comparison, the Fama MacBeth coefficients on the anomaly portfolios with small size EDGAR filers decline by 0.704% per month upon the EDGAR implementation.

[45] To test the statistical difference between the high and low information availability groups (in other words, to test the statistical difference between $\delta_{LowInfoGroup}$ and $\delta_{HighInfoGroup}$), I run the following second pass: $\widehat{\beta_{a,p,t}} = \gamma_t + \delta_1 * Post_{a,p,t} + \delta_2 * LowInfo_{a,p,t} + \delta_3 * Post_{a,p,t} * LowInfo_{a,p,t} + \epsilon_{a,t}$, where $LowInfo_{a,p,t}$ is an indicator variable that equals one if classified as low information availability group and zero otherwise.

One thing to note is that the accounting-based anomaly portfolios constructed with large size EDGAR filers do show a marginal, albeit statistically insignificant, reduction in return predictability. This suggests that the information on large stocks could have been only partially available to investors or that it was still costly to acquire relevant information in the pre-EDGAR period. Therefore EDGAR helped investors access some of the less readily available information on large stocks as well.

Overall, these results show how the variation of information cost-saving effect has a differential attenuation effect on the accounting anomaly portfolios. This differential effect is consistent with the predictions of my hypotheses and economic intuition.

5. Robustness Tests

A. Non-random Assignment of the EDGAR filers into Implementation Groups

The results from the previous sections rest on the assumption that the SEC has randomly assigned the firms to one of the ten implementation phases. However, SEC Release No. 33-6944 (Proposed Rulemaking for EDGAR System) states that the smaller firms were assigned to the latter phases to give the smaller firms ample time to prepare for EDGAR adoption.[46] This potentially violates the random assignment assumption of my analysis and raises the issue of endogeneity: the results may be driven by firm characteristics that are correlated with the assignment to phase. To address this issue, I run the Fama-MacBeth regression after controlling for firm size and the bid-ask spread.

In a similar spirit to the Fama-MacBeth regression in Section 4.D, I first estimate a cross-section regression for every stock, for all 234 anomalies, for every EDGAR implementation phase, and for every month from January 1992 to December 1997 (i.e., by anomaly, by month, by phase) using the following equation:

$$R_{i,a,p,t+1} = \beta_0 + \beta_1 * SignalPercentile_{i,a,p,t} + \beta_2 * Treated_{i,p} + \beta_3 * SignalPercentile_{i,a,p,t} * Treated_{i,p} + \beta_4 * \ln(MarketCap)_{i,t} + \beta_5 * BidAskSpread_{i,t} + IndustryFixedEffect \quad (6)$$

[46] SEC Release No. 33-6944 (Proposed Rulemaking for EDGAR System, 1992) states that "the Commission has designed the EDGAR system to accommodate small entities to the greatest degree possible while still carrying out its mandate to develop a system for the electronic dissemination of information to the public. Small companies will be the last group phased into the system, allowing them to take advantage of the substantial body of experience gained by those who precede them."

where i is a stock (either an EDGAR filer or a non-EDGAR filer) assigned to a given EDGAR implementation phase p for a given anomaly a in the month t; $R_{i,a,p,t+1}$ is the simple monthly return of the following month for the given stock i; $SignalPercentile_{i,a,p,t}$ is the stock i's signal (characteristic) percentile within all the stocks (both treated and controlled stocks) for given month t, anomaly a, and implementation phase p; $Treated_{i,p}$ is an indicator variable that equals one if the stock i is an EDGAR filer for a given implementation phase p and equals zero otherwise. Industry fixed effect is defined as the first two digits of the stocks SIC code. $ln(MarketCap)_{i,t}$ is the log of market capitalization, and $BidAskSpread_{i,t}$ is the bid-ask spread for the stock i in month t. Note that $Treated_{i,p}$ is invariant to anomaly and month. Another thing to note is that I create signal percentiles using *all* the stocks (both treated and controlled stocks) in a given month.[47]

To estimate how the time-series of beta estimates from the first pass respond to the EDGAR implementation, I estimate the following second-pass regression separately for the accounting-based anomalies and the non-accounting anomalies:

$$\widehat{\beta_{a,p,t}} = \gamma_t + \delta * Post_{a,p,t} + \epsilon_{a,t}, \tag{7}$$

where a is the anomaly portfolio formed on stocks in a given implementation phase p; $\widehat{\beta_{a,p,t}}$ are the monthly beta estimates (the beta estimates of interest are $\hat{\beta}_3$ for the treatment group and $\hat{\beta}_1$ for the control group) from the first-pass regression; γ_t is the monthly time fixed effect; $Post_{a,p,t}$ is an indicator variable that equals one if the month is after the effective date for implementation phase p and equals zero otherwise; and δ is the coefficient of primary interest. Again, the standard errors are clustered by month and by anomaly to address the potential correlation in errors. I drop four outlier anomalies which distort the results.[48]

[47] This allows me to address the concern that the two following methods I use in this study might yield different results. (Method 1) sorting only the treated stocks based on their signal to construct treated portfolios, and then sorting controlled stocks separately to construct a controlled portfolio. (Method 2) sorting all the stocks together (including both the treated and the controlled stocks) to create percentile signals. The difference-in-difference analysis in the previous session follows Method 1. I follow Method 2 in this section.

[48] The anomalies are "Order Back Log", "Quarterly RD sales", "Probability of Informed Trading", and "Citations RD". The results remain unchanged even if these anomalies are not dropped. However, the statistical significance of the difference in $Post_{a,p,t}$ coefficient between the accounting and non-accounting treatment portfolios becomes less significant due to excessively large standard errors induced these outliers.

To test the statistical difference between the accounting and the non-accounting anomalies, in other words, to test the statistical difference between $\delta_{accounting,1}$ and $\delta_{non-accounting,1}$ estimated above, I run the following second pass as well:

$$\widehat{\beta_{a,p,t}} = \gamma_t + \delta_1 * Post_{a,p,t} + \delta_2 * ACC_{a,p,t} + \delta_3 * Post_{a,p,t} * ACC_{a,p,t} + \epsilon_{a,t}, \tag{8}$$

where $ACC_{a,p,t}$ is an indicator variable that equals one if anomaly a is classified as an accounting-based anomaly.

Table A.11 presents the results for the above regressions. Consistent with the results in the previous sections, only the accounting anomalies consisting of EDGAR filers (the treatment group) weaken in response to EDGAR implementation. Moreover, the difference in attenuation between accounting and non-accounting anomalies is statistically significant even after controlling for firm size, bid-ask spread, and industry fixed effect. Specifically, in Table A.11 Column (1), the coefficient on $Post_{a,p,t}$ for the accounting anomalies in the treatment group is -0.784 and statistically significant at 1% level, whereas the other coefficients are either positive or statistically insignificant. This result shows that the accounting anomalies constructed with EDGAR filers (treatment group) significantly weaken in response to the EDGAR introduction even after controlling for firm characteristics. In contrast, the non-accounting anomalies and the control group remain mostly unchanged.

B. Understanding the Mechanism behind the Attenuation of Anomaly Profitability

In this section, I investigate how investors' trading activity measures respond to the EDGAR implementation with an aim to better understand the mechanism behind the attenuation of accounting anomaly portfolio profitability. Suppose the arbitrageurs trade the anomalies in response to the EDGAR implementation. In that case, the stocks captured by those trades should exhibit higher trading volume, lower bid-ask spread, lower idiosyncratic volatility, and lower Amihud illiquidity measure.[49] In contrast, changes in institutional ownership and short interest

[49] The data are from CRSP and Compustat. Idiosyncratic volatility is the standard deviation of the regression residual from the Fama-French three-factor model. Daily return data in the past month are used for the estimation. Amihud illiquidity measure (Amihud (2002)) is defined as the past twelve-month average of daily return divided by dollar volume. The bid-ask spread is estimated following the methodology in Shane and Schultz (2012).

change in response to EDGAR introduction are harder to predict. Therefore, it remains an empirical question.

To examine the responses, I first estimate a cross-section regression for every stock that is captured by 1-10 decile anomaly portfolios (captured by both the short and the long leg of the decile portfolio), for all 234 anomalies, for every EDGAR implementation phase, and for every month from January 1992 to December 1997 (i.e., estimate cross-sectional regressions by anomaly, by phase, by month) using the following equation:

$$Measure_{i,a,p,t} = \beta_0 + \beta_1 * Treated_{i,a,p} \,, \tag{9}$$

where $Measure_{i,a,p,t}$ is the trading related measure of interest, and $Treated_{i,a,p}$ equals one if the stock *i* is an EDGAR filer in the decile treatment portfolio (group) for anomaly *a* for a given implementation phase *p*, and zero if stock *i* is a non-EDGAR filer in the decile control portfolio (group) for anomaly *a.* I re-estimate regression equations (7) and (8) introduced in the previous subsection to estimate the changes in time-series average of beta estimates from the first pass and test the statistical difference between the accounting and the non-accounting anomalies.

Table A.12 reports the results of the estimations. In line with the expectations, the EDGAR filer stocks in the decile accounting anomaly portfolios (treatment group) exhibit higher trading volume, lower bid-ask spread, lower idiosyncratic volatility, and lower Amihud illiquidity index in response to EDGAR introduction. An interesting point to note is that the EDGAR filer stocks in the *non*-accounting anomaly decile portfolios also exhibit statistically significant economic effects. For example, the coefficient on $Post_{a,p,t}$ for idiosyncratic volatility in Panel C is -0.0010 for the non-accounting anomalies in the controlled group, and the coefficient is statistically significant at 1% level. However, the effect is more pronounced for the accounting anomaly portfolio stocks across all the measures and the magnitude of the difference in coefficient is statistically significant for almost all of them.[50]

After restricting the samples, I analyze changes in institutional ownership and short interest in more detail by repeating the above test. I limit the samples to (i) the EDGAR filers that are

[50] If I define the dependent variable as the raw value of trading volume (instead of the log of trading volume) the difference in coefficients for the treatment group (the coefficient in the first row and the last column) becomes statistically significant at 10% level.

captured by the *long* leg of the decile anomaly portfolios in the treatment group and (ii) the EDGAR filers that are captured by the *short* leg of the decile anomaly portfolios in the treatment group.[51] Note that this allows me to compare the long leg to the short leg of the treated anomaly portfolios, where I set the reference group as the short leg.

Table A.13 (in the appendix) shows the changes in institutional ownership and short interest. Table A.13 shows that the stocks in the short leg of the accounting-based decile anomaly portfolios in the treatment group are mostly affected by the EDGAR implementation. In other words, the institutional ownership increases, and the short interest decreases for the EDGAR filer stocks in the short leg. By contrast, the stocks in the long leg of the treated decile portfolios remain significantly less affected.

At first glance, these results seem inconsistent with the economic intuition. However, these results are not surprising once we consider the analysis presented in the next section, Section 5.C (In Section 5.C, I show that the short leg of the accounting anomaly portfolios is driving the attenuation of profitability). The average short interest for the EDGAR filer stocks in the short leg of the treated accounting portfolios drops by 19 percent because shorting these stocks becomes less profitable after the EDGAR implementation. Moreover, this decline in the short interest (i.e., short covering) could be driving the observed post-EDGAR increase in institutional ownership of the stocks captured by the short leg of the accounting anomaly portfolios. Therefore, the results presented in Table A.13 are consistent with those in Section 5.C.

C. Long Leg versus Short Leg Accounting Anomaly Portfolios

This section examines the dynamics of the short legs versus the long legs of the accounting-based anomaly portfolio. The main results indicate that the profitability of the *long-short* accounting anomaly portfolio attenuates in response to the adoption of EDGAR. However, whether the accounting anomaly portfolio's short leg or long leg drives the attenuation is an empirical question. The attenuation of portfolio profitability depends on how less sophisticated investors, such as retail investors, respond to the EDGAR implementation. Suppose retail investors' net buying of the

[51] The indicator variable $Treated_{i,a,p}$ in the Equation (9) is now defined as one if the stock *i* is an EDGAR filer in the *long* leg of the decile treatment portfolio for anomaly *a* in a given implementation phase *p*, and zero if stock *i* is in the *short* leg of the decile treatment portfolio for anomaly *a* in a given implementation phase *p*.

EDGAR filer stocks increases in response to EDGAR as evidenced by an increase in trading volume in Section 5.B. Also, suppose that retail investors are slower to trade on the stocks than the arbitrageurs are since they are less sophisticated. If these two conditions hold, then retail investors would exert upward pressure on the price of the EDGAR filer stocks that the arbitrageurs have already constructed their portfolios with immediately following the EDGAR introduction. This upward price pressure would increase the profitability of the long leg portfolio and, at the same time, decrease the profitability in the short leg. Likewise, if retail investors were to unload their stocks while still responding slowly to the EDGAR information, the supply pressure would decrease the profitability of arbitrageurs' long leg portfolio while increasing the profitability of the short leg. While the stock prices would have adjusted according to the increase in arbitrage trades immediately following the implementation of EDGAR, the ensuing retail investors' response could have impacted the arbitrage profit in the longer term. Note that the main analysis shows that the long-short portfolio profitability (the net effect) attenuates in response to the adoption of EDGAR. Therefore, the net effect of the two changes in profitability should be negative regardless of retail investors' reaction.

Table A.14 presents the baseline difference-in-difference analysis for the long leg versus the short leg of the accounting-based anomaly portfolios.[52] Table A.14 Panel A shows that the profitability in the long leg increased, albeit marginal. By contrast, Table A.14 Panel B confirms that the short leg of the portfolios mostly drives the attenuation of profitability. For example, the difference-in-difference coefficient for the long leg of the value-weighted quintile Fama-French five-factor alpha portfolio (Column "FF5 Alpha 1-10 VW" in Panel A) shows an increase of 5.2 basis points per month or 0.84% per annum. On the other hand, the coefficient for the same value-weighted decile short leg portfolio (Column "FF5 Alpha 1-10 VW" in Panel B) shows an attenuation of 45.0 basis points per month or -5.54% per year. These results strongly suggest that retail investors tended to increase the net buying of the EDGAR filer stocks that the arbitrageurs shorted as EDGAR rolled out. As a result, the arbitrage profit on the short leg of the portfolio attenuated, forcing the arbitrageurs to cover their short position. This can also explain the change in institutional ownership discussed in the previous section, Section 5.B.

[52] Table A.15 in the appendix shows the result for the Fama-French three-factor alphas.

D. Concerns Regarding the First Implementation Phase

Before EDGAR began rolling out in April 1993, the SEC called for volunteers to file electronically. The goal of this test was to check the integrity of the EDGAR system before engaging in a full-fledged implementation. The firms that volunteered were called the transitional filers, and the SEC assigned these firms to the first implementation phase group. Although these transitional filers were required to submit all the filings via EDGAR in April 1993, these circumstances might have introduced a certain degree of selection bias to the sample because the transitional firms have self-selected into the first phase.

In addition, some researchers document that EDGAR became available for free to the public (via a National Science Foundation grant to New York University) only after January 17, 1994.[53] According to these studies, investors had to pay a high fee discussed in Section 2.A before January 1994 because investors only had access to EDGAR via Mead Data Central. Given the standard three-month information acquisition lag assumption I introduce, if EDGAR were indeed not available for free prior to January 1994, the first implementation phase would have little cost-saving effect to investors because the effective date (October 1, 1993) of the first implementation phase falls before January 1994. Note that the remaining implementation phases (from CF-02 to CF-09) are unaffected because the respective effective date falls on or after January 1994.

To address these concerns, I repeat the baseline difference-in-difference analysis for the accounting and non-accounting-based anomaly portfolios after dropping the first implementation phase. Table A.16 presents the estimates for the repeated regression analysis. The results confirm that the attenuation of profitability is marginally stronger across all portfolio specifications compared to the results in Table 4. For example, the difference-in-difference coefficient for the Fama-French five-factor alpha value-weighted decile portfolio (the "FF5 Alpha 1-10 EW" Column of Panel A, Table A.14) is -0.324. The same coefficient in Table 4 (the "FF5 Alpha 1-10 EW" Column of Table 4) is -0.316. Therefore, the overall inferences remain mostly unchanged even if I exclude the first phase from my main analysis.

E. Pre-trends and Falsification Tests

[53] For example, Goldstein, Yang, and Zuo (2020) and Chang, Ljungqvist, and Tseng (2020).

One of the key assumptions of the difference-in-difference analysis is the parallel trend assumption. In Section 4.A, I have already discussed how the results in Table 6 suggest that this assumption likely holds. Given the importance of this assumption, I formally test whether it is indeed the case. To this end, following the methodology in Gao and Huang (2020), I re-estimate the baseline difference-in-difference regression in Equation (1) over a four-year period *prior to* the actual EDGAR implementation, using pseudo-event dates. The pseudo-events of each EDGAR implementation phase are assumed to take place two years *before* the actual phase dates. Naturally, the indicator variable $Post_{p,t}$ in Equation (1) is redefined as a variable that equals one if the month is after the first pseudo-event date on which investors presumably trade on the latest EDGAR information related to the new EDGAR filers and zero if the month is before that pseudo-event date. Table A.17 presents the result for the pre-trend test. As expected, the difference-in-difference coefficient is a mixture of positive and negative numbers that are very close to zero. These results show that the parallel trend assumption is likely to hold in my difference-in-difference setting.

In addition to the parallel trend test, I also run a falsification test. Again, applying the methodology in Gao and Huang (2020), I re-estimate the baseline difference-in-difference regression in Equation (1) over a four-year period *following* the actual EDGAR implementation, using the pseudo-event dates from two years *after* the actual phase dates. The indicator variable $Post_{p,t}$ in Equation (1) is adjusted accordingly. Table A.18 reports the results for the falsification test. The numbers are very similar to those in Table A.17, suggesting no meaningful attenuation of alphas around the pseudo-events, both before and after the EDGAR implementation.

6. Conclusion

In this paper, I investigate the causal effect of the information acquisition costs on the anomaly portfolio returns using the SEC's EDGAR implementation as an exogenous shock that lowers the costs of acquiring information for publicly traded stocks. Using the staggered difference-in-difference framework, I find that the profitability of the accounting-based anomaly portfolios constructed with EDGAR filer stocks (the treatment group) attenuates 32.6 basis points per month (4.0% per year) on average in response to the EDGAR introduction. This explains over one-half of pre-EDGAR accounting-based anomaly portfolio alphas. The reduction in profitability is economically and statistically significant. Moreover, as expected, the profitability of the non-accounting anomaly portfolios remains unchanged as EDGAR disseminates only accounting information.

The average profitability attenuation of 32.6 basis points per month (4.0% per year) has a very special meaning because the 4.0% translates to the costs of acquiring accounting information that investors had to bear in the absence of the EDGAR system. To the best of my knowledge, this is the first paper to present a clean-cut estimation of investors' information acquisition costs and show that information acquisition costs can be as important as the transaction or short sale costs.

I also find that the profitability of the accounting-based anomalies consisting of EDGAR filer stocks attenuates more in the first month. In addition, in a series of tests exploiting the cross-sectional variation in information cost-saving effect at the anomaly portfolio level, I show that the accounting-anomalies (1) with higher sensitivity to recent information, (2) constructed using EDGAR filer stocks with a lower number of analyst coverage, and (3) constructed with smaller EDGAR filer stocks exhibit greater attenuation.

Finally, I demonstrate that (1) the results remain unchanged even after controlling for the firm size, the bid-ask spread, and industry effects (2) trading volume, liquidity, volatility, short interest, and institutional ownership respond as expected (3) the short leg portfolios are mostly driving the attenuation of profitability. Moreover, I show that my results are robust to any potential selection bias issues associated with the first implementation phase, and the post-publication effect is not driving the results (Appendix Section B). Lastly, I show that the results pass the pre-trend and the falsification tests.

References

Abarbanell, J. S., & Bushee, B. J. (1998). Abnormal returns to a fundamental analysis strategy. *Accounting Review*, 19-45.

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

Beaver, W., McNichols, M., & Price, R. (2007). Delisting returns and their effect on accounting-based market anomalies. *Journal of Accounting and Economics*, *43*(2-3), 341-368.

Blankespoor, E., deHaan, E., & Marinovic, I. (2020). Disclosure processing costs, investors' information choice, and equity market outcomes: A review. *Journal of Accounting and Economics*, *70*(2-3), 101344.

Chang, Y. C., Ljungqvist, A., & Tseng, K. (2020). Do corporate disclosures constrain strategic analyst behavior? *Available at SSRN*.

Chen, D., Ma, Y., Martin, X., & Michaely, R. (2021). On the fast track: information acquisition costs and information production. *Journal of Financial Economics*.

Chen, A. Y. (2020). The limits of p-hacking: Some thought experiments. *The Journal of Finance* (Forthcoming).

Chen, A. Y., & Zimmermann, T. (2020). Open source cross-sectional asset pricing. *Available at SSRN*.

Chu, Y., Hirshleifer, D., & Ma, L. (2020). The causal effect of limits to arbitrage on asset pricing anomalies. *The Journal of Finance*, *75*(5), 2631-2672.

Corwin, S. A., & Schultz, P. (2012). A simple way to estimate bid-ask spreads from daily high and low prices. *The Journal of Finance*, *67*(2), 719-760.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of political Economy*, *98*(4), 703-738.

Dichev, I. D. (1998). Is the risk of bankruptcy a systematic risk?. *the Journal of Finance*, *53*(3), 1131-1147.

D'Souza, J. M., Ramesh, K., & Shen, M. (2010). The interdependence between institutional ownership and information dissemination by data aggregators. *The Accounting Review*, *85*(1), 159-193.

Easley, D., & O'hara, M. (2004). Information and the cost of capital. *The journal of finance*, *59*(4), 1553-1583.

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

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *the Journal of Finance*, *47*(2), 427-465.

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

Fama, E. F., & French, K. R. (1995). Size and book-to-market factors in earnings and returns. *The journal of finance*, *50*(1), 131-155.

Fama, E. F., & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *The journal of finance*, *51*(1), 55-84.

Fama, E. F., & French, K. R. (1998). Value versus growth: The international evidence. *The journal of finance*, *53*(6), 1975-1999.

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

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics*, *116*(1), 1-22.

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

Francis, J., LaFond, R., Olsson, P. M., & Schipper, K. (2004). Costs of equity and earnings attributes. *The accounting review*, *79*(4), 967-1010.

Gao, M., & Huang, J. (2020). Informing the market: The effect of modern information technologies on information production. *The Review of Financial Studies*, *33*(4), 1367-1411.

Goldstein, I., Yang, S., & Zuo, L. (2020). *The real effects of modern information technologies* (No. w27529). National Bureau of Economic Research.

Gomez, E. (2020). The Effect of Mandatory Disclosure Dissemination on Information Asymmetry: Evidence from the Implementation of the EDGAR System. *Available at SSRN 3577162*.

Green, J., Hand, J. R., & Zhang, X. F. (2017). The characteristics that provide independent information about average US monthly stock returns. *The Review of Financial Studies*, *30*(12), 4389-4436.

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American economic review*, *70*(3), 393-408.

Han, Y., Lu, Y. J., Xu, W., & Zhou, G. (2020). Mispricing and Anomalies: An Exogenous Shock to Short Selling from the Dividend Tax Law Change. *Available at SSRN*.

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

Hirshleifer, D., Hou, K., & Teoh, S. H. (2012). The accrual anomaly: risk or mispricing?. *Management Science*, *58*(2), 320-335.

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

Kern, B. B., & Morris, M. H. (1994). Differences in the Compustat and expanded Value Line databases and the potential impact on empirical research. *Accounting Review*, 274-284.

Lam, F. E. C., & Wei, K. J. (2011). Limits-to-arbitrage, investment frictions, and the asset growth anomaly. *Journal of Financial Economics*, *102*(1), 127-149.

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.

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics, 47*(1), 13-37. doi:10.2307/1924119

Love, J. P. (1992, January 14). Public Access to SEC information. *IU Bio-Archive. http://www.bio.net/bionet/mm/ag-forst/1992-January/000187.html*

Merton, R. C. (1987) A simple model of capital market equilibrium with incomplete information. *The Journal of Finance*, 42(3): 483–510

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

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica, 34*(4), 768-783. doi:10.2307/1910098

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

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of financial studies*, *22*(1), 435-480.

Porcano, T. (1986). Corporate tax rates: Progressive, proportional, or regressive. *Journal of the American Taxation Association*, *7*(2), 17-31.

Richards, D. (1988). Compact Disclosure. *RQ*, *27*(3), 413-415.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, *19*(3), 425-442.

Shapiro, A. (2002). The investor recognition hypothesis in a dynamic general equilibrium: Theory and evidence. *The Review of Financial Studies*, *15*(1), 97-141.

Shumway, T. (1997). The delisting bias in CRSP data. *The Journal of Finance*, *52*(1), 327-340.

Verrecchia, R. E. (1982). Information Acquisition in a Noisy Rational Expectations Economy, *Econometrica* 50, 1415−1430.

Whyte, A. (2020, July 9). How Much Are Managers Paying for Data? *Institutional Investors.* https://www.institutionalinvestor.com/article/b1mf7j9918j9lg/How-Much-Are-Managers-Paying-for-Data (accessed 25 July 2021).

Table 1: The SEC EDGAR Implementation Schedule.

This table shows the SEC's implementation timeline of EDGAR as recorded in the SEC Release documents. The announced implementation date is the date which the SEC mandated the assigned firms of a given implementation phase to start filing their financial statements electronically via EDGAR. The effective date is defined as the first date as of which investors start trading the EDGAR filers stocks (for a given implementation phase) using the latest financial information on the EDGAR filers retrieved from EDGAR. Following the standard one-quarter lag assumption of the quarterly accounting data availability in the anomaly literature, investors must wait for one quarter until the effective date for the recent information to become available on EDGAR before they can start trading on the EDGAR filer stocks using the newly acquired information. The number of EDGAR filer stocks is the number of stocks that are successfully matched to the Compustat and CRSP database.

Implement Phase	Implementation Date	The Year-Quarter of the Phase	Effective Date	Number of EDGAR Filer Stocks
1	4/26/1993	1993Q2	10/1/1993	149
2	7/19/1993	1993Q3	1/1/1994	541
3	10/4/1993	1993Q4	4/1/1994	564
4	12/6/1993	1993Q4	4/1/1994	737
5	8/1/1994	1994Q3	1/1/1995	1,033
6	11/1/1994	1994Q4	4/1/1995	866
7	5/1/1995	1995Q2	10/1/1995	858
8	8/1/1995	1995Q3	1/1/1996	756
9	11/1/1995	1995Q4	4/1/1996	386
10	5/1/1996	1996Q2	10/1/1996	2,723

Table 2: Summary Statistics of the Most Cited Accounting-based Anomalies

This table presents the time-series averages of the Fama-French five-factor alpha for the decile (denoted by "1-10") accounting-based anomaly portfolios studied in this paper. Out of a total of 320 anomalies documented in Chen and Zimmermann (2020), only 234 survive the set of filters I apply: stock price, common stocks, market capitalization, pair-wise correlation, and alpha filters. Out of 234 remaining anomalies, 126 belong to the accounting-based category, and 20 accounting anomalies with the most citations are presented in this table (126 accounting-based anomalies are sorted on the total number of times the original paper has been cited as of February 2021). The sample period is from January 1992 to December 1997. The unit is in percentage. The quarterly versions (denoted by "(q)") of the original anomalies are generated following Chen and Zimmermann (2020). The equal-weighted portfolio is denoted by "EW", whereas the value-weighted portfolio is denoted by "VW". The citation numbers are collected via Google Scholar. See the Internet Appendix for the summary statistics of all the anomalies, including those of the Fama-French five-factor alphas for quintile portfolios.

Accounting-based Anomaly	FF5 Alpha 1-10 EW		FF5 Alpha 1-10 VW	
	Mean	*t*-stat	Mean	*t*-stat
Book leverage (q)	-0.195	-1.746	-0.244	-1.753
Annual sales growth	0.661	4.735	0.422	2.502
Cash flow to market	0.344	1.547	-0.043	-0.182
Accruals	1.757	11.521	1.720	9.731
Earnings persistence	-0.412	-1.799	-0.221	-0.914
Earnings predictability	1.698	6.618	1.826	6.788
Value relevance of earnings	-0.135	-0.682	-0.254	-1.223
Book to market using December ME	0.487	3.625	-0.508	-3.174
Failure probability	0.243	1.959	0.257	1.610
Failure probability (June)	0.017	0.136	0.189	1.191
Advertising Expense	-0.453	-2.237	-0.599	-2.669
R&D to sales	0.573	2.225	0.624	2.167
Kaplan-Zingales index	1.140	8.196	0.562	3.444
Gross profits/total assets	0.373	2.672	0.873	5.516
Gross profits/total assets (lag)	0.010	0.069	0.567	3.273
Investment to revenue	0.270	1.802	0.093	0.538
Abnormal Accruals	0.326	2.144	0.303	1.624
Change in current operating assets	0.688	5.631	0.647	4.267
Change in equity to assets	0.855	6.255	0.691	4.147
Change in financial liabilities	0.106	0.840	0.074	0.481

Table 3: Summary Statistics of the Most Cited Non-Accounting-based Anomalies

This table presents the time-series averages of the Fama-French five-factor alpha for the decile (denoted by "1-10") non-accounting-based anomaly portfolios studied in this paper. Out of a total of 320 anomalies documented in Chen and Zimmermann (2020), only 234 survive the set of filters I apply: stock price, common stocks, market capitalization, pair-wise correlation, and alpha filters. Out of 234 remaining anomalies, 108 belong to the non-accounting-based category, and 20 non-accounting anomalies with most citations are presented in this table (108 non-accounting-based anomalies are sorted on the total number of times the paper that first documents the anomaly has been cited as of February 2021). The sample period is from January 1992 to December 1997. The unit is in percentage. The quarterly versions (denoted by "(q)") of the original anomalies are generated following Chen and Zimmermann (2020). The equal-weighted portfolio is denoted by "EW", whereas the value-weighted portfolio is denoted by "VW". The citation numbers are collected via Google Scholar. See the Internet Appendix for the summary statistics of all the anomalies, including those of the Fama-French five-factor alphas for quintile portfolios.

Non-Accounting Anomaly	FF5 Alpha 1-10 EW		FF5 Alpha 1-10 VW	
	Mean	*t*-stat	Mean	*t*-stat
Momentum (12 month)	1.115	7.416	0.774	3.986
Momentum (6 month)	0.852	6.178	0.284	1.634
Momentum-Reversal	0.616	4.553	0.320	1.825
Governance Index	-0.083	-0.583	-0.037	-0.224
Size	0.669	7.218	0.632	6.718
Bid-ask spread	0.806	8.194	0.792	7.158
Pastor-Stambaugh liquidity beta	0.089	0.680	0.318	1.824
Initial Public Offerings	0.176	1.250	-0.209	-1.362
IPO and age	-0.788	-2.771	-0.660	-2.275
Systematic volatility	0.760	6.726	1.220	7.970
Earnings-to-Price Ratio	-0.642	-3.684	-0.546	-2.940
Illiquidity-illiquidity beta	-0.219	-1.185	-0.389	-1.889
Net liquidity beta	-0.427	-2.389	-0.414	-2.180
Return-market illiquidity beta	0.616	3.448	0.616	3.150
Short term reversal	0.098	0.785	-0.083	-0.522
Coskewness	0.035	0.366	-0.328	-2.458
Earnings announcement return	1.099	7.651	0.859	4.703
Earnings forecast revisions	0.922	7.670	0.555	3.637
Share repurchases	0.192	3.000	0.204	2.691
Probability of Informed Trading	1.873	6.136	2.012	6.621

Table 4: The Baseline Staggered Difference-in-Difference for the Accounting-based Anomalies.

This table presents the coefficients from the baseline difference-in-difference regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post_{p,t} + \beta_3 * EDGAR_p * Post_{p,t} + \epsilon_{p,t}$ for the accounting anomaly portfolios. $\alpha_{p,t}$ is the Fama-French five-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group for one of the nine EDGAR implementation phases), zero if the portfolio is constructed with the non-EDGAR filers (the control group for each of nine treatment groups); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first date on which investors start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR (i.e., the effective date), zero if the month is before that date. The dependent variables are the Fama-French five-factor alphas for either decile (1-10) or quintile (1-5) value-weighted (VW) or equal-weighted (EW) portfolios. The sample period is from January 1992 to December 1997. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

The Fama-French Five-Factor Alphas for the Accounting-based Anomalies

	FF5 Alpha 1-5 EW	FF5 Alpha 1-5 VW	FF5 Alpha 1-10 EW	FF5 Alpha 1-10 VW
Post	-0.164**	-0.111	-0.134	-0.237**
	(-2.55)	(-1.37)	(-1.48)	(-2.14)
EDGAR	-0.012	0.134	0.004	0.133
	(-0.12)	(1.59)	(0.03)	(1.28)
Post # EDGAR	-0.253**	-0.343***	-0.316**	-0.371***
	(-2.47)	(-3.49)	(-2.53)	(-3.04)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	126	126	126
Mean of Dep. Var.	0.409	0.299	0.467	0.398

Table 5: The Baseline Staggered Difference-in-Difference for the Non-Accounting-based Anomalies.

This table presents the coefficients from the baseline difference-in-difference regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post_{p,t} + \beta_3 * EDGAR_p * Post_{p,t} + \epsilon_{p,t}$ for the non-accounting anomaly portfolios. $\alpha_{p,t}$ is the Fama-French five-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group for one of the nine EDGAR implementation phases), zero if the portfolio is constructed with the non-EDGAR filers (the control group for each of nine treatment groups); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first date on which investors start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR (i.e., the effective date), zero if the month is before that date. The dependent variables are the Fama-French five-factor alphas for either decile (1-10) or quintile (1-5) value-weighted (VW) or equal-weighted (EW) portfolios. The sample period is from January 1992 to December 1997. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

The Fama-French Five-Factor Alphas for the Non-accounting-based Anomalies

	FF5 Alpha 1-5 EW	FF5 Alpha 1-5 VW	FF5 Alpha 1-10 EW	FF5 Alpha 1-10 VW
Post	-0.093	0.001	-0.114	-0.045
	(-1.48)	(0.02)	(-1.61)	(-0.57)
EDGAR	-0.204***	-0.068	-0.235***	-0.104
	(-2.95)	(-0.80)	(-2.76)	(-1.06)
Post # EDGAR	-0.003	-0.059	-0.016	-0.034
	(-0.03)	(-0.57)	(-0.16)	(-0.29)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	108	108	108	108
Mean of Dep. Var.	0.370	0.224	0.414	0.258

Table 6: The Mean Anomaly Alphas for EDGAR and Non-EDGAR stocks in the Pre-EDGAR and Post-EDGAR periods.

This table shows the average alphas of anomaly portfolios constructed with the EDGAR filers (treatment group) and the non-EDGAR filers (control group) in the pre-EDGAR and post-EDGAR periods. The difference-in-difference coefficients (denoted by "DiD Coeff.") are from the baseline regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post_{p,t} + \beta_3 * EDGAR_p * Post_{p,t} + \epsilon_{p,t}$ for the accounting and non-accounting anomaly portfolios. $\alpha_{p,t}$ is either the Fama-French five or three-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group for one of the nine EDGAR implementation phases, denoted by "EDGAR"), zero if the portfolio is constructed with the non-EDGAR filers (the control group for each of nine treatment groups, denoted by "Non-ED." or "Non-EDGAR"); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first date on which investors start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR (i.e., the effective date), zero if the month is before that date. In Panel A and C, the dependent variables are the Fama-French three-factor alphas for either decile (1-10) or quintile (1-5) value-weighted (VW) or equal-weighted (EW) portfolios. In Panel B and D, the dependent variables are the Fama-French five-factor alphas for the respective portfolio specifications. The sample period is from January 1992 to December 1997. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors.

Panel A: The Average Fama-French Three-Factor Alphas for the Accounting-based Anomalies

	Control Group				Treatment Group				
Acc. FF3	Non-ED. Pre (1)	Non-ED. Post (2)	**Diff. Non-EDGAR (3)=(2)-(1)**		EDGAR Pre (4)	EDGAR Post (5)	**Diff. EDGAR (6)=(5)-(4)**		**DiD Coeff. (7)=(6)-(3)**
1-5 EW	0.558	0.610	0.052		0.502	0.286	-0.216		-0.268
1-5 VW	0.376	0.482	0.106		0.464	0.233	-0.231		-0.337
1-10 EW	0.645	0.698	0.053		0.592	0.305	-0.286		-0.339
1-10 VW	0.547	0.555	0.007		0.641	0.264	-0.377		-0.384
Average	0.532	0.586	0.054		0.550	0.272	-0.278		-0.332

Panel B: The Average Fama-French Five-Factor Alphas for the Accounting-based Anomalies

	Control Group				Treatment Group				
Acc. FF5	Non-ED. Pre (1)	Non-ED. Post (2)	**Diff. Non-EDGAR (3)=(2)-(1)**		EDGAR Pre (4)	EDGAR Post (5)	**Diff. EDGAR (6)=(5)-(4)**		**DiD Coeff. (7)=(6)-(3)**
1-5 EW	0.482	0.473	-0.010		0.471	0.208	-0.263		-0.253
1-5 VW	0.291	0.344	0.053		0.425	0.134	-0.291		-0.343
1-10 EW	0.552	0.535	-0.016		0.556	0.223	-0.333		-0.316
1-10 VW	0.454	0.393	-0.061		0.587	0.155	-0.432		-0.371
Average	0.445	0.436	-0.008		0.510	0.180	-0.330		-0.321

Panel C: The Average Fama-French Three-Factor Alphas for the Non-Accounting-based Anomalies

	Control Group				Treatment Group				
Non-Acc. FF3	Non-ED. Pre (1)	Non-ED. Post (2)	**Diff. Non-EDGAR (3)=(2)-(1)**		EDGAR Pre (4)	EDGAR Post (5)	**Diff. EDGAR (6)=(5)-(4)**		**DiD Coeff. (7)=(6)-(3)**
1-5 EW	0.474	0.615	0.141		0.257	0.352	0.095		-0.045
1-5 VW	0.263	0.442	0.179		0.177	0.231	0.054		-0.124
1-10 EW	0.552	0.677	0.125		0.303	0.366	0.063		-0.061
1-10 VW	0.352	0.489	0.137		0.228	0.247	0.019		-0.117
Average	0.410	0.556	0.146		0.241	0.299	0.058		-0.087

Panel D: The Average Fama-French Five-Factor Alphas for the Non-Accounting-based Anomalies

	Control Group				Treatment Group				
Non-Acc. FF5	Non-ED. Pre (1)	Non-ED. Post (2)	**Diff. Non-EDGAR (3)=(2)-(1)**		EDGAR Pre (4)	EDGAR Post (5)	**Diff. EDGAR (6)=(5)-(4)**		**DiD Coeff. (7)=(6)-(3)**
1-5 EW	0.410	0.528	0.118		0.207	0.322	0.115		-0.003
1-5 VW	0.200	0.342	0.142		0.132	0.215	0.083		-0.059
1-10 EW	0.485	0.580	0.095		0.250	0.328	0.078		-0.017
1-10 VW	0.278	0.355	0.077		0.174	0.217	0.043		-0.034
Average	0.343	0.451	0.108		0.191	0.271	0.080		-0.028

Table 7: A Comparison between the Accounting Anomalies and the Non-accounting Anomalies.

This table reports the coefficients from the triple difference-in-difference regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_{p,t} + \beta_2 * ACC_{p,t} + \beta_3 * Post_{p,t} + \beta_4 * EDGAR_{p,t} * ACC_{p,t} + \beta_5 * ACC_{p,t} * Post_{p,t} + \beta_6 * Post_{p,t} * EDGAR_{p,t} + \beta_7 * EDGAR_{p,t} * ACC_{p,t} * Post_{p,t} + \epsilon_{p,t}$. $\alpha_{p,t}$ is the Fama-French five-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group for one of the nine EDGAR implementation phases), zero if the portfolio is constructed with the non-EDGAR filers (the control group for each of nine treatment groups); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first date on which investors start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR (i.e., the effective date), zero if the month is before that date; $ACC_{p,t}$ is an indicator variable that equals one if the anomaly portfolio is constructed using accounting variables and zero otherwise. The dependent variables are the Fama-French five-factor alphas for the respective portfolio specifications. The sample period is from January 1992 to December 1997. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

The Fama-French Five-Factor Alphas

	FF5 1-5 EW	FF5 1-5 VW	FF5 1-10 EW	FF5 1-10 VW
EDGAR	-0.204***	-0.068	-0.235***	-0.104
	(-2.96)	(-0.80)	(-2.76)	(-1.06)
Accounting	0.073	0.091	0.068	0.177*
	(0.90)	(1.08)	(0.67)	(1.72)
Post	-0.062	-0.011	-0.064	-0.074
	(-0.72)	(-0.11)	(-0.64)	(-0.62)
Post # EDGAR	-0.003	-0.060	-0.017	-0.034
	(-0.04)	(-0.57)	(-0.16)	(-0.29)
Post # Acc	-0.129	-0.090	-0.112	-0.140
	(-1.13)	(-0.73)	(-0.86)	(-0.96)
EDGAR # Acc	0.192*	0.202**	0.239*	0.237*
	(1.91)	(2.00)	(1.86)	(1.86)
Post # EDGAR # Acc	-0.250**	-0.284**	-0.299**	-0.336**
	(-2.22)	(-2.34)	(-2.16)	(-2.20)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	234	234	234	234
Mean of Dep. Variable	0.391	0.265	0.444	0.335

Table 8: The Dynamics of the Profitability Attenuation.

This table reports the coefficient from the regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post1_{p,t} + \beta_3 * Post2_{p,t} + \beta_5 * EDGAR_p * Post1_{p,t} + \beta_6 * EDGAR_p * Post2_{p,t} + \epsilon_{p,t}$, where $\alpha_{p,t}$ is the Fama-French five-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group for a given EDGAR implementation phases), zero if the portfolio is constructed with the non-EDGAR filers (the control group); $Post1_{p,t}$ is an indicator variable that equals one if the month is the first month that investors start trading on the EDGAR filer stocks for a given implementation phase and zero otherwise; and $Post2_{p,t}$ is an indicator variable that equals one if the month falls under the period after the first month investors start trading on the EDGAR filer stocks for a given implementation phase and zero otherwise. The dependent variables are the Fama-French five-factor alphas for the respective portfolio specifications. *Test Beta Diff.* and *Diff. t-stat.* in the bottom section show the results of testing the statistical difference between the coefficient on the *EDGAR # Post1* term and on the *EDGAR # Post2* term. The sample period is from January 1992 to December 1997. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

The Fama-French Five-Factor Alphas

	FF5 Alpha 1-5 EW	FF5 Alpha 1-5 VW	FF5 Alpha 1-10 EW	FF5 Alpha 1-10 VW
EDGAR	-0.012	0.134	0.004	0.133
	(-0.12)	(1.59)	(0.03)	(1.28)
Post1	-0.180**	0.070	-0.244**	-0.100
	(-2.26)	(0.58)	(-2.17)	(-0.63)
Post2	-0.159**	-0.108	-0.121	-0.229**
	(-2.36)	(-1.29)	(-1.28)	(-2.03)
EDGAR # Post1	-0.411	-1.066**	-0.517	-1.119***
	(-0.94)	(-2.35)	(-1.32)	(-2.93)
EDGAR # Post2	-0.249**	-0.323***	-0.310**	-0.349***
	(-2.39)	(-3.23)	(-2.47)	(-2.83)
Test Beta Diff.	0.162	0.743*	0.206	0.770**
Diff. *t*-stat.	0.36	1.63	0.53	1.96
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	118	118	118	118
Mean of Dep. Var	0.409	0.299	0.467	0.398

Table 9: The Information Turnover Ratio and the Attenuation in Profitability of the Accounting-based Anomaly Portfolios

This table shows the difference-in-difference coefficients from the baseline difference-in-difference regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post_{p,t} + \beta_3 * EDGAR_p * Post_{p,t} + \epsilon_{p,t}$ for the accounting anomaly portfolios with high information turnover ratios and with low information turnover ratios. The information turnover ratio is defined as the total number of new incoming stocks divided by the total number of stocks in the existing portfolio when the anomaly portfolio updates its signal and rebalances its stocks. Information turnover ratios for the accounting-based anomalies are computed for the pre-EDGAR period from October 1, 1983, to September 30, 1993. The anomaly portfolios used to compute the information turnover ratios include all the stocks in the NYSE-AMEX-NASDAQ universe that pass the stock-level filters mentioned in Section 3.B. After categorizing the anomalies into accounting-based and non-accounting-based anomalies, I sort the anomalies in each category with respect to their information turnover ratio rank percentile. The anomalies that are over 50% rank percentile are defined as the "High Turnover" anomalies and the rest as the "Low Turnover". $\alpha_{p,t}$ is the Fama-French five-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group), zero if the portfolio is constructed with the non-EDGAR filers (the control group); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first date on which investors start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR (i.e., the effective date), zero if the month is before that date. The dependent variables are the Fama-French five-factor alphas for the respective portfolio specifications. Three outlier anomalies from each turnover category in each portfolio specification which distort results are dropped from the sample. The sample period is from January 1992 to December 1997. *Low-High* and *Diff. t-stat* in the bottom section show the results of testing statistical difference between the difference-in-difference coefficients on the *High Turnover* and those on *Low Turnover* using the triple difference-in-difference Equation (2) introduced in Section 4.A, after replacing the $ACC_{p,t}$ term in Equation (2) with an indicator variable that identifies high versus low turnover anomaly. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust t-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

The Fama-French Five-Factor Alphas

Accounting Anomalies	FF5 1-5 EW DiD Coeff.	FF5 1-5 VW DiD Coeff.	FF5 1-10 EW DiD Coeff.	FF5 1-10 VW DiD Coeff.
High Turnover	-0.424***	-0.448***	-0.502***	-0.520***
	(-3.45)	(-3.59)	(-3.07)	(-3.10)
Low Turnover	-0.087	-0.223*	-0.156	-0.228*
	(-0.77)	(-1.95)	(-1.18)	(-1.81)
Low-High	0.337***	0.226*	0.346**	0.292*
Diff. *t*-stat.	2.99	1.79	2.21	1.79
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	120	120	120	120

Table 10: The Information Turnover Ratio and the Attenuation in Profitability of the Non-Accounting-based Anomaly Portfolios

This table shows the difference-in-difference coefficients from the baseline difference-in-difference regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post_{p,t} + \beta_3 * EDGAR_p * Post_{p,t} + \epsilon_{p,t}$ for the non-accounting anomaly portfolios with high information turnover ratios and with low information turnover ratios. The information turnover ratio is defined as the total number of new incoming stocks divided by the total number of stocks in the existing portfolio when the anomaly portfolio updates its signal and rebalances its stocks. Information turnover ratios for non-accounting anomalies are computed for the pre-EDGAR period from October 1, 1983, to September 30, 1993. The anomaly portfolios used to compute the information turnover ratios include all the stocks in the NYSE-AMEX-NASDAQ universe that pass the stock-level filters mentioned in Section 3.B. After categorizing the anomalies into accounting-based and non-accounting-based anomalies, I sort the anomalies in each category with respect to their information turnover ratio rank percentile. The anomalies that are over 50% rank percentile are defined as the "High Turnover" anomalies and the rest as the "Low Turnover". $\alpha_{p,t}$ is the Fama-French five-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group), zero if the portfolio is constructed with the non-EDGAR filers (the control group); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first date on which investors start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR (i.e., the effective date), zero if the month is before that date. The dependent variables are the Fama-French five-factor alphas for the respective portfolio specifications. The sample period is from January 1992 to December 1997. *Low-High* and *Diff. t-stat* in the bottom section show the results of testing statistical difference between the difference-in-difference coefficients on the *High Turnover* and those on *Low Turnover* using the triple difference-in-difference Equation (2) introduced in Section 4.A, after replacing the $ACC_{p,t}$ term in Equation (2) with an indicator variable that identifies high versus low turnover anomaly. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

The Fama-French Five-Factor Alphas

Non-Accounting Anomalies	FF5 1-5 EW DiD Coeff.	FF5 1-5 VW DiD Coeff.	FF5 1-10 EW DiD Coeff.	FF5 1-10 VW DiD Coeff.
High Turnover	0.060	0.106	0.029	0.084
	(0.49)	(0.74)	(0.20)	(0.50)
Low Turnover	-0.060	-0.216*	-0.052	-0.142
	(-0.59)	(-1.68)	(-0.45)	(-1.03)
Low-High	-0.119	-0.321*	-0.080	-0.225
Diff. S.E.	-0.85	-1.83	-0.47	-1.14
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	108	108	108	108

Table 11: The Information Availability and the Accounting-based Anomaly Portfolio Profitability

This table shows the Fama-MacBeth regression results for different information availability groups of accounting-based anomaly portfolios. To estimate the regression coefficients, I first compute the average number of analyst coverage and the firm size (the market capitalization of equity) for all EDGAR filers from January 1990 to December 1992, and then I sort the firms into their respective rank percentile of the two measures. I assign the stocks in the top 50% percentile rank to the high analyst coverage (large size) group and the bottom 50% to the low analyst coverage (small size) group. For each information availability group, I run separate Fama-MacBeth regression. In the first pass, I estimate a cross-section regression for 126 accounting-based anomalies, for every EDGAR implementation phase, and for every month from January 1992 to December 1997 (i.e., estimate cross-sectional regressions by-anomaly, by-group, by-month) using the following equation $R_{i,a,p,t+1} = \alpha + \beta * SignalPercentile_{i,a,p,t} + \epsilon_{i,a,p,t}$, where i is an EDGAR-filer assigned to a given EDGAR implementation phase p for a given accounting-based anomaly a in the month *t*; $R_{i,a,p,t+1}$ is the simple monthly return of the following month for the given EDGAR-filer (stock) *i*; β is the factor loading; $SignalPercentile_{i,a,p,t}$ is the stock i's signal (characteristic) percentile within the EDGAR filers assigned to implementation phase *p* for anomaly *a* in month *t*. The first pass regression creates a panel of monthly beta estimates for each accounting-based anomaly portfolio constructed with stocks in a given implementation phase. To obtain the time-series average of these beta estimates and to measure the effect of EDGAR implementation on all the $\hat{\beta}$, I estimate the following second-pass regression $\widehat{\beta_{a,p,t}} = \gamma_t + \ \delta * Post_{a,p,t} + \epsilon_{a,t}$, where a is the anomaly portfolio formed on stocks in a given implementation phase *p*; $\widehat{\beta_{a,p,t}}$ are the monthly beta estimates from the first-pass; γ_t is the monthly time fixed effect; $Post_{a,p,t}$ is an indicator variable that equals one if the month is after the first date (the effective date for implementation phase *p*) on which investors start to trade on the anomaly a using the latest information on new EDGAR filers that they obtain from EDGAR and equals zero otherwise; and δ is the coefficient of primary interest. Finally, to test the statistical difference between the high and low information availability groups, in other words, to test the statistical difference between $\delta_{LowInfoGroup}$ and $\delta_{HighInfoGroup}$, I run the following second pass $\widehat{\beta_{a,p,t}} = \gamma_t + \delta_1 * Post_{a,p,t} + \delta_2 * LowInfo_{a,p,t} + \ \delta_3 * Post_{a,p,t} * LowInfo_{a,p,t} + \epsilon_{a,t}$, where $LowInfo_{a,p,t}$ is an indicator variable that equals one if classified as low information availability group and zero otherwise. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

Panel A: High Analyst coverage stocks vs low coverage EDGAR filers

	(1) High Analyst Coverage EDGAR Filers	(2) Low Analyst Coverage EDGAR Filers	(1) – (2) Difference
Post	0.046	-0.728***	0.774***
	(0.24)	(-3.52)	(2.81)
Mon. FE	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	126	126

Panel B: Large firm stocks vs low coverage EDGAR filers

	(1) Large Size EDGAR Filers	(2) Small Size EDGAR Filers	(1) – (2) Difference
Post	-0.099	-0.704***	0.605**
	(-0.47)	(-3.29)	(2.03)
Mon. FE	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	126	126

Figure 1: The Attenuation of the Accounting-based Anomaly Profitability in Response to the EDGAR Implementation.

This figure shows the changes in average monthly Fama-French alphas (both the three-factor and five-factor alphas) for the accounting-based anomaly portfolios in response to the staggered EDGAR implementation. The two orange horizontal lines represent the pre-EDGAR period versus the post-EDGAR period average alphas of the accounting anomaly portfolios constructed with EDGAR filers (treatment group) over the sample period with respect to all nine implementation phases studied in this paper. Similarly, the two blue horizontal lines show the pre-EDGAR period versus the post-EDGAR average alphas of accounting anomaly portfolio consisting of non-EDGAR filers (the control group) over the sample period with respect to all nine implementation phases. The grey vertical dotted line in the middle represents the nine effective dates for implementation phase 1 (CF-01) to implementation phase 9 (CF-09). Therefore, the left side of the grey vertical dotted line shows the pre-EDGAR period, while the right side of the grey vertical dotted line shows the post-EDGAR period. The sample period is from January 1992 to December 1997.

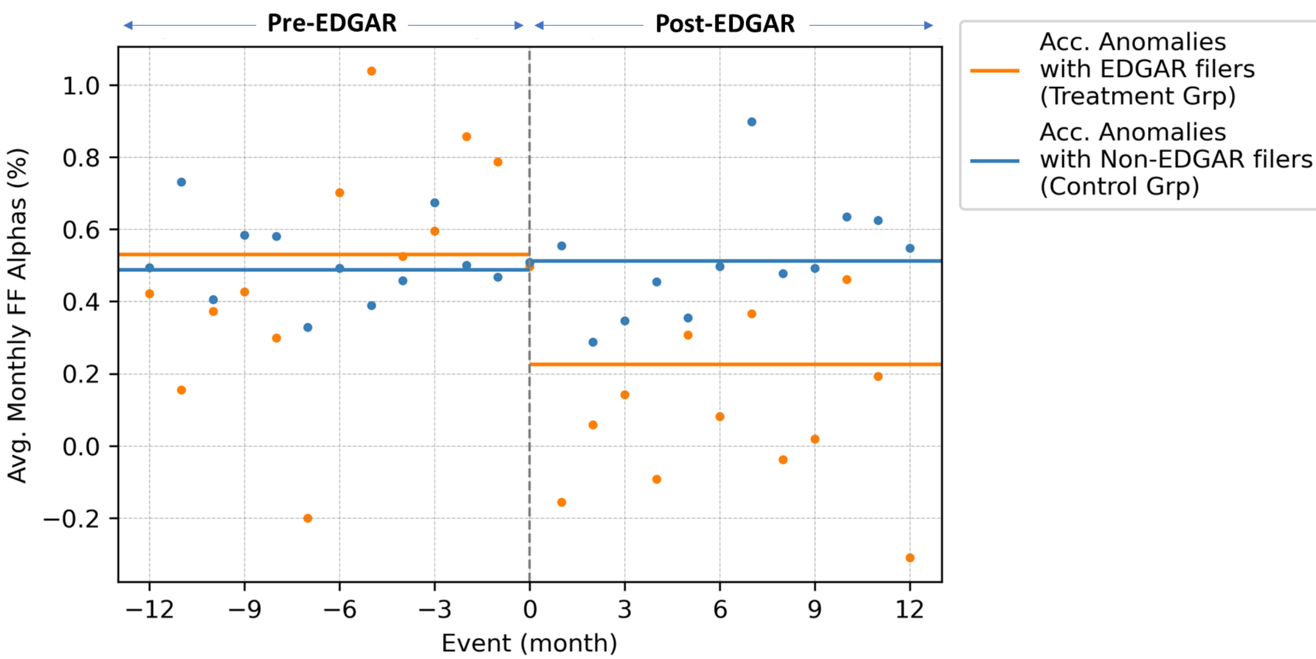
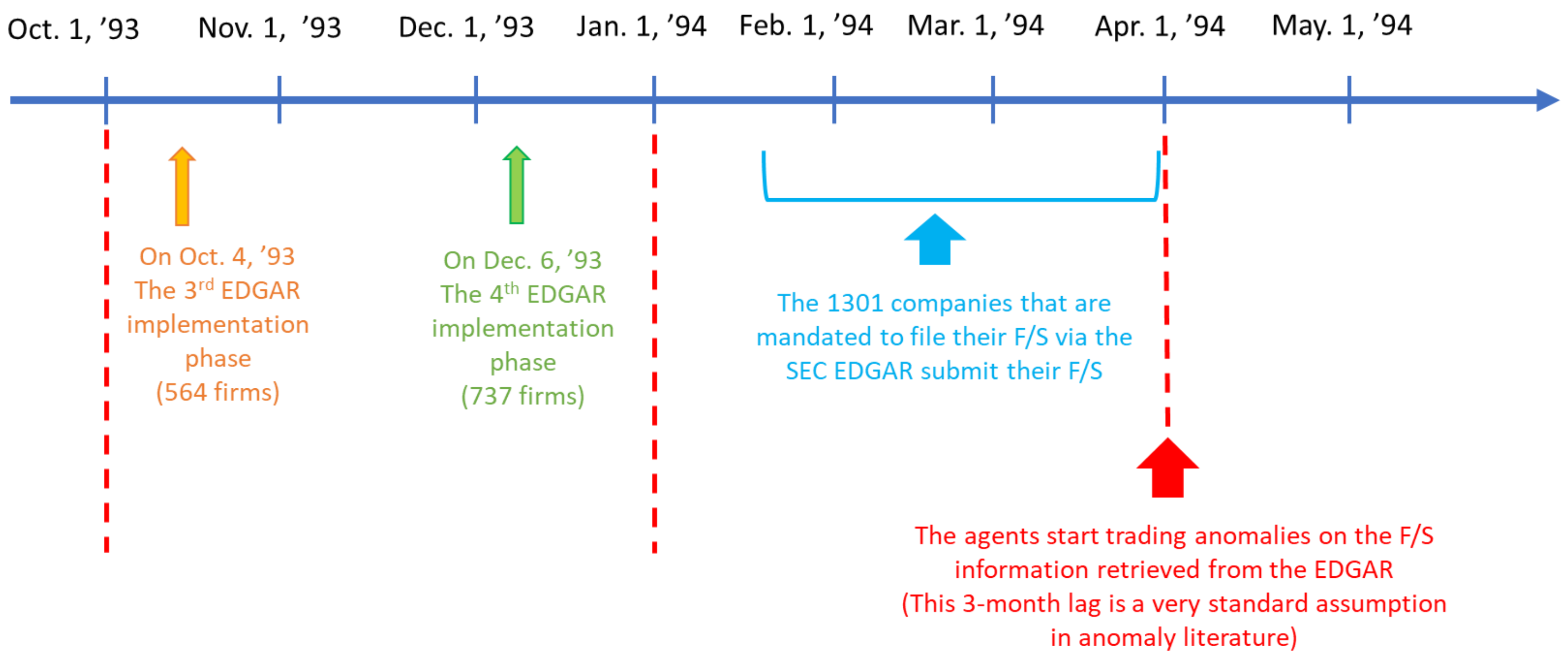


Figure 2: An Example of Timeline of Events.

This figure shows an example of how the timeline of events is constructed. I assume that investors must wait for three months before they can start trading on the EDGAR filer stocks, using the previous quarter's financial information collected from the EDGAR. Note that this is a standard assumption in this literature.



Appendix

Table A.1: Summary Statistics of the Most Cited Accounting-based Anomalies (The Average Fama-French Three-Factor Alphas)

This table presents the time-series averages of the Fama-French three-factor alpha for the decile (denoted by "1-10") accounting-based anomaly portfolios studied in this paper. Out of a total of 320 anomalies documented in Chen and Zimmermann (2020), only 234 survive the set of filters I apply: stock price, common stocks, market capitalization, pair-wise correlation, and alpha filters. Out of 234 remaining anomalies, 126 belong to the accounting-based category, and 20 accounting anomalies with the most citations are presented in this table (126 accounting-based anomalies are sorted on the total number of times the paper that first documents the anomaly has been cited as of February 2021). The sample period is from January 1992 to December 1997. The unit is in percentage. The quarterly versions (denoted by "(q)") of the original anomalies are generated following Chen and Zimmermann (2020). The equal-weighted portfolio is denoted by "EW", whereas the value-weighted portfolio is denoted by "VW". The citation numbers are collected via Google Scholar. See the Internet Appendix for the summary statistics of all the anomalies, including those of the Fama-French three-factor alphas for quintile portfolios.

Accounting-based Anomaly	FF3 Alpha 1-10 EW		FF3 Alpha 1-10 VW	
	Mean	*t*-stat	Mean	*t*-stat
Book leverage (q)	-0.757	-6.485	-0.657	-4.530
Annual sales growth	0.506	3.604	0.205	1.209
Cash flow to market	0.540	2.381	0.370	1.527
Accruals	1.672	10.847	1.686	9.402
Earnings persistence	-0.214	-0.929	0.214	0.875
Earnings predictability	1.855	6.932	1.965	6.990
Value relevance of earnings	0.082	0.400	0.034	0.156
Book to market using December ME	0.837	6.098	-0.319	-1.942
Failure probability	0.590	4.609	0.680	4.165
Failure probability (June)	0.368	2.789	0.566	3.473
Advertising Expense	-0.061	-0.297	0.076	0.327
R&D to sales	-0.328	-1.227	-0.287	-0.961
Kaplan Zingales index	1.294	9.084	0.722	4.339
Gross profits/total assets	0.776	5.377	1.316	8.046
Gross profits/total assets (lag)	0.230	1.493	0.625	3.541
Investment to revenue	0.352	2.307	0.170	0.967
Abnormal Accruals	0.249	1.625	0.187	0.990
Change in current operating assets	0.630	5.103	0.641	4.168
Change in equity to assets	0.729	5.207	0.521	3.043
Change in financial liabilities	0.252	1.969	0.341	2.170

Table A.2: Summary Statistics of the Most Cited Non-Accounting-based Anomalies (The Average Fama-French Three-Factor Alphas)

This table presents the time-series averages of the Fama-French three-factor alpha for the decile (denoted by "1-10") non-accounting-based anomaly portfolios studied in this paper. Out of a total of 320 anomalies documented in Chen and Zimmermann (2020), only 234 survive the set of filters I apply: stock price, common stocks, market capitalization, pair-wise correlation, and alpha filters. Out of 234 remaining anomalies, 108 belong to the non-accounting-based category, and 20 non-accounting anomalies with most citations are presented in this table (108 non-accounting-based anomalies are sorted on the total number of times the paper that first documents the anomaly has been cited as of February 2021). The sample period is from January 1992 to December 1997. The unit is in percentage. The quarterly versions (denoted by "(q)") of the original anomalies are generated following Chen and Zimmermann (2020). The equal-weighted portfolio is denoted by "EW", whereas the value-weighted portfolio is denoted by "VW". The citation numbers are collected via Google Scholar. See the Internet Appendix for the summary statistics of all the anomalies, including those of the Fama-French three-factor alphas for quintile portfolios.

Non-Accounting Anomaly	FF3 Alpha 1-10 EW		FF3 Alpha 1-10 VW	
	Mean	*t*-stat	Mean	*t*-stat
Momentum (12 month)	0.939	6.186	0.593	3.016
Momentum (6 month)	0.857	6.124	0.286	1.614
Momentum-Reversal	0.736	5.382	0.602	3.395
Governance Index	-0.133	-0.911	0.188	1.128
Size	0.625	6.569	0.608	6.273
Bid-ask spread	0.578	5.756	0.619	5.495
Pastor-Stambaugh liquidity beta	-0.130	-0.989	0.228	1.293
Initial Public Offerings	0.009	0.063	-0.308	-1.982
IPO and age	-0.303	-1.046	-0.058	-0.197
Systematic volatility	0.574	4.937	0.984	6.271
Earnings-to-Price Ratio	-0.448	-2.540	-0.451	-2.401
Illiquidity-illiquidity beta	-0.406	-2.172	-0.582	-2.791
Net liquidity beta	-0.478	-2.666	-0.554	-2.902
Return-market illiquidity beta	0.715	3.952	0.736	3.730
Short term reversal	-0.106	-0.837	-0.292	-1.807
Coskewness	0.200	2.043	-0.171	-1.264
Earnings announcement return	1.345	9.306	1.119	6.077
Earnings forecast revisions	0.992	8.116	0.760	4.877
Share repurchases	0.221	3.419	0.230	2.978
Probability of Informed Trading	1.497	4.850	1.610	5.243

Table A.3: Descriptive Statistics for the Treatment Group versus the Control Group Portfolios

This table shows the descriptive statistics for the sample of stocks in the treatment and control groups. Panel A and B report the statistics for decile (1-10) anomaly portfolios. Panel C and D report the statistics for the quintile (1-5) anomaly portfolio. *Amihud illiquid.* is the past twelve-month average of daily return divided by turnover. *Bid-ask spread* is estimated following Shane and Schultz (2012). The bid-ask spread estimates are from Shane Corwin's website. *BM* is the book value of common equity divided by the market value of common equity. *CapEx* is the ratio of capital expenditure to total asset. *Firm age* is the number of years since the first trading date on CRSP. The earliest possible year is set to 1925. *Firm size* is the firm's market value (stock price times shares outstanding). *Idio. volatility* is the standard deviation of residuals from the Fama-French three-factor model using the past month of daily data. *Inst. own* is the number of shares held by institutional investors divided by the number of shares outstanding. *Op. leverage* is the sum of administrative expenses and cost of goods sold divided by total assets. *Price* is the end-of-month closing price of the stock. *ROA* is the ratio of quarterly net income to lagged total assets. *Sales growth* is the ratio of quarterly sales relative to previous sales. *Trad. volume* is the total number of shares of a stock traded in a given month during the sample period. *Turnover ratio* is the monthly trading volume divided by the shares outstanding.

Panel A: Decile Treated Portfolio

	Accounting Anomalies				Non-Accounting Anomalies			
	Pre-EDGAR		Post-EDGAR		Pre-EDGAR		Post-EDGAR	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Amihud illiquid.	0.037	0.170	0.023	0.107	0.041	0.172	0.021	0.099
Bid-ask spread	0.012	0.010	0.009	0.007	0.012	0.010	0.009	0.008
BM	0.514	0.427	0.486	0.399	0.532	0.403	0.497	0.372
CapEx	0.067	0.071	0.067	0.067	0.067	0.064	0.068	0.062
Firm age	22.22	14.87	24.55	15.28	22.60	15.05	25.25	15.46
Firm size	14.85	46.94	31.85	104.9	14.89	46.61	31.68	98.60
Idio. volatility	0.022	0.012	0.020	0.011	0.021	0.011	0.019	0.011
Inst. own	0.419	0.209	0.452	0.218	0.423	0.210	0.457	0.215
Op. leverage	1.079	0.866	1.049	0.894	1.080	0.824	1.053	0.837
Price	31.76	322.3	51.20	888.0	30.06	271.8	49.92	845.1
ROA	0.005	0.076	0.009	0.056	0.009	0.055	0.011	0.043
Sales growth	0.445	5.307	0.533	20.46	0.287	3.626	0.348	15.09
Short interest	0.183	0.326	0.212	0.344	0.162	0.303	0.198	0.327
Trad. volume	3.667	7.464	7.066	18.68	3.308	6.770	6.556	16.31
Turnover ratio	0.101	0.127	0.104	0.147	0.092	0.120	0.097	0.137

Panel B: Decile Controlled Portfolio

	Accounting Anomalies				Non-Accounting Anomalies			
	Pre-EDGAR		Post-EDGAR		Pre-EDGAR		Post-EDGAR	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Amihud illiquid.	0.044	0.158	0.035	0.152	0.050	0.173	0.036	0.148
Bid-ask spread	0.016	0.012	0.013	0.010	0.015	0.012	0.013	0.010
BM	0.449	0.410	0.407	0.356	0.478	0.386	0.432	0.345
CapEx	0.070	0.082	0.074	0.090	0.072	0.079	0.076	0.086
Firm age	15.47	15.21	14.79	14.70	14.46	15.67	14.33	15.06
Firm size	6.857	21.01	9.158	36.12	7.156	20.17	9.849	35.03
Idio. volatility	0.026	0.013	0.026	0.014	0.025	0.013	0.025	0.014
Inst. own	0.374	0.216	0.398	0.230	0.384	0.216	0.403	0.231
Op. leverage	1.000	0.841	0.949	0.835	1.000	0.801	0.945	0.793
Price	21.19	124.7	24.12	264.9	21.80	100.4	25.34	280.2
ROA	-0.005	0.108	-0.003	0.081	0.005	0.074	0.005	0.063
Sales growth	0.974	20.31	1.183	31.66	0.524	12.06	0.693	21.92
Short interest	0.199	0.357	0.231	0.402	0.166	0.318	0.204	0.359
Trad. volume	2.736	5.436	4.238	10.94	2.459	4.839	3.864	9.696
Turnover ratio	0.126	0.152	0.149	0.185	0.116	0.146	0.136	0.175

Panel C: Quintile Treated Portfolio

	Accounting Anomalies				Non-Accounting Anomalies			
	Pre-EDGAR		Post-EDGAR		Pre-EDGAR		Post-EDGAR	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Amihud illiquid.	0.036	0.156	0.021	0.098	0.039	0.161	0.021	0.094
Bid-ask spread	0.012	0.010	0.009	0.007	0.012	0.010	0.009	0.007
BM	0.517	0.408	0.488	0.379	0.531	0.396	0.494	0.363
CapEx	0.067	0.067	0.068	0.064	0.067	0.064	0.069	0.062
Firm age	22.89	14.99	25.27	15.34	22.91	15.04	25.53	15.45
Firm size	15.08	47.33	31.18	100.0	15.17	46.86	32.21	99.43
Idio. volatility	0.021	0.011	0.019	0.011	0.020	0.011	0.018	0.010
Inst. own	0.423	0.209	0.455	0.215	0.425	0.209	0.459	0.214
Op. leverage	1.099	0.839	1.063	0.863	1.079	0.822	1.050	0.834
Price	30.54	286.4	48.48	824.3	30.27	272.2	48.58	812.4
ROA	0.008	0.065	0.011	0.048	0.009	0.052	0.012	0.042
Sales growth	0.354	4.430	0.416	17.13	0.279	3.580	0.340	15.04
Short interest	0.167	0.308	0.196	0.322	0.158	0.296	0.193	0.318
Trad. volume	3.426	6.973	6.506	16.85	3.251	6.603	6.398	15.74
Turnover ratio	0.096	0.122	0.099	0.140	0.090	0.117	0.095	0.133

Panel D: Quintile Controlled Portfolio

	Accounting Anomalies				Non-Accounting Anomalies			
	Pre-EDGAR		Post-EDGAR		Pre-EDGAR		Post-EDGAR	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Amihud illiquid.	0.042	0.146	0.034	0.137	0.048	0.165	0.035	0.141
Bid-ask spread	0.015	0.011	0.013	0.009	0.015	0.011	0.013	0.009
BM	0.466	0.399	0.424	0.346	0.480	0.383	0.434	0.343
CapEx	0.071	0.079	0.075	0.087	0.071	0.078	0.076	0.085
Firm age	16.64	15.58	15.65	15.04	14.89	15.73	14.66	15.10
Firm size	7.380	21.67	9.540	36.06	7.309	20.59	9.993	35.34
Idio. volatility	0.025	0.013	0.025	0.013	0.024	0.013	0.025	0.014
Inst. own	0.383	0.215	0.402	0.229	0.386	0.215	0.405	0.230
Op. leverage	1.013	0.818	0.953	0.812	1.002	0.800	0.945	0.793
Price	22.03	126.5	25.16	289.2	21.95	101.4	25.47	280.3
ROA	0.001	0.090	0.002	0.070	0.005	0.071	0.005	0.061
Sales growth	0.701	15.72	0.901	26.67	0.499	11.24	0.658	20.91
Short interest	0.180	0.333	0.212	0.370	0.161	0.310	0.201	0.352
Trad. volume	2.697	5.357	4.078	10.45	2.439	4.802	3.806	9.496
Turnover ratio	0.121	0.148	0.141	0.180	0.011	0.143	0.133	0.172

Table A.4: The Baseline Staggered Difference-in-Difference for the Accounting-based Anomalies from Chen and Zimmermann (2020)

This table presents the coefficients from the baseline difference-in-difference regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post_{p,t} + \beta_3 * EDGAR_p * Post_{p,t} + \epsilon_{p,t}$ for the accounting anomaly portfolios that are listed in Chen and Zimmermann (2020). $\alpha_{p,t}$ is the Fama-French three-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group), zero if the portfolio is constructed with the non-EDGAR filers (the control group); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first date on which investors start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR (i.e., the effective date), zero if the month is before that date. In Panel A, the dependent variables are the Fama-French three-factor alphas for either decile (1-10) or quintile (1-5) value-weighted (VW) or equal-weighted (EW) portfolios. In Panel B, the dependent variables are the Fama-French five-factor alphas for the respective portfolio specifications. The sample period is from January 1992 to December 1997. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust t-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

Panel A: The Fama-French Three-Factor Alphas for the Accounting-based Anomalies

	FF3 Alpha 1-5 EW	FF3 Alpha 1-5 VW	FF3 Alpha 1-10 EW	FF3 Alpha 1-10 VW
Post	-0.128*	-0.104	-0.121	-0.237**
	(-1.98)	(-1.32)	(-1.33)	(-2.14)
EDGAR	-0.052	0.053	-0.073	0.020
	(-0.56)	(0.69)	(-0.64)	(0.20)
Post # EDGAR	-0.260**	-0.284***	-0.306**	-0.288**
	(-2.51)	(-3.10)	(-2.45)	(-2.41)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	127	127	127	127
Mean of Dep. Var.	0.459	0.375	0.519	0.478

Panel B: The Fama-French Five-Factor Alphas for the Accounting-based Anomalies

	FF5 Alpha 1-5 EW	FF5 Alpha 1-5 VW	FF5 Alpha 1-10 EW	FF5 Alpha 1-10 VW
Post	-0.167**	-0.136*	-0.152	-0.267**
	(-2.54)	(-1.68)	(-1.64)	(-2.38)
EDGAR	-0.012	0.094	-0.018	0.052
	(-0.12)	(1.14)	(-0.16)	(0.50)
Post # EDGAR	-0.248**	-0.293***	-0.290**	-0.278**
	(-2.40)	(-3.09)	(-2.33)	(-2.29)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	127	127	127	127
Mean of Dep. Var.	0.380	0.282	0.423	0.364

Table A.5: The Baseline Staggered Difference-in-Difference for the Accounting-based Anomalies. (FF3 Alphas)

This table presents the coefficients from the baseline difference-in-difference regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post_{p,t} + \beta_3 * EDGAR_p * Post_{p,t} + \epsilon_{p,t}$ for the accounting anomaly portfolios. $\alpha_{p,t}$ is the Fama-French three-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group), zero if the portfolio is constructed with the non-EDGAR filers (the control group); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first date on which investors start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR (i.e., the effective date), zero if the month is before that date. The dependent variables are the Fama-French three-factor alphas for either decile (1-10) or quintile (1-5) value-weighted (VW) or equal-weighted (EW) portfolios. The sample period is from January 1992 to December 1997. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

The Fama-French Three-Factor Alphas for the Accounting-based Anomalies

	FF3 Alpha 1-5 EW	FF3 Alpha 1-5 VW	FF3 Alpha 1-10 EW	FF3 Alpha 1-10 VW
Post	-0.125*	-0.083	-0.103	-0.215*
	(-1.98)	(-1.04)	(-1.15)	(-1.96)
EDGAR	-0.056	0.088	-0.054	0.093
	(-0.62)	(1.12)	(-0.47)	(0.93)
Post # EDGAR	-0.268**	-0.337***	-0.339***	-0.384***
	(-2.61)	(-3.53)	(-2.68)	(-3.21)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	126	126	126
Mean of Dep. Var.	0.490	0.389	0.561	0.502

Table A.6: The Baseline Staggered Difference-in-Difference for the Non-Accounting-based Anomalies. (FF3 Alphas)

This table presents the coefficients from the baseline difference-in-difference regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post_{p,t} + \beta_3 * EDGAR_p * Post_{p,t} + \epsilon_{p,t}$ for the non-accounting anomaly portfolios. $\alpha_{p,t}$ is the Fama-French three-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group), zero if the portfolio is constructed with the non-EDGAR filers (the control group); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first date on which investors start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR (i.e., the effective date), zero if the month is before that date. The dependent variables are the Fama-French three-factor alphas for either decile (1-10) or quintile (1-5) value-weighted (VW) or equal-weighted (EW) portfolios. The sample period is from January 1992 to December 1997. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

The Fama-French Three-Factor Alphas for the Non-accounting-based Anomalies

	FF3 Alpha 1-5 EW	FF3 Alpha 1-5 VW	FF3 Alpha 1-10 EW	FF3 Alpha 1-10 VW
Post	-0.104	0.008	-0.122*	-0.045
	(-1.64)	(0.10)	(-1.74)	(-0.57)
EDGAR	-0.217***	-0.086	-0.249***	-0.125
	(-3.26)	(-1.01)	(-3.06)	(-1.31)
Post # EDGAR	-0.045	-0.124	-0.061	-0.117
	(-0.53)	(-1.19)	(-0.61)	(-0.99)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	108	108	108	108
Mean of Dep. Var.	0.428	0.280	0.478	0.331

Table A.7: A Comparison between the Accounting Anomalies and the Non-accounting Anomalies. (FF3 Alphas)

This table reports the coefficients from the triple difference-in-difference regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_{p,t} + \beta_2 * ACC_{p,t} + \beta_3 * Post_{p,t} + \beta_4 * EDGAR_{p,t} * ACC_{p,t} + \beta_5 * ACC_{p,t} * Post_{p,t} + \beta_6 * Post_{p,t} * EDGAR_{p,t} + \beta_7 * EDGAR_{p,t} * ACC_{p,t} * Post_{p,t} + \epsilon_{p,t}$. $\alpha_{p,t}$ is the Fama-French three-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group), zero if the portfolio is constructed with the non-EDGAR filers (the control group); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first date on which investors start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR (i.e., the effective date), zero if the month is before that date; $ACC_{p,t}$ is an indicator variable that equals one if the anomaly portfolio is constructed using accounting variables and zero otherwise. The dependent variables are the Fama-French three-factor alphas for either decile (1-10) or quintile (1-5) value-weighted (VW) or equal-weighted (EW) portfolios. The sample period is from January 1992 to December 1997. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

Panel A: The Fama-French Three-Factor Alphas

	FF3 1-5 EW	FF3 1-5 VW	FF3 1-10 EW	FF3 1-10 VW
EDGAR	-0.217***	-0.086	-0.249***	-0.125
	(-3.26)	(-1.01)	(-3.07)	(-1.31)
Accounting	0.085	0.114	0.094	0.196*
	(1.04)	(1.36)	(0.91)	(1.92)
Post	-0.067	-0.003	-0.072	-0.068
	(-0.78)	(-0.03)	(-0.71)	(-0.57)
Post # EDGAR	-0.045	-0.124	-0.061	-0.117
	(-0.53)	(-1.19)	(-0.62)	(-0.99)
Post # Acc	-0.089	-0.072	-0.073	-0.129
	(-0.76)	(-0.58)	(-0.53)	(-0.87)
EDGAR # Acc	0.161*	0.174*	0.195	0.218*
	(1.67)	(1.75)	(1.58)	(1.80)
Post # EDGAR # Acc	-0.223*	-0.213*	-0.277**	-0.267*
	(-1.98)	(-1.72)	(-2.00)	(-1.76)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	234	234	234	234
Mean of Dep. Variable	0.462	0.340	0.524	0.426

Table A.8: The Dynamics of the Profitability Attenuation. (FF3 Alphas)

This table reports the coefficients from the regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post^1_{p,t} + \beta_3 * Post^2_{p,t} + \beta_5 * EDGAR_p * Post^1_{p,t} + \beta_6 * EDGAR_p * Post^2_{p,t} + \epsilon_{p,t}$, where $\alpha_{p,t}$ is the Fama-French three-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group), zero if the portfolio is constructed with the non-EDGAR filers (the control group); $Post^1_{p,t}$ is an indicator variable that equals one if the month is the first month that investors start trading on the EDGAR filer stocks for a given implementation phase and zero otherwise; and $Post^2_{p,t}$ is an indicator variable that equals one if the month falls under the period after the first month investors start trading on the EDGAR filer stocks for a given implementation phase and zero otherwise. The dependent variables are the Fama-French three-factor alphas for either decile (1-10) or quintile (1-5) value-weighted (VW) or equal-weighted (EW) portfolios. *Test Beta Diff.* and *Diff. t-stat* in the bottom section show the results of testing statistical difference between the coefficient on the *EDGAR # Post1* term and on the *EDGAR # Post2* term. The sample period is from January 1992 to December 1997. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

Panel A: The Fama-French Three-Factor Alphas

	FF3 Alpha 1-5 EW	FF3 Alpha 1-5 VW	FF3 Alpha 1-10 EW	FF3 Alpha 1-10 VW
EDGAR	-0.056	0.088	-0.054	0.093
	(-0.62)	(1.12)	(-0.47)	(0.93)
Post1	-0.145*	0.120	-0.221**	-0.071
	(-1.87)	(0.90)	(-2.08)	(-0.44)
Post2	-0.120*	-0.080	-0.092	-0.212*
	(-1.84)	(-0.98)	(-0.99)	(-1.91)
EDGAR # Post1	-0.400	-1.104**	-0.420	-0.967***
	(-0.87)	(-2.08)	(-1.18)	(-2.98)
EDGAR # Post2	-0.264**	-0.315***	-0.336**	-0.367***
	(-2.52)	(-3.23)	(-2.63)	(-3.00)
Test Beta Diff.	0.136	0.789	0.084	0.600*
Diff. *t*-stat	0.29	1.46	0.23	1.73
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	118	118	118	118
Mean of Dep. Variable	0.490	0.389	0.561	0.502

Table A.9: The Information Turnover Ratio and the Attenuation in Profitability of the Accounting-based Anomaly Portfolios. (FF3 Alphas)

This table shows the difference-in-difference coefficients from the baseline difference-in-difference regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post_{p,t} + \beta_3 * EDGAR_p * Post_{p,t} + \epsilon_{p,t}$ for the accounting anomaly portfolios with high information turnover ratios and with low information turnover ratios. The information turnover ratio is defined as the total number of new incoming stocks divided by the total number of stocks in the existing portfolio when the anomaly portfolio updates its signal and rebalances its stocks. Information turnover ratios for the accounting-based anomalies are computed for the pre-EDGAR period from October 1, 1983, to September 30, 1993. The anomaly portfolios used to compute the information turnover ratios include all the stocks in the NYSE-AMEX-NASDAQ universe that pass the stock-level filters mentioned in Section 3.B. After categorizing the anomalies into accounting-based and non-accounting-based anomalies, I sort the anomalies in each category with respect to their information turnover ratio rank percentile. The anomalies that are over 50% rank percentile are defined as the "High Turnover" anomalies and the rest as the "Low Turnover". $\alpha_{p,t}$ is the Fama-French three-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group), zero if the portfolio is constructed with the non-EDGAR filers (the control group); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first date on which investors start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR (i.e., the effective date), zero if the month is before that date. The dependent variables are the Fama-French three-factor alphas for the respective portfolio specifications. Three outlier anomalies from each turnover category in each portfolio specification which distort results are dropped from the sample. The sample period is from January 1992 to December 1997. *Low-High* and *Diff. t-stat* in the bottom section show the results of testing statistical difference between the difference-in-difference coefficient on the *High Turnover* anomaly portfolios and those of the *Low Turnover*. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

The Fama-French Three-Factor Alphas

Accounting	FF3 1-5 EW DiD Coeff.	FF3 1-5 VW DiD Coeff.	FF3 1-10 EW DiD Coeff.	FF3 1-10 VW DiD Coeff.
High Turnover	-0.435***	-0.431***	-0.519***	-0.529***
	(-3.53)	(-3.50)	(-3.17)	(-3.23)
Low Turnover	-0.101	-0.193*	-0.150	-0.225*
	(-0.89)	(-1.77)	(-1.13)	(-1.83)
Low-High	0.334***	0.238*	0.369**	0.304*
Diff. *t*-stat	2.98	1.90	2.40	1.88
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	120	120	120	120

Table A.10: The Information Turnover Ratio and the Attenuation in Profitability of the Non-Accounting-based Anomaly Portfolios (FF3 Alphas)

This table shows the difference-in-difference coefficients from the baseline difference-in-difference regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post_{p,t} + \beta_3 * EDGAR_p * Post_{p,t} + \epsilon_{p,t}$ for the non-accounting anomaly portfolios with high information turnover ratios and with low information turnover ratios. The information turnover ratio is defined as the total number of new incoming stocks divided by the total number of stocks in the existing portfolio when the anomaly portfolio updates its signal and rebalances its stocks. Information turnover ratios for non-accounting anomalies are computed for the pre-EDGAR period from October 1, 1983, to September 30, 1993. The anomaly portfolios used to compute the information turnover ratios include all the stocks in the NYSE-AMEX-NASDAQ universe that pass the stock-level filters mentioned in Section 3.B. After categorizing the anomalies into accounting-based and non-accounting-based anomalies, I sort the anomalies in each category with respect to their information turnover ratio rank percentile. The anomalies that are over 50% rank percentile are defined as the "High Turnover" anomalies and the rest as the "Low Turnover". $\alpha_{p,t}$ is the Fama-French three-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group), zero if the portfolio is constructed with the non-EDGAR filers (the control group); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first date on which investors start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR (i.e., the effective date), zero if the month is before that date. The dependent variables are the Fama-French three-factor alphas for the respective portfolio specifications. The sample period is from January 1992 to December 1997. *Low-High* and *Diff. t-stat* in the bottom section show the results of testing statistical difference between the difference-in-difference coefficient on the *High Turnover* anomaly portfolios and those of the *Low Turnover*. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

The Fama-French Three-Factor Alphas

Non-Accounting	FF3 1-5 EW DiD Coeff.	FF3 1-5 VW DiD Coeff.	FF3 1-10 EW DiD Coeff.	FF3 1-10 VW DiD Coeff.
High Turnover	0.0008	0.0050	-0.0396	-0.0507
	(0.01)	(0.04)	(-0.28)	(-0.31)
Low Turnover	-0.084	-0.244*	-0.071	-0.170
	(-0.81)	(-1.86)	(-0.60)	(-1.21)
Low-High	-0.084	-0.247	-0.03	-0.117
Diff. t-stat	(-0.61)	(-1.43)	(-0.19)	(-0.60)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	108	108	108	108

Table A.11: Fama-MacBeth Regression after Controlling for Firm Size and Bid-Ask Spread.

This table shows the results for the Fama-MacBeth regression after controlling for firm size and bid-ask spread. I first estimate a cross-section regression for every stock, for all 234 anomalies, for every EDGAR implementation phase, and for every month from January 1992 to December 1997 using the following equation: $R_{i,a,p,t+1} = \beta_0 + \beta_1 * SignalPercentile_{i,a,p,t} + \beta_2 * Treated_{i,p} + \beta_3 * SignalPercentile_{i,a,p,t} * Treated_{i,p} + \beta_4 \ln(MarketCap)_{i,t} + \beta_5 * BidAskSpread_{i,t} + IndustryFixedEffect$, where i is a stock (either an EDGAR filer or a non-EDGAR filer) assigned to a given EDGAR implementation phase p for a given anomaly a in the month t; $R_{i,a,p,t+1}$ is the simple monthly return of the following month for the given stock i; $SignalPercentile_{i,a,p,t}$ is the stock i's signal (characteristic) percentile within all the stocks for given month t and given anomaly a. $Treated_{i,p}$ is an indicator variable that equals one if the stock i is an EDGAR filer for a given implementation phase p, and zero otherwise. Industry fixed effect which is defined as the first two digits of the stock's SIC code is also controlled for. The second-pass is estimated separately for the accounting-based anomalies and the non-accounting anomalies using the equation $\widehat{\beta_{a,p,t}} = \gamma_t + \delta * Post_{a,p,t} + \epsilon_{a,t}$, where a is the anomaly portfolio formed on stocks in a given implementation phase p; $\widehat{\beta_{a,p,t}}$ are the monthly beta estimates from the first-pass regression (the beta estimates of interest are $\hat{\beta}_3$ for the treatment group and $\hat{\beta}_1$ for the control group); γ_t is the monthly time fixed effect; $Post_{a,p,t}$ is an indicator variable that equals one if the month is after the effective date for implementation phase p and equals zero otherwise; and δ is the coefficient of primary interest. Note that I drop four outlier anomalies which distort the results. Finally, to test the statistical difference between the accounting and the non-accounting anomalies, in other words, to test the statistical difference between $\delta_{accounting,1}$ and $\delta_{non-accounting,1}$, I run the following second pass $\widehat{\beta_{a,p,t}} = \gamma_t + \delta_1 * Post_{a,p,t} + \delta_2 * ACC_{a,p,t} + \delta_3 * Post_{a,p,t} * ACC_{a,p,t} + \epsilon_{a,t}$, where $ACC_{a,p,t}$ is an indicator variable that equals one if anomaly a is classified as an accounting-based anomaly. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

EDGAR-filers (Treatment Group) versus Non-EDGAR filers (Control Group)

	(1) Accounting Anomalies	(2) Non-Accounting Anomalies	(1) – (2) Difference
A. Treatment Group			
$Post_{a,p,t}$	-0.784***	0.055	-0.839***
	(-3.55)	(0.23)	(-2.47)
B. Control Group			
$Post_{a,p,t}$	-0.136	0.222	-0.359
	(-1.66)	(0.77)	(-1.32)
Controls	**Yes**	**Yes**	**Yes**
Mon. & Ind. FE	**Yes**	**Yes**	**Yes**
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	124	106	230

Table A.12: Investor Response to EDGAR Implementation

This table shows how the stock-related measures responded to EDGAR implementation. In order to examine the responses, I first estimate a cross-section regression for every stock that is captured by 1-10 decile anomaly portfolios, for all 234 anomalies, for every EDGAR implementation phase, and for every month from January 1992 to December 1997 (i.e., estimate cross-sectional regressions by-anomaly, by-group, by-month) using to the following equation: $Measure_{i,a,p,t} = \beta_0 + \beta_1 * Treated_{i,a,p}$ where $Measure_{i,a,p,t}$ is the trading related measure of interest, and $Treated_{i,a,p}$ equals one if the stock i is an EDGAR filer in the decile treatment portfolio for anomaly a for a given implementation phase p, and zero if stock i is a non-EDGAR filer in the decile control portfolio for anomaly a. The second-pass is estimated separately for the accounting-based anomalies and the non-accounting anomalies using the equation $\widehat{\beta_{a,p,t}} = \gamma_t + \delta * Post_{a,p,t} + \epsilon_{a,t}$, where a is the anomaly portfolio formed on stocks in a given implementation phase p; $\widehat{\beta_{a,p,t}}$ are the monthly beta estimates (the beta estimates of interest are $\hat{\beta}_3$ for the treatment group and $\hat{\beta}_1$ for the control group) from the first-pass regression; γ_t is the monthly time fixed effect; $Post_{a,p,t}$ is an indicator variable that equals one if the month is after the effective date for implementation phase p and equals zero otherwise; δ is the coefficient of primary interest. Finally, to test the statistical difference between the accounting and the non-accounting anomalies, I run the following second pass $\widehat{\beta_{a,p,t}} = \gamma_t + \delta_1 * Post_{a,p,t} + \delta_2 * ACC_{a,p,t} + \delta_3 * Post_{a,p,t} * ACC_{a,p,t} + \epsilon_{a,t}$, where $ACC_{a,p,t}$ is an indicator variable that equals one if anomaly a is classified as an accounting-based anomaly. The sample period is from January 1992 to December 1997. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

Panel A: Log of Trading Volume

	(1) Ln(Trading Vol) Acc. Anomalies	(2) Ln(Trading Vol) Non-Acc. Anomalies	(1) – (2) Difference
A. Treatment Group			
$Post_{a,p,t}$	0.848***	0.830***	0.017
	(12.06)	(12.05)	(0.32)
B. Control Group			
$Post_{a,p,t}$	-0.098***	0.007	-0.105***
	(-5.36)	(0.22)	(-2.84)
Mon. FE	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	108	

Panel B: Bid-ask spread

	(1) Bid-ask Spread Acc. Anomalies	(2) Bid-ask Spread Non-Acc. Anomalies	(1) – (2) Difference
A. Treatment Group			
$Post_{a,p,t}$	-0.0061***	-0.0050***	-0.0012***
	(-15.86)	(-13.85)	(-4.58)
B. Control Group			
$Post_{a,p,t}$	-0.0011***	-0.0010***	0.0001
	(-5.82)	(-3.72)	(-0.45)
Mon. FE	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	108	

Panel C: Idiosyncratic Volatility

	(1) Idiosyn. Volatility Acc. Anomalies	(2) Idiosyn. Volatility Non-Acc. Anomalies	(1) – (2) Difference
A. Treatment Group			
$Post_{a,p,t}$	-0.0067***	-0.0047***	-0.0020***
	(-16.62)	(-11.97)	(-6.03)
B. Control Group			
$Post_{a,p,t}$	-0.0014***	-0.0010***	-0.0004
	(-6.25)	(-3.79)	(-1.21)
Mon. FE	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	108	

Panel D: Amihud Illiquidity

	(1) Amihud Illiquidity Acc. Anomalies	(2) Amihud Illiquidity Non-Acc. Anomalies	(1) – (2) Difference
A. Treatment Group			
$Post_{a,p,t}$	-0.283***	-0.253***	-0.030
	(-12.65)	(-11.21)	(-1.44)
B. Control Group			
$Post_{a,p,t}$	0.009	-0.025	0.034**
	(1.19)	(-1.61)	(1.99)
Mon. FE	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	108	

Table A.13: Changes in Institutional Ownership and Short Interest

This table shows how institutional ownership and short interest in the long and the short leg of the anomaly portfolios in the treatment group respond to EDGAR implementation. I first estimate a cross-section regression for every stock that is captured by 1-10 decile anomaly portfolios in the treatment group, for all 234 anomalies, for every EDGAR implementation phase, and for every month from January 1992 to December 1997 (i.e., estimate cross-sectional regressions by-anomaly, by-group, by-month) using to the following equation: $Measure_{i,a,p,t} = \beta_0 + \beta_1 * Treated_{i,a,p}$ where $Measure_{i,a,p,t}$ is the trading related measure of interest, and $Treated_{i,a,p}$ is an indicator variable that equals one if the stock *i* is an EDGAR filer in the long leg of the decile treatment portfolio for anomaly a in a given implementation phase *p* and zero if stock *i* is in the short leg of the decile treatment portfolio for anomaly a in a given implementation phase p. The second-pass is estimated separately for the accounting-based anomalies and the non-accounting anomalies using the equation $\widehat{\beta_{a,p,t}} = \gamma_t + \delta * Post_{a,p,t} + \epsilon_{a,t}$, where a is the anomaly portfolio formed on stocks in a given implementation phase p; $\widehat{\beta_{a,p,t}}$ are the monthly beta estimates (the beta estimates of interest are $\hat{\beta}_3$ for the treatment group and $\hat{\beta}_1$ for the control group) from the first-pass regression; γ_t is the monthly time fixed effect; $Post_{a,p,t}$ is an indicator variable that equals one if the month is after the effective date for implementation phase *p* and equals zero otherwise; δ is the coefficient of primary interest. Finally, to test the statistical difference between the accounting and the non-accounting anomalies, I rerun the following second pass $\widehat{\beta_{a,p,t}} = \gamma_t + \delta_1 * Post_{a,p,t} + \delta_2 * ACC_{a,p,t} + \delta_3 * Post_{a,p,t} * ACC_{a,p,t} + \epsilon_{a,t}$, where $ACC_{a,p,t}$ is an indicator variable that equals one if anomaly *a* is classified as an accounting-based anomaly. The sample period is from January 1992 to December 1997. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

A. Institutional Ownership: Long versus Short Leg of 1-10 Treatment Portfolio

Treatment Portfolio	(1) Inst. Ownership Acc. Anomalies	(2) Inst. Ownership Non-Acc. Anomalies	(1) – (2) Difference
A. Long Leg	0.010^{*}	0.017^{**}	-0.007^{*}
$Post_{a,p,t}$	(1.79)	(2.34)	(-0.79)
B. Short Leg			
$Post_{a,p,t}$	0.137^{***}	0.123^{***}	0.014^{*}
	(16.84)	(13.98)	(1.65)
Mon. FE	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	108	

B. Short Interest: Long versus Short Leg of 1-10 Treatment Portfolio

Treatment Portfolio	(1) Short Interest Acc. Anomalies	(2) Short Interest Non-Acc. Anomalies	(1) – (2) Difference
A. Long Leg			
$Post_{a,p,t}$	0.076^{***}	0.031^{**}	0.045^{*}
	(3.14)	(2.10)	(1.66)
B. Short Leg			
$Post_{a,p,t}$	-0.190^{***}	-0.071^{***}	-0.119^{***}
	(-4.70)	(-3.91)	(-3.50)
Mon. FE	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	108	

Table A.14: Dynamics of Long and Short Leg Accounting Anomaly Portfolios (FF5 Alphas)

This table presents the coefficients from the baseline difference-in-difference regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post_{p,t} + \beta_3 * EDGAR_p * Post_{p,t} + \epsilon_{p,t}$ for the long and short leg of the accounting anomaly portfolios. $\alpha_{p,t}$ is the Fama-French five-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group), zero if the portfolio is constructed with the non-EDGAR filers (the control group); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first date on which investors start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR (i.e., the effective date), zero if the month is before that date. Panel A shows the results for the long leg of the accounting anomaly portfolios, whereas Panel B shows the results for the short leg. The dependent variables are the Fama-French five-factor alphas for either decile (1-10) or quintile (1-5) value-weighted (VW) or equal-weighted (EW) portfolios. The sample period is from January 1992 to December 1997. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

Panel A: Long Leg Accounting Anomaly Portfolios

	FF5 Alpha 1-5 EW	FF5 Alpha 1-5 VW	FF5 Alpha 1-10 EW	FF5 Alpha 1-10 VW
Post	-0.090	-0.074	-0.049	-0.099
	(-0.68)	(-0.58)	(-0.34)	(-0.72)
EDGAR	-0.081	0.021	-0.026	0.001
	(-0.63)	(0.14)	(-0.19)	(0.00)
Post # EDGAR	0.005	0.033	-0.017	0.052
	(0.03)	(0.19)	(-0.10)	(0.27)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	126	126	126
Mean of Dep. Var.	0.691	0.575	0.666	0.587

Panel B: Short leg Accounting Anomaly Portfolios

	FF5 Alpha 1-5 EW	FF5 Alpha 1-5 VW	FF5 Alpha 1-10 EW	FF5 Alpha 1-10 VW
Post	-0.079	-0.054	-0.088	-0.153
	(-0.56)	(-0.37)	(-0.54)	(-0.91)
EDGAR	0.079	0.137	0.030	0.155
	(0.51)	(0.73)	(0.17)	(0.78)
Post # EDGAR	-0.297	-0.400*	-0.337	-0.450*
	(-1.55)	(-1.82)	(-1.59)	(-1.93)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	126	126	126
Mean of Dep. Var.	-0.203	-0.237	-0.112	-0.142

Table A.15: Dynamics of Long and Short Leg Accounting Anomaly Portfolios (FF3 Alphas)

This table presents the coefficients from the baseline difference-in-difference regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post_{p,t} + \beta_3 * EDGAR_p * Post_{p,t} + \epsilon_{p,t}$ for the long and short leg of the accounting anomaly portfolios. $\alpha_{p,t}$ is the Fama-French three-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group), zero if the portfolio is constructed with the non-EDGAR filers (the control group); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first date on which investors start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR (i.e., the effective date), zero if the month is before that date. Panel A shows the results for the long leg of the accounting anomaly portfolios, whereas Panel B shows the results for the short leg. The dependent variables are the Fama-French three-factor alphas for either decile (1-10) or quintile (1-5) value-weighted (VW) or equal-weighted (EW) portfolios. The sample period is from January 1992 to December 1997. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

Panel A: Long Leg Accounting Anomaly Portfolios

	FF3 Alpha 1-5 EW	FF3 Alpha 1-5 VW	FF3 Alpha 1-10 EW	FF3 Alpha 1-10 VW
Post	0.014	0.035	0.043	-0.009
	(0.10)	(0.26)	(0.28)	(-0.06)
EDGAR	-0.047	0.039	0.004	0.035
	(-0.36)	(0.26)	(0.03)	(0.21)
Post # EDGAR	0.049	0.114	0.043	0.129
	(0.30)	(0.64)	(0.25)	(0.64)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	126	126	126
Mean of Dep. Var.	0.542	0.413	0.493	0.406

Panel B: Short leg Accounting Anomaly Portfolios

	FF3 Alpha 1-5 EW	FF3 Alpha 1-5 VW	FF3 Alpha 1-10 EW	FF3 Alpha 1-10 VW
Post	-0.141	-0.132	-0.148	-0.219
	(-0.93)	(-0.83)	(-0.85)	(-1.22)
EDGAR	-0.001	0.072	-0.056	0.085
	(-0.01)	(0.39)	(-0.32)	(0.43)
Post # EDGAR	-0.355*	-0.468**	-0.420*	-0.533**
	(-1.82)	(-2.08)	(-1.93)	(-2.24)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	126	126	126
Mean of Dep. Var.	0.007	-0.011	0.137	0.117

Table A.16: Excluding the First Implementation Phase

This table presents the coefficients from the baseline difference-in-difference regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post_{p,t} + \beta_3 * EDGAR_p * Post_{p,t} + \epsilon_{p,t}$ for the accounting anomaly portfolios, after dropping the first EDGAR implementation phase (Group CF-01). $\alpha_{p,t}$ is the Fama-French five-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group for one of the eight EDGAR implementation phases), zero if the portfolio is constructed with the non-EDGAR filers (the control group for each of eight treatment groups); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first date on which investors start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR (i.e., the effective date), zero if the month is before that date. Note that the $EDGAR_p$ indicator variable does not include the EDGAR filers assigned to the first EDGAR implementation phase, Group CF-01. The dependent variables are the Fama-French five-factor alphas for either decile (1-10) or quintile (1-5) value-weighted (VW) or equal-weighted (EW) portfolios. The sample period is from January 1992 to December 1997. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

Panel A: The Fama-French Five-Factor Alphas for the Accounting-based Anomalies

	FF5 Alpha 1-5 EW	FF5 Alpha 1-5 VW	FF5 Alpha 1-10 EW	FF5 Alpha 1-10 VW
Post	-0.115*	-0.074	-0.065	-0.171
	(-1.70)	(-0.86)	(-0.67)	(-1.47)
EDGAR	0.003	0.164*	0.020	0.171
	(0.03)	(1.84)	(0.16)	(1.54)
Post # EDGAR	-0.258**	-0.381***	-0.324**	-0.421***
	(-2.43)	(-3.61)	(-2.51)	(-3.32)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	126	126	126
Mean of Dep. Var.	0.440	0.326	0.496	0.432

Panel B: The Fama-French Three-Factor Alphas for the Accounting-based Anomalies

	FF3 Alpha 1-5 EW	FF3 Alpha 1-5 VW	FF3 Alpha 1-10 EW	FF3 Alpha 1-10 VW
Post	-0.086	-0.050	-0.045	-0.162
	(-1.32)	(-0.60)	(-0.48)	(-1.42)
EDGAR	-0.043	0.116	-0.039	0.129
	(-0.45)	(1.39)	(-0.33)	(1.22)
Post # EDGAR	-0.276**	-0.379***	-0.349***	-0.442***
	(-2.59)	(-3.66)	(-2.70)	(-3.58)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	126	126	126
Mean of Dep. Var.	0.517	0.415	0.587	0.533

Table A.17: Pre-trends Test

This table presents the coefficients from the baseline difference-in-difference regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post_{p,t} + \beta_3 * EDGAR_p * Post_{p,t} + \epsilon_{p,t}$ for the accounting anomaly portfolios. Following Gao and Huang (2020), the regression is estimated over a four-year period *prior to* the actual EDGAR implementation, and the pseudo-event dates are assumed to take place 2 years *before* the actual dates. $\alpha_{p,t}$ is the Fama-French five-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group), zero if the portfolio is constructed with the non-EDGAR filers (the control group); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first pseudo-event date on which investors are assumed to start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR, zero if the month is before that pseudo-event date. The dependent variables are the Fama-French five-factor alphas for either decile (1-10) or quintile (1-5) value-weighted (VW) or equal-weighted (EW) portfolios. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

Panel A: The Fama-French Five-Factor Alphas for the Accounting-based Anomalies

	FF5 Alpha 1-5 EW	FF5 Alpha 1-5 VW	FF5 Alpha 1-10 EW	FF5 Alpha 1-10 VW
Post	-0.134**	-0.142*	-0.101	-0.146
	(-2.03)	(-1.74)	(-1.23)	(-1.31)
EDGAR	-0.147**	-0.192**	-0.159*	-0.235**
	(-2.56)	(-2.47)	(-1.88)	(-2.36)
Post # EDGAR	-0.061	0.006	-0.066	0.028
	(-0.80)	(0.07)	(-0.59)	(0.25)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	126	126	126
Mean of Dep. Var.	0.360	0.285	0.403	0.333

Panel B: The Fama-French Three-Factor Alphas for the Accounting-based Anomalies

	FF3 Alpha 1-5 EW	FF3 Alpha 1-5 VW	FF3 Alpha 1-10 EW	FF3 Alpha 1-10 VW
Post	-0.130*	-0.171**	-0.117	-0.227**
	(-1.92)	(-2.03)	(-1.42)	(-2.03)
EDGAR	-0.132**	-0.154*	-0.139*	-0.201*
	(-2.42)	(-1.93)	(-1.74)	(-2.01)
Post # EDGAR	-0.111	-0.079	-0.129	-0.067
	(-1.38)	(-0.80)	(-1.16)	(-0.57)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	126	126	126
Mean of Dep. Var.	0.422	0.363	0.476	0.428

Table A.18: Falsification Test

This table presents the coefficients from the baseline difference-in-difference regression $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post_{p,t} + \beta_3 * EDGAR_p * Post_{p,t} + \epsilon_{p,t}$ for the accounting anomaly portfolios. Following Gao and Huang (2020), the regression is estimated over a four-year period *following* the actual EDGAR implementation, and the pseudo-event dates are assumed to take place 2 years *after* the actual dates. $\alpha_{p,t}$ is the Fama-French five-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group), zero if the portfolio is constructed with the non-EDGAR filers (the control group); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first pseudo-event date on which investors are assumed to start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR, zero if the month is before that pseudo-event date. The dependent variables are the Fama-French five-factor alphas for either decile (1-10) or quintile (1-5) value-weighted (VW) or equal-weighted (EW) portfolios. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust t-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

Panel A: The Fama-French Five-Factor Alphas for the Accounting-based Anomalies

	FF5 Alpha 1-5 EW	FF5 Alpha 1-5 VW	FF5 Alpha 1-10 EW	FF5 Alpha 1-10 VW
Post	-0.089	-0.254**	-0.096	-0.268**
	(-1.05)	(-2.61)	(-1.06)	(-2.57)
EDGAR	0.063	-0.107	0.236	-0.010
	(0.41)	(-0.43)	(1.30)	(-0.04)
Post # EDGAR	0.014	-0.079	-0.085	-0.212
	(0.08)	(-0.30)	(-0.40)	(-0.80)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	126	126	126
Mean of Dep. Var.	0.487	0.364	0.588	0.506

Panel B: The Fama-French Three-Factor Alphas for the Accounting-based Anomalies

	FF3 Alpha 1-5 EW	FF3 Alpha 1-5 VW	FF3 Alpha 1-10 EW	FF3 Alpha 1-10 VW
Post	-0.138	-0.225**	-0.166	-0.244**
	(-1.55)	(-2.30)	(-1.61)	(-2.37)
EDGAR	-0.102	-0.176	-0.062	-0.135
	(-0.61)	(-0.73)	(-0.32)	(-0.55)
Post # EDGAR	0.197	0.053	0.250	-0.031
	(0.96)	(0.20)	(1.07)	(-0.11)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Num. Anomalies	126	126	126	126
Mean of Dep. Var.	0.566	0.424	0.658	0.559

Appendix B. The Post-Publication Effect

McLean and Pontiff (2016) argue that the academic publication attracts arbitrageurs and show that publication-informed trading lowers anomaly portfolio returns by 32%. I address the concern that the post-publication effect could be driving the main results of my analysis.

To this end, I first define the publication year as the year during which the paper first appears on the SSRN, and then compare the publication year to the year in which the effective date for a given implementation phase falls. This allows for a determination of whether a given anomaly was known to investors at the time of trading. Following McLean and Pontiff (2016), I assume an average of 12-month lag between the SSRN posting date and the journal publication date if the paper could not be found on the SSRN.[1] In addition, I manually search Google Scholar to check if any paper published in 1998 had a version posted on the SSRN in 1996. I find this to be true of two papers: Abarbanell and Bushee (1998), and Dichev (1998). Therefore, the anomalies documented in these papers are assumed to have been known to investors at the time of the implementation phase 9 (CF-09).

I re-estimate the baseline regression described in Equation (1) separately for the published and for the non-published accounting anomalies. Table B.1 shows the results for Fama-French three-factor alphas.[2] Table B.1 Panel A indeed confirms the post-publication effect documented in McLean and Pontiff (2016) is present. The difference-in-difference coefficients for the anomalies that were published at the time EDGAR was introduced exhibit a greater attenuation of profitability across all portfolio and profitability specifications. In addition, Table B.1 confirms that the post-publication effect is not driving the results I document in this paper. The average number of published anomalies prior to and during the EDGAR implementation period is only 13, which is a small portion of the total number of accounting anomalies I study. Moreover, the magnitudes of the coefficients in Table B.1 Panel B are not drastically different from those reported in Table 4. Therefore, the post-publication effect does not significantly affect the results I document in this paper.

[1] McLean and Pontiff (2016) finds that the average length of time between the end-of-the sample and publication dates is 56 months, whereas the average time between the end-of- the sample and the SSRN dates is 44 months. Therefore, I take the difference between the durations of the two periods as the lag between the SSRN and the publication dates.
[2] The results are very similar for the Fama-French five-factor alphas; however, the publication effect is less clear for the equal-weighted quintile portfolio. Please see Table B.2 in the appendix.

Table B.1: The Post-Publication Effect (McLean and Pontiff (2016)) and the Accounting Anomaly Portfolio Profitability (FF3 Alphas)

This table shows the attenuation of profitability in the published versus the unpublished accounting-based anomalies in response to the implementation of the EDGAR. The publication year is defined as the year in which the paper first appears on the SSRN. The publication year is compared to the year in which each implementation phase date falls, to decide whether the anomaly investors were aware of the anomaly at the time of trading. If the paper is not found on the SSRN, an average 12-month lag between the SSRN posting date and the journal publication date is assumed following McLean and Pontiff (2016). In addition, if any paper published in 1998 had a version posted on the SSRN during 1996, these papers were assumed to be known to investors at the time of implementation phase 9. The baseline difference-in-difference regression is estimated separately for the published and the unpublished accounting-based anomalies. The regression is specified as $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post_{p,t} + \beta_3 * EDGAR_p * Post_{p,t} + \epsilon_{p,t}$. $\alpha_{p,t}$ is the Fama-French three-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group), zero if the portfolio is constructed with the non-EDGAR filers (the control group); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first date on which investors start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR (i.e., the effective date), zero if the month is before that date. The dependent variables are the Fama-French three-factor alphas for either decile (1-10) or quintile (1-5) value-weighted (VW) or equal-weighted (EW) portfolios. The sample period is from January 1992 to December 1997. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust t-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

Panel A. Published Accounting Anomalies with the Fama-French Three-Factor Alphas

	FF3 Alpha 1-5 EW	FF3 Alpha 1-5 VW	FF3 Alpha 1-10 EW	FF3 Alpha 1-10 VW
Post	-0.063	0.032	0.097	-0.033
	(-0.41)	(0.22)	(0.43)	(-0.14)
EDGAR	0.083	0.367**	0.175	0.579**
	(0.45)	(2.06)	(0.69)	(2.15)
Post # EDGAR	-0.372*	-0.618***	-0.584**	-0.849***
	(-1.70)	(-2.68)	(-2.21)	(-2.89)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Avg. Num. Anomalies	13	13	13	13
Mean of Dep. Var.	0.404	0.167	0.502	0.262

Panel B. Non-published Accounting Anomalies with the Fama-French Three-Factor Alphas

	FF3 Alpha 1-5 EW	FF3 Alpha 1-5 VW	FF3 Alpha 1-10 EW	FF3 Alpha 1-10 VW
Post	-0.155**	-0.115	-0.159*	-0.270**
	(-2.53)	(-1.46)	(-1.82)	(-2.41)
EDGAR	-0.074	0.053	-0.083	0.031
	(-0.84)	(0.68)	(-0.75)	(0.30)
Post # EDGAR	-0.254**	-0.301***	-0.308**	-0.324***
	(-2.58)	(-3.20)	(-2.50)	(-2.65)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Avg. Num. Anomalies	113	113	113	113
Mean of Dep. Var.	0.500	0.415	0.568	0.530

Table B.2: The Post-Publication Effect (McLean and Pontiff (2016)) and the Accounting Anomaly Portfolio Profitability (FF5 Alphas)

This table shows the attenuation of profitability in the published versus the unpublished accounting-based anomalies in response to the implementation of the EDGAR. The publication year is defined as the year in which the paper first appears on the SSRN. The publication year is compared to the year in which each implementation phase date falls, to decide whether the anomaly investors were aware of the anomaly at the time of trading. If the paper is not found on the SSRN, an average 12-month lag between the SSRN posting date and the journal publication date is assumed following McLean and Pontiff (2016). In addition, if any paper published in 1998 had a version posted on the SSRN during 1996, these papers were assumed to be known to investors at the time of implementation phase 9. The baseline difference-in-difference regression is estimated separately for the published and the unpublished accounting-based anomalies. The regression is specified as $\alpha_{p,t} = \gamma_t + \beta_1 * EDGAR_p + \beta_2 * Post_{p,t} + \beta_3 * EDGAR_p * Post_{p,t} + \epsilon_{p,t}$. $\alpha_{p,t}$ is the Fama-French five-factor alpha of the anomaly portfolio p in month t; $EDGAR_p$ is an indicator variable that equals one if the portfolio is constructed with the EDGAR filers (the treatment group), zero if the portfolio is constructed with the non-EDGAR filers (the control group); $Post_{p,t}$ is an indicator variable that equals one if the month is after the first date on which investors start trading the new EDGAR filer stocks based on the latest information they collect from EDGAR (i.e., the effective date), zero if the month is before that date. The dependent variables are the Fama-French three-factor alphas for either decile (1-10) or quintile (1-5) value-weighted (VW) or equal-weighted (EW) portfolios. The sample period is from January 1992 to December 1997. The portfolio alphas and the coefficients are all in percentage terms. The standard errors are clustered by anomaly and by month to address the potential correlation in errors. Robust t-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

Panel A. Published Accounting Anomalies with the Fama-French Five-Factor Alphas

	FF5 Alpha 1-5 EW	FF5 Alpha 1-5 VW	FF5 Alpha 1-10 EW	FF5 Alpha 1-10 VW
Post	-0.113	0.057	0.018	-0.094
	(-0.80)	(0.41)	(0.08)	(-0.41)
EDGAR	0.169	0.462**	0.279	0.655**
	(0.89)	(2.29)	(1.10)	(2.34)
Post # EDGAR	-0.288	-0.629***	-0.480*	-0.800***
	(-1.37)	(-2.79)	(-1.89)	(-2.69)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Avg. Num. Anomalies	13	13	13	13
Mean of Dep. Var.	0.355	0.097	0.436	0.182

Panel B. Non-published Accounting Anomalies with the Fama-French Five-Factor Alphas

	FF5 Alpha 1-5 EW	FF5 Alpha 1-5 VW	FF5 Alpha 1-10 EW	FF5 Alpha 1-10 VW
Post	-0.192***	-0.149*	-0.184**	-0.286**
	(-3.01)	(-1.80)	(-2.05)	(-2.52)
EDGAR	-0.035	0.092	-0.031	0.067
	(-0.38)	(1.10)	(-0.27)	(0.62)
Post # EDGAR	-0.246**	-0.306***	-0.292**	-0.314**
	(-2.48)	(-3.13)	(-2.38)	(-2.52)
Mon. FE	Yes	Yes	Yes	Yes
S.E. Cluster	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.	Anomal.&Mon.
Avg. Num. Anomalies	113	113	113	113
Mean of Dep. Var.	0.415	0.322	0.471	0.423