Causal effect of information costs on asset pricing anomalies[☆]

March 16, 2024

Abstract

The SEC's EDGAR introduction slashed the costs of acquiring and trading on accounting information, especially for smaller investors. We both causally identify and assess how these information costs affect stock anomalies. Using the staggered EDGAR introduction, we show that average alphas for 125 accounting anomalies decline substantially, and that the decline explains most of the pre-EDGAR alphas. By contrast, alphas for 80 non-accounting anomalies do not change significantly. Information costs are as substantial as the other limits to arbitrage.

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

JEL Codes: G12, G14

[☆]We thank Andrew Chen, Tony Cookson, Stephen Dimmock, Mariassunta Giannetti, Charles Hadlock, Jiekung Huang, Mark Huson, Joshua Pollet, Jeffrey Pontiff, Seth Pruitt, Andrea Rossi, Stefano Rossi, Andrei Simonov, David Solomon, Scott Weisbenner, Mao Ye, as well as the participants at Michigan State University, the University of Western Ontario, the University of Illinois at Urbana-Champaign, the 2022 China International Conference in Finance (especially the discussant, Si Cheng), the 11th MSUFCU Conference, the 2022 Financial Research Association (FRA) Conference (especially the discussant, David McLean), the 2023 Midwestern Finance Association Conference (especially the discussant, Sina Ehsani), and the 2023 SFS Cavalcade (especially the discussant, Alexander Chinco) for comments and suggestions.

1. Introduction

Traditional asset pricing theories such as the CAPM (Sharpe (1964), Lintner (1965), Mossin (1966)) and the APT (Ross (1976)) assume frictionless markets, including costless trading and information gathering. Theoretical contributions by Grossman and Stiglitz (1980) and Verrecchia (1982) are quick to point out that costly information acquisition, an inevitable reality of financial markets, affects investor decisions and market outcomes. Investors, especially large institutional ones, identify and purchase data, wait for it to arrive, and process it; each of these steps is costly. Retail investors and smaller institutional investors may find it too costly to pursue data acquisition. Intuitively, as a particular type of data becomes more affordable to acquire and more widely available, it becomes less profitable to trade on it. Nonetheless, most prior studies ignore information costs and instead focus on trading costs or short sale costs.[1] Consequently, there continues to be a lack of empirical studies that assess the costs of acquiring information.

This paper fills that gap in the literature by estimating information costs in the U.S. equity markets. Information costs include *direct* and *indirect* costs of gathering the information, waiting for the data to arrive, compensating analysts processing the information, and acquiring the computing resources. These costs are particularly taxing and often beyond reach for smaller investors, be they retail investors or small institutional investors. We use a quasi-natural experiment—the SEC's staggered implementation of the Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system from February 1993 to May 1996—to study how stock anomalies responded to the EDGAR shock. EDGAR has reduced information costs substantially by making

[1] Keim and Madhavan (1997), Korajczyk and Sadka (2004), Lesmond, Schill, and Zhou (2004), Novy-Marx and Velikov (2016), Frazzini, Israel, and Moskowitz (2018), Patton and Weller (2020) and Chen and Velikov (2023) study how transaction costs affect stock anomalies, while Geczy, Musto, and Reed (2002), Drechsler and Drechsler (2014), Chu, Hirshleifer, and Ma (2020), and Muravyev, Pearson, and Pollet (2022) focus on the effect of short sale costs.

corporate filings readily searchable and accessible from anywhere with internet connectivity at any time. Finally, if the return patterns generated by accounting anomalies were to become less profitable after the introduction of EDGAR, these anomalies would be more likely to have represented mispricing than risk premium.

The introduction of EDGAR is well-suited for our inquiry for several reasons. First, most documented stock anomalies rely on accounting information. Second, EDGAR is an online system that enables companies to report their corporate filings electronically and investors to access and search them freely from anywhere. Its adoption lowered investors' costs and delays in acquiring this information, especially for smaller investors.[2] Third, the SEC's adoption design allows us to harness a staggered difference-in-difference framework. Finally, Grossman and Stiglitz (1980) establish that, in a competitive equilibrium, the drop in trading profitability matches the decrease in information costs. Thus, studying anomaly profitability lets us assess information costs.

We estimate the effect of EDGAR introduction on anomaly alphas in a staggered difference-in-difference framework. The SEC adopted EDGAR following a phase-in schedule over three years, assigning largely randomly each public firm to one of ten implementation phases. This adoption design helps us identify a causal effect of the decline in information costs on anomaly profitability. The identification comes from firms entering EDGAR at different times and from anomalies requiring—or not requiring—accounting information. We analyze a comprehensive set of anomalies documented by Chen and Zimmermann (2020). Our baseline results are grounded in a panel of monthly returns for long-short portfolios for ten implementation

[2] Before EDGAR, a comprehensive analysis of a broad cross-section of stocks was cost-prohibitive or impractical for most investors (Chang, Ljungqvist, and Tseng (2021)). An investor could physically visit one of the SEC's reference rooms in Washington DC, New York, or Chicago and read paper financial statements; pursue costly subscriptions to commercial data vendors such as Compustat, Value Line, or Dialog, which were often delayed and contradicted one another (Kern and Morris (1994)); or request companies to mail the filing documents. EDGAR also made information available to investors much faster: "within an hour of submitting a document on EDGAR, it could well be on an analyst's screen in Hong Kong, London, Frankfurt, Los Angeles or Chicago" (Star Tribune (1993)).

phases and 205 asset-pricing anomalies, 125 of which require accounting information and 80 do not.[3] The analysis is at the anomaly-by-phase level, giving us enough statistical power despite the relatively short implementation window.

We find that the Fama-French six-factor (the Fama and French (2015) five factors and the Carhart (1997) momentum factor) alphas for the accounting-based anomaly portfolios decline on average by 47 to 62 basis points per month in response to the EDGAR introduction, accounting for most of the pre-EDGAR alphas for these anomalies. By contrast, EDGAR has not lowered the costs of gathering non-accounting information. Indeed, the alphas for non-accounting anomalies have not been significantly affected by the EDGAR introduction. The difference between the effect of EDGAR on the profitability of accounting and non-accounting anomalies, 42 to 50 basis points per month, could be interpreted as the estimate of information costs.[4]

The results are robust to using alternative specifications and factor models and are also robust to controlling for differences in stock characteristics, including firm size, across the implementation phases. They are not driven by one or a few phases. In fact, alphas for accounting anomalies decrease for every phase. We extensively validate our difference-in-difference analysis.

These results reinforce the simple, but pivotal point that, as a particular dataset becomes easier to access, its value for alpha-generation drops. We are the first to demonstrate the point for a comprehensive set of anomalies by studying the EDGAR introduction in difference-in-difference settings. Prior literature documents that limits to arbitrage (such as noise trader risk, trading costs, and short sale costs) partially explain anomaly returns, but no study of which we are aware

[3] We follow the anomaly classification by Chen and Zimmermann (2020). The results hold if we restrict accounting anomalies to rely only on accounting variables (and do not rely on stock price), as in McLean and Pontiff (2016).

[4] Lower information costs reduce mispricing due to accounting anomalies. As a result, the annual alphas associated with trading on such anomalies that require EDGAR information decreased by 6-7%. Grossman and Stiglitz (1980) state in general, and prove for CARA utility functions and normal return distributions, that a decline in the information cost prompts an increase in the fraction of informed traders participating in the markets (relative to uninformed traders), whose trading activities, in turn, make the price system more informative.

examines the effect of information costs per se on anomaly returns. This paucity of research is all the more glaring in light of the simple fact that investors must identify which stocks to trade even before they pay trading costs.

We test several mechanisms for the ways the change in the information environment prompted by EDGAR could lead to the accounting anomalies' profitability decline. EDGAR makes accounting information easier to acquire, especially for smaller investors. Profitability could decline due to more aggressive trading by arbitrageurs, less mispricing caused by noise traders (such as retail investors), or both. Moreover, post-EDGAR, investors act on accounting information faster, leading to faster price discovery. The effects of arbitrageurs and noise traders are especially challenging to separate because they are endogenous to the market equilibrium. Nonetheless, we strive to distinguish between them and find evidence more consistent with the less mispricing due to noise trading.

To test the mechanisms, we first focus on information availability. The profitability of accounting anomalies should decline more among the stocks for which the information was harder to gather pre-EDGAR. We use two proxies for information availability—analyst coverage and market capitalization (e.g., Kelly and Ljungqvist (2012))—to show that, indeed, the accounting anomalies' profitability decline is driven by stocks with low information availability. By contrast, the EDGAR-prompted profitability decline for high information availability stocks is statistically indistinguishable from zero. Small-cap firms and firms with low analyst coverage tend to have a greater fraction of retail and small institutional traders, who become better informed after EDGAR and, consequently, cause less mispricing. These firms are also more lucrative for arbitrage trading.

Splitting each long-short anomaly-phase portfolio reveals that profit attenuation is concentrated among the short legs of the accounting anomalies. This finding is consistent with many anomalies being concentrated in the short legs (Stambaugh, Yu, and Yuan (2012)). Because of short sale constraints, noise traders are more likely to cause overpricing than underpricing (Miller (1977)). If EDGAR makes noise traders less uninformed, overpricing should decline more compared to underpricing, which is what we observe. Short sale costs also discourage arbitrageurs.

To distinguish between the arbitrageur and noise trader channels, we explore how the EDGAR introduction affects participation by different types of investors as well as measures of price efficiency and liquidity. First, consistent with theoretical predictions, lower costs of acquiring information improved price efficiency and liquidity, and reduced information asymmetry. Second, investor participation increased once firms started filling electronically via EDGAR, especially for retail investors and smaller institutional investors.[5] Finally, because pre-EDGAR anomaly alphas and their attenuation due to EDGAR were concentrated in the short legs of anomaly portfolios, we study participation by short sellers, an important class of informed investors. The EDGAR introduction decreased participation by short sellers. These results are consistent with noise traders causing less anomaly-related mispricing for informed investors to explore.

We next focus on the speed of information dissemination as EDGAR made information spread faster. Our first test contrasts the high-turnover and low-turnover accounting anomalies, defined on the basis of the fraction of stocks that rotate in the top and bottom portfolios. Because the high-turnover anomalies require more up-to-date information, the alpha decline post-EDGAR should be higher for high-turnover anomalies. The difference in alpha decline is economically significant and consistent with this hypothesis, although it lacks statistical significance.

Our second test explores the lag between the timing of anomaly signals and returns. Intuitively, if investors receive accounting information faster post-EDGAR, anomaly alphas will

---

[5] We measure participation by retail investors as trading volume outside regular trading hours, by smaller institutional investors as average position size and holding concentration, and by short-sellers as short interest.

migrate closer to the signal in time. Thus, alphas for accounting anomalies should decline less for months closer to the signal. We compare portfolio returns to the pure accounting anomalies formed with a standard timeline—allowing a 6-month (3-month) lag for the strategies based on annual (quarterly) portfolio formation—with portfolio returns to the same anomalies formed with a lag shorter by one month, a conservative speed-up in portfolio formation facilitated by EDGAR. The post-EDGAR portfolio alpha decline is smaller for the strategies forming portfolios one month sooner—by a margin of 15 to 23 basis points per month, establishing that information delay costs are indeed economically significant. Overall, these two tests suggest that faster information dissemination contributes to the profitability decline.

Our results are quite distinct from McLean and Pontiff (2016), who show that the dissemination of *academic research* affects anomalies. By contrast, we show that the dissemination and availability of *data* affect anomalies. Because most anomalies were published after the EDGAR introduction, we do not assume that investors trade the exact signals suggested by academic anomalies. Rather, the academic signals span relevant dimensions of potential mispricing. The mispricing decreased in part because EDGAR made information more accessible for small investors. Finally, some arbitrageurs could have discovered accounting anomalies that are sufficiently close to the ones later popularized by academics.

Our results remain relevant for today's markets. Indeed, while utilizing accounting information in active portfolio management was innovative in the early 1990s, it quickly became commoditized post-EDGAR, reducing its alpha-generating ability. In response, active investment managers expanded into new types of data. Such data are presently expensive and hard to process, similar to the status of accounting information pre-EDGAR.[6] Thus, the same principles we uncovered for circumstances surrounding the EDGAR introduction likely apply to these data.

2. Brief review of related literature

This paper contributes to four strands of the literature. First, it contributes to the literature concerning information costs and market outcomes. Merton (1987) and Shapiro (2002) point out that costly information constraints compel investors to trade only the securities regarding which they possess adequate information and show how these constraints affect the general equilibrium process and outcomes. Grossman and Stiglitz (1980) show that perfect market efficiency is elusive because information is costly to collect. Kadan and Manela (2019) derive a general expression for the value of information and estimate it for macroeconomic announcements. Our paper contributes to this literature by estimating information costs in the context of U.S. equity markets.

Second, we primarily contribute to the stock anomaly literature by identifying the causal effect of information constraints on anomaly returns. Only a few papers use exogenous shocks to study the effect of limits to arbitrage on anomalies (e.g., Albuquerque, Song, and Yao (2020) and Ben-David et al. (2021)). However, unlike our study, they do not explore costly information constraints as limits to arbitrage. McLean and Pontiff (2016) document that portfolio alphas decline by 58% on average after publication. Whereas they argue that investors learn about anomalies from academic research, we point out that even investors who discover similar anomalies before academics incur substantial information costs of computing the anomaly signals. Also, the EDGAR introduction helped smaller investors make more informed trades and thereby

[6] For example, PanAgora‘s CIO George Mussalli notes: “We stay away from over-marketed data purely curated for hedge fund consumption, such as satellite data, credit card transactions, and email receipts. These data sources are overused, and we have seen a marked deterioration in their predictive power.” (Bloomberg, 11/29/2018).

contribute less to mispricing. Also, Green, Hand, and Zhang (2017) find that stock anomalies became substantially less profitable in the early 2000s. We show that this profitability decline started even earlier for accounting anomalies and associate it with the EDGAR introduction.[7]

Third, a strand of literature focuses on the effect of EDGAR on information production and its accuracy. Gao and Huang (2020) show that the EDGAR introduction enhances information production by individual investors and sell-side analysts. Post-EDGAR, the amount and accuracy of analysts' information increase—more analysts start covering a firm, their forecasts are more accurate, and stock prices react stronger to their revisions. As for individual investors, Gao and Huang (2020) show that their net purchases become *more* predictive of future stock returns once the firm begins filing through EDGAR. On the other hand, we find that a broad set of accounting anomaly signals become *less* predictive of future stock returns post-EDGAR. This difference in results could be due to the difference in return periods— their predictability analysis focuses on the post-earnings announcement period, while we study regular monthly returns. Also, some retail investors could analyze EDGAR filings beyond the basic accounting variables used in anomalies, and their informed trading in stocks could be diversifiable at the anomaly portfolio level. In a broader sense, we explore the effects of EDGAR adoption aggregated at the level of mispricing embedded in anomalies, a type of price inefficiency that cannot readily be arbitraged away by diversification.[8] Our paper is the first to study how EDGAR adoption affects anomaly profitability.

[7] Using Compustat Snapshot, which contains precise release dates of accounting information, Bowles *et al.* (2020) show that anomaly returns are concentrated early, within the first 30 days after the information release. They conclude that anomalies are not spurious because anomaly returns should not depend on their proximity to the information release. Consistent with Bowles *et al.* (2020), we find that the profitability of accounting anomalies declined quickly in response to the EDGAR introduction. However, our focus is on identifying and quantifying the effect of data availability on information costs rather than on assessing the speed of market reaction to information releases.

[8] The EDGAR adoption also improves equity financing (Goldstein, Yang, and Zuo (2023)), reduces the information asymmetry between managers and investors (Gomez (2023)), reduces investor disagreement, and mitigates crash risk (Chang, Ljungqvist, and Tseng (2022)).

Finally, we contribute to the literature that considers the effects of changes in a firm's information environment. Dong *et al.* (2016) show that stock return synchronicity decreases for firms that file with the SEC using a machine-friendly format XBRL. XBRL made it easier to process subtle accounting information such as footnotes. However, none of the accounting anomalies in our sample require such information. Thus, the EDGAR introduction affects accounting anomalies, whereas the XBRL introduction should not. Chen, Kelly, and Wu (2020) show that a reduction in analyst coverage prompts hedge funds to acquire more information by searching EDGAR. The hedge fund participation mitigates the impairment of market efficiency caused by coverage reductions. Whereas both studies focus on the change in the information environment, unlike our study, neither seeks to quantify the change in information costs in the aftermath of the information environment change in terms of changes in average stock returns.

3. Implementation of the EDGAR system

3.1. Costs of information acquisition before EDGAR

Prior to the EDGAR adoption in the mid-1990s, direct and indirect costs of acquiring the information contained in corporate filings were large. Investors were mostly limited to three options, all of which were prohibitively costly for smaller investors. The first option was to 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 Compustat, Disclosure, Value Line, or Dialog. Lastly, current shareholders could request that the companies mail their filing documents to them.

Anecdotal evidence confirms that the first option was costly and unreliable. 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. Occasionally, investors could not even access the information they needed because some paper files in the SEC's reference rooms were lost.[9]

The second option was also costly because the pre-EDGAR data aggregators charged high fees and, more importantly, were slow. A petition filed to the SEC and the U.S. House of Representatives in 1992 documents the related complaints. The petition demands free public access to corporate filings, pointing out that the Compustat CD-ROM database with historical filings for just 7,200 companies cost $18,000 (Love (1992)).[10] Depending on the coverage, annual subscription fees ranged between $5,000 and $50,000.[11] Value Line Database cost $1,700 per quarter and covered only 1,650 companies. Mead Data Central was only available for a fee that consisted of a $125 monthly fixed fee, a $39 hourly connection fee, and a search fee ranging from $6 to $51 per search.[12] Whereas all these costs could have been acceptable to large institutional investors, retail investors and smaller institutions realistically could not afford them. Accordingly, trading on accounting information—particularly in a timely fashion—was out of many investors' reach. Whatever trades they carried out pre-EDGAR, those smaller investors would be more accurately described as noise traders, rather than informed traders.

Aside from high fees, Compustat suffered from production lag and inaccuracy, which also increased the costs of acquiring accurate financial information. D'Souza, Ramesh, and Shen (2010)

---

[9] A *Wall Street Journal* article reports in 1991 that "…nowadays 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."

[10] According to Love (1992), the CD-ROM was called "COMPUSTAT PC Plus." A less expensive product, "COMPUSTAT Corporate Text," was available for $9,000, but was limited in its coverage to only 3,200 firms.

[11] SEC: Oversight of the Edgar System (March 14, 1985), pp. 51.

[12] The petition also reveals that Dialog charged $84 per hour on top of a $1 per page search fee. Compact Disclosure was another popular commercial database at the time. Richards (1988) documents that Compact Disclosure had quarterly updated financial and management information on 10,150 public companies, and cost around $4,500 per year for commercial institutions. However, Richards (1988) notes that Compact Disclosure's access software had technical issues retrieving time-series data, and was missing information on brokerage houses, foreign companies, and microcap stocks with less than $5 million in assets.

find that Compustat had an average pre-EDGAR dissemination lag of 24.7 weekdays, which dropped by almost 50% post-EDGAR. Such a delay has made the trades placed even by those traditionally regarded as informed investors either significantly delayed or largely uninformed.

Moreover, even if investors had subscribed to commercial data vendor services, their accuracy was not particularly high. There existed a significant mismatch between their databases, thus obfuscating whatever information content they may have offered to their subscribers. Kern and Morris (1994) compare two popular commercial databases, Value Line and Compustat, and find material disagreements between the two datasets from 1985 to 1990. They also show that empirical results by Porcano (1986) could have had different outcomes, depending on the database used. Kothari, Shanken, and Sloan (1995) explore the implications of a selection bias in the Compustat data for return predictability. Finally, accounting restatements are instantly available on EDGAR but take a while to be reflected by Compustat. Therefore, the costs of obtaining *accurate* financial information were still substantial, even after paying the data-vendor fees.

Lastly, in principle, investors could have requested that companies directly mail them the financial documents. Besides the costs of a long wait, this was not a viable option for any investor intending to perform cross-sectional analyses on firm characteristics because such analyses require simultaneous availability of financial information concerning many public companies.

3.2. Introduction of EDGAR

Responding to the call for more transparency and easier accessibility of corporate filings by publicly traded companies, the SEC harnessed the advances in information technology by developing and introducing the EDGAR system. The SEC began developing the system in 1983. Eventually, after extensive testing, on February 23, 1993, the Commission issued four releases adopting the rules that required filers to file electronically. The process began on April 26, 1993, gradually bringing all filers onto the EDGAR system. EDGAR allows public firms to disclose their financial information electronically, and investors or any information consumers to access the filed corporate information instantaneously via the internet without charge.

The introduction of EDGAR significantly lowered the costs of information acquisition by expediting electronic filing and information dissemination 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." Furthermore, EDGAR's search function and other interface features allowed the users to retrieve specific information in electronic documents that may not have been available in commercial databases.

A feature of EDGAR implementation, central to our empirical design, is that the SEC adopted EDGAR following a phase-in schedule. The schedule assigned each public firm that required filing to one of ten phases. Each phase had a designated date on which electronic filling was mandated (SEC Release No. 33-6977). The firms in the first group were mandated to start uploading filings through EDGAR on April 26, 1993, and those in the last group were required to do so on May 1, 1996. Table 1 shows the implementation schedule.

<TABLE 1 ABOUT HERE>

We estimate the extent to which the investor information costs decreased. The staggered nature of EDGAR implementation helps us better identify the effect of information costs, alleviate alternative explanations, and control for other confounding factors. For example, one alternative explanation could be that the equity market is becoming increasingly efficient, and non-information costs decrease over time. However, to explain our results, these trends would have to discontinuously change for each firm at exactly the time it starts to file with EDGAR, a highly implausible set of circumstances.

4. Data and methodology

4.1. The SEC EDGAR implementation data

To construct anomaly portfolios for firms in each implementation phase, we first identify the date each firm becomes an EDGAR filer by examining SEC Release No. 33-6977. We also incorporate all the subsequent changes and corrections to the initial phase-in list.[13] The SEC Release documents provide the list of company names and their Central Index Key (CIK). We manually match each firm to their record in Compustat using the company name and CIK. We then use the linking file provided by WRDS to link Compustat with CRSP. The last column of Table 1 reports the number of firms in each phase that we were able to match to the two databases.

4.2. The anomalies

We start by examining a total of 320 anomalies replicated and shared by Chen and Zimmermann (2020), covering most return signals that academic researchers have reported to date.[14] By analyzing a comprehensive set of anomalies, we capture the full ramification of the

[13] The subsequent changes and corrections to the initial EDGAR phase-in list reported in SEC Release No. 33-6977 can be found in 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.

[14] Specifically, Chen and Zimmermann (2020) document all the anomalies in Hou, Xue, and Zhang (2020), 98% of the anomalies in McLean and Pontiff (2016), 90% of anomalies in Green, Hand, and Zhang (2017), and 90% of the

information cost-saving effect of the EDGAR introduction on the anomalies' profitability. We follow Chen and Zimmermann (2020), who in turn follow the original academic papers that introduced each anomaly, their filters, and datasets including CRSP, Compustat, IBES, the SEC's Form 13Fs, and the Federal Reserve Economic Data (FRED). Chen and Zimmermann (2020) provide quarterly versions of the anomalies by modifying the original characteristics to incorporate quarterly instead of annual information (assuming the standard one-quarter lag for quarterly data availability). Following this approach, we convert nine additional anomalies from annual to quarterly versions.[15]

We exclude penny stocks, that is, firms with a market capitalization below $50 million or a stock price lower than $5. Applying these two stock-level filters also mitigates the concern that microcap stock returns shape our results (Hou, Xue, and Zhang (2020)).

We eliminate anomalies that rely on binary signals, are unprofitable pre-EDGAR, or are too correlated with each other (we keep one of the two). We first compute the Fama and French three-factor alphas (Fama and French (1992, 1993)) and the pairwise return correlation of the decile equal-weighted anomaly portfolio returns over ten years pre-EDGAR, from October 1983 to September 1993. Then, in the spirit of Green, Hand, and Zhang (2017), we exclude the 58 anomalies that have negative pre-EDGAR alphas.[16] Next, to ensure that we focus on relatively independent anomalies, we identify "twin" anomalies that have a pairwise return correlation above 0.9 and eliminate 28 anomalies by dropping one of the twins. Finally, we drop 29 anomalies with

anomalies in Harvey, Liu, and Zhu (2016). We thank Andrew Chen and Tom Zimmerman for sharing the anomaly signal generating codes.

[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, and net income/book equity.

[16] Arbitrageurs would have been unlikely to trade anomalies without positive alphas, and it is not clear how to capture alpha attenuation caused by EDGAR for such anomalies. Nonetheless, we confirm that our main results remain robust to keeping these 58 negative-alpha anomalies in the sample (Table A.1 in the Appendix).

binary signals, such as a dividend-paid indicator, because we cannot form quintile or decile portfolios for such anomalies. Our final sample includes 205 anomalies.

Next, we compute the anomaly monthly abnormal returns over the period from January 1992 to December 1997 for the final sample of 205 anomalies. To compute the monthly portfolio returns, we first adjust the daily stock returns for delisting return bias following the approach of Shumway (1997) and then aggregate them to compute monthly returns. We compute the Fama-French (2015) five-factor alphas adjusted for momentum (Carhart (1997))—henceforth the Fama-French six-factor alpha—for the equal-weighted and value-weighted decile and quintile portfolio returns. Jensen, Kelly, and Pedersen (2021) emphasize the importance of focusing on anomaly alphas instead of anomaly returns. We focus on alphas, but our main results broadly hold if we do not risk-adjust anomaly returns (Table A.2 in the Appendix).

5. Baseline difference-in-difference results

EDGAR provides free and instant online access to SEC filings and thus lowers information costs, making it easier for arbitrageurs to identify mispriced stocks and for noise traders to make more informed trading decisions. Accordingly, the profitability of anomaly portfolios constructed from stocks that started to file with EDGAR should weaken. However, that attenuation should take place only for the anomalies that rely on accounting information from EDGAR.

We first compute the alpha for the long-short portfolio for a given anomaly, implementation phase, and month in two steps. First, we compute the difference between the top and bottom decile (quintile) portfolio returns, aggregated in the equal-weighted (value-weighted) manner, for each anomaly, phase, and month. Second, we calculate the alpha (or abnormal return) in a standard way as the sum of the residuals and the average alpha (intercept) from a regression of top-minus-bottom portfolio return on Fama-French factors, estimated over the sample period.

The results are robust to using raw returns without factor adjustment or to estimating factor betas using pre-EDGAR data (Table A.2 in the Appendix).

Our baseline specification estimates the effect of EDGAR introduction on the anomaly portfolio profitability using a standard difference-in-difference framework:

$$\widehat{\alpha_{a,p,t}} = \gamma_t + \gamma_a + \beta_1 * Post_{p,t} + \beta_2 * Post_{p,t} * ACC_a + \epsilon_{a,p,t}. \quad (1)$$

The dependent variable, $\widehat{\alpha_{a,p,t}}$, is the Fama-French six-factor alpha of the anomaly a top-minus-bottom portfolio for phase p in month t; γ_t are monthly fixed effects; γ_a are anomaly fixed effects; $Post_{p,t}$ is an indicator variable equal to one if month t is on or after the effective date for phase p, and equal to zero before that date; and ACC_a is an indicator variable equal to one if a is an accounting anomaly, and equal to zero if a is a non-accounting anomaly. The interaction $Post_{p,t} * ACC_a$ is the main variable of interest. It reflects the effect of EDGAR on accounting anomalies relative to non-accounting anomalies. Equation (1) can be viewed as a panel regression with anomaly by phase as one dimension and month as the other dimension. Standard errors are clustered by anomaly and month to address the potential correlation in errors (Petersen (2009)). Results remain largely unaffected if we cluster by month only or use no anomaly fixed effects.

In our difference-in-difference setup, the effect of EDGAR on the profitability of accounting anomalies is primarily identified from stocks in the ten implementation phases entering EDGAR at different times (i.e., before versus after). The entire universe of stocks is split into ten phases, and anomalies are computed separately for stocks within each phase. This gives us the statistical power to draw precise inferences. Indeed, with the 72 months of data, nine distinct phases (phases 3 and 4 became effective on the same date), and more than two hundred anomalies, estimation over the full sample relies on nearly 130,000 observations.

The difference-in-difference setup can also be interpreted cross-sectionally; we compare anomaly profitability for the phases already in EDGAR with the phases that will enter EDGAR later. The secondary identification comes from comparing the profitability of accounting and non-accounting anomalies. $Post_{p,t}$ equals one once a given phase enters EDGAR, but $Post_{p,t} * ACC_a$ then equals one only for accounting anomalies, and thus quantifies the profitability reduction for accounting anomalies relative to non-accounting anomalies.

Presented in Panel A of Table 2, the results are remarkably consistent across alternative specifications for the dependent variable—the Fama-French six-factor alphas for decile or quintile, equal-weighted or value-weighted portfolios. The portfolio alphas and the coefficients are expressed in percentages. As shown at the bottom of Table 2, Panel A, accounting-based anomaly alphas declined by 47 to 62 basis points per month (or 5.7% to 7.4% per year) because of the EDGAR introduction. This decline completely offsets the average accounting anomaly alphas of 44.9 basis points per month from the pre-EDGAR period. By contrast, as shown in the top row of Table 2, Panel A, non-accounting alphas do not decline post-EDGAR. The difference between the two, captured by the difference-in-difference coefficient, β_2, is between 42 and 49 basis points per month across different specifications, statistically significant at the one-percent level. It can be interpreted as the amount of information costs investors face in the absence of EDGAR.

These findings show that the information costs can be as important as other limits to arbitrage. There is a debate about the extent to which short sale costs affect anomaly profitability. Geczy, Musto, and Reed (2002) show that stock borrow fees explain a small portion of anomaly returns. By contrast, Chu, Hirshleifer, and Ma (2020) show that relaxed short sale constraints by Regulation SHO reduce abnormal returns of 11 anomalies by 72 basis points per month. A similar debate is ongoing about the effect of trading costs. Using TAQ data, Novy-Marx and Velikov

(2016) show that the average trading costs range from 20 to 57 basis points for the mid-turnover anomalies. By contrast, Frazzini, Israel, and Moskowitz (2018) argue that institutional trading costs are much smaller than the effective bid-ask spreads in TAQ. Although our results do not speak to the two debates, the 42 to 49 basis point per month information costs that we estimate around EDGAR are comparable to the upper bounds for the trading and short sale costs. Also, investors need to acquire information to identify which stocks to buy or sell before they start trading. Thus, investors incur information costs even before they pay trading or short sale costs.

One potential concern is that stock characteristics differ across EDGAR implementation phases, which could lead to changes in return predictability. To address this concern, Panel B of Table 2 provides estimates of Equation (1) with a range of additional controls, especially the logarithm of firm size.[17] The results in Panel B of Table 2 closely mirror those from Panel A, suggesting that additional controls do not affect our results.

<TABLE 2 ABOUT HERE>

Figure 1 highlights the discontinuity in how the average anomaly alphas of the treatment and control groups responded to EDGAR implementation in a two-year window centered around the effective dates for all implementation phases. Each point represents an average of alphas for nine phases and four risk-adjustment specifications (as in Table 2), separately for accounting

[17] The control variables include log market capitalization, Amihud illiquidity, book-to-market ratio, book leverage, return on assets, sales growth, and capital expenditures to total asset. For all implementation phases, we compute the monthly equal-weighted (value-weighted) average of the control variables for all the stocks captured by the long and the short leg of the equal-weighted (value-weighted) quintile and equal-weighted (value-weighted) decile anomaly portfolios, respectively. Then the control variables with a given specifications are used to control for the portfolio returns or alphas that corresponds to the same specifications. For example, in Column 2 of Table 2, monthly value-weighted average of the control variables for all the EDGAR-filer stocks in implementation phase 5 captured by the long and the short leg of the value-weighted average quintile anomaly portfolios are used to control for the FF6 alphas of “1-5 VW” anomaly portfolios constructed using EDGAR-filer stocks in implementation phase 5.

anomalies (black points) and non-accounting anomalies (white points) in a given month relative to the phase implementation date. The lines show average alphas in the year before and the year after phase implementation.

The figure reiterates the salient features of our regression results from Table 2. The treatment group—accounting anomalies—experiences a sharp decline in average alphas, from 0.73 and 0.78 percent two and one months before EDGAR implementation, to 0.58 percent at the effective date, to the substantially lower values of -0.13, 0.23, and 0.22 percent per month during the first three months following the effective date. The accounting alphas dropped from 0.44 to 0.08 percent per month in a year after the phases' effective date. By contrast, the average alphas of non-accounting anomalies—the control group—did not experience a decline; they remained steady at around 0.2 percent per month. The estimates in Figure 1 and Table 2 differ slightly because of the differences in methodology (such as a lack of fixed effects for the figure).

<FIGURE 1 ABOUT HERE>

In Section 9, we report additional robustness tests, including standard difference-in-difference diagnostic tests, alternative fixed effects specifications, alternative anomaly samples, and individual implementation phase effects.

Overall, these results show that, as accounting information became easier and cheaper to access, it became less profitable to trade on it. The difference-in-difference approach helps us cleanly identify this effect and rule out many alternative explanations.

6. EDGAR effects and information availability costs

A key facet of information costs is associated with information availability. EDGAR prompted a decline in profitability of accounting anomalies because accounting information became more easily and readily available. The profitability decline should be more pronounced among the stocks for which the information was more difficult to gather in the pre-EDGAR period. To test this hypothesis, we use two empirical proxies for information availability—analyst coverage and market capitalization (e.g., Kelly and Ljungqvist (2012))—to classify stocks into high or low information-availability groups. For example, full-service broker-dealers provided their clients with analysts' research and opinions in addition to executing trades as part of an overall package of services (the so-called "soft" dollar arrangements). Thus, information for stocks with high analyst coverage is easier to acquire.

We also confirm the main results from the previous section (based on decile/quintile portfolio sorts) using a two-stage approach inspired by Fama-MacBeth regression methodology (Fama and MacBeth (1973)). In the first stage, we estimate a cross-sectional regression of monthly returns on an anomaly signal for the stocks in each phase and month. In the second stage, we estimate the standard difference-in-difference regression in Equation (1), except the dependent variable is the linear slope from the first stage instead of the top-minus-bottom portfolio alpha. We conduct this analysis separately for stocks with high and low information availability.

We first outline the methodology for this test. To gauge information availability, we compute the average analyst coverage and market capitalization pre-EDGAR, from January 1990 to December 1992, and then classify each stock i as high-information, h (above-median analyst coverage; above-median market capitalization of equity), or low-information, l (below-median analyst coverage; below-median market capitalization of equity). For each of the 125 accounting anomalies, each implementation phase, and every month from 1992 to 1997, we estimate two first-pass regressions of the form:

$$R_{i,a,p,t+1} = \alpha + \beta_{a,p,t} * SignalPercentile_{i,a,p,t} + \epsilon_{i,a,p,t}, \quad (2)$$

separately for high and low information availability stock groups. For each group, firm i is assigned to phase p for accounting anomaly a in month t; $R_{i,a,p,t+1}$ is the next-month return for stock i; $SignalPercentile_{i,a,p,t}$ is anomaly signal's percentile within stocks in phase p for anomaly a in month t. Finally, $\beta_{a,p,t}$ is the coefficient of interest. This first-pass regression step creates a panel of $\widehat{\beta_{h/l,a,p},_t}$, monthly beta estimates for information availability groups *h*/*l* (high or low). Next, we estimate the second-pass panel regression, similar to Equation (1):

$$\widehat{\beta_{h/l,a,p}},_t = \gamma_t + \gamma_a + \delta_1 * Post_{p,t} + \delta_2 * LoInfo_{a,p,t} + \delta_3 * Post_{p,t} * LoInfo_{a,p,t} + \epsilon_{h/l,a,p,t}, \quad (3)$$

where $\widehat{\beta_{h/l,a,p}},_t$ are the monthly beta estimates from the first-pass regressions; γ_t and γ_a are monthly and anomaly fixed effects; $Post_{p,t}$ is an indicator variable equal to one if month *t* is after the effective date for phase *p*, and equal to zero otherwise; and $LoInfo_{a,p,t}$ is an indicator variable equal to one for all $\widehat{\beta_{h/l,a,p}},_t$ associated with low information availability groups, and equal to zero for all $\widehat{\beta_{h/l,a,p}},_t$ associated with high information availability groups.

The results, presented in Table 3, confirm our hypothesis for both analyst coverage and market capitalization. The EDGAR-prompted profitability decline for accounting anomalies is solely concentrated in low information availability stocks. For these stocks, monthly alphas decline by 70 to 71 basis points per month, or about 8.5% per year, statistically significant at the one-percent level. By contrast, the EDGAR-prompted profitability decline for accounting anomalies among high information availability stocks is statistically indistinguishable from zero. These findings confirm the intuition that the effects of EDGAR introduction are particularly pronounced for the stocks regarding which information was particularly costly to acquire pre-EDGAR.

<TABLE 3 ABOUT HERE>

7. EDGAR effects and various market outcomes

In this section, we study how the change in the information environment caused by EDGAR propagates through the equity market and affects various outcomes. We first explore an asymmetry between the short and long legs. Next, we study the effect on price efficiency, information asymmetry, and liquidity. Finally, we turn our attention to measures of investor participation, particularly for retail investors and smaller institutional investors.

7.1. Long and short anomaly portfolio legs

We estimate the baseline difference-in-difference regression from Equation (1), separately for the long and short legs of the 205 anomaly portfolios. The results, presented in Table 4, confirm that profit attenuation effects are concentrated among the short legs of the accounting anomaly portfolios. Across all four columns of Table 4, the difference-in-difference coefficient estimates for the long-leg accounting anomaly portfolios (Panel A) are small and statistically indistinguishable from zero. By contrast, the difference-in-difference coefficient estimates associated with short-leg anomaly portfolios (Panel B) are 32 to 50 basis points per month, statistically significant at the one-percent level across all columns. These estimates are comparable to the difference-in-difference coefficient estimates of 42 to 49 basis points from the baseline specification from Table 2. These results are consistent with both arbitrageur and noise trading mechanisms.

<TABLE 4 ABOUT HERE>

7.2. Price efficiency, liquidity, and investor participation

In this section, we study the ways the EDGAR introduction affected investor participation, price efficiency, and related market variables. Intuitively, better access to information, especially for smaller investors, should make stock prices more efficient, reduce information asymmetry, and, ultimately, improve liquidity. The predictions about participation by investor types are less straightforward, but the participation results help us separate the effects of arbitrageurs and noise traders.

We first compute $M_{a,p,t}$, an average market outcome of interest over stocks in the top and bottom decile portfolios for each accounting anomaly, phase, and month. We then estimate for each of the average measures a difference-in-difference regression similar to Equation (1):

$$M_{a,p,t} = \gamma_t + \gamma_a + \delta_1 * Post_{p,t} + \epsilon_{a,p,t}. \tag{4}$$

The dependent variable, $M_{a,p,t}$ is the average market outcome; γ_t and γ_a are monthly and anomaly fixed effects; $Post_{p,t}$ is an indicator variable equal to one if month t is on or after the effective date for phase p, and equal to zero otherwise; δ_1 estimates the effect of EDGAR introduction on the market outcome. The standard errors are clustered by anomaly and month.

Table 5 reports the results. Panel A focuses on three measures of price efficiency and liquidity. Kyle's (1985) lambda is a common measure of liquidity and asymmetric information.

Amihud's (2002) illiquidity is defined as the past twelve-month average of daily absolute return divided by dollar volume. The variance ratio is based on the ratio of one-minute and 15-second return variances, effectively measuring autocorrelation in returns. A higher ratio corresponds to less liquidity and price efficiency. Kyle's lambda and variance ratio come from the WRDS Intraday Indicators database, which aggregates the TAQ database to the stock-day level.

The results from Panel A of Table 5 showcase economically large and highly statistically significant decreases for all market variables. Once stocks join EDGAR, their Kyle's lambda, Amihud's illiquidity, and variance ratio decrease by 0.54 ($*10^{-6}$), 0.22, and 0.012, respectively, or by about 10% relative to the pre-EDGAR levels. These three measures span different dimensions of price efficiency, asymmetric information, and liquidity. Thus, the EDGAR introduction prompted stocks in top and bottom portfolios of accounting anomalies to be more efficiently priced, and more liquid.

We next examine the effect of EDGAR introduction on participation by different types of investors. Total dollar trading volume captures activity by all investors, whereas overnight dollar trading volume captures primarily trading by retail investors because institutional investors trade predominantly during regular trading hours. We study participation by institutional investors using measures from the WRDS Thomson Reuters Institutional (13F) Holdings database, including total institutional ownership (as the fraction of shares outstanding), average ownership per institution, and ownership concentration (Herfindahl index), all of which are available quarterly. Finally, short interest (as the fraction of shares outstanding) from Compustat reflects participation by short sellers, an important class of arbitrageurs, especially given anomaly alphas are concentrated in the short leg.[18]

As shown in Panel B of Table 5, the EDGAR introduction prompted an increase in total trading volume, but overnight volume increased even more, indicating that retail investors substantially increased their participation. Total institutional ownership also increased, but the increase is primarily driven by more intense small institutional investor participation post-EDGAR, as reflected by lower ownership per investor and lower ownership concentration. These findings suggest that the noise trader channel may have been at play to a higher extent than the arbitrageur channel. Higher participation by retail and small institutional investors once a stock joins EDGAR is consistent with the noise trading channel. Finally, while participation by small investors increased, short interest declined for stocks that joined EDGAR. This result is consistent with short sellers finding less mispricing in top and bottom portfolios of accounting anomalies for stocks that already joined EDGAR. Overall, these results are more consistent with noise traders causing less mispricing rather than with arbitrageurs increasing their efforts to eliminate mispricing.

<TABLE 5 ABOUT HERE>

8. EDGAR effects and information delay costs

Our baseline results from Table 2 show that the EDGAR introduction reduced the profitability of accounting anomalies but did not affect the profitability of non-accounting anomalies. In this section, we focus on information delay costs as a potential mechanism to explain

[18] For a given implementation stage, we compute the monthly portfolio mean of the measures for all the stocks captured by the long and the short legs of equal-weight decile anomaly portfolios, and then analyze how portfolios' monthly measures respond to EDGAR implementation. We also apply this methodology to the three measures of price efficiency and liquidity.

the decline in profitability, supplementing the analyses of information availability costs and participation by arbitrageurs and noise traders from the previous sections.

While Compustat and other databases provide comprehensive data that help back-test trading strategies, they are delayed by at least a month as data providers accumulate, enter, and ship the information. EDGAR cut this lag because the information is almost instantly available after a company submits its report.

We implement two tests that seek to understand information delay costs by exploring whether faster information dissemination through EDGAR can explain our results. The first test centers on the notion that faster dissemination would affect anomalies differently. Anomalies that have high turnover in the top/bottom portfolios are particularly affected by faster dissemination because stocks in top/bottom portfolios can be identified only with the most up-to-date information. By contrast, for "low turnover" anomalies, portfolio assignment is persistent, so even stale information suffices to identify stocks in top/bottom portfolios. If faster dissemination speed is partially responsible for our main result, the accounting anomalies requiring high turnover should exhibit a more pronounced alpha decline post-EDGAR than the accounting anomalies requiring low turnover.

We define the turnover ratio as the share of new stocks that enter the top and bottom portfolios once the anomaly portfolio is rebalanced. We compute the pre-EDGAR turnover ratio for each accounting anomaly from October 1983 to September 1993 and rank the anomalies based on their average turnover ratio. The anomalies that exceed the median are classified as high turnover accounting anomalies. We then estimate a regression akin to Equation (1), with the accounting anomaly indicator variable split into two indicator variables for low- and high-turnover accounting anomalies, respectively.

Table 6 shows that the point estimates of alpha decline are 13 to 18 basis points per month larger for high-turnover accounting anomalies than for low-turnover accounting anomalies. In relative terms, the 13 to 18 basis point alpha decline corresponds to 33% to 44% of the overall decline of alpha for accounting anomalies due to EDGAR in Table 2. This result is consistent with the hypothesis that the post-EDGAR alpha decline is associated with the faster dissemination of accounting information. These differences, documented in the bottom row of Table 6, although sizeable, do not reach conventional levels of statistical significance. Thus, faster information dissemination plays a role but cannot solely explain our results because the profitability declines for both high- and low-turnover anomaly groups. For example, the profitability decline ranges from 33 to 41 basis points per month for low-turnover accounting anomalies.

<TABLE 6 ABOUT HERE>

The other test explores the timing of the anomaly signals relative to the portfolio formation. Intuitively, accounting information is received with less delay through EDGAR and, thus, is priced faster, reducing mispricing. Therefore, anomaly alphas should decay less for returns that are closer to the signal in time. This test takes advantage of a standard lag introduced into the anomaly portfolio formation being quite conservative to allow sufficient time for the information to reach those forming the portfolios. Indeed, anomalies based on annual (quarterly) portfolio formation are subject to the standard 6-month (3-month) lag from the date the information became available. To model a faster portfolio formation induced by EDGAR, we introduce a faster implementation of pure accounting anomaly portfolios by cutting the standard portfolio formation lag by one month. Introducing a two-month portfolio speed-up produces qualitatively similar results.

If the speed of information dissemination is at play, the value of the signals embedded in accounting information should be less stale if the delay until portfolio information is shorter. This, in turn, should result in a less pronounced attenuation of pure accounting anomaly profits for the portfolios formed with a one-month shorter delay. Also, to avoid the confounding effects of the stock price, we focus on the 103 pure accounting anomalies, that is, the anomalies that rest solely upon accounting information. For example, the anomaly associated with the book-to-market ratio is an accounting anomaly but not a pure accounting anomaly.

Table 7 presents the results of the canonical analyses of post-EDGAR pure accounting anomaly profitability for the standard portfolio timing (Panel A) and for the portfolio timing accelerated by a month (Panel B). The coefficients associated with $Post_{p,t}$ pertaining to standard portfolio formation timing (Panel A) are similar to those reported for accounting anomalies in Table 2 (the small differences stem from the fact that Table 7 features only *pure* accounting anomalies), ranging from 45 to 65 basis points. By contrast, the coefficients associated with $Post_{p,t}$ pertaining to one-month accelerated portfolio formation (Panel B) range from 29 to 42 basis points. Therefore, as shown at the bottom of Table 7, the ability to form portfolios a month sooner is associated with post-EDGAR pure accounting anomaly reduction by a statistically significant margin of 15 to 23 basis points. These results suggest that information acquisition costs encompass information delay costs, a significant indirect cost.

<TABLE 7 ABOUT HERE>

9. Robustness tests

9.1. Anomaly-time fixed effects

Exploiting the staggered introduction of EDGAR is at the root of our identification strategy. At every point in time during the implementation period, there are multiple versions of the same anomaly, populated with stocks from different EDGAR cohorts.

The key prediction we are testing is that, assuming similar pre-trends, the EDGAR version of the anomaly should have lower abnormal returns than the non-EDGAR version, achieved by including anomaly fixed effects and time fixed effects in our main specification. An even stronger test would replace them with still more stringent anomaly-time fixed effects to compare directly EDGAR and non-EDGAR versions of anomalies in the same month. The beauty of the staggered introduction of EDGAR, as opposed to a one-time change, is that it allows us to control for differential trends in individual anomaly returns. The downside of using anomaly-time fixed effects is much lower statistical power.

We report in Table A.3 in the Appendix our baseline results from Table 2, replicated with anomaly-month fixed effects instead of separate time and anomaly fixed effects. The results are qualitatively similar, especially for the post-estimation tests associated with accounting anomalies. EDGAR prompted alphas for accounting anomalies to decrease by between 0.31% and 0.49% per month, depending on the specification. The difference-in-difference coefficient that reflects the decline in profitability for accounting anomalies relative to non-accounting anomalies ranges from -0.18% to -0.27%. The results are broadly consistent with the main results in Table 2, but sharper fixed effects make it harder to produce a non-EDGAR explanation for our results because we are comparing the same anomaly in the same month across stocks that joined or not joined EDGAR.

9.2. EDGAR and anomaly publications

McLean and Pontiff (2016) show that anomalies become less profitable once academic literature makes them publicly known. In our next robustness check, we address the potential concern that the McLean and Pontiff (2016) post-publication effect may drive our results. Using journal publication dates, we find that only 16 accounting anomalies were published before or during the EDGAR introduction, while the remaining 109 accounting anomalies were published after EDGAR. The anomaly publication dates do not coincide with dates for EDGAR implementation phases; thus, our results are not explained by the post-publication effect.

We further explore whether the EDGAR introduction affects published and unpublished accounting anomalies differently. We estimate a regression akin to Equation (1), slightly altered so that the accounting anomaly indicator is split into two indicator variables, one for the accounting anomalies published before or during the EDGAR implementation period, and the other for the accounting anomalies published after EDGAR.

Table 8 shows that, for the 109 accounting anomalies published post-EDGAR, the alphas declined by 40 to 43 basis points per month in response to the EDGAR introduction. These magnitudes are very close to the alpha decline for the entire sample of 125 accounting anomalies in Table 2. Thus, these 109 accounting anomalies are not affected by the post-publication effect. For the 16 accounting anomalies published pre-EDGAR, the alphas declined by 45 to 95 basis points per month in response to EDGAR depending on the specification. The decline in alphas, if anything, is larger for the accounting anomalies published before or during EDGAR implementation, although the difference is only statistically significant in one out of four specifications in Table 8. These results confirm that our findings are quite distinct from the McLean and Pontiff (2016) post-publication effect.

A plausible explanation for the existence of the post-EDGAR alpha decline for the accounting anomalies published before or during EDGAR implementation is that awareness of such anomalies was a necessary, but not a sufficient condition for successfully implementing (and thus commoditizing) them before EDGAR. It was only post-EDGAR, after the data became more readily available, that investors both knew about such anomalies and had the requisite data to implement them, thus resulting in the decline in alpha documented in the top row of Table 8.

<TABLE 8 ABOUT HERE>

A broad perspective concerning the ultimate fate of accounting anomalies is that, in no small part because of the publication effect documented in McLean and Pontiff (2016), they are believed to have all but disappeared by the early 2000s. Our results do not speak to this issue because many of the accounting anomalies in our sample were not identified in the academic literature by the time EDGAR had been implemented. Rather, our results suggest that a drastic decline in the information costs of timely access to accounting information was an important catalyst in the process of attenuation of accounting anomaly profits.

9.3. Concerns regarding individual implementation phases

Randomized assignment of firms into implementation phases is crucial for the difference-in-difference methodology. If phase assignment is not fully random, a cross-phase comparison could be affected. The assignment was not perfectly random for the first phase. Before the EDGAR rollout in April 1993, the SEC called for volunteers to file electronically. This trial confirmed the integrity of the EDGAR system before engaging in a full-fledged implementation. The volunteer firms were subsequently assigned primarily to the first phase.

Also, accounting information for the first phase was delayed. The public could freely access EDGAR only after January 17, 1994 (Chang, Ljungqvist, and Tseng (2021)); before that date, investors had access to EDGAR through Mead Data Central. Given the standard three-month information lag assumption we introduce, if EDGAR were not easily available before January 1994, the first phase would have a humble cost-saving effect because its effective date (October 1, 1993) would fall before January 1994. The remaining implementation phases are unaffected by these issues because their effective dates are after January 1994. We address these concerns by repeating the baseline difference-in-difference analysis after dropping the first phase. The results, presented in Table A.4 in the Appendix, are very similar to the main results reported in Table 2, thus alleviating concerns about the first phase.

A related concern is that a small subset of implementation phases could be driving the results. Table 9 documents the contribution of each phase. Specifically, we augment the specification from Equation (1) by interacting the difference-in-difference indicator $Post_{p,t} * ACC_a$ with indicators for each implementation phase except the first (the reference phase). Whereas the effect of EDGAR introduction on anomaly returns varies across phases, the difference-in-difference coefficient is negative for every phase. The coefficients are not always statistically significant because each phase has relatively little statistical power. The 8th phase features the largest alpha attenuation of 112 basis points, while the alphas declined the least (by 2 basis points) for the last phase. This variation is natural given that we are trying to estimate changes in alphas that are inherently "noisy." These results also imply that excluding any one phase has a negligible effect on the overall results.

<TABLE 9 ABOUT HERE>

9.4. Pre-trends and falsification tests

The difference-in-difference analysis assumes parallel trends before the treatment. We formally test this assumption following the methodology from Gao and Huang (2020). Specifically, we estimate the baseline difference-in-difference regression over ten years *before* the actual EDGAR implementation, using pseudo-event dates. The pseudo-events of each EDGAR phase are assumed to take place five years *before* the actual phase dates. Accordingly, the indicator variable $Post_{p,t}$ is redefined to equal one if month *t* is after the first *pseudo*-event date on which investors presumably trade on the latest EDGAR information.

Panel A of Table A.5 in the Appendix presents the result of the pre-trend test and confirms that the parallel trend assumption is likely to hold in our setting. The difference-in-difference coefficient switches signs and becomes positive, though small in magnitude and statistically insignificant (*t*-statistics range from 0.30 to 0.55). That is, accounting anomalies become slightly more profitable relative to non-accounting anomalies on the pseudo-implementation dates, the opposite of the effect for the actual EDGAR implementation.

We also implement a falsification test, similarly as the pre-trends test. Once again following Gao and Huang (2020), we estimate the baseline difference-in-difference regression over ten years *following* the actual EDGAR implementation, using the pseudo-event dates from five years *after* the actual phase dates. The indicator variable $Post_{p,t}$ is redefined accordingly. Panel B of Table A.5 in the Appendix reports the results of the falsification test. Similar to the pre-trends test, the difference-in-difference coefficient sometimes switches signs and becomes positive, but it is small and statistically insignificant (*t*-statistics range from -0.15 to 0.33).

These results indicate that accounting anomaly alphas decline shortly around the EDGAR implementation, rather than a long time before or after EDGAR. Figure 1 further shows that accounting alphas decline discontinuously around the actual phase implementation dates, thus further supporting the difference-in-difference identification.

9.5. Dropping annual siblings, thin portfolios

In this section, we address two potential issues: annual and quarterly versions of the same anomaly conceivably could be highly correlated, and some anomaly portfolios could contain only a few stocks. First, as discussed in Section 4.2, we constructed the sample of 205 anomalies by following the process of Chen and Zimmermann (2020). That process resulted in 23 anomalies based on both annual and quarterly portfolio formation, introducing the issue of potential double-counting. At the outset, returns for these “sibling” anomalies pass the correlation filter described in Section 4.2 and, thus, contain independent information. Nonetheless, we further address the issue of potential double-counting by estimating our baseline results from Table 2 on the sample of 182 anomalies, obtained from the full sample of 205 anomalies by dropping the annual “sibling” anomalies. The results, presented in Panel A of Table A.6 in the Appendix, are virtually identical to those from Table 2. Thus, the presence of annual sibling anomalies does not drive our results.

Second, the portfolio construction of long and short legs of an anomaly in each phase could result in “thin” portfolios, consisting of relatively few stocks. These thin portfolios could make our estimates more variable and thus imprecise. However, this issue affects only a few observations because the median number of stocks in the top/bottom portfolio is 25. To alleviate this concern, we replicate our baseline results from Table 2 with the added step of dropping all the observations based on “thin” portfolios consisting of fewer than five stocks in either long or short portfolio leg. This step creates a gently unbalanced panel (13.4% of observations are affected). Once again, the results, presented in Panel B of Table A.6 of the Appendix, are virtually identical to those from Table 2. Therefore, the issue of thin portfolios does not affect our results.

10. Conclusion

In this paper, we investigate the causal effect of the information costs on the anomaly portfolio returns. We use the SEC’s EDGAR introduction as a quasi-exogenous shock that lowers the costs of acquiring accounting information. Using the difference-in-difference framework, we find that alphas of accounting anomalies attenuate on average by 47 to 62 basis points per month in response to the EDGAR introduction. This decline explains away most of the pre-EDGAR accounting anomaly alphas. By contrast, the profitability of the non-accounting anomalies remains largely unaffected by the EDGAR introduction. Overall, as accounting data become easier to acquire and process, the profitability of trading strategies that rely on these data declines.

The profitability decline translates to the costs of acquiring accounting information that investors had to bear in the absence of the EDGAR system. Thus, by lowering the accounting information costs, the EDGAR introduction increased price informativeness, which, in turn, eroded the profitability of strategies based on accounting signals. The results of our empirical tests support multiple channels. The profitability decline is driven partially by more aggressive trading by arbitrageurs and partially by less mispricing caused by noise traders. Finally, post-EDGAR, investors receive and trade on accounting information faster, leading to faster price discovery.

Our results remain highly relevant for today's markets. To generate alpha, investors strive to establish an information advantage by exploiting novel data. Pre-EDGAR, accounting data was at the cutting edge of investors' data exploration efforts. EDGAR made accounting data widely available and thus less profitable to trade on it. Arbitrageurs move on to other, more costly, and thus less explored data. Our conclusions likely extrapolate to alternative data.

Data, as a source of information, are central to arbitrageurs' success. For example, Citadel CEO Ken Griffin notes that "our ability to leverage big data effectively in our investment process is critical to our success as a firm" (Randle (2018)). Many hedge funds have introduced the Chief Data Officer position to highlight the importance of data. Our paper offers an insight into the role that information costs play in the investment process.

References

Albuquerque, R., Song, S., and Yao, C. (2020). The price effects of liquidity shocks: A study of the SEC's tick size experiment. *Journal of Financial Economics* 138(3), 700-724.

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

Ben-David, I., Li, J., Rossi, A. and Song, Y., 2021. Discontinued Positive Feedback Trading and the Decline of Return Predictability (No. w28624). National Bureau of Economic Research.

Bowles, B., A. V. Reed, M. C. Ringgenberg, and J. R. Thornock, 2020, Anomaly Time. Working Paper. *Available at SSRN*.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of finance, 52*(1), 57-82.

Chang, Y. C., Ljungqvist, A., and Tseng, K. (2021). Do corporate disclosures constrain strategic analyst behavior? *Available at SSRN*.

Chang, Y. C., Hsiao, P. J., Ljungqvist, A., and Tseng, K. (2022). Testing disagreement models. *The Journal of Finance* 77(4), 2239-2285.

Chen, Y., Kelly, B., and Wu, W. (2020). Sophisticated investors and market efficiency: Evidence from a natural experiment. *Journal of Financial Economics* 138, 316–341.

Chen, A.Y. and Velikov, M. (2023). Zeroing in on the expected returns of anomalies. *Journal of Financial and Quantitative Analysis* 58(3), 968-1004.

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

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

Dong, Y., Li, O. Z., Lin, Y., and Ni, C. (2016). Does Information-Processing Cost Affect Firm-Specific Information Acquisition? Evidence from XBRL Adoption. *Journal of Financial and Quantitative Analysis* 51(2), 435-462.

Drechsler, I. and Drechsler, Q. F. (2014). *The shorting premium and asset pricing anomalies* (No. w20282). National Bureau of Economic Research.

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

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

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

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

Frazzini, A., Israel, R., and Moskowitz, T. J. (2018). Trading costs. *Available at SSRN 3229719*.

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

Geczy, C. C., Musto, D. K., and Reed, A. V. (2002). Stocks are special too: An analysis of the equity lending market. *Journal of Financial Economics*, *66*(2-3), 241-269.

Goldstein, I., Yang, S. and Zuo, L. (2023). The real effects of modern information technologies: Evidence from the EDGAR implementation. Journal of Accounting Research 61(5), 1699-1733.

Gomez, E.A.(2023). The Effect of Mandatory Disclosure Dissemination on Information Asymmetry among Investors: Evidence from the Implementation of the EDGAR System. The Accounting Review, pp.1-23.

Green, J., Hand, J. R., and 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., and Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review* 70(3), 393-408.

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

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

Jensen, T. I., Kelly, B. T., and Pedersen, L. H. (2021). *Is there a replication crisis in finance?* (No. w28432). National Bureau of Economic Research.

Kadan, O. and Manela, A., 2019. Estimating the value of information. *The Review of Financial Studies* 32(3), 951-991.

Kelly, B. and Ljungqvist, A. (2012). Testing Asymmetric-Information Asset Pricing Models. *The Review of Financial Studies* 25(5), 1366–1413.

Keim, D. B., and Madhavan, A. (1997). Transactions costs and investment style: an inter-exchange analysis of institutional equity trades. *Journal of Financial Economics* 46(3), 265-292.

Kern, B. B., and Morris, M. H. (1994). Differences in the Compustat and expanded Value Line databases and the potential impact on empirical research. *Accounting Review* 69(1), 274-284.

Korajczyk, R. A., and Sadka, R. (2004). Are momentum profits robust to trading costs? *The Journal of Finance* 59(3), 1039-1082.

Kothari, S. P., Shanken, J., and Sloan, R. G. (1995). Another look at the cross-section of expected stock returns. *The Journal of Finance* 50(1), 185-224.

Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica* 53(6), 1315–1335.

Lesmond, D. A., Schill, M. J., and Zhou, C. (2004). The illusory nature of momentum profits. *Journal of Financial Economics* 71(2), 349-380.

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* (accessed 25 July 2021).

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., and 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

Muravyev, D., Pearson, N. D., and Pollet, J. M. (2022). Anomalies and Their Short-Sale Costs. Available at SSRN 4266059.

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

Patton, A. J. and Weller, B. M. (2020). What you see is not what you get: The costs of trading market anomalies. *Journal of Financial Economics*, *137*(2), 515-549.

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.

Randle, C. (2018). Alternative Data: A New Source for Alpha. Available from www.futuresmag.com/2018/02/15/alternative-data-new-source-alpha .

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

Ross, S. A. (1976). The Arbitrage Theory of Capital Asset Pricing. *Journal of Economic Theory* 13, 341–360.

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.

Stambaugh, R. F., Yu, J., and Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics*, *104*(2), 288-302.

Star Tribune (1993, April 26). "EDGAR on-line - 240 Firms Start Filing Documents Electronically with the SEC Today." By Glenn Howatt. [online] NewsBank: Access World News. (accessed 25 July 2021)

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
EDGAR implementation schedule

This table shows the EDGAR implementation timeline. EDGAR was implemented in ten phases over three years. The anomaly literature assumes a one-quarter lag before accounting information is available to investors. “Effective date” accounts for this lag and is the first date when investors start trading the EDGAR filers' stocks using the latest information retrieved from EDGAR. The last column reports the number of stocks in our sample for each phase that we match successfully with Compustat and CRSP.

Implementation phase	Implementation date	Effective date	Number of stocks
1	4/26/1993	10/1/1993	149
2	7/19/1993	1/1/1994	541
3	10/4/1993	4/1/1994	564
4	12/6/1993	4/1/1994	737
5	8/1/1994	1/1/1995	1,033
6	11/1/1994	4/1/1995	866
7	5/1/1995	10/1/1995	858
8	8/1/1995	1/1/1996	756
9	11/1/1995	4/1/1996	386
10	5/1/1996	10/1/1996	2,723

Table 2
Difference-in-difference estimates of EDGAR effect on anomalies

This table presents the coefficients from the main difference-in-difference regression from Equation (1). The coefficient associated with $Post_{p,t} * ACC_a$ captures the gap in the extent to which anomaly portfolio alphas change in response to EDGAR for accounting anomalies relative to non-accounting anomalies. We also report the coefficient sum ($Post_{p,t} + Post_{p,t} * ACC_a$) and the mean of the dependent variable (portfolio alpha). The main regression is estimated on an anomaly-phase-month panel, where $Post_{p,t}$ is an indicator equal to one if month t is on or after the effective date for phase p, and equal to zero if month t is before that date; ACC_a is an indicator variable that equals one if a is an accounting anomaly, and equals zero if a is a non-accounting anomaly. The dependent variables $\widehat{\alpha_{a,p,t}}$ in the four columns of the table are the Fama-French six-factor alphas for quintile (1-5) or decile (1-10), equal-weighted (EW) or value-weighted (VW) portfolios. All specifications contain anomaly and month fixed effects. The sample extends over 205 anomalies from January 1992 to December 1997. Panel A provides estimates without additional control variables. Panel B provides estimates with the full set of control variables including firm size, book-to-market, and illiquidity. The full list is described in Section 5. The portfolio alphas and the coefficients are expressed in percentages. The standard errors are clustered by anomaly and month. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	FF6 alpha 1-5 EW	FF6 alpha 1-5 VW	FF6 alpha 1-10 EW	FF6 alpha 1-10 VW
Panel A: EDGAR effect on anomalies, no additional controls				
$Post_{p,t}$	-0.066	-0.042	-0.120	-0.123
(non-accounting anomalies)	(-0.63)	(-0.34)	(-0.95)	(-0.88)
$Post_{p,t} * ACC_a$	-0.419***	-0.429***	-0.421***	-0.494***
(difference-in-difference)	(-3.55)	(-3.69)	(-3.11)	(-3.47)
N	129,893	129,893	129,893	129,893
R-squared	0.0110	0.0084	0.0097	0.0078
Post-estimation test:				
$Post_{p,t} + Post_{p,t} * ACC_a$	-0.485***	-0.471***	-0.541***	-0.617***
(accounting anomalies)	(-4.85)	(-4.03)	(-3.78)	(-4.09)
Mean of dependent variable	0.301	0.231	0.332	0.281
Panel B: EDGAR effect on anomalies, full set of controls				
$Post_{p,t}$	0.0380	0.0294	-0.0482	-0.0939
(non-accounting anomalies)	(0.36)	(0.24)	(-0.38)	(-0.67)
$Post_{p,t} * ACC_a$	-0.417***	-0.425***	-0.407***	-0.476***
(difference-in-difference)	(-3.58)	(-3.68)	(-3.03)	(-3.35)
N	128,902	128,902	128,903	128,903
R-squared	0.0114	0.0085	0.0100	0.0080
Post-estimation test:				
$Post_{p,t} + Post_{p,t} * ACC_a$	-0.379***	-0.395***	-0.455***	-0.570***
(accounting anomalies)	(-4.22)	(-3.32)	(-3.51)	(-3.86)
Mean of dependent variable	0.304	0.234	0.335	0.285

Table 3
Information availability and EDGAR effect on accounting anomalies

This table reports the coefficients from the regression in Equation (3), estimated separately for stocks with high and low information availability. Analyst coverage (Panel A) and market capitalization (Panel B) proxy for information availability. A stock is classified as low information if its analyst coverage (or market capitalization) is below the cross-sectional median pre-EDGAR (January 1990 to December 1992). For each of the 125 accounting anomalies, phase, and month, we estimate the first-pass regression from Equation (2) separately for high- and low-information availability stocks. This step creates an anomaly-phase-month panel of beta estimates for how well an anomaly signal predicts future stock returns. Next, we estimate the second-pass regression from Equation (3) that estimates how the EDGAR introduction affects anomaly predictability. The table reports the coefficients of interest for each regression and the difference between the two groups (last column). All specifications contain anomaly and month fixed effects. The standard errors are clustered by anomaly and by month. Robust t-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

Panel A: Analyst coverage			
	(1) Low analyst coverage $\delta_1 + \delta_2 + \delta_3$	(2) High analyst coverage δ_1	(1) – (2) Difference $\delta_2 + \delta_3$
$Post_{p,t}$	-0.706***	0.032	-0.739***
	(-3.27)	(0.17)	(-2.74)
N	78,715	78,388	157,103
R-squared	0.0115	0.0080	0.0090

Panel B: Market capitalization			
	(1) Small stocks $\delta_1 + \delta_2 + \delta_3$	(2) Large stocks δ_1	(1) – (2) Difference $\delta_2 + \delta_3$
$Post_{p,t}$	-0.701***	-0.100	-0.608*
	(-3.23)	(-0.46)	(-1.91)
N	78,885	78,134	157,019
R-squared	0.0104	0.0094	0.0089

Table 4
EDGAR effect on long and short anomaly 0ortfolio legs
This table presents the coefficients from the baseline difference-in-difference regression from Equation (1), estimated separately for the long (Panel A) and short (Panel B) legs of the 205 anomaly portfolios. $Post_{p,t}$ is an indicator equal to one if month t is on or after the effective date for phase p, and equal to zero if month t is before that date; ACC_a is an indicator variable that equals one if a is an accounting anomaly and equals zero if a is a non-accounting anomaly. The dependent variables $\widehat{\alpha_{a,p,t}}$ in the four columns of the table are the Fama-French six-factor long (Panel A) or short (Panel B) alphas for quintile (1-5) or decile (1-10), equal-weighted (EW) or value-weighted (VW) portfolios. All specifications contain anomaly and month fixed effects. The sample extends over 205 anomalies in the period from January 1992 to December 1997. The portfolio alphas and the coefficients are expressed in percentages.

	FF6 alpha 1-5 EW	FF6 alpha 1-5 VW	FF6 alpha 1-10 EW	FF6 alpha 1-10 VW
Panel A: Long Leg Anomaly Portfolios				
$Post_{p,t}$	-0.221	0.031	-0.238	-0.032
(non-accounting anomalies)	(-1.29)	(0.25)	(-1.28)	(-0.21)
$Post_{p,t} * ACC_a$	-0.0004	-0.126	0.059	-0.0445
(difference-in-difference)	(-0.00)	(-1.26)	(0.51)	(-0.35)
N	129,893	129,893	129,893	129,893
R-squared	0.0339	0.0214	0.0246	0.0185
Post-estimation test:				
$Post_{p,t} + Post_{p,t} * ACC_a$	-0.221	-0.095	-0.179	-0.076
(accounting anomalies)	(-1.18)	(-0.71)	(-0.85)	(-0.49)
Mean of dependent variable	0.674	0.566	0.668	0.569
Panel B: Short Leg Anomaly Portfolios				
$Post_{p,t}$	0.128	-0.088	0.099	-0.100
(non-accounting anomalies)	(0.70)	(-0.56)	(0.48)	(-0.55)
$Post_{p,t} * ACC_a$	-0.431***	-0.324***	-0.499***	-0.482***
(difference-in-difference)	(-4.07)	(-3.16)	(-3.77)	(-3.44)
N	129,893	129,893	129,893	129,893
R-squared	0.0395	0.0266	0.0303	0.0238
Post-estimation test:				
$Post_{p,t} + Post_{p,t} * ACC_a$	-0.303	-0.412**	-0.400	-0.582***
(accounting anomalies)	(-1.45)	(-2.30)	(-1.64)	(-2.68)
Mean of dependent variable	-0.291	-0.288	-0.250	-0.230

Table 5
EDGAR effect on market quality and participation
This table shows the response of market quality and participation measures to the EDGAR implementation. We first compute an average market outcome of interest over stocks in the top and bottom decile portfolios for each accounting anomaly, phase, and month. We then estimate for each of the average measures a difference-in-difference regression from Equation (4). Panel A focuses on measures of price efficiency and liquidity. Kyle's lambda, a measure of market impact. Amihud's illiquidity is defined as the past twelve-month average of daily return divided by dollar volume (Amihud (2002)). Variance ratio is the absolute value of the ratio of the variance of one-minute log returns and four times 15-second log returns minus one. Panel B focuses on measures of participation. Natural logarithm of volume (overnight volume) is the logarithm of average monthly dollar trading volume (overnight trading dollar volume). Institutional ownership is the ratio between the total institutional ownership percentage of shares outstanding and the number of 13-F institutional owners, expressed in percentages. Ownership concentration is the Herfindahl index measure of institutional ownership concentration. Short interest is the number of shares held short as of the settlement date divided by shares outstanding. The sample period is from January 1993 to December 1997. The regressions contain anomaly and month fixed effects. The standard errors are clustered by anomaly and month. Robust t-statistics are presented along with the regression coefficient estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

Panel A. Price efficiency and liquidity

	$Post_{p,t}$	t-statistic	N	R-squared	Mean of dep.variable
Kyle's lambda (x 10^{-6})	-0.5420***	-4.80	63,712	0.2108	4.96
Amihud illiquidity	-0.2192***	-10.36	63,712	0.0969	3.54
Variance ratio	-0.0120***	-5.18	63,712	0.4211	0.13

Panel B. Participation by investor type

	$Post_{p,t}$	t-statistic	N	R-squared	Mean of dep.variable
ln(Dollar volume)	0.9942***	10.67	63,712	0.1985	17.16
ln(Overnight dollar volume)	1.1163***	7.88	63,712	0.1646	12.08
Institutional ownership	0.0929***	8.92	63,427	0.1210	0.40
Institutional ownership per investor	-0.0035***	-13.21	63,427	0.1665	0.0094
Ownership concentration (HHI)	-0.0526***	-13.23	63,430	0.1472	0.13
Short interest	-0.0973***	-5.45	48,950	0.0914	0.24

Table 6
Baseline difference-in-difference for accounting anomalies by turnover ratio

This table documents the way that attenuation of accounting anomalies affected by EDGAR varies with the turnover ratio of the anomalies. Specifically, the table presents the coefficients from the difference-in-difference anomaly portfolios regression in Equation (1), except the accounting anomaly indicator is split into two indicators for low and high turnover accounting anomalies. The table also presents the estimates of the statistical differences of the difference-in-difference coefficients between the high- and the low-turnover groups. The turnover ratio is defined as the total number of new incoming stocks divided by the total number of stocks in the existing portfolio (for the long and the short leg) when the accounting anomaly portfolio updates its signal and rebalances its stocks. We compute the pre-EDGAR turnover ratio for all the accounting anomalies from October 1, 1983, to September 30, 1993, using the stocks in the sample period. We then sort the accounting anomalies based on their turnover ratio rank percentile. The accounting anomalies that exceed the 50th percentile are classified as "High Turnover Accounting Anomalies," and the remaining accounting anomalies are classified as "Low Turnover Accounting Anomalies." Non-accounting anomalies serve as the control group (benchmark). The standard errors are clustered by anomaly and by month. Robust t-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

	FF6 alpha 1-5 EW	FF6 alpha 1-5 VW	FF6 alpha 1-10 EW	FF6 alpha 1-10 VW
$Post_{p,t} \times Acc\_HighTurnover_a$	-0.487***	-0.521***	-0.510***	-0.579***
	(-4.08)	(-4.07)	(-3.39)	(-3.31)
$Post_{p,t} \times Acc\_LowTurnover_a$	-0.349**	-0.336**	-0.330**	-0.409***
	(-2.63)	(-2.57)	(-2.18)	(-2.69)
N	129,893	129,893	129,893	129,893
R-squared	0.0110	0.0084	0.0098	0.0078
Post-estimation test:				
$Post_{p,t} \times HighAcc_a - Post_{p,t} \times LowAcc_a$	-0.137	-0.185	-0.180	-0.169
	(-1.52)	(-1.60)	(-1.32)	(-1.04)
Mean of dependent variable	0.301	0.231	0.332	0.281

Table 7
Difference-in-difference estimates of EDGAR effect on pure accounting anomalies: Standard portfolio formation and one-month accelerated portfolio formation

This table presents the coefficients from the main difference-in-difference regression from Equation (1), restricted to pure accounting anomalies. The coefficient associated with $Post_{p,t}$ reflects the decline in information costs post-EDGAR. It captures the gap in the extent to which pure accounting anomaly portfolio alphas change in response to EDGAR relative to non-accounting anomalies. The regression is estimated on an anomaly-phase-month panel, where $Post_{p,t}$ is an indicator equal to one if month t is on or after the effective date for phase *p*, and equal to zero if month *t* is before that date. The dependent variables $\widehat{\alpha_{a,p,t}}$ in the four columns of the table are the Fama-French six-factor alphas for quintile (1-5) or decile (1-10), equal-weighted (EW) or value-weighted (VW) portfolios. All specifications contain anomaly and month fixed effects. The sample extends over 103 pure accounting anomalies from January 1992 to December 1997. Panel A provides estimates involving portfolio formation timing analogous to that from the baseline tests. Panel B provides estimates involving portfolio formation timing accelerated by one month relative to the portfolio formation timing from the baseline tests. The portfolio alphas and the coefficients are expressed in percentages. The standard errors are clustered by anomaly and month. Robust t-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	FF6 alpha 1-5 EW	FF6 alpha 1-5 VW	FF6 alpha 1-10 EW	FF6 alpha 1-10 VW
Panel A: EDGAR effect on pure accounting anomalies, standard portfolio formation timing				
$Post_{p,t}$	-0.452***	-0.493***	-0.505***	-0.645***
	(-4.08)	(-3.78)	(-3.00)	(-3.52)
N	66,625	66,625	66,625	66,625
R-squared	0.0124	0.0094	0.0113	0.0090
Panel B: EDGAR effect on pure accounting anomalies, one-month accelerated portfolio formation timing				
$Post_{p,t}$	-0.302***	-0.291**	-0.296**	-0.417**
	(-2.84)	(-2.19)	(-2.04)	(-2.41)
N	66,553	66,553	66,553	66,553
R-squared	0.0134	0.0120	0.0120	0.0118
Post-estimation test:				
$Post_{p,t}$(Panel A) $-$ $Post_{p,t}$(Panel B)	-0.150***	-0.202***	-0.209***	-0.228***
	(-3.18)	(-4.57)	(-3.21)	(-3.66)

Table 8
Baseline difference-in-difference pre- and post-publication

This table documents the way that attenuation of accounting anomalies affected by EDGAR varies relative to the anomaly publication date. The table presents the coefficients from the difference-in-difference anomaly portfolios regression in Equation (1), except the accounting anomaly indicator is split into two indicator variables, one for the accounting anomalies published before or during the EDGAR implementation period ($Acc\_PubPreEDGAR_a$), and the other for the accounting anomalies published after EDGAR ($Acc\_PubAfterEDGAR_a$). The table also presents the estimates of the differences between the difference-in-difference coefficients associated with the two publication groups. Following McLean and Pontiff (2016). The publication date is defined as the year and month on the cover of the journal. Of the 125 accounting anomalies we investigate, 16 anomalies were published before or during the EDGAR implementation period. The standard errors are clustered by anomaly and by month. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

	FF6 alpha 1-5 EW	FF6 alpha 1-5 VW	FF6 alpha 1-10 EW	FF6 alpha 1-10 VW
$Post_{p,t} \times Acc\_PubPreEDGAR_a$	-0.446***	-0.589***	-0.590***	-0.950***
	(-2.73)	(-3.31)	(-2.96)	(-4.87)
$Post_{p,t} \times Acc\_PubAfterEDGAR_a$	-0.415***	-0.405***	-0.396***	-0.427***
	(-3.53)	(-3.51)	(-2.93)	(-3.03)
N	129,893	129,893	129,893	129,893
R-squared	0.0110	0.0084	0.0097	0.0079
Post-estimation test:				
$Post_{p,t} \times AccPubPreEDGAR_a - Post_{p,t} \times Acc\_PubAfterEDGAR_a$	-0.031	-0.184	-0.194	-0.522***
	(-0.27)	(-1.30)	(-1.23)	(-3.79)
Mean of dependent variable	0.301	0.231	0.332	0.281

Table 9
Baseline difference-in-difference by implementation phase
This table presents the coefficients from the difference-in-difference anomaly portfolios regression similar to that from Equation (1). The estimated specification features additional interaction terms of the form $Post_{p,t} * ACC_a$* Phase i, thus enabling the estimation of the difference-in-difference coefficient separately for each phase (Phase 3 has the same effective date as Phase 4). $Post_{p,t}$ is an indicator equal to one if month t is on or after the effective date for phase p, and equal to zero if month t is before that date; ACC_a is an indicator variable that equals one if a is an accounting anomaly and equals zero if a is a non-accounting anomaly. The dependent variables $\widehat{\alpha_{a,p,t}}$ in the four columns of the table are the Fama-French six-factor alphas for quintile (1-5) or decile (1-10), equal-weighted (EW) or value-weighted (VW) portfolios. The sample extends over 205 anomalies in the period from January 1992 to December 1997. The portfolio alphas and the coefficients are expressed in percentages. All specifications contain anomaly and month fixed effects. The standard errors are clustered by anomaly and month. Robust t-statistics are presented in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	FF6 alpha 1-5 EW	FF6 alpha 1-5 VW	FF6 alpha 1-10 EW	FF6 alpha 1-10 VW
$Post_{p,t}$	0.0136	-0.0682	0.0466	-0.150
	(0.07)	(-0.29)	(0.20)	(-0.49)
$Post_{p,t} * ACC_a$* Phase 1	-0.286	0.0492	-0.371	0.154
	(-1.54)	(0.21)	(-1.61)	(0.52)
$Post_{p,t} * ACC_a$* Phase 2	-0.347	-0.234	-0.252	-0.209
	(-1.58)	(-1.22)	(-1.09)	(-1.05)
$Post_{p,t} * ACC_a$* Phase 3/4	-0.0585	-0.122	-0.0308	-0.0492
	(-0.41)	(-0.77)	(-0.19)	(-0.26)
$Post_{p,t} * ACC_a$* Phase 5	-0.171	-0.0310	-0.160	-0.0869
	(-0.91)	(-0.13)	(-0.61)	(-0.26)
$Post_{p,t} * ACC_a$* Phase 6	-0.428*	-0.438	-0.478	-0.785**
	(-1.94)	(-1.61)	(-1.65)	(-2.19)
$Post_{p,t} * ACC_a$* Phase 7	-0.748**	-1.026***	-0.940**	-1.434***
	(-2.19)	(-2.65)	(-2.23)	(-3.16)
$Post_{p,t} * ACC_a$* Phase 8	-1.122**	-0.615	-1.160*	-0.701
	(-2.33)	(-1.10)	(-1.85)	(-1.12)
$Post_{p,t} * ACC_a$* Phase 9	-0.775*	-1.230***	-1.008*	-1.302**
	(-1.84)	(-3.19)	(-1.93)	(-2.45)
$Post_{p,t} * ACC_a$* Phase 10	-0.0233	-0.562**	0.0742	-0.352
	(-0.10)	(-2.24)	(0.24)	(-1.13)
N	129,893	129,893	129,893	129,893
R-squared	0.0123	0.0093	0.0107	0.0089
Mean of dependent variable	0.301	0.231	0.332	0.281

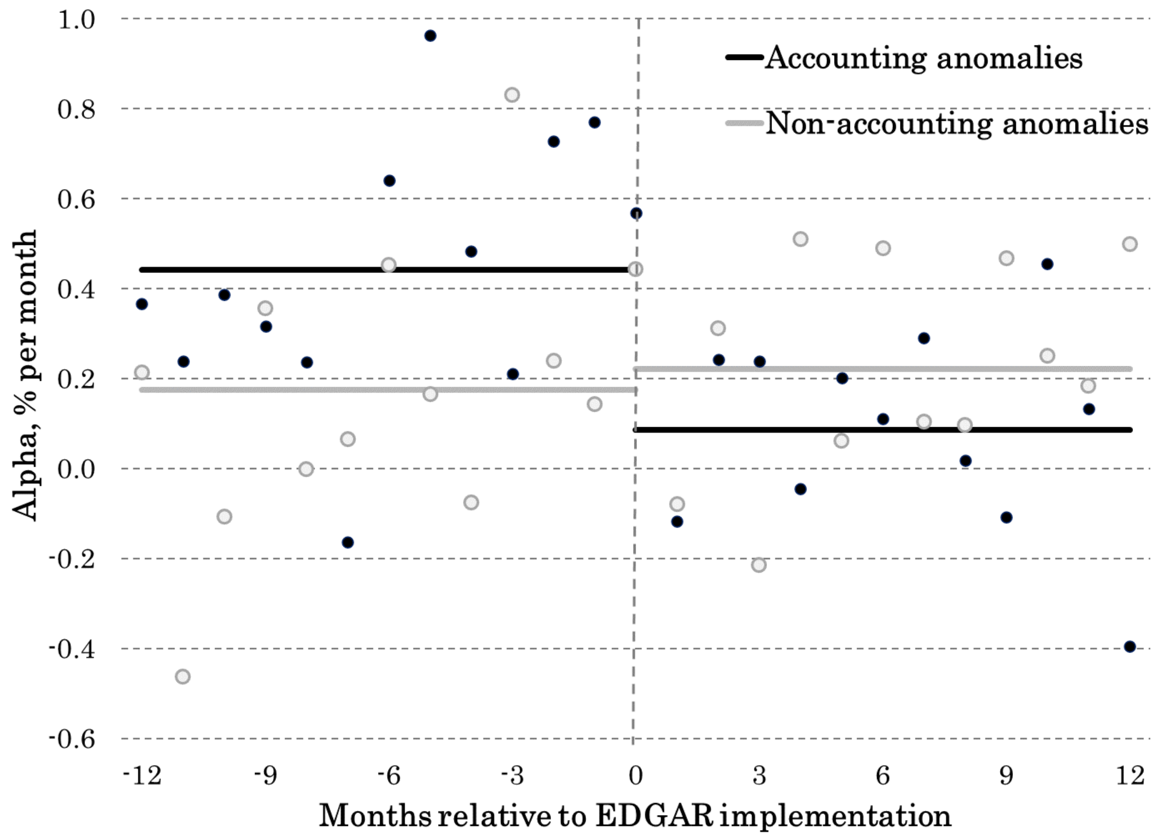

Figure 1: Effect of EDGAR on anomaly profitability

This figure shows the average alphas for anomaly portfolios around the staggered EDGAR implementation. The black and grey lines show the average alphas for accounting and non-accounting anomalies, respectively, in 12 months pre- and post-EDGAR effective dates. Because implementation dates differ for each of the ten phases, we center each phase's implementation date at zero and average over phases. Dots show alphas for individual months relative to implementation dates. Black dots represent accounting anomalies, whereas grey dots represent non-accounting anomalies. The Y-axis reports the average monthly alphas estimated using the six-factor Fama-French model across all four portfolio specifications in the main analyses reported in Table 2.

Appendix

Table A.1
Robustness: Difference-in-difference estimates of EDGAR effect on anomalies *keeping* anomalies with *negative* alphas in the pre-EDGAR period
This table presents the coefficients from the main difference-in-difference regression from Equation (1) when anomalies with negative alphas in the pre-EDGAR period are not dropped. The coefficient associated with $Post_{p,t} * ACC_a$ captures the gap in the extent to which anomaly portfolio alphas change in response to EDGAR for accounting anomalies relative to non-accounting anomalies. To facilitate interpretation, we also report the coefficient sum ($Post_{p,t} + Post_{p,t} * ACC_a$) and the mean of the dependent variable (portfolio alpha). All the variables are identical to those utilized in Table 2. The sample extends over 205 anomalies from January 1992 to December 1997. The portfolio alphas and the coefficients are expressed in percentages. The standard errors are clustered by anomaly and month. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	FF6 alpha 1-5 EW	FF6 alpha 1-5 VW	FF6 alpha 1-10 EW	FF6 alpha 1-10 VW
$Post_{p,t}$	0.0156	0.0211	-0.0705	-0.0873
(non-accounting anomalies)	(0.18)	(0.20)	(-0.67)	(-0.73)
$Post_{p,t} * ACC_a$	-0.305***	-0.298***	-0.280**	-0.315***
(difference-in-difference)	(-3.06)	(-3.02)	(-2.66)	(-2.96)
N	162,243	162,243	162,246	162,246
R-squared	0.0107	0.0077	0.0096	0.0075
Post-estimation test:				
$Post_{p,t} + Post_{p,t} * ACC_a$	-0.289***	-0.277***	-0.351***	-0.403***
(accounting anomalies)	(-3.52)	(-2.71)	(-2.94)	(-3.20)
Mean of dependent variable	0.247	0.185	0.282	0.248

Table A.2
Robustness: Difference-in-difference estimates of EDGAR effect on anomalies with raw returns

This table presents the coefficients from the main difference-in-difference regression from Equation (1) using raw portfolio returns instead of FF6 alphas. The coefficient associated with $Post_{p,t} * ACC_a$ captures the gap in the extent to which anomaly portfolio returns change in response to EDGAR for accounting anomalies relative to non-accounting anomalies. We also report the coefficient sum ($Post_{p,t} + Post_{p,t} * ACC_a$) and the mean of the dependent variable (portfolio return). The main regression is estimated on an anomaly-phase-month panel, where $Post_{p,t}$ is an indicator equal to one if month t is on or after the effective date for phase *p*, and equal to zero if month *t* is before that date; ACC_a is an indicator variable that equals one if *a* is an accounting anomaly, and equals zero if a is a non-accounting anomaly. The dependent variables $\widehat{r_{a,p,t}}$ in the four columns of the table are the anomaly portfolio raw returns for quintile (1-5) or decile (1-10), equal-weighted (EW) or value-weighted (VW) portfolios. All specifications contain anomaly and month fixed effects. The sample extends over 205 anomalies from January 1992 to December 1997. Panel A provides estimates without additional control variables. Panel B provides estimates with the full set of control variables including firm size, book-to-market, and illiquidity. The full list is described in the main text (Section 5). The portfolio returns and the coefficients are expressed in percentages. The standard errors are clustered by anomaly and month. Robust *t*-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	1-5 EW	1-5 VW	1-10 EW	1-10 VW
Panel A: EDGAR effect on anomalies, no additional controls				
$Post_{p,t}$	-0.177	-0.155	-0.270*	-0.270
(non-accounting anomalies)	(-1.36)	(-1.09)	(-1.74)	(-1.61)
$Post_{p,t} * ACC_a$	-0.304**	-0.305**	-0.297*	-0.357*
(difference-in-difference)	(-2.07)	(-2.10)	(-1.71)	(-1.96)
N	130,089	130,089	130,089	130,089
R-squared	0.0158	0.0114	0.0132	0.0105
Post-estimation test:				
$Post_{p,t} + Post_{p,t} * ACC_a$	-0.482***	-0.460***	-0.567***	-0.627***
(accounting anomalies)	(-4.12)	(-3.67)	(-3.67)	(-3.80)
Mean of dependent variable	0.420	0.343	0.484	0.416
Panel B: EDGAR effect on anomalies, full set of controls				
$Post_{p,t}$	0.00885	0.000203	-0.103	-0.161
(non-accounting anomalies)	(0.06)	(0.00)	(-0.64)	(-0.92)
$Post_{p,t} * ACC_a$	-0.305**	-0.299**	-0.280	-0.334*
(difference-in-difference)	(-2.04)	(-2.03)	(-1.59)	(-1.82)
N	128,902	128,902	128,903	128,903
R-squared	0.0167	0.0120	0.0140	0.0110
Post-estimation test:				
$Post_{p,t} + Post_{p,t} * ACC_a$	-0.296***	-0.299**	-0.383***	-0.495***
(accounting anomalies)	(-2.66)	(-2.25)	(-2.66)	(-2.94)
Mean of dependent variable	0.424	0.346	0.488	0.419

Table A.3
Robustness: Difference-in-difference estimates of EDGAR effect on anomalies with anomaly-by-month fixed effects

This table presents the coefficients from the main difference-in-difference regression from Equation (1) with anomaly-by-month fixed effects instead of anomaly and month fixed effects. The coefficient associated with $Post_{p,t} * ACC_a$ captures the gap in the extent to which anomaly portfolio alphas change in response to EDGAR for accounting anomalies relative to non-accounting anomalies. To facilitate interpretation, we also report the coefficient sum ($Post_{p,t} + Post_{p,t} * ACC_a$) and the mean of the dependent variable (portfolio alpha). All the variables are identical to those utilized in Table 2. The sample extends over 205 anomalies from January 1992 to December 1997. The portfolio alphas and the coefficients are expressed in percentages. The standard errors are clustered by anomaly and month. Robust t-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	FF6 alpha 1-5 EW	FF6 alpha 1-5 VW	FF6 alpha 1-10 EW	FF6 alpha 1-10 VW
$Post_{p,t}$	-0.0429	-0.0531	-0.177	-0.210
(non-accounting anomalies)	(-0.44)	(-0.44)	(-1.45)	(-1.59)
$Post_{p,t} * ACC_a$	-0.265**	-0.275**	-0.180	-0.276
(difference-in-difference)	(-2.12)	(-2.11)	(-1.08)	(-1.52)
N	128,892	128,892	128,893	12,8893
R-squared	0.1543	0.1433	0.1436	0.1375
Post-estimation test:				
$Post_{p,t} + Post_{p,t} * ACC_a$	-0.308***	-0.328**	-0.357**	-0.486***
(accounting anomalies)	(-3.12)	(-2.61)	(-2.46)	(-2.94)
Mean of dependent variable	0.301	0.231	0.332	0.281

Table A4
Robustness: Excluding the first implementation phase

This table presents the coefficients from the baseline difference-in-difference anomaly portfolios regression from Equation (1). $Post_{p,t}$ is an indicator equal to one if month t is on or after the effective date for phase p, and equal to zero if month t is before that date; ACC_a is an indicator variable that equals one if a is an accounting anomaly and equals zero if a is a non-accounting anomaly. The dependent variables $\widehat{\alpha_{a,p,t}}$ in the four columns of the table are the Fama-French six-factor alphas for quintile (1-5) or decile (1-10), equal-weighted (EW) or value-weighted (VW) portfolios. All specifications contain anomaly and monthly fixed effects. The sample extends over 205 anomalies in the period from January 1992 to December 1997, with the observations associated with the first implementation phase excluded from the sample. The portfolio alphas and the coefficients are expressed in percentages. The standard errors are clustered by anomaly and month. Robust t-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	FF6 alpha 1-5 EW	FF6 alpha 1-5 VW	FF6 alpha 1-10 EW	FF6 alpha 1-10 VW
$Post_{p,t}$	-0.035	-0.0265	-0.095	-0.095
(non-accounting anomalies)	(-0.33)	(-0.21)	(-0.74)	(-0.67)
$Post_{p,t} * ACC_a$	-0.426***	-0.462***	-0.413***	-0.534***
(difference-in-difference)	(-3.45)	(-3.78)	(-2.86)	(-3.57)
N	115,560	115,560	115,560	115,560
R-squared	0.0120	0.0094	0.0111	0.0091
Post-estimation test:				
$Post_{p,t} + Post_{p,t} * ACC_a$	-0.461***	-0.489***	-0.508***	-0.625***
(accounting anomalies)	(-4.74)	(-4.01)	(-3.62)	(-4.19)
Mean of dependent variable	0.327	0.253	0.358	0.304

Table A.5
Robustness: Pre-trends test, falsification test

This table presents the coefficients from the pre-trends and falsification tests of the baseline difference-in-difference regression results reported in Table 2. The regression reported in Panel A (Panel B) is estimated over a 5-year period prior to (following) the actual EDGAR implementation, and the pseudo-event dates are also 5 years earlier (later) than the actual dates. $Post_{p,t}$ is an indicator equal to one if month t is on or after the pseudo-event effective date for phase p, and equal to zero if month t is before that date; ACC_a is an indicator variable that equals one if a is an accounting anomaly and equals zero if a is a non-accounting anomaly. The dependent variables $\widehat{\alpha_{a,p,t}}$ in the four columns of the table are the Fama-French six-factor alphas for quintile (1-5) or decile (1-10), equal-weighted (EW) or value-weighted (VW) portfolios. The portfolio alphas and the coefficients are expressed in percentages. All specifications contain anomaly and month fixed effects. The standard errors are clustered by anomaly and month. Robust t-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	FF6 alpha 1-5 EW	FF6 alpha 1-5 VW	FF6 alpha 1-10 EW	FF6 alpha 1-10 VW
Panel A: Pre-trends test (placebo dates five years earlier)				
$Post_{p,t}$	0.120	-0.0396	0.270*	0.0832
(non-accounting anomalies)	(0.85)	(-0.27)	(1.67)	(0.53)
$Post_{p,t} * ACC_a$	0.0461	0.0828	0.0854	0.103
(difference in difference)	(0.30)	(0.53)	(0.46)	(0.55)
N	126,246	126,246	126,246	126,295
R-squared	0.0059	0.0060	0.0053	0.0051
Post-estimation test:				
$Post_{p,t} + Post_{p,t} * ACC_a$	0.166	0.043	0.356*	0.186
(accounting anomalies)	(1.07)	(0.30)	(1.95)	(1.13)
Mean of dependent variable	0.152	0.102	0.118	0.0816
Panel B: Falsification test (placebo dates five years later)				
$Post_{p,t}$	-0.276*	-0.230	-0.262	-0.239
(non-accounting anomalies)	(-1.88)	(-1.23)	(-1.49)	(-1.11)
$Post_{p,t} * ACC_a$	-0.0221	-0.00504	0.0592	0.0377
(difference in difference)	(-0.15)	(-0.03)	(0.33)	(0.18)
N	132,242	132,242	132,242	132,242
R-squared	0.0091	0.0068	0.0073	0.0062
Post-estimation test:				
$Post_{p,t} + Post_{p,t} * ACC_a$	-0.298***	-0.235	-0.203	-0.201
(accounting anomalies)	(-2.12)	(-1.22)	(-1.19)	(-0.89)
Mean of dependent variable	0.385	0.225	0.433	0.289

Table A.6
Robustness: Drop sibling anomalies, thin portfolios

This table presents the coefficients from the two robustness tests of the baseline difference-in-difference regression results reported in Table 2. Panel A features a modified sample of anomalies. We first identify the 23 sibling anomalies in our sample of 205 anomalies, that is, pairs of anomalies that exploit the same investment idea, but have portfolios formed based on annual signals and quarterly signals, respectively. We then drop the 23 annual siblings and estimate Equation (1) on the sample of 183 anomalies. Panel B features a full sample of 205 anomalies, but thin portfolio observations are dropped from the anomaly-phase-month panel. That is, observations are removed from the sample if the portfolio construction of either the long leg or the short leg for a given anomaly and implementation phase was based on fewer than five stocks. $Post_{p,t}$ is an indicator equal to one if month t is on or after the effective date for phase p, and equal to zero if month t is before that date; ACC_a is an indicator variable that equals one if a is an accounting anomaly, and equals zero if a is a non-accounting anomaly. The dependent variables $\widehat{\alpha_{a,p,t}}$ in the four columns of the table are the Fama-French six-factor alphas for quintile (1-5) or decile (1-10), equal-weighted (EW) or value-weighted (VW) portfolios. The portfolio alphas and the coefficients are expressed in percentages. All specifications contain anomaly and month fixed effects. The standard errors are clustered by anomaly and month. Robust t-statistics are presented in parentheses below the estimates. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	FF6 alpha 1-5 EW	FF6 alpha 1-5 VW	FF6 alpha 1-10 EW	FF6 alpha 1-10 VW
Panel A: Excluding annual siblings				
$Post_{p,t}$	-0.070	-0.049	-0.136	-0.140
(non-accounting anomalies)	(-0.68)	(-0.40)	(-1.09)	(-1.02)
$Post_{p,t} * ACC_a$	-0.424***	-0.419***	-0.439***	-0.508***
(difference in difference)	(-3.77)	(-3.52)	(-3.43)	(-3.57)
N	115,065	115,065	115,065	115,065
R-squared	0.0108	0.0084	0.0097	0.0079
Post-estimation test:				
$Post_{p,t} + Post_{p,t} * ACC_a$	-0.494***	-0.468***	-0.575***	-0.648***
(accounting anomalies)	(-4.69)	(-3.92)	(-3.81)	(-4.23)
Mean of dependent variable	0.256	0.182	0.300	0.225
Panel B: Excluding thin portfolios (with < 5 stocks)				
$Post_{p,t}$	-0.214**	-0.181	-0.169	-0.141
(non-accounting anomalies)	(-2.38)	(-1.65)	(-1.56)	(-1.07)
$Post_{p,t} * ACC_a$	-0.223**	-0.240**	-0.269**	-0.312**
(difference in difference)	(-2.13)	(-2.19)	(-2.20)	(-2.26)
N	112,476	112,476	99,014	99,014
R-squared	0.0158	0.0109	0.0147	0.0108
Post-estimation test:				
$Post_{p,t} + Post_{p,t} * ACC_a$	-0.437***	-0.421***	-0.438***	-0.453***
(accounting anomalies)	(-4.61)	(-3.93)	(-4.06)	-3.41
Mean of dependent variable	0.464	0.172	0.556	0.235