# Venture Fraud*

Alexander Dyck, Freda Fang, Camille Hebert, and Ting Xu†

November 14, 2025

**Abstract**

We assemble the first comprehensive sample of fraud cases in U.S. venture capital (VC)-backed startups to study venture fraud. Using a sample of newly public firms involved in class action suits to align detection rates, we first show that VC-backed firms are more likely to commit fraud than non-VC-backed firms. This difference is particularly pronounced after 2010, coincident with a rise in founder-friendliness in the VC market. In a panel prediction model covering all venture fraud cases on pre-IPO startups, we find that proxies for agency frictions strongly predict fraud, while founder characteristics matter little. Fraud is more likely in startups with strong founders' control rights and weak yet more convex founder cash flow rights. It also rises with coordination frictions among investors and the presence of non-traditional venture investors. Hot funding conditions at the initial round predict future fraud. Fraudulent entrepreneurs continue to found new VC-backed firms unharmed relative to non-fraudulent entrepreneurs, suggesting a lack of market discipline. Overall, our results highlight rising agency costs in VC-backed firms and suggest that a founder-friendly VC market may lead to misallocation and broader social costs.

JEL: G21, G23, G32, J15, J16
Keywords: Fraud, Misconduct, Venture Capital, Entrepreneurship, Agency Issues, Corporate Governance, Securities Regulation

---

*We thank Michael Ewens, Thomas Hellmann, Ramana Nanda, Tracy Wang, and seminar participants at HBS, University of Toronto, and Auburn University for helpful comments. We also thank our research assistants Aman Aggarwal, Eric Au-Yeung, Noémie Bucourt, Maryann Chen, Peter Guo, Mariya Grynevych, Jiecheng Huang, Weixuan Jiang, Ting-Hsuan Kung, Gary Liang, Nivedha Madhivanan, Fiona Trepanier, and Lingyi Zhou. Funding from the Social Sciences and Humanities Research Council (SSHRC) and the Management Analytics Research Cluster is gratefully acknowledged.

†Dyck: University of Toronto (alexander.dyck@rotman.utoronto.ca), Fang: University of Toronto (freda.fang@rotman.utoronto.ca), Hebert: University of Toronto (camille.hebert@rotman.utoronto.ca), Xu: University of Toronto & NBER (tingxu.xu@rotman.utoronto.ca) (corresponding author).

“*All securities transactions, even exempt transactions, are subject to the anti-fraud provisions of the federal securities laws.*” – Mary Jo White, SEC Commissioner

## 1 Introduction

Venture capitalists (VCs) use control and cash flow rights to enhance startup performance and align incentives (Kaplan and Strömberg, 2004; Ewens et al., 2022). These contracting choices mitigate, but do not eliminate, agency costs in VC-backed firms, with founder fraud being an important symptom of the residual frictions. Founder fraud involves misrepresentation and reflects a founder's breach of fiduciary duty to shareholders. It is also illegal under anti-fraud laws that govern both public and private companies. Although fraud may be privately optimal for founders and VCs, it may exceed the socially optimal level if it imposes externalities on other stakeholders, leading to capital misallocation. In this paper, we measure venture fraud and examine its drivers and consequences.

Early papers highlight the strength of the venture capital model in addressing agency problems (Jensen, 1989; Kaplan and Stromberg, 2003; Ewens and Marx, 2018), with the implication that fraud in VC-backed firms should be low. Yet this perspective may be overly optimistic. With fat-tailed return distributions, VCs have always prioritized upside capturing over downside mitigation (Kerr et al., 2014; Denes et al., 2023). Two recent trends may have further weakened the emphasis on corporate governance in VC-backed firms (Lerner and Nanda, 2020). First, startups are staying private longer and raising larger rounds from a broader set of investors, including many non-traditional investors (Ewens and Farre-Mensa, 2020; Chernenko et al., 2021). This phenomenon creates horizontal conflicts among investors with different incentives and payoffs, which may undermine their collective monitoring of founders (Pollman, 2019). Second, the VC market has shifted toward founder-friendliness since the mid-2000s, with founders increasingly retaining control and becoming better shielded from investor pressure (Janeway et al., 2021; Malenko and Ewens, 2024). These trends could create greater room for fraud.

We ask three research questions: How prevalent is fraud among VC-backed companies, and how has it evolved over time? What are the determinants of venture fraud—in particular, the relative importance of founder characteristics (“nature”) versus contracts and incentives (“nurture”)? Is there ex-post market discipline for fraudulent founders?

To answer these questions, we compile a comprehensive database of VC-backed firms and founders who faced fraud-related lawsuits from 2000 to 2024. We use data from PitchBook to identify VC-backed startups and their founders. We then search for these founders and firms in lawsuits that allege fraud. These lawsuits are initiated by public enforcers, i.e., the SEC and

DOJ, and private enforcers, i.e., shareholder lawsuits in state and federal courts (collected from Westlaw), and shareholder security class action suits (from the Stanford Security Class Action Clearinghouse (SCAC) database). In all of these cases, founders are alleged to have made material misrepresentations or omissions that harm investors, including misrepresentations about financials, operations, technology (e.g. Holmes in the Theranos fraud), or customers (e.g. Javice in the Frank fraud), etc.[1] To mitigate concerns about frivolous lawsuits, we filter our core sample to cases with merit, i.e., non-dismissed cases, cases involving civil and criminal sanctions, or cases with financial or non-financial settlements. We identify 550 such cases involving VC-backed startups founded post-2000. We manually code the attributes of each case, including fraud period, fraud type, charges, case outcome, and victims. We then complement this data with information on founders, investors, and term sheets from PitchBook, as well as board characteristics from Malenko and Ewens (2024) and Form D.

One potential concern is that our sample may miss fraud cases that took place but were never brought to light. Undetected fraud is a problem in publicly traded firms, and is likely a larger issue in private firms. The SEC and DOJ do police fraud in private firms, but their enforcement resources are limited, and they may face detection challenges, partly because there are fewer external monitors of private firms. Investors can also police fraud, but their cost-benefit calculus in bringing suits may be more challenging for private firms than publicly-traded firms. Nonetheless, examining fraud cases with formal legal actions has the advantage of focusing on cases that are economically consequential, as immaterial frauds are seldom pursued by regulators or litigants.

We mitigate the undetected fraud problem in three ways. First, we relax our definition of fraud to include dismissed cases from Westlaw and allegations identified from news feeds from Crunchbase and Pitchbook.[2] This extended sample includes 907 cases and mitigates concerns about selective detection. Second, we employ a sample of newly public firms to align the detection rate and benchmark fraud rate for VC-backed firms with those for non-VC-backed firms. At IPO, compliance and disclosure jump sharply, raising the likelihood of detection for both VC- and non-VC-backed firms, which face the same intensified enforcement environment. The key assumption is that fraud revealed shortly after IPO is related to the pre-IPO governance approach baked into the firm while it was private. We also focus on variation in this gap over time, as differences in detection rates between the two firm types are likely to be persistent. Third, in our predictive analysis of all venture-fraud cases, we focus on predicting the start of fraud commission, rather than the timing of fraud detection (i.e, when the suit was filed). We also control for determinants of detection

[1] We exclude cases where founder's misconduct does not involve fraud allegations, such as anti-trust, bribery, insider trading, labor violations, and personal misconduct. We also exclude lawsuits where founders accuse investors of abuse or those focusing on horizontal agency conflicts among investors.

[2] Atanasov et al. (2012) argue that many dismissed suits are not meritless given the legal expenses, emotional stress, and potential reputational harm from launching such suits. Dismissals also often happen for technical and procedural reasons.

likelihood, such as firm size, valuation, and opacity (proxied by missing PitchBook fields).

We have four primary findings. First, using a sample of IPO firms, we find that VC-backed firms have higher fraud rates than non-VC-backed firms, a difference that is particularly strong in the recent decade that coincides with the rise of founder-friendliness and the increase of frictions among investors. Controlling for industry-specific shocks as well as pre-IPO firm size and growth, VC-backed firms are 4.8 percentage points more likely to face class-action suits within two years of IPO—a difference that is 59% of the overall mean. This result is robust to alternative post-IPO windows and to different definitions of suit materiality. Notably, when we split the sample period in half, we find no statistically significant difference between VC- and non-VC-backed IPOs during 2002-2012, but a pronounced gap during 2013-2023. These results suggest that VC-backed firms are more prone to fraud, especially in the recent decade when founder-friendliness was most prominent.

Second, we find that contractual incentives and investor-side frictions strongly predict fraud, whereas founder characteristics matter little. We employ a firm-level panel regression to predict the start of fraud among VC-backed startups. The sample is a hazard-style panel that tracks all US VC-backed firms founded since 2000, with fraud firms remaining in the sample until the year of fraud start, and non-fraud startups remaining until closure/exit. To mitigate concerns about undetected fraud, we capture variation in detection likelihood using firm size, industry-year fixed effects, and measures of opacity.

We find that, conditional on firm age, valuation, and the amount raised, fraud is more likely when the board is founder-controlled rather than VC-controlled and when investors' converted ownership is low, suggesting that founder control rights increase fraud likelihood. In particular, startups with VC-controlled boards are two times less likely to commit fraud than those with founder-controlled boards, while those with shared-control boards (i.e., same seats by founders and investors) are 50% less likely. A 10% decrease in investor ownership increases the likelihood of fraud by 17% relative to the mean. We further find that contracts that reduce founders' cash-flow rights while increasing payoff convexity—such as high liquidation preferences and participating preferred—are associated with higher fraud incidence, consistent with reduced skin-in-the-game and stronger incentives to gamble for extreme outcomes.

Investor syndicate structures that induce greater horizontal frictions among investors further exacerbate vertical conflicts between founders and investors. In particular, a larger number of investors on the cap table increases the likelihood of fraud, suggesting that coordination frictions hinder monitoring. A greater presence of non-traditional investors (e.g., angels, CVCs, mutual funds) also predicts fraud, as these investors tend to have weaker monitoring capabilities. Finally, lead investor's reputation, measured by the number of past successful exits, mitigates fraud risk.

In contrast, founder characteristics have limited explanatory power for venture fraud. Prior founder experience, education, age, gender, and immigrant status do not significantly predict fraud. The only significant predictor is being a solo founder, consistent with co-founders who provide checks and balances. The lack of explanatory power of founder traits for fraud contrasts with the large literature on the importance of behavioral traits in explaining entry into entrepreneurship (Åstebro et al., 2014; Kerr et al., 2018)

The above results are robust to a subsample test on a cross-section of VC-backed IPO firms, which exhibit high and homogeneous detection rates. They are also robust to restricting to startups that raised large ($10M) and very large ($50M) amounts of capital, or using an expanded sample of frauds that include dismissed cases and allegations from the news.

Third, we show that hot market conditions, in which investors chase founders, are associated with greater fraud, consistent with weaker screening and more founder-friendly governance. At the market level, time series in aggregate VC market valuation and founder-friendliness conditions closely track aggregate venture fraud rates. In a cross-sectional firm-level regression, we show that hot and founder-friendly market conditions at a startup's initial financing strongly predict future fraud. We use two market condition variables at the industry-year level: the average valuation multiple and the share of founder-controlled boards at a startup's initial financing year. We find that a one-standard-deviation increase in the industry valuation multiple (share of founder-controlled board) at a startup's initial round is associated with a 24% (20%) increase in the future fraud rate relative to the mean. This finding holds even controlling for the startup's own valuation and board control at the first round, suggesting that hot markets weaken screening.

Fourth, we analyze the consequences of fraud for founders to determine whether market discipline in the VC market mitigates the likelihood of fraud. If fraud carries a long-lasting stigma, involved founders should find it harder to raise capital for the next startup. If, on the other hand, "fake it till you make it" is the norm, and being caught for fraud is seen as just bad luck, then there would be little punishment.

We find that the VC market does not discipline founders ex post. Specifically, we conduct a matched event study at the person level. For each treated founder in the year before the initial fraud charge, we match her to three similar control founders that were never involved in fraud, but had otherwise similar past founding experience, tenure at current startup, sector, education, gender, and age. Using this matched panel, we estimate a dynamic stacked DID (Cengiz et al., 2019) tracing these individuals' founding of new VC-backed startups. Overall, the cumulative number of startups founded post-fraud is no different from that of the control founders; if anything, fraudulent entrepreneurs founded 0.04 more startups over the next 6 years. The lack of ex post market discipline provides another ex ante incentive for founders to commit fraud in a market where founders have substantial bargaining power.

Taken together, these findings point to sizable and rising agency costs in VC-backed startups, driven in part by founder-friendly contracting and growing frictions among investors as more money chases deals. A narrow takeaway from the paper is that it quantifies one side of the tradeoff in VC investing: generating upside versus mitigating downside. With a fat-tailed return distribution, investors optimally prioritize upside capture over downside mitigation (Kerr et al. (2014), Denes et al. (2023)), and a founder-friendly approach further tilts this balance. Our evidence on increased fraud sheds light on the downside of this tradeoff, documenting an important cost that investors experience and are willing to tolerate.

It may also be the case that the level of fraud is not privately optimal for investors themselves: when deceptive types become pervasive, they can crowd out non-deceptive firms that could deliver the very right-tail outcomes VCs seek, leading to capital misallocation. Horizontal frictions among investors also undermine their collective effort to optimally prevent fraud. Regardless of private optimality, the level of fraud could be higher than socially optimal when negative externalities are not internalized. The Theranos case illustrates this point: founder Elizabeth Holmes's fabricated claims about breakthrough blood-testing technology misled not only investors but also patients and physicians, resulting in misdiagnoses, delayed treatments, and other harms. If there is too much fraud, this creates a case for strengthening the information or resources available to public enforcers – such as the SEC and DOJ – who may be better positioned to offset investors' limited incentives to mitigate fraud.[3]

This paper contributes to several strands of literature. First, we contribute to the literature on corporate fraud, which largely focuses on public firms. This literature is interested in firm and industry characteristics that predict fraud and measure the likelihood of fraud (Beneish, 1999; Dechow et al., 1996; Beneish et al., 2013; Povel et al., 2007; Wang, 2013; Dyck et al., 2024). It also explores the economic consequences for executives of committing fraud (e.g. Karpoff et al. (2008a,b)). Our paper contributes by examining these issues for VC-backed firms, an understudied group with agency problems distinct from those in public firms. Related to our paper, Atanasov et al. (2012) examine lawsuits against VCs (which are not necessarily about portfolio company fraud) and show that litigated VCs suffer declines in future businesses. Tian et al. (2015) show that financial market disciplines VCs who are associated with IPO fraud by their portfolio companies. Our paper studies fraud by all VC-backed startups regardless of their IPO exit status, using a comprehensive sample of fraud cases new to the literature. We also link the drivers of venture fraud to founder-friendly contracting and document a lack of market discipline for fraudulent founders.

We also add to the literature on governance and contracting in VC-backed firms. Analyzing

[3]The increasing participation of retail and non-institutional investors in private markets strengthens the case for more public oversight and enforcement, as these investors are less equipped to conduct due diligence and bear the losses associated with fraud (Phalippou and Magnuson, 2025).

VC contracts, Kaplan and Stromberg (2003) demonstrates how VCs allocate cash flow, control, and liquidation rights to incentivize entrepreneurs while protecting themselves from risk. Ewens et al. (2022) use a dynamic model with contract data to show that VCs' superior bargaining power results in more investor-friendly contracts, impacting the equity split and value creation for startups. Legal scholars highlight that startups face governance issues that are different and more complex than those in public firms (Bartlett III, 2006; Pollman, 2019, 2020)—not only are there vertical conflicts between shareholders, boards, and founders, there are also horizontal conflicts between shareholders with different payoff structures (e.g., preferred vs common). Malenko and Ewens (2024) study the evolution of startup boards over the firm lifecycle and highlight the mediating role of independent directors in balancing the power of founders and VCs. Bian et al. (2023) show that a common-maximizing legal regime reduces the likelihood of startup fire sales by VCs under liquidity pressure. Our paper contributes by using fraud to understand the extent and drivers of agency problems in VC-backed firms, and how they have increased over time.

Within the above literature, several papers document the rise of founder-friendliness, potentially driven by private market deregulation, increasing competition for deals, and the participation of non-traditional venture investors (Nanda and Rhodes-Kropf, 2013; Ewens and Farre-Mensa, 2020; Lerner and Nanda, 2020; Chernenko et al., 2021; Janeway et al., 2021). Several theory papers examine the optimality of founder-friendliness (Banerjee and Szydlowski, 2024; Broughman and Wansley, 2023), arriving at different views. Our paper is the first to link the rise of founder-friendliness to that of venture fraud, highlighting potential misallocation and social costs associated with it.

## 2 Legal Framework and Institutional Background

We measure fraud in VC-backed companies by identifying cases where there is a lawsuit initiated by an investor, regulator, or stakeholder that alleges illegal conduct and fraud by a founder or firm. These suits can allege misconduct under three legal bases: violations of state or federal laws, violations of fiduciary duty obligations in state corporate laws, and violations of contract under state contract laws.

The most important regulatory laws are the federal anti-fraud provisions in the 1933 Securities Act, which relate to security issuance, and the 1934 Securities Exchange Act, primarily in Section 10b, which addresses fraud related to already issued securities. Elements of fraud include misrepresentations, materiality, intent, and reliance. Misrepresentations, which includes omissions, are broadly defined and can relate to financial performance as well as to business operations, risk exposure, legal compliance, partnerships, forward-looking guidance, and registration statements, etc.

Suits under federal securities laws can be brought by the SEC and by private actors. Private suits are mostly class-action lawsuits against publicly traded companies where the fraud-on-the-market theory applies, and investors who rely upon the integrity of the market price are presumed to have relied on the misstatement. States have their own securities laws, which include anti-fraud provisions, and investors can also sue in state courts that can have easier enforcement than in federal courts (e.g., not requiring the proof of intent). Fraud in securities regulations cases also often overlaps with allegations of wire and mail fraud in criminal laws (18 U.S.C. Sections 1341, 1343), as the mail or wires are used to further the fraud. If the DOJ pursues a criminal case, it invokes these statutes.

Second, investors can appeal to corporate law at the state level and make claims that directors and officers (which includes founders) violated their fiduciary duty, which encompasses both a duty of care and a duty of loyalty (this requires fairness in self-interested transactions, or disclosure and approval). Fiduciary duties are owed to the corporation and, in some cases to shareholders. Fiduciary duty suits are often derivative suits brought on behalf of the corporation. In Delaware, where most VC-backed firms are incorporated and where most suits are filed, derivative suits need to satisfy a demand requirement that shareholders must first make a demand of the board or plead why the board is too conflicted to act.

Third, investors can allege breach of contract and pursue cases in state courts. Claims can be brought for many of the same reasons as in securities laws (e.g., failure to disclose material information, misrepresentation, or fraudulent financial statements), along with other reasons (e.g., failing to follow investment terms, failure to provide required information (e.g., books and records).

## 2.1 Data Sources for Alleged Fraud

We use multiple data sources to identify cases of alleged fraud in VC-backed companies. The same fraud can be present in multiple sources or might only appear in one. Hence, using multiple data sources helps maximize the chance we capture each fraud case. Two criteria to evaluate a data source are: how compelling is the source in identifying fraud; and what is the omission rate in using this source in identifying frauds.

Our first data source consists of civil cases brought by the SEC, alleging federal securities law violations. We focus on those cases that were adjudicated or settled, with most cases involving penalties for founders and/or firms, although rarely requiring admitting of wrongdoing. This source is a compelling one, given that the SEC chose to pursue the case and the defendants settled. It is important to note that settled cases, for both public or privates suits, do not mean that the case had no merit or that the defendant was innocent. It is simply an agreement between the plaintiff(s) and the defendant(s) to end the case without further litigation. Defendants settle precisely because the defendant wants to avoid a judicial finding of guilt.

Our second data source consists of criminal cases that involve fraud allegations initiated by the DOJ, which are resolved in its favor, resulting in defendants being required to serve time and/or pay financial penalties. This is a compelling source for fraud cases, as there is judicial certification of fraud and criminal cases are typically more severe than civil cases.

Both the SEC and DOJ data can omit undetected frauds. The SEC and DOJ have limited resources and have challenges monitoring private firms, as external monitors typical for publicly traded firms are not active in private markets (e.g., short-sellers and security analysts). They may also have difficulties identifying fraud due to the general lack of information on private firms. That said, relative to non-VC-backed private firms, VC-backed firms are more visible, since they typically exist in financial databases and are required to file regulatory filings (e.g., Form D) for their exempt transactions.

Our third data source is security class action suits under the anti-fraud provisions of the federal securities laws. As a minimum test of materiality, we only include non-dismissed cases. This is another compelling source for fraud cases, albeit somewhat weaker than SEC or DOJ cases. There is the possibility that a defendant settled when there was no wrongdoing just to make the case go away.[4] For private firms, security class actions have a high rate of omission, since private firms do not have a daily stock price that is efficient and incorporates all information. Thus, each investor bringing a suit must prove their reliance on the misinformation, which prevents them from forming and suing as a class (Winship, 2024). However, once these private firms become public, security class actions have a low omission rate with many plaintiff law firms motivated to launch such cases (Karpoff et al., 2008a). This is a feature we exploit in our analysis of VC-backed IPO firms.

Our fourth data source is Westlaw, a comprehensive legal database of both adjudicated and non-adjudicated suits filed at both federal and local courts. The Westlaw database broadens the legal scope beyond regulatory and securities class actions by capturing common-law fraud suits, contract and commercial disputes, and shareholder derivative suits (e.g., breach of fiduciary duties), thus extending the sample to private and state court cases that would otherwise be missed. These cases are compelling because we restrict to those with published judicial opinions as well as cases that are resolved through summary judgment or private settlement. In our main analysis, we exclude suits that were withdrawn by plaintiffs or dismissed by court, typically for insufficient evidence or procedural deficiencies.[5] In our robustness analysis, we include dismissed cases to address concern of omitted fraud, as these cases still indicate concerns for fraud given the cost to launch a suit and

[4] In some specifications, we further restrict security class-action cases to those with financial penalties higher than $3 million, a common threshold used in the law and economics literature.

[5] Many lawsuits are dismissed on technical grounds. For example, derivative suits alleging breaches of fiduciary duty are often dismissed when shareholders fail to show that presenting the complaint to directors would be futile.

the fact that many dismissals were for technical reasons.

Lastly, to mitigate concern of omitted fraud, we use news feeds on VC-backed companies from PitchBook and Crunchbase to identify allegations of fraud, even if the allegation has not led to legal actions or enforcement. This sample has lower omission rates than enforced cases, with the tradeoff being that they may be less compelling.

## 2.2 Examples of Alleged Fraud in VC-backed Firms

We now consider a few cases of alleged fraud in VC-backed firms to illustrate the legal bases for suits and which databases identify such cases.

First, consider the case of Theranos. This case appears in our SEC, DOJ and Westlaw sample, with a March 2018 SEC suit charging Theranos, Elizabeth Holmes, and Ramesh Balwani for violations of Section 10(b) of the Exchange Act and Section 17(a) of the Securities Act, and a June 2018 DOJ suit charging Holmes and Balwani with wire fraud and conspiracy to commit wire fraud. There were also class action suits against Theranos for consumer fraud in California and Arizona, which were dismissed following Theranos' bankruptcy, but highlighted consumer losses that contributed to the DOJ suit.

Theranos was also subject to multiple private suits, including direct suits from partner Walgreens (2016) for breach of contract and Partner Fund Management (2016) for fraud. Theranos was also sued by the Arizona Attorney General for consumer fraud. These suits were settled out of court. There were attempts at security class-action suits, but these proved ineffective as "the Court firmly agrees with Defendants that the fraud-on-the-market presumption of reliance cannot apply here, because Theranos securities were not sold in an efficient market." (Winship, 2024). Elizabeth Holmes was sentenced to 11 years and 3 months of prison and $452 million in restitution.

As a second example, consider the VC-backed firm Marrone Bio Innovations, which went public in 2013. This case is present in our SEC, DOJ, and class-action sample. The firm was accused of violating securities laws around its IPO relating to improper revenue recognition, misleading financial statements, expense misconduct, and failure of disclosure obligations. The SEC signaled that it thought such problems were common with VC-backed high-growth firms: "Rapidly growing enterprises present significant risks if the appropriate control structure is not in place. Time and again, we have seen companies go public and grow at a pace that exceeds their control structure. For example, just last month, the Commission brought charges against a company and a former executive for inflating financial results to meet projections that it would double revenues in its first year as a public company. Because, in part, of insufficient internal controls, the executive was able to direct his subordinates to obtain false sales and shipping documents and intentionally ship the wrong product to book sales."[6] Marrone Bio's former COO Hector Absi was sentenced to two years of prison.

# 3 Data and Sample

## 3.1 Identifying Venture Fraud from Data Sources

To identify securities fraud and litigation events related to start-ups, we rely on five data sources. We link companies involved in these cases to PitchBook to get information on startup financing, founding teams, and term sheets. In addition, we use PitchBook and SEC Form D filings to extend the startup board composition data from Malenko and Ewens (2024). Below, we describe how we identify venture fraud cases from each data source. Appendix B provides more details. Table 1, Panel A shows the case count by source and how the sources overlap.

**SEC enforcement actions**

To identify SEC enforcement actions involving VC-backed startups and their founders, we begin by scraping three types of SEC releases from 1995 to 2023: litigation releases, administrative proceedings, and Accounting and Auditing Enforcement Releases (AAERs). From each of these releases, we extract the respondent (i.e., violator) name, which could be either an individual name or a company name.

We employ two approaches to identify cases by VC-backed firms that we can link to Pitchbook. Our first approach links SEC respondents to PitchBook companies and founders using a combination of exact and fuzzy name matching on the respondent name. We retain matches where the release date occurs between a company's founding and five years after its last financing round or IPO. For individuals, we ensure the release date falls within five years of their departure from the VC-backed company. Our second approach applies keyword-based filters to litigation release texts to flag potential cases involving VC-backed startups or founders (e.g., "venture capital", "startup", "palo alto", "Silicon Valley"). We then manually verify and match flagged cases to PitchBook. After these two steps, we further exclude cases involving non-relevant violations such as delinquent filings, unregistered broker activity, or market manipulation. After imposing filters relevant for our analysis (e.g., founded post 2000, has a non-missing fraud start year), we obtain 126 SEC cases involving VC-backed firms.

[6] See SEC release: `www.sec.gov/newsroom/press-releases/2016-32`.

### DOJ enforcement actions

We collect information on criminal enforcement actions from the U.S. Department of Justice (DOJ) by scraping all publicly available releases from 2013 to 2024. DOJ enforcement of white-collar crime is more sparse before 2013. Nevertheless, we use Westlaw (see below) to extend the DOJ sample back to before 2013. 65% of these cases include subject-matter tags, out of which“Financial Fraud” and “Securities, Commodities, & Investment Fraud” are our focus. Using the tagged sample as a training sample, we apply a machine learning classifier (BERT) to tag the remaining releases with missing tags.

To identify cases involving startups, we proceed in several steps. First, we filter for cases containing startup-related keywords (e.g., “founder”, “venture capital", “Silicon Valley"). We then manually verify the matching to VC-backed startups and founders in PitchBook. Using the manually verified sample as the training sample, we then train a BERT classifier to identify potential cases that may involve startups, and repeat the manual verification. We iterate several times between BERT classification and manual check to arrive at our final sample. After imposing various filters, we obtain 91 DOJ cases.

### Securities class action lawsuits

We obtain a comprehensive dataset of securities class action lawsuits from the Stanford Securities Class Action Clearinghouse (SCAC). We restrict the sample to non-dismissed cases filed in the U.S. between 1999 and 2023. The majority of these cases concern public firms, with less than 3% concerning private firms. We determine whether each company was ever VC-backed using PitchBook’s deal history, supplemented with Jay Ritter’s VC flag in IPO data. Our main analysis further restricts to cases with a class period starting within 2 years of IPO, yielding 276 cases by VC-backed firms.

### Westlaw court cases

To complement cases from the SEC, DOJ, and SCAC, we additional use Thomson Reuters’ Westlaw database, a comprehensive legal research platform. This database contains all U.S. federal and state court cases and dockets, including lawsuits that are settled, withdrawn, or dismissed.

We use three modules of Westlaw to complement our SEC/DOJ/SCAC sample: court cases, dockets, and administrative rulings. This expands fraud-related litigation coverage in three ways. First, the court cases— adjudicated suits that went in front of the judge with issued legal opinions— broadens our sample by capturing common-law fraud suits, contract and commercial disputes, and shareholder derivative suits by both private parties and state regulators. Second, Westlaw dockets cover cases that never yield a published opinion, such as suits that were settled out of court, withdrawn, or dismissed. The dockets also chronicle the entire history of each case, allowing us to trace the timing of the initial suit or charge regardless of case outcome. Third, the administrative ruling module provides comprehensive coverage of SEC and DOJ news releases, which allows us to extend our DOJ sample back to 2000.

We pre-screen adjudicated cases as well as dockets using legal classifications and keyword searches of fraud and fiduciary duty-related terms. Using textual analysis tools, we first identified defendant names and match them to VC-backed startups in PitchBook. We then applied GPT-4o to analyze case texts and extract relevant case attributes. We then manually read each case to verify that (1) the match to the PitchBook company was correct, (2) the suit had not been dismissed, and (3) the filing provided some information regarding the alleged period of fraudulent activity. We also manually verify and correct the case attributes extracted by GPT. After imposing relevant filters, we obtain 186 non-dismissed cases involving VC-backed companies.

#### Crunchbase and PitchBook News Feeds

Finally, we use news articles from CrunchBase and PitchBook to construct a dataset that includes alleged fraud involving VC-backed startups. The advantage of this database is that each news article is already linked to firms by PitchBook and CrunchBase, minimizing matching errors. News also contains credible allegations and ongoing investigations, which provides a broader set of “likely fraud" or “concerns for fraud", allowing us to mitigate concerns about omitted frauds. A downside of this source is that the news feed data begins primarily in 2017, and for this reason we only include it in robustness checks.

We aggregate all news articles linked to startups or their founders. We then screen articles using a combination of keyword filters (see Appendix for more information), manual checks by research assistants, and a machine learning classifier trained on over 6,000 hand-labeled articles to identify articles about startup fraud. After imposing relevant filters, 551 news-based cases enters into our final sample. 193 of these cases also exist in our other data sources, while 358 cases are fraud allegations.

### 3.2 Estimation samples

We rely on two main samples to study the likelihood and determinants of venture fraud.

Our first sample includes the universe of newly public firms that are either VC-backed or non-VC-backed. We identify the VC-backed status using VC deals using PitchBook as well as the VC indicator available in the Jay Ritter's database. We use class-action cases with the class period starting within two years of the IPO as a measure of fraud. This sample aligns the enforcement environment faced by VC-backed and non-VC-backed firms, as both types of firms face the same public market enforcement, disclosure requirements and intense scrutiny by regulators, underwriters, and other market participants (i.e., analysts, auditors, and investors). As a result, this sample helps mitigate the issue of low and differential detection rates in venture fraud. The key assumption is that fraud detected soon after IPO is likely to be related to pre-IPO governance choices and is hence tied to the incentives in place while the firm was private. This sample contains 5,206 IPOs over the period for 2002 to 2024.

The second sample, used for our panel predictive analysis, includes all enforced fraud cases against VC-backed firms in the US from 2000-2023. This sample combines data from the SEC, DOJ, class actions, and Westlaw and thus includes frauds in VC-backed firms while they are private, and for VC-backed firms that are publicly-traded within 2 years of their IPO. The sample also includes all non-fraudulent VC-backed startups in Pitchbook over the same period. We construct a panel of private firm-years to study the incidence of fraud using a hazard-style analysis. In our panel, we trace each fraud firm from its founding to the year the fraud started, and each non-fraud firm from its founding to its closure or exit. After imposing non-missing founding year and at least one round of finance, our sample includes 989,139 firm years for 105,342 unique VC-backed firms, out of which 550 were involved in fraud identified from SEC, DOC, Westlaw, and SCAC, and 907 if including fraud allegations from news.

# 4 Descriptive Statistics

## 4.1 Fraud Cases by Source

Table 1 reports summary statistics for the fraud cases in our sample. Panel A reports case count by sources. We obtain 133 cases from SEC, 91 from DOJ, 207 from Westlaw, and 224 from class action suits. We additionally sourced 550 cases and allegations from news. Note that these sources can overlap, with overlapped case counts indicated in off-diagonal numbers. Together, the four sources excluding news give us 550 enforced fraud cases, which serve as our main sample. Additionally including news-based cases gives us 907 cases and allegations. Panels B to E report specific case attributes for SEC, DOJ, Westlaw, and class action suits separately.

Among SEC cases (Panel B), frauds last on average 29 months, though with significant variation (median at 24 months and 95th percentile at 72 months). Monetary sanctions are substantial, with a mean fine of \$11.4 million. A majority of cases result in bans (70%), disgorgement orders (68%), and fines (58%), while prison sentences are rare (4%), as the SEC only enforces civil cases. Most cases involve financial misrepresentation (72%), with the remaining misrepresenting products (28%) and use of funds (21%). Investors are the most frequent victims (88%), followed by the general public (8%), government (4%), and other entities (7%).

The DOJ cases (Panel C) have a longer average fraud duration, at 41 months. Regarding fraud types, wire fraud makes up 85% of cases, securities fraud 33%, corporate fraud 10%, and bank fraud 8%. Sanctions are also more severe: prison sentences occur in 43% of cases, with an average duration of 84 months. Monetary fines are large, with a mean fine of $18.7 million. Additional penalties include forfeiture (15% of cases), and supervised release (34% of cases). Criminal cases, hence, represent more severe cases than civil cases with harsher punishment. Similar to SEC cases, investors are the most frequent victims, representing 66% of cases. However, the public, government, and other entities are also frequent victims, representing 22%, 19%, and 11% of the cases respectively, consistent with DOJ's broader enforcement scope than SEC and the social externalities these frauds generated.

For Westlaw cases (Panel D), which include criminal and civil cases in both federal and state courts, the average fraud duration is 28 months, similar to SEC cases. Corporate fraud account for 44% of the cases, with fiduciary duty violations accounting for 43% and securities fraud 35%. Misrepresentation typically involves firm financials (53%), products (30%), and use of funds (8%). Sanctions are less systematically reported, as many are privately settled. Investors are again the most frequent victims (59%), followed by the public (16%).

Finally, in Panel E, among the 224 class action cases, about half are settled outside court, with an average settlement amount of $16.1 million. The most common grounds for suit are the anti-fraud provision (1934 Act Section 10(b)), control person liability (1934 Act Section 20(a), 1933 Act Section 15), and misleading IPO disclosure and oral communications (1934 Act Sections 11 and 12(a)2).

In Tables A.1 and A.2, we list the top cases in our sample with detailed case information.

## 4.2 Timing and Incidence of Venture Fraud

Figure 3 describes the timing of fraud based on our main sample of fraud cases excluding news-based allegations. Panel A shows the distribution of cases by fraud start year. There are increasing numbers of cases over time, especially after the mid-2010, with a spike in 2021. The dwindling of frauds in 2022-2023 may reflect truncation, as some cases are yet to be revealed. We observe a similar pattern in fraud committing years (i.e., from fraud start to fraud end years) in Panel B. Panel C shows the distribution of initial charge year, which exhibits an increasing trend over time especially post 2015.

Figure 4 shows variation in fraud likelihood among US VC-backed firms founded post 2000. Panel A shows the trend in the fraction of startups engaging in fraud in a given year (in percentage points) based on the fraud-committing period. To minimize concerns about truncation, we end the period in 2021. We find an overall increasing trend, with a small dip during the 2008 financial crisis years. The incidence rate increases with firm size, as proxied by the cumulative amount raised. For example, among firms raised at least %5M, the fraud rate is around 0.37% in the post-2013 period. Panel B confirms this increasing relationship with size when binning by post-money valuation at the first VC round. In particular, fraud rate rises to 1.9% among firms with $100M-$500M valuation, and to 7.2% among those with at least $500M valuation. However, we caveat that this positive relationship between firm size and fraud incidence could reflect either higher detection rate among larger (and hence more visible) firms, or higher underlying true fraud rate.

Panels B examines variation by geography and industry. Among states with at least 2000 startups, the venture fraud rate is highest in New Jersey, Pennsylvania, California, and Virginia, and is lowest in Maryland, Ohio, and Delaware. Using Pitchbook industry categories, we find that fraud is the most prevalent in financial services, energy, and healthcare, and is the least frequent in the B2B and B2C sectors.

### 4.3 Class Actions Against IPO Firms

Table 2 presents summary statistics for the IPO and class action samples over the 2002–2024 period. Panel A shows that out of 4,094 IPOs, about 35% are VC-backed. 8.5% of IPO firms faced a non-dismissed securities class action suit at some point, with 4% experiencing it within the first year and 6% within three years of IPO.[7]

Panel B focuses on the sample of 429 all non-dismissed class actions between 2002 and 2024. A majority of these cases involved securities fraud allegations (79%). 59% were settled by June 2024 (the date at which we obtained the data). The mean settlement amount is $21 million, though the distribution is skewed and the amount is missing for ongoing cases. More than half of the class action defendants are VC-backed IPOs (58%), higher than the fraction of VC-backed IPOs over the same period (35%). Regarding timing, 50% of cases have class periods starting within one year of the IPO and 70% within three years, with an average of 2.3 years between IPO and class period start date.

Panel C restricts the sample to the 276 cases that start within two years of IPO, i.e., start of the class period. In this subsample, 59% involve VC-backed firms. Taken together, the descriptive statistics suggest that securities class actions are associated with VC-backed firms, particularly in

[7]The full sample of class action suits includes 917 cases between 2002–June 2024. 54% of class actions end up being dismissed. We exclude these dismissed suits from our main sample.

the immediate post-IPO period. This motivates our subsequent analysis of venture fraud.

### 4.4 Prediction Sample

Table 3 provides summary statistics for the prediction panel. All statistics in this table are based on observed values, whereas in regression we impute missing values with zeros to preserve sample size and include various missing-value indicators. The mean of the fraud start indicator in our hazard-style panel is 0.045%. This number is low because the panel stops at the fraud start year, rather than including subsequent years when fraud was committed. The average firm-year is 5.8 years old, with a valuation of $20.5 million, and total raised capital of $2.2 million. The average firm-year had a board of 3.2 members, an 18% likelihood of being VC-controlled, a 29% likelihood of shared control (i.e., founders and VCs have the same number of seats), and a 53% (=1-18%-29%) likelihood of being founder-controlled. The average investor ownership on a converted basis is 46%. Weighted by the raised amount, 27% of the rounds have a high liquidation multiple above 1, and/or investor participation rights. The average firm-year has 3.8 unique investors, out of which 61% are non-independent VC investors (e.g., angels, accelerators, CVCs, PEs, hedge/mutual funds, etc.). The lead investor has had 8.6 successful exits. In terms of founding teams, 49% were founded by solo founders, 20% have serial entrepreneurs, and 26% of them graduated from top schools.[8] The average founder age is 42.

Finally, we also keep track of unreported information in PitchBook, which is a common pattern in private firm databases, but arguably captures important information about a firm's opacity. Missing values (at the firm-year level) are most frequent for valuation, board composition, and term sheets, and less so for investors and founder information.

## 5 Benchmarking Fraud Incidence by VC-Backed Firms: Evidence from Newly Public Firms

A natural way to benchmark fraud rates among VC-backed firms is to compare them with similar non-VC-backed firms. A direct comparison with private, non-VC-backed firms is problematic because data on their size and financing is scarce and often not publicly available. Moreover, fraud detection rates increase with firm visibility, and VC-backed firms are inherently more visible than non-VC-backed private firms: they raise funds through exempt transactions that require regulatory filings (e.g., Form D with SEC), appear in commercial databases, and draw more attention from regulators and investors.

[8] We define top schools as Harvard, Yale, MIT, UPenn, Stanford, Columbia, NYU, Chicago, Cornell, Berkeley, Oxford, and Cambridge.

To address this comparability issue, we focus on newly public firms and we compare VC-backed and non-VC-backed firms within a setting where size, disclosure requirements, and public visibility are more closely aligned. The IPO event serves as a natural benchmark: it imposes standardized disclosure requirements, and similar market and regulatory oversight; it also represents a discrete jump in fraud detection rate. In our sample, 50% of securities class actions against newly public firms have a start of the class period within two years of the IPO, and 48% have a filing date within 3 years of IPO (see Table 2, Panel B and Wang (2013)). This supports the idea that frauds revealed in this window originate from governance incentives in the pre-IPO period. This approach also ensures that we study fraud events that are both economically significant and comparable across VC- and non-VC-backed firms.

## 5.1 Descriptive Evidence

Figures 1 and 2 describe securities class actions among newly public firms, highlighting differences between VC- and non-VC-backed IPOs.

We start with all class actions filed after the IPO by filing year, as shown in Figure 1. Both VC- and non-VC-backed cases rise in the mid-2000s. Overall, VC-backed firms represent 47% of firms involved in class actions, much higher than their overall IPO share over the same period, which is 35% (see Figure A.3).

Figure 1 (b) shows the distribution of cases by the number of years between IPO and the start of the alleged class period. Cases tend to cluster around IPO year, but this clustering is more pronounced for VC-backed firms than non-VC-backed firms. Nearly half of the cases against VC-backed firms begin in the IPO year, and a substantial fraction start within 1-2 years after. In contrast, class actions against non-VC-backed firms are less shifted toward the IPO year. To the extent that fraud revealed close to IPO year reflects wrongdoings committed before or during the IPO, the evidence suggests that VC-backed firms are more likely to commit fraud than non-VC-backed firms.

Motivated by Figure 1 (b), to better isolate wrongdoings committed before the IPO, Figure 1 (c) plots the evolution of cases with the class period starting within two years of IPO. 64% of these cases involve VC-backed firms, which is higher than the 47% share in Figure 1 (b), and much higher than their 37% IPO share. This underscores that VC-backed firms are particularly prone to fraud, which is disproportionately revealed at IPO.

Figure 1 (d) breaks down these IPO class actions by industry. Technology and healthcare dominate, accounting for the bulk of cases, while finance and other sectors contribute smaller shares. VC-backed firms are disproportionately represented in these innovation-intensive industries.

Taken together, Figure 1 shows that class actions involving VC-backed firms are more concentrated around the IPO event. These patterns motivate our focus on venture fraud in the private space. They also motivate our empirical approach to use class actions started within two years of IPO as a proxy for wrongdoings initiated during the pre-IPO period.

Figure 2 compares the likelihood of class actions between VC-backed and non-VC-backed IPOs. Panel A plots the percentage of firms that face a class action within two years of IPO, based on the class period start date. VC-backed firms show a higher likelihood of litigation relative to non-VC-backed firms over time, with the exception of 2008-2009, when very few VC-backed firms went public. The divergence between the two groups widens notably from 2009 to 2015, coinciding with the rise of founder-friendly market conditions. Over the entire period of 2002-2022, the class action rate is 10% among VC-backed firms, compared with 5% among non-VC-backed firms.

Panel (b) of Figure 2 compares the share of VC-backed firms among IPOs that face class actions versus those that do not. Consistent with prior findings, we find that the VC-backed rate is higher among IPOs that face a class action within two years of the IPO than among those that do not, except again in financial crisis years due to a lack of VC-backed IPOs. Both figures highlight a widening difference in fraud rates between VC- and non-VC-backed IPOs from 2009 to 2015, coinciding with the rise of founder-friendly conditions in private markets. We examine this link more directly in Sections 6 and 7.

## 5.2 Regression Results

Table 4 reports regression estimates comparing the likelihood of securities class actions between VC-backed and non-VC-backed IPO firms. The specification controls for industry-year fixed effects and pre-IPO firm size and revenue growth – two characteristics that could correlate with fraud commission and detection.

Panel A covers the full 2002-2023 sample of non-dismissed class actions. The dependent variable equals one if a firm is sued within one year (columns 1 and 4), two years (columns 2 and 5), or three years (columns 3 and 6) of IPO, based on the start of the class period. VC-backed firms are 4.3–5.2 percentage points more likely to face litigation than non-VC-backed firms, a large difference representing 80–90% of the sample mean. We obtain similar results when controlling for pre-IPO financials (columns 4 to 6).[9] Both pre-IPO assets and revenue growth are positively associated with having a class starting immediately after the IPO.

When we split the sample period in the middle, we find that this difference is larger in the recent decade (2013-2023) (Panel B) than the decade post the Dot-com bubble (2002-2012) (Panel

[9]The reduction in sample size reflects limited availability of pre-IPO asset and revenue data in Compustat.

C). In Panel B, we fail to find a robustly significant difference in class action likelihood between VC and non-VC backed firms, though the percentage effect relative to the mean is still substantial. In contrast, over the recent decade of 2013 to 2023, the effects are both statistically and economically more significant than the full sample effects in Panel A.

Appendix Table A.3 reports several robustness tests. Panel A shows that results are similar when defining the post-IPO window using filing date rather than class period start. Panel B includes dismissed suits, which may still signal credible fraud concerns. Panel C restricts to more serious cases with settlement amounts above $3 million. Panel D (to be added) addresses sample selection issues by ending the window in 2019.

Overall, we find evidence that VC-backed firms are more likely to be involved in fraud than non-VC-backed firms before or around IPO. Importantly, this difference is not driven by public versus private ownership status, disclosure requirements, size, or growth differences. Rather, the unique nature of VC-backed startups and VC ownership drives this difference.

# 6 What Predicts Fraud by VC-Backed Startups?

## 6.1 Empirical Strategy

In this section, we seek to understand the determinants of fraud among VC-backed startups. One approach to studying this question is to use a predictive analysis of fraud in a startup panel. The empirical challenge is that we can only predict detected fraud, which is committed fraud multiplied by detection likelihood. Conditioning on enforced fraud has the advantage of focusing on material cases with significant economic impact, as this is also the criterion regulators use to allocate enforcement resources. The downside, however, is that we must also worry about predictors of fraud correlating with drivers of enforcement or detection likelihoodâ-an omitted-variable problem.

We take four approaches to address this challenge. First, rather than predicting fraud detection, we predict the timing of fraud start, which predates detection. Second, we remove potential confounders that affect fraud detection by controlling for a set of drivers of enforcement likelihood while ensuring that our remaining predictors do not impact detection. Specifically, we control for firms' time-varying age, valuation, and total raised amount, as older, larger, and more visible firms are more likely to attract attention from the public or regulators. The likelihood of enforcement may also be higher in certain industries or periods, such as the cryptocurrency industry from 2020 to 2024. We therefore control for industry-year fixed effects.[10] Finally, missing values in various firm characteristics contain important information about firm's visibility, which could also drive

[10]For our predictive analysis, we define industry by PitchBook's primary industry code, which contains 214 industries.

detection, which we include as controls.

Our third approach is similar to the approach used in Section 5, where we focus on newly public firms that just went through an IPO—a subsample where detection likelihood is high and uniform across firms. In particular, we focus on VC-backed IPO firms and measure fraud using class-action suits within two years of IPO (Tian et al., 2015). The key assumption is that fraud revealed shortly after the IPO can be attributed to governance failure prior to the IPO that was baked into the firm while private.

Our final approach relaxes our focus on enforced fraud cases to additionally include dismissed cases as well as allegations from news. Starting a lawsuit or an investigation is costly, these allegations from the new can indicate legitimate concerns for fraud, even though they do not result in formal enforcement actions for various reasons (often technical reasons to dismiss the suit). Using this broadened sample of fraud allegations helps mitigate concerns about selective enforcement when focusing on enforced cases.

Finally, we note that our predictive analysis is correlational not causal, as many of the predictor variables are endogenous, such as governance contracts. In robustness (Section 6.3), we mitigate time-invariant endogeneity concerns by including firm fixed effects to exploit only within-firm variation in the timing of fraud start.

**Panel analysis of VC-Backed startups** We use a panel regression to predict the start of fraud among VC-backed startups, leveraging the fraud commitment period we hand-collected. The sample is a hazard-style panel of US VC-backed firms founded since 2000. For fraud firms, we track them from their founding year to the year the fraud starts. For non-fraud firms, we trace all years from a firm's founding to its closure/exit. Hence, we exploit not just the cross-sectional variation in fraud incidence across firms, but also variation in the timing of fraud. Specifically, we estimate the following specification:

$$\begin{aligned}\mathbb{1}(FraudStart)_{i,t} \times 100 = {} & \alpha_{j,t} + \beta_1 FirmAge_{i,t} + \beta_2 Ln(valuation)_{i,t} + \beta_3 Ln(raised)_{i,t} \\ & + \boldsymbol{\beta_4 Board_{i,t}} + \boldsymbol{\beta_5 TermSheet_{i,t}} + \boldsymbol{\beta_6 Investor_{i,t}} + \boldsymbol{\beta_7 Team_i} \\ & + \boldsymbol{\beta_8 \mathbb{1}(MissingVal)_{i,t}} + \epsilon_{i,t} \end{aligned} \tag{1}$$

The dependent variable is an indicator for the fraud start year, multiplied by 100 to interpret changes in percentage points. For our main analysis, we consider fraud cases sourced from SEC, DOJ, Westlaw, and class action suits.[11] In robustness, we also expand to include cases and allegations from news. All independent variables other than team characteristics are time-varying at the

[11] We do not know the exact fraud start time for class action cases, as class action suits focus on issues during the class period (typically at or right after IPO). For simplicity, we set the fraud start year to the IPO year for class action cases.

firm-year level.

We consider several sets of determinants of fraud motivated by theories and the literature. First, Wang et al. (2010) show that industry business conditions affect the incentives to commit fraud. We capture these conditions with industry-year fixed effects ($\alpha_{j,t}$). We account for potential life-cycle patterns in fraud likelihood by controlling for $FirmAge$ relative to founding. We also control for startup size with $Ln(valuation)$ and $Ln(raised)$—time-varying post-money valuation and cumulative raised amount.[12] These two variables together also capture the growth expectation of a startup implied from its valuation multiple (i.e., valuation/total raised). As discussed above, all these controls also absorb firm and industry-level variation in detection likelihood, so that our remaining predictors pick up only variation in true fraud.

Next, we consider proxies for vertical conflicts between investors and founders. In particular, how control and cash flow rights are allocated between them. We first consider three time-varying board characteristics in $\boldsymbol{Board_{i,t}}$: board size, an indicator for VC-controlled board, and an indicator for board with shared control, with founder-controlled board being the omitted group.[13] The board control variables capture the extent to which founders can unilaterally make decisions, or if investors or independent directors have enough control power to discipline founders.

We also consider several important term sheet features that affect entrepreneurs' incentives or ability to commit fraud. $\boldsymbol{TermSheet_{i,t}}$ includes 1) investor ownership on a fully converted basis; 2) a founder payoff convexity index that is the sum of the dollar-weighted fraction of rounds with high liquidation multiple (>1) and the dollar-weighted fraction of rounds with participation rights. Because we do not observe contingent payoffs, converted investor ownership is calculated based on all preferred shares converted to common and hence does not capture investors' actual cash flow rights. However, it is a close proxy for investors' voting rights vis-à-vis founders, as most shares (common or preferred) have one vote per share, and dual-class shares are not typically not created untill late stages before or at IPO. High liquidation multiple and participation rights reduce entrepreneurs' cash flow rights and hence their "skin-in-the-game"; they also make entrepreneurs' payoff more convex, increasing their incentive to gamble for extreme outcomes.[14] Figure A.4 illustrates these incentive effects by plotting founder's payoff diagram under three scenarios: a baseline case of non-participating convertible preferred with 1x liquidation preference, non-participating preferred with 2x liquidation preference, and participating preferred with 1x liquidation preference. Relative to the

[12] Following Ewens et al. (2024), we interpolate valuation between rounds.

[13] Following Malenko and Ewens (2024), we define a board as "VC-controlled" if VC directors hold strictly more than 50% of seats, or if they hold exactly 50% of seats and executives hold strictly less. We define a board as "founder-controlled" if executives hold strictly more than 50% of seats, or if they hold exactly 50% and VCs hold strictly less. The remaining cases are deemed "shared control", i.e., VC and executive directors have the same number of seats. Appendix A details how we construct these variables and extend the sample in Malenko and Ewens (2024).

[14] We do not include anti-dilution provisions as they are less important for founders (Feld and Mendelson, 2019) and are rarely exercised (Metrick and Yasuda, 2021).

baseline case, both high liquidation preference and participation rights by investors reduce founders' payoffs while increasing the payoff convexity.

It is important to note that various control and cash flow rights measures can correlate with each other, and it is important to interpret the partial effect of one variable conditioning on the others. For example, controlling for investor converted ownership allows us to focus on control rights that deviate from one-share-one-vote (through board voting rights or superior voting shares), as well as the non-linearities in cash flow rights.

We also include proxies for horizontal conflicts among investors. In $\boldsymbol{Investor_{i,t}}$, we include the number of unique investors to measure coordination frictions among investors and potential free-riding incentives in monitoring. We also include the fraction of investors who are not independent VCs (IVCs), such as angels, CVCs, mutual funds, banks, and hedge funds. These investors are known to be weaker monitors than IVCs due to their liquidity pressures and/or a lack of expertise (Chernenko et al., 2021; Shane, 2008). They also tend to come in either very early (e.g., angels) or late (e.g., mutual funds) in firm's financing rounds, creating horizontal conflicts between investors across rounds. These horizontal frictions can exacerbate the vertical conflicts between founders and investors (Pollman, 2019). Lastly, we control for investor reputation using lead investors' past number of successful exits. These variables are time-varying and measured as of year $t$.

Finally, we consider founder characteristics. $\boldsymbol{Team_i}$ includes an indicator for solo founder, an indicator for the presence of serial entrepreneurs in the founding team, the fraction of founders graduated from top universities (defined in footnote 8), and average founder age.[15] Solo founder may find it easier to commit fraud due to the absence of checks and balances from co-founders. As such, we consider this as a governance variable. Founders' experience, education, and age may also shape their propensity to commit fraud.

Many of the above variables are missing for a sizable share of the firm-year observations in our sample. We explicitly model these missing values, as they are often driven by firms' disclosure choices or their visibility, which could affect fraud incentives or the likelihood of detection. Another benefit is that it helps to preserve sample size when we simultaneously include all predictors. To this end, we fill missing values with zero while controlling for a vector of missing-value indicators in $\mathbb{1}\boldsymbol{(MissingVal)_{i,t}}$. Table 3 provides summary statistics for these variables with missing values unfilled.

The key identifying assumption in this analysis is that, conditioning on variables that drive enforcement likelihood (i.e., age, valuation, size, industry-year FE, and missing values), the remaining predictors, such as internal governance and founder demographics, should largely explain variations in true fraud, because these variables do not directly impact enforcement likelihood.

[15] We infer founder age from BA degree year.

**Cross-Sectional Analysis of VC-Backed IPOs** Our second approach addresses variation in detection likelihood by exploiting a subsample of newly public VC-backed firms that face the same heightened and homogeneous enforcement environment. We estimate the following specification on a cross-section of VC-backed IPOs:

$$\begin{aligned}\mathbb{1}(ClassAction)_i \times 100 = \alpha_{j,y} &+ \beta_1 FirmAge_i + \beta_2 Ln(valuation)_i + \beta_3 Ln(raised)_i + \beta_4 DualClass_i \\ &+ \boldsymbol{\beta_5 Board_i} + \boldsymbol{\beta_6 TermSheet_i} + \boldsymbol{\beta_7 Investor_i} + \boldsymbol{\beta_8 Team_i} \\ &+ \boldsymbol{\beta_9 \mathbb{1}(MissingVal)_i} + \epsilon_i \end{aligned} \tag{2}$$

The sample is at the IPO-firm level, indicated by $i$. The dependent variable is an indicator (multiplied by 100) for whether the firm was involved in a class action suit within 2 years of IPO (defined based on class period start date). As shown in Figure 1(b), the majority of class action suits against newly public firms happen within 2 years of IPO, suggesting that fraud revealed at IPO likely originated beforehand.[16] We include fixed effects for the firm's industry ($j$) interacted with IPO year ($y$). All independent variables are the same as those in Equation 1, but measured in the year before IPO. In addition and specifically for this test, we include an indicator for dual-class IPOs.[17]

## 6.2 Predictive Regression Results

Table 5 presents the panel prediction results. We employ a hazard-style panel to predict the initiation of fraud, focusing on fraud cases identified from SEC, DOJ, Westlaw, and class action suits.

We find that firm age does not significantly predict fraud. However, high valuation and total raised amount strongly predict venture fraud, but with very different magnitudes. Based on Column 1, a 10% increase in post-money valuation increases fraud likelihood by 30% relative to the mean, while a 10% increase in total raised amount only increases fraud likelihood by 2.3% relative to the mean. This suggest that the majority of the valuation effect reflects valuation multiple rather than book value.[18] To the extent that cost of capital varies largely by industry and is absorbed by our industry-year fixed effects, remaining variation in valuation multiple largely reflects firm-specific expectations on future growth or profitability. This finding suggests that high expectations by the market incentivizes fraud, likely through the pressure on founders to meet such expectations. It is also possible that these two variables affect enforcement likelihood, as larger firms are more likely to be targeted by regulators, particularly those with more assets in place to pay a penalty or settlement.

[16] Consistent with this, Wang (2013) also find that the average fraud at IPO takes two years to reveal.

[17] A large literature (Masulis et al., 2009; Smart and Zutter, 2003; Xu, 2021) highlight the agency issues in dual-class firms, which may drive fraud.

[18] The implied coefficient on ln(multiple) is the difference between the coefficients on ln(valuation) and ln(raised), i.e., 0.143-0.011=0.132 based on Column 1.

We find strong predictive power in various proxies for agency conflicts within startups. Starting with board characteristics, we find that a larger board is associated with a higher likelihood of fraud. Yet board composition matters much more. Holding constant board size, VC-controlled boards are 1.8 times less likely to commit fraud than founder-controlled boards (the omitted group), while shared-control boards are 55% less likely. This suggests that board control rights by the founders are a strong determinant of venture fraud. Confirming the importance of control rights in predicting venture fraud, we also find that a 10% higher investor ownership (on a converted basis) reduces founder fraud by 17% relative to mean.

Founders' payoff structure matters too. In particular, high liquidation multiples (>1) and/or investor participation rights, captured in *Founders' payoff convexity*, is associated with a higher fraud likelihood. Specifically, a one-standard-deviation increase in *Founders' payoff convexity* is associated with a 17.4% increase in fraud likelihood relative to the mean. This is consistent with these contracting features reducing founders' cash flow rights and hence their "skin-in-the-game"; they also make founders' payoff more convex, incentivizing them to gamble for extreme outcomes. [19]

We then turn to horizon conflicts among investors in Column 2 of Table 5. We find that a greater number of unique investors and a higher fraction of non-IVC investors positively predict fraud. In particular, one additional investor in the cap table increases the fraud likelihood by 9.3% relative to the mean, while a 10% higher fraction of non-IVC investors increases it by 6.2%.[20] Investor reputation also matters. A one-standard-deviation increase in the lead investor's track record, measured by past exits, reduces fraud likelihood by 24%, a large effect.

Table 5 Column 3 examines founder characteristics. We find that solo founders are much more likely to commit fraud than teams, with a magnitude of 24%. This is consistent with solo founders being free from the checks and balances from co-founders. Founders' past experience and education do not significantly predict fraud, while age has a precisely estimated zero effect on fraud likelihood.[21]

Column 4 combines all predictors into a single regression. The above estimated effects persist, with even larger coefficients for investor ownership, founders' payoff convexity, and the number of unique investors.

Finally, as reported in Table A.4, startups with greater opaqueness, proxied by missing information about their valuation, board, investors, or founding team in PitchBook, are associated with

[19] Payoff convexity should not be considered as a firm-level founder-friendliness measure, since founders typically value control rights more than cash flow rights when trading off the two (Berger et al., 2025).

[20] Based on Table 3, the mean investor count is 3.8. Hence the percentage effect of one additional investor is (ln(4.8)-ln(3.8))*0.018/0.045=9.3%.

[21] We explore more founder characteristics in the next section and continue to find insignificant effects.

higher fraud likelihood. This is consistent with transparency facilitating public monitoring. Note that much of this information is sourced from startups' Form D filings, yet many startups choose not to file them to fly under the radar (Malenko and Ewens, 2024; Hanley and Yu, 2023). In other words, missing data is not random, but reflects a choice by startups that is predictive of fraud. As such, it is important to model this missing information in our analysis.

Table 6 shows the cross-sectional prediction results on class action suits against VC-backed IPO firms. This analysis further addresses unobserved variation in detection likelihood by focusing on a subsample with a uniform enforcement environment. We find that younger IPO firms are more likely to face class actions. A likely explanation is that these firms experienced faster growth and took fewer years to achieve exit. As in our panel analysis, both valuation and cumulative raised amount pre-IPO positively predict class action suits after IPO. The coefficient is again much larger for valuation than raised amount, suggesting the effect of a high valuation multiple.

We continue to find that, relative to founder-controlled boards, VC-controlled boards pre-IPO are associated with 61% lower fraud likelihood relative to the mean, while boards with shared control pre-IPO have 36% lower fraud likelihood relative to the mean, though statistically insignificant. Dual-class IPOs are 81% more likely to be involved in class action suit immediately after IPO. This is consistent with the idea that excess control rights give founders, who typically hold the superior voting class shares, the ability and incentive to exploit minority public shareholders. Similar to the panel regression, we also find evidence that lower investor ownership and higher founder payoff convexity pre-IPO positively predict fraud, though these effects are not statistically significant due to low power. The coefficients on the number of unique investors, the fraction of non-IVC investors, and lead investor reputation all have consistent signs, with statistical significance for non-IVC investors in Column 2 and for unique investor count in Column 4.

Finally, similar to the panel regression, we find insignificant coefficients on founder traits such as experience, education, and age. Solo founder loads positively but is statistically weak.

Taken together, the results from this section suggest that venture fraud is strongly shaped by agency frictions and contracting environments, which are distinct and more complex than those of public firms. Founder backgrounds, in contrast, play a much smaller role. These findings suggest that "nurture" is more important than "nature" in explaining venture fraud.

### 6.3 Robustness

We conduct several robustness tests for our main prediction analysis.

**Exploiting only within-firm variations.** Finally, we conduct a predictive analysis including firm fixed effects. This specification only exploits variation in the timing of fraud start within firm, not the cross-sectional variation of whether a firm is ever involved in fraud. This addresses endogeneity from time-invariant firm-level unobservables—for example, both contracts and fraud select on unobserved founder or business characteristics. Importantly, it also addresses selective enforcement based on fixed firm characteristics, since this specification only exploits timing of fraud start (not fraud detection) within enforced firms.

**Additional team characteristics.** Table A.5 adds more founder characteristics, such as immigrant status and gender. These variables are defined as the average across founders within the team. We identify immigrants as those whose undergraduate school is outside US. We continue to find weak to insignificant effects from these team characteristics, either by themselves or conditional on other controls in Table 5. This confirms our finding that venture fraud is not driven by founder backgrounds or any inherent tendencies. Rather, incentives and contracts shape the likelihood of fraud. The similar coefficients on team variables with and without other controls also suggest that founders do not sort into different governance environments based on their demographics.

**Alternative samples or definitions of fraud.** Our results are robust to alternative samples or definitions of fraud cases. First, Table 7 expands the fraud sample by including cases identified from news. These news-based cases include allegations and investigations, and may not all result in legal actions or penalties. Nevertheless, to the extent that media rarely report news without merits and investigations are costly, these cases represent legitimate concerns about fraud. Importantly, by not conditioning on formally enforced cases, this sample mitigates the concern that our results reflect variation in enforcement likelihood rather than underlying fraud. We find very similar results in Table 7. To address the concern that our news-based cases have much shorter coverage (i.e., mainly after 2017) than other sources of cases, we show in Table A.6 that our results remain similar using only news-based cases. Next, Table A.7 considers restricted subsamples of startup-years with at least $1M, $5M, or $10M of cumulative funding. This accounts for the fact that startups without meaningful traction will attract little public or regulatory attention and face a low likelihood of enforcement. Fraud by these firms, detected or not, are also unlikely to be economically consequential. Finally, there was a large enforcement wave by the SEC against cryptocurrency companies. Table A.8 shows that our baseline predictive results are robust to dropping firms in cryptocurrencies or blockchain Pitchbook verticals.

# 7 Founder-Friendly Market Conditions and Venture Fraud

A central theme emerging from the analyses above is that a founder-friendly governance environment, such as high control rights to founders, breeds fraud. To further examine this insight, we investigate the effect of aggregate VC market conditions at firms' initial financing on firms' future probability of being involved in fraud.

We measure funding conditions at firms' first VC round (Nanda and Rhodes-Kropf, 2013), as this is the round when major screening occurs. We hypothesize that a founder-friendly VC market would lead to lax screening, increasing the chance that VCs bet on fraudulent startups. Further, regardless of screening intensity, founder-friendly contracting in the initial round may also lay the foundation for lax future monitoring, thereby increasing the risk of fraud.

We employ two proxies for founder-friendliness at the industry-initial-funding-year level: average valuation multiple and the fraction of founder-controlled boards, both measured from first VC-round deals.[22] To make sure our funding condition measures do not pick up general industry heterogeneity or local shocks, we control for industry fixed effects and state-year fixed effects. As such, we exploit variation in market conditions within an industry across years and within a year across industries. We also control for firm-level age, post-money valuation, and total amount raised, all measured at the firm's first VC round.

Table 8 reports the results. We find that firms initially financed in hotter markets–those with higher valuation multiples or a higher fraction of founder-controlled boards—are more likely to commit fraud in the future. A one standard deviation higher market multiple (=1.39) is associated with 18% higher future fraud likelihood relative to the mean, and a 10% higher fraction of founder-controlled boards is associated with 14% higher future fraud likelihood relative to the mean (column 1). These effects are similar when controlling for the firm's age at initial round. Importantly, the effects are similar even when we control for the startup's own valuation and board composition at initial round. These results suggest that a founder-friendly VC market is more likely to finance fraud-prone startups through lax screening or monitoring.

We then examine time-series correlations at the macro level. Figure 5 plots the trend in venture fraud rate against various proxies for founder-friendliness, all at the annual level. To mitigate truncation, we end the sample period at 2021. The blue dashed line represents the fraud rate among all VC-backed firms that raised at least $1M. We observe a sharp increase in the fraud rate from 2008 to 2016, followed by a dip.

Strikingly, the trend in venture fraud rate mirrors the trends in various proxies for founder-friendliness in the VC market. For instance, valuation multiple exhibits a similar pattern, rising

[22] We define industry as the 42 primary industry groups in PitchBook.

sharply until 2015, dipping slightly before peaking again. A similar pattern is observed for the fraction of founder-controlled boards. The opposite trends are observed for proxies for investor control or bargaining power vis-à-vis founders, such as the fraction of VC-controlled boards (Panel C) and average investor ownership (Panel D). These patterns provide a strong indication that the rise of founder-friendliness in the VC market post-financial crisis is responsible for the rise in venture fraud.

# 8 Does the VC Market Discipline Fraudulent Founders? Evidence from Future Founding Activities

Our results show that startups' internal governance incentives, rather than founder characteristics, shape founders' propensity to commit fraud. However, external discipline from the labor and financing markets, if strong enough, could mitigate these internal incentive effects. In this section, we investigate whether such a market mechanism exists by studying how fraud detection affects entrepreneurs' future founding activities.

To isolate the treatment effect of fraud detection on entrepreneurs, we conduct a matched-event study at the individual level. For each treated founder in the year before the initial fraud charge, we match him/her to three similar control founders who were never involved in any fraud event. Specifically, we match on the exact same number of past startups, tenure at the current startup, the current startup's sector, gender, immigrant status, and the top school indicator. Within this set, we then select those closest in age to the treated founder. Matching on past founding experience, age, and tenure at current startup is important, as it ensures that the treatment and control founders are on similar stage of their career, since founding activities typically follow a life-cycle pattern Azoulay et al. (2020). We then construct a person-year level panel by tracing these individuals' founding activities in PitchBook. Because firms in PitchBook VC universe are all VC-backed, such founding represents both launching a new startup and successfully raising VC financing.

Using this matched panel, we estimate a dynamic DID. To account for potential bias from staggered treatment events, we implement a stacked DID following Cengiz et al. (2019). This method explicitly pairs each treated founder with never-treated control founders to form a "stack". The specification includes founder fixed effects, stack $\times$ event year fixed effects, and stack $\times$ calendar year fixed effects. As an alternative, we also implement the imputation-based DID from (Borusyak et al., 2024).

Figure 6 shows the event study plots, with Panel (a) using stacked DID and Panel (b) using imputation-based DID. The dependent variable is the number of new VC-backed startups founded by a person in a year. Overall, we do not find any significantly negative effect of fraud detection on fraudulent entrepreneurs' future founding activities. If anything, these individuals found slightly more new startups after fraud revelation, especially in the initial three years. Over the six years post fraud charge, treated founders started about 0.008 more new startups than control founders, a negligible difference. Baseline DID estimates in Table 9 confirm these findings.

The above results are striking, since one would expect that the legal and reputational damage from the fraud charges would negatively affect the founders. This suggests that these fraudulent founders can continue unharmed, launch their next startup, and raise VC financing. This provides evidence that the VC market does not discipline fraudulent founders ex post. This finding is consistent with rising founder-friendliness, in which VCs compete to fund a limited supply of startups and are thus willing to give up bargaining power and relax screening. It is also consistent with Silicon Valley culture, which embraces failure regardless of the reason. The lack of ex post market discipline provides another ex ante incentive for founders to commit fraud in a market where founders have all the bargaining power.

# 9 Conclusion

There is a trade-off in VC investing when investor choices have different impacts on generating (and capturing) upside versus mitigating downside risk. A shift towards founder-friendly governance in VC investment, including ceding control to founders and a "spray-and-pray" approach, has been advocated as a way to increase the potential for upside. This paper contributes to understanding this tradeoff by examining the impact, if any, of choices that reduce investor power on the downside.

We assemble a comprehensive database of fraud in US VC-backed companies from 2002–2023. We collect data on enforcement actions by the SEC and DOJ, and through private suits, including class actions, from the Stanford Securities Class Action database and Westlaw. We test whether fraud has increased during the founder-friendly period and whether the likelihood of fraud is associated with founder-friendly choices, other contractual incentives, and the characteristics of founders and investors. We also examine whether there are market penalties for venture fraud for founders.

We have four primary findings. Among firms that recently had an IPO, VC-backed firms are significantly more likely to have fraud than non-VC-backed firms, consistent with weak control environments established in the pre-IPO period persisting after IPO. The difference is particularly strong in the past decade, which saw a prevalence of founder-friendly structures. Second, we use panel regressions of startups to investigate the importance of founder-friendly structures on fraud incidence. Firms with founder-controlled boards have 2 times higher fraud likelihood than VC-controlled boards. Fraud rate is also substantially higher with lower VC ownership. Third, we find that incentives arising from contractual features and investor composition significantly affect fraud incidence, whereas founder characteristics have a much smaller impact. More investors, and more non-VC investors with weak monitoring skills, are associated with more fraud, while VCs with a stronger reputation are associated with fewer frauds. Fourth, we find that the VC market does not discipline fraudulent founders ex post, providing another ex-ante incentive for founders to commit fraud.

A founder-friendly approach that prioritizes upside over downside risk may be privately optimal if the left-tail losses are outweighed by the right-tail gains. But the equilibrium with weak screening raises serious concerns. If deceptive types are pervasive, they could crowd out non-deceptive firms that could deliver the very right-tail outcomes VCs seek, leading to capital misallocation. Moreover, venture fraud can be socially costly and impose negative externalities on non-fraud firms through pooling. Our evidence raises the question of whether the pendulum may have swung too far. Because ex-post discipline is weak, credible contract design and targeted public enforcement are necessary to mitigate the negative externalities of venture fraud.

## References

Åstebro, T., H. Herz, R. Nanda, and R. A. Weber (2014). Seeking the roots of entrepreneurship: Insights from behavioral economics. *Journal of Economic Perspectives 28*(3), 49–70.

Atanasov, V., V. Ivanov, and K. Litvak (2012). Does reputation limit opportunistic behavior in the vc industry? evidence from litigation against vcs. *The Journal of Finance 67*(6), 2215–2246.

Azoulay, P., B. F. Jones, J. D. Kim, and J. Miranda (2020). Age and high-growth entrepreneurship. *American Economic Review: Insights 2*(1), 65–82.

Banerjee, S. and M. Szydlowski (2024). Founder friendly vcs.

Bartlett III, R. P. (2006). Venture capital, agency costs, and the false dichotomy of the corporation. *UCLA Law Review 54*, 37.

Beneish, M. D. (1999). The detection of earnings manipulation. *Financial Analysts Journal 55*(5), 24–36.

Beneish, M. D., C. M. Lee, and D. C. Nichols (2013). Earnings manipulation and expected returns. *Financial Analysts Journal 69*(2), 57–82.

Berger, M., T. Doherr, S. Gottschalk, and M. Pellens (2025). The private value of entrepreneurial control: Evidence from a discrete choice experiment. *The Review of Corporate Finance Studies*, cfaf019.

Bian, B., Y. Li, and C. A. Nigro (2023). Conflicting fiduciary duties and fire sales of vc-backed start-ups.

Borusyak, K., X. Jaravel, and J. Spiess (2024). Revisiting event-study designs: robust and efficient estimation. *Review of Economic Studies 91*(6), 3253–3285.

Broughman, B. J. and M. T. Wansley (2023). Risk-seeking governance. *Vand. L. Rev. 76*, 1299.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics 134*(3), 1405–1454.

Chernenko, S., J. Lerner, and Y. Zeng (2021). Mutual funds as venture capitalists? evidence from unicorns. *The Review of Financial Studies 34*(5), 2362–2410.

Dechow, P. M., R. G. Sloan, and A. P. Sweeney (1996). Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the sec. *Contemporary accounting research 13*(1), 1–36.

Denes, M., S. T. Howell, F. Mezzanotti, X. Wang, and T. Xu (2023). Investor tax credits and entrepreneurship: Evidence from us states. *The Journal of Finance 78*(5), 2621–2671.

Dyck, A., A. Morse, and L. Zingales (2024). How pervasive is corporate fraud? *Review of Accounting Studies 29*(1), 736–769.

Ewens, M. and J. Farre-Mensa (2020). The deregulation of the private equity markets and the decline in ipos. *The Review of Financial Studies 33*(12), 5463–5509.

Ewens, M., A. Gorbenko, and A. Korteweg (2022). Venture capital contracts. *Journal of Financial Economics 143*(1), 131–158.

Ewens, M. and M. Marx (2018). Founder replacement and startup performance. *The Review of Financial Studies 31*(4), 1532–1565.

Ewens, M., K. Xiao, and T. Xu (2024). Regulatory costs of being public: Evidence from bunching estimation. *Journal of Financial Economics 153*, 103775.

Feld, B. and J. Mendelson (2019). *Venture deals: Be smarter than your lawyer and venture capitalist.* John Wiley & Sons.

Hanley, K. W. and Q. Yu (2023). Strategic regulatory non-disclosure: The case of the missing form d. *Available at SSRN 4363027*.

Janeway, W. H., R. Nanda, and M. Rhodes-Kropf (2021). Venture capital booms and start-up financing. *Annual Review of Financial Economics 13*(1), 111–127.

Jensen, M. C. (1989, Sep–Oct). Eclipse of the public corporation. *Harvard Business Review 67*(5), 61–74. From the Magazine.

Kaplan, S. N. and P. Stromberg (2003). Financial contracting theory meets the real world: an empirical analysis of venture capital contracts. *The Review of Economic Studies 70*(2), 281–315.

Kaplan, S. N. and P. E. Strömberg (2004). Characteristics, contracts, and actions: Evidence from venture capitalist analyses. *The journal of finance 59*(5), 2177–2210.

Karpoff, J. M., D. S. Lee, and G. S. Martin (2008a). The consequences to managers for financial misrepresentation. *Journal of Financial Economics 88*(2), 193–215.

Karpoff, J. M., D. S. Lee, and G. S. Martin (2008b). The cost to firms of cooking the books. *Journal of financial and quantitative analysis 43*(3), 581–611.

Kerr, S. P., W. R. Kerr, T. Xu, et al. (2018). Personality traits of entrepreneurs: A review of recent literature. *Foundations and Trends® in Entrepreneurship 14*(3), 279–356.

Kerr, W. R., R. Nanda, and M. Rhodes-Kropf (2014). Entrepreneurship as experimentation. *Journal of Economic Perspectives 28*(3), 25–48.

Lerner, J. and R. Nanda (2020). Venture capitalâs role in financing innovation: What we know and how much we still need to learn. *Journal of Economic Perspectives 34*(3), 237–261.

Malenko, N. and M. Ewens (2024). Board dynamics over the startup life cycle. *Journal of Finance, forthcoming*.

Masulis, R. W., C. Wang, and F. Xie (2009). Agency problems at dual-class companies. *The Journal of Finance 64*(4), 1697–1727.

Metrick, A. and A. Yasuda (2021). *Venture capital and the finance of innovation*. John Wiley & Sons.

Nanda, R. and M. Rhodes-Kropf (2013). Investment cycles and startup innovation. *Journal of financial economics 110*(2), 403–418.

Phalippou, L. and W. J. Magnuson (2025). Private equity, public capital, and litigation risk. *SSRN Electronic Journal*. Working paper, available at `https://ssrn.com/abstract=5748424`.

Pollman, E. (2019). Startup governance. *University of Pennsylvannia Law Review 168*, 155.

Pollman, E. (2020). Private company lies. *Georgetown Law Journal 109*, 353.

Povel, P., R. Singh, and A. Winton (2007). Booms, busts, and fraud. *The Review of Financial Studies 20*(4), 1219–1254.

Shane, S. (2008). *Fool's Gold?: The truth behind angel investing in America*. Oxford University Press.

Smart, S. B. and C. J. Zutter (2003). Control as a motivation for underpricing: a comparison of dual and single-class ipos. *Journal of Financial Economics 69*(1), 85–110.

Tian, X., G. F. Udell, and X. Yu (2015). Disciplining delegated monitors: when venture capitalists fail to prevent fraud by their ipo firms. *Journal of Accounting and Economics 61*(2), 526–544.

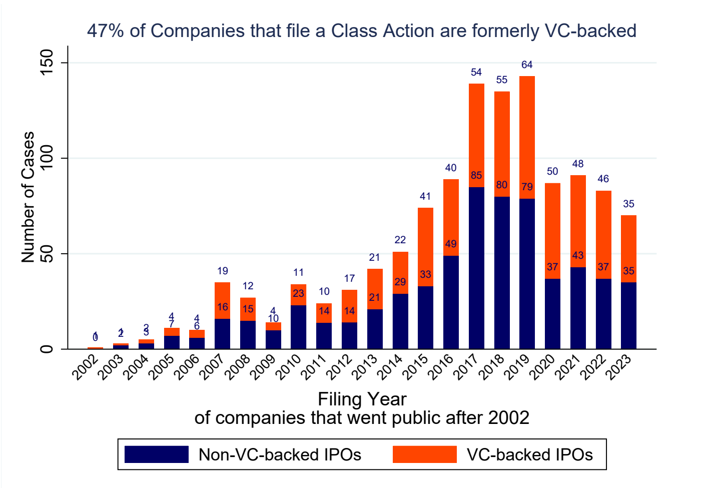
Wang, T. Y. (2013). Corporate securities fraud: Insights from a new empirical framework. *The Journal of Law, Economics, & Organization 29*(3), 535–568.

Wang, T. Y., A. Winton, and X. Yu (2010). Corporate fraud and business conditions: evidence from ipos. *The Journal of Finance 65*(6), 2255–2292.

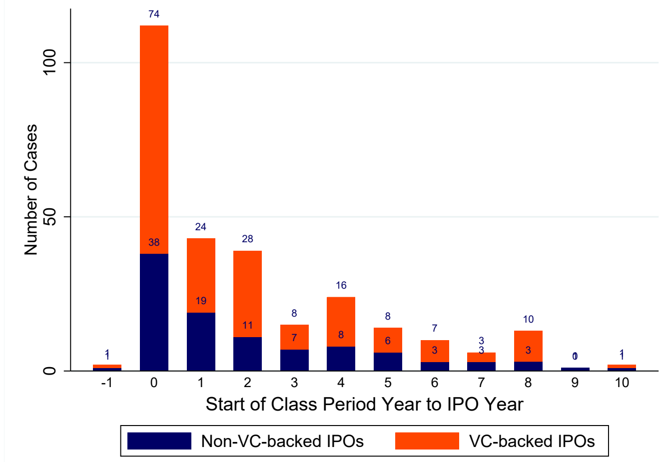
Winship, V. (2024). Unicorn shareholder suits. *Indiana Law Journal 100*, 24–03.

Xu, T. (2021). Do excess control rights benefit creditors? evidence from dual-class firms. *Journal of Financial and Quantitative Analysis 56*(3), 821–852.

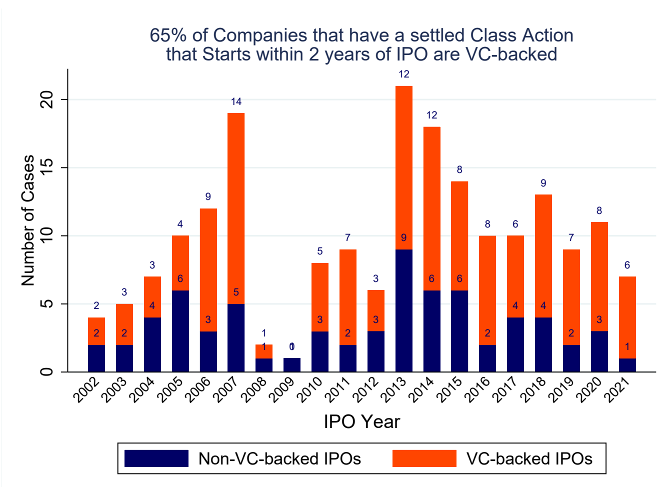
Figure 1: Class Actions by VC-backed Status



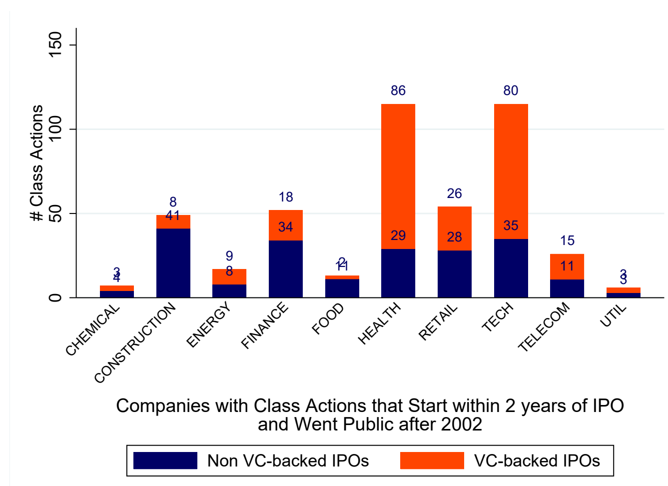

(a) Class Actions by VC Status




(b) Start of Class Period to IPO



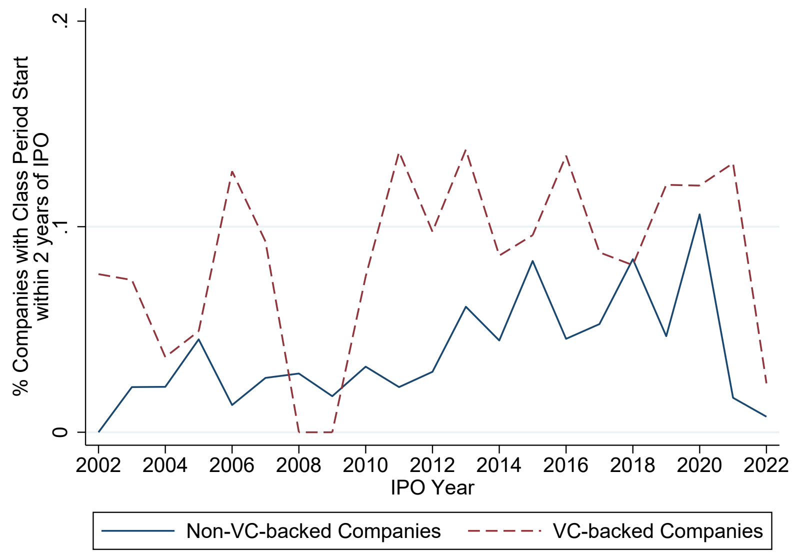

(c) Class Actions Started within 2 Years of IPO



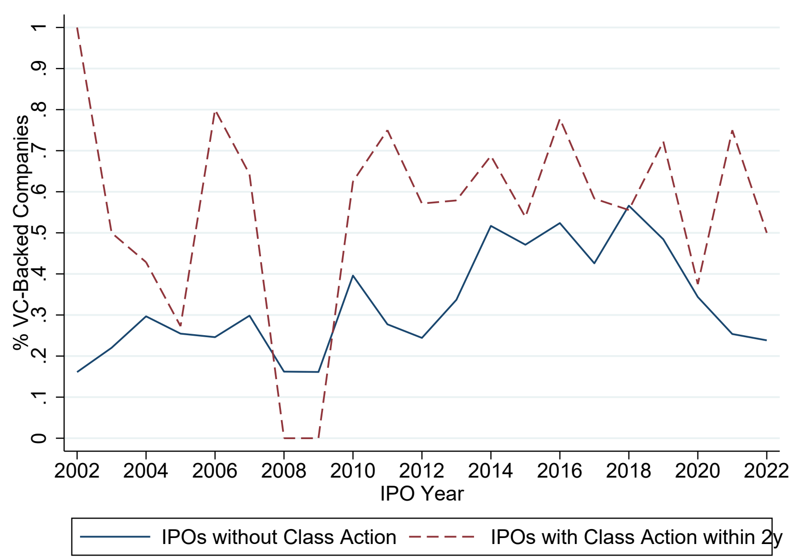

(d) Class Actions Started within 2 Years of IPO: By Industry

These figures display the distribution of securities class action lawsuits over time and by VC status, for VC-backed and non-VC-backed firms. The sample is restricted to companies that went public after 2002. Only non-dismissed class actions are included in the sample. Figure (a) shows the number of class action filings by filing year. Figure (b) reports the number of class actions by the number of years between the IPO year and the start of the class period year. Figure (c) displays the number of cases by IPO year, in which the class period began within two years of the IPO. Figure (d) shows the number of class actions by industry, restricting to cases where the class period began within two years of the IPO.

Figure 2: Class Action Likelihood and VC-Backing Among IPO Firms



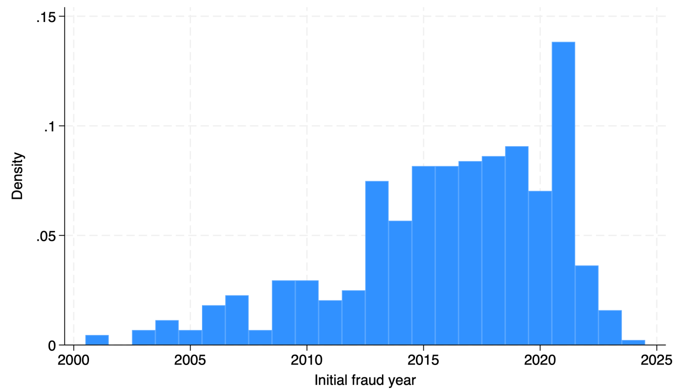

(a) Class Action Likelihood: VC-Backed vs Non-VC-Backed



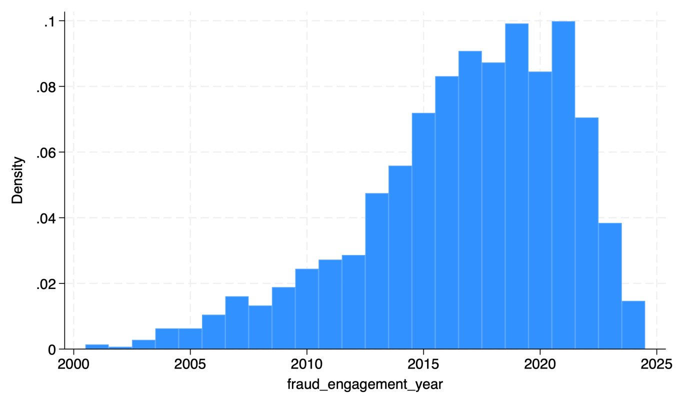

(b) Fraction of VC-Backed Firms: All IPOs vs IPOs with Class Actions

These figures plot the likelihood of securities class action lawsuits among VC-backed versus non-VC-backed IPOs. Figure (a) compares the class action rates of VC-backed and non-VC-backed IPO firms, when considering only lawsuits where the class period begins within two years of the IPO. Figure (b) shows the percentage of VC-backed IPOs among all IPOs over time (solid line) and the percentage of VC-backed IPOs that are subject to a securities class action lawsuit within two years of the IPO, measured by the start date of the class period (dashed line). The sample of class actions includes only non-dismissed class actions.

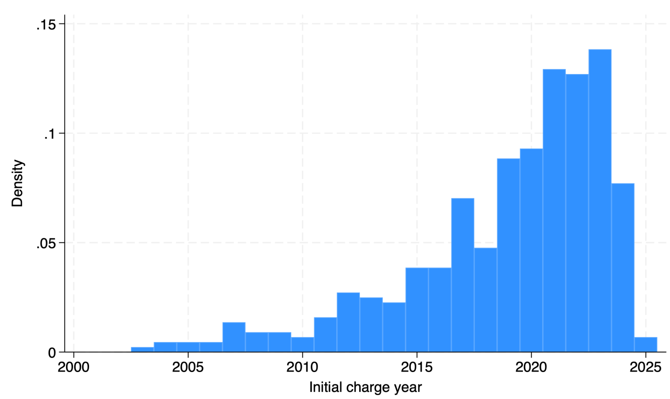
Figure 3: Timing of Fraud Among US VC-Backed Startups




(a) Fraud Start Year



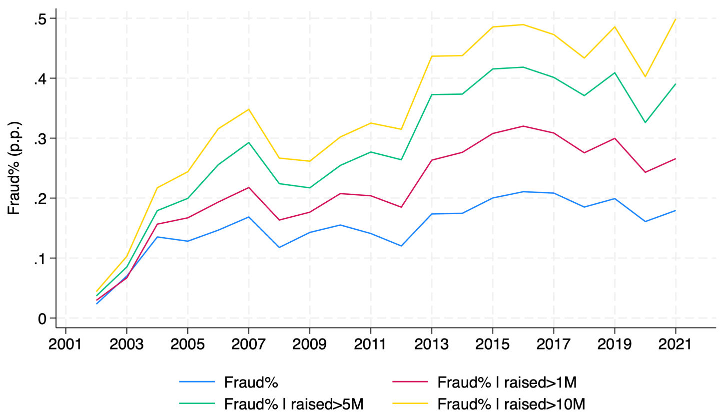

(b) Fraud Committing Years



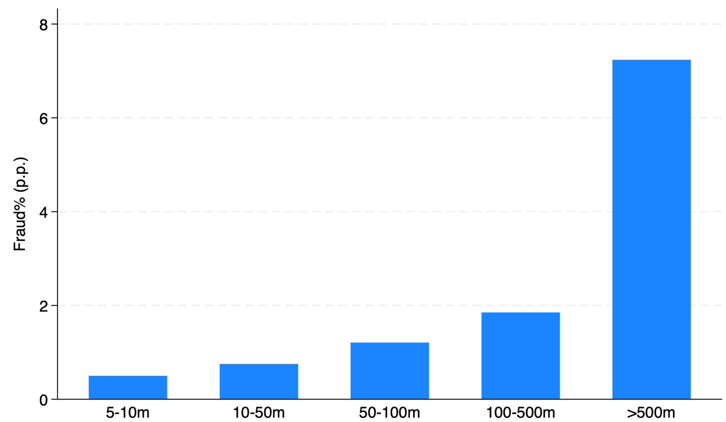

(c) Initial Charge Year

This figure shows the timing of fraud among US VC-backed startups founded post-2000. Panel A shows the histogram for fraud start year, Panel B shows the distribution for fraud committing years (i.e., years form fraud start to fraud end), and Panel C shows the distribution of initial charge year.

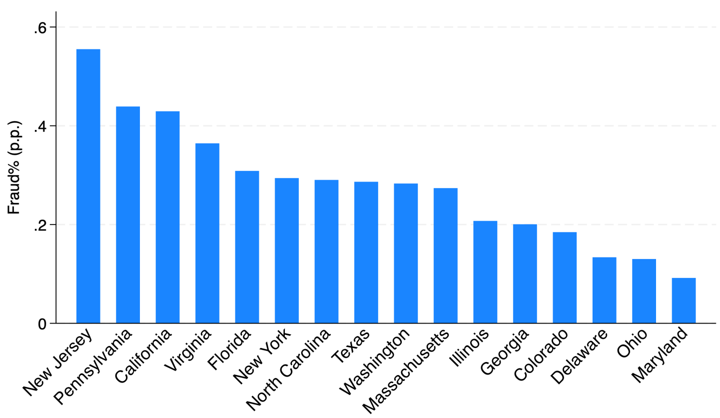
Figure 4: Variation in Fraud Likelihood Among US VC-Backed Startups



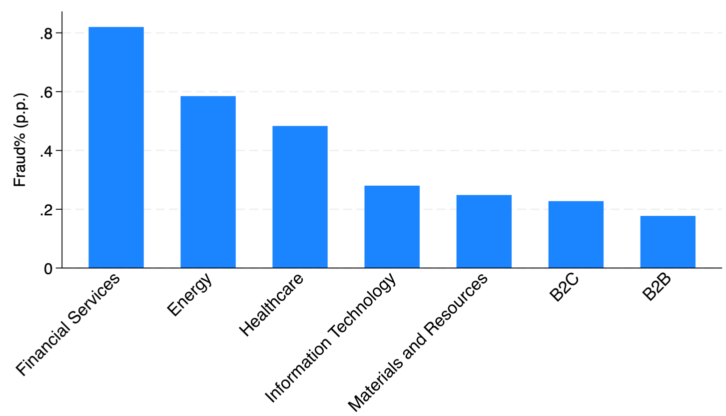

(a) By cumulative raised amount




(b) By post-money valuation at first round



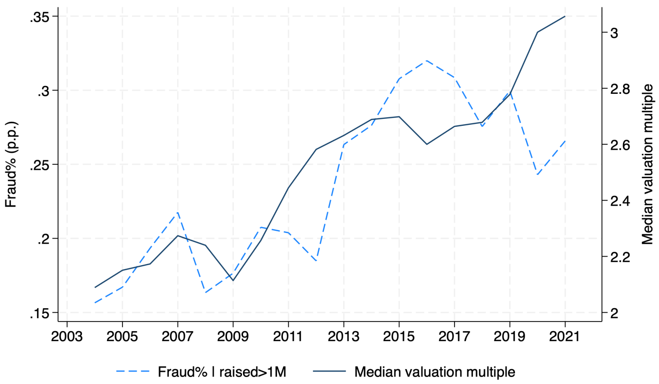

(c) By headquarter state



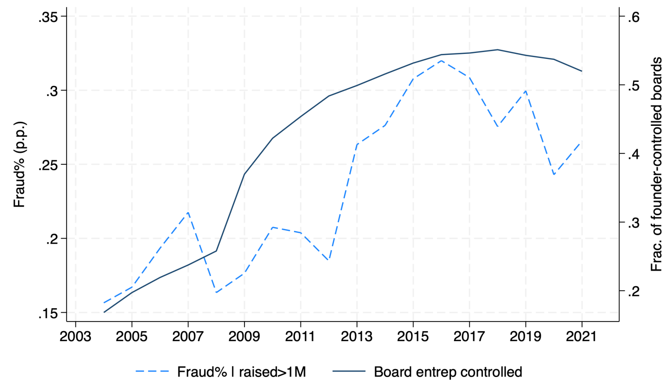

(d) By sector

This figure shows fraud likelihood among US VC-backed Startups along with cumulative raised amount (Panel A), post-money valuation at first VC round (Panel B), headquarter state (Panel C), and sector (Panel D)

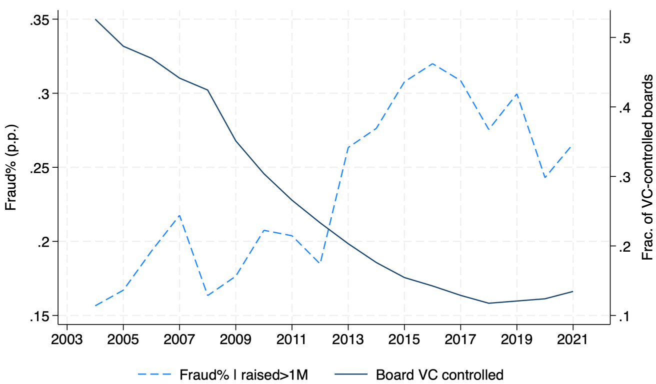
Figure 5: Trends in Fraud Likelihood and Founder Friendliness



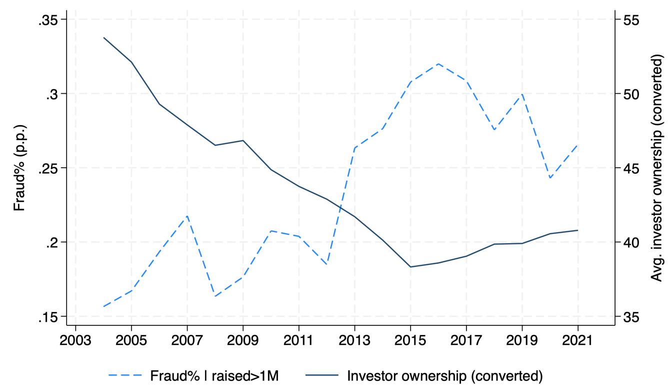

(a) Valuation multiple




(b) Founder-controlled boards



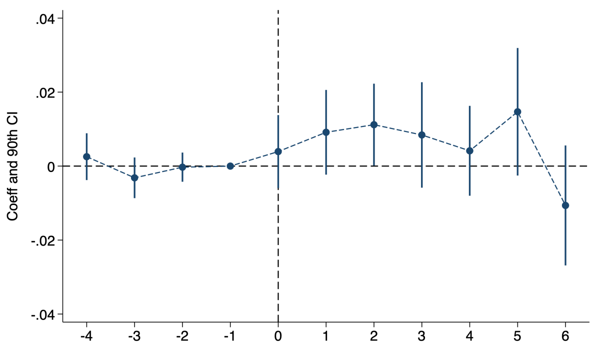

(c) VC-controlled boards



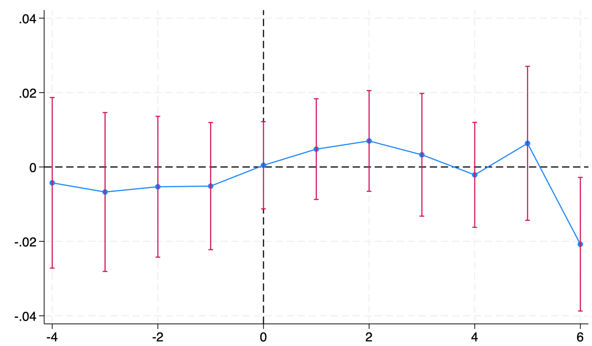

(d) Investor ownership (converted)

This figure plots the annual probability of fraud (in percentage points) by US VC-backed startups founded since 2000, against various measures of founder-friendliness in the VC market. In all plots, the blue dashed line is fraud likelihood among startups that raised at least $1M. The solid navy line is the median valuation multiple (post-valuation divided/total raised amount) in Figure (a), the fraction of founder-controlled boards in Figure (b), the fraction of VC-controlled boards in Figure (c), and the average investor ownership (converted) in Figure (d).

Figure 6: Effect of Fraud on Entrepreneurs' Future Founding: Dynamics Around Charge Year




(a) Annual # of startups founded - stacked DID




(b) Annual # of startups founded - imputation DID

This figure shows the dynamic DID effect of fraud charge on entrepreneurs' subsequent founding activities, estimated through a matched DID. The sample is at the person-year level. Panel (a) estimates a stacked DID following Cengiz et al. (2019), with fixed effects for founder, stack × event year, and stack × calendar year. Panel (b) uses the imputation-based DID from Borusyak et al. (2024). The dependent variable is the number of new startups founded by a person in a year. For each treated (fraud) founder in the year before charge, we match him/her to 3 control founders, based on the same number of past founded startups, tenure at current startup, gender, immigrant indicator (US or foreign BA degree), top school indicator, sector of current startup, and closest in age. Standard errors are clustered at the person level.

Table 1: Summary Statistics on Fraud Cases

Panel A. Case Count by Source

| Case count | SEC | DOJ | WestLaw | SCAC | News |
|---|---|---|---|---|---|
| SEC | 133 | 39 | 17 | 15 | 43 |
| DOJ | | 91 | 11 | 7 | 40 |
| WestLaw | | | 207 | 30 | 37 |
| SCAC | | | | 224 | 130 |
| News | | | | | 550 |
| All excluding news: 550 | | | | | |
| All including news: 907 | | | | | |

Panel B. SEC Cases

| | Obs | Mean | Std Dev | P5 | P50 | P95 |
|---|---|---|---|---|---|---|
| Fraud duration (months) | 126 | 28.778 | 24.230 | 1.000 | 24.000 | 72.000 |
| Fined amount ($k) | 58 | 11410.5 | 37693.8 | 35.0 | 1131.9 | 65000.0 |
| Judgement: fine | 100 | 0.580 | 0.496 | 0.000 | 1.000 | 1.000 |
| Judgement: prison | 100 | 0.040 | 0.197 | 0.000 | 0.000 | 0.000 |
| Judgement: disgorgement | 100 | 0.680 | 0.469 | 0.000 | 1.000 | 1.000 |
| Judgement: ban | 100 | 0.700 | 0.461 | 0.000 | 1.000 | 1.000 |
| Misrepresented: financials | 100 | 0.720 | 0.451 | 0.000 | 1.000 | 1.000 |
| Misrepresented: product | 100 | 0.280 | 0.451 | 0.000 | 0.000 | 1.000 |
| Misrepresented: use of funds | 100 | 0.210 | 0.409 | 0.000 | 0.000 | 1.000 |
| Victim: investor | 100 | 0.880 | 0.327 | 0.000 | 1.000 | 1.000 |
| Victim: government | 100 | 0.040 | 0.197 | 0.000 | 0.000 | 0.000 |
| Victim: public | 100 | 0.080 | 0.273 | 0.000 | 0.000 | 1.000 |
| Victim: other | 100 | 0.070 | 0.256 | 0.000 | 0.000 | 1.000 |

Panel C. DOJ Cases

| | Obs | Mean | Std Dev | P5 | P50 | P95 |
|---|---|---|---|---|---|---|
| Fraud duration (months) | 78 | 40.603 | 40.062 | 3.000 | 33.000 | 92.000 |
| Fraud type: bank | 91 | 0.077 | 0.268 | 0.000 | 0.000 | 1.000 |
| Fraud type: wire | 91 | 0.846 | 0.363 | 0.000 | 1.000 | 1.000 |
| Fraud type: mail | 91 | 0.055 | 0.229 | 0.000 | 0.000 | 1.000 |
| Fraud type: corporate | 91 | 0.099 | 0.300 | 0.000 | 0.000 | 1.000 |
| Fraud type: federal | 91 | 0.132 | 0.340 | 0.000 | 0.000 | 1.000 |
| Fraud type: securities | 91 | 0.330 | 0.473 | 0.000 | 0.000 | 1.000 |
| Fraud type: tax | 91 | 0.022 | 0.147 | 0.000 | 0.000 | 0.000 |
| Fraud type: other | 91 | 0.022 | 0.147 | 0.000 | 0.000 | 0.000 |
| Sentence: months of prison | 39 | 84.179 | 90.675 | 12.000 | 51.000 | 340.000 |
| Sentence: fined amount ($k) | 47 | 18702.9 | 81012.3 | 50.0 | 1750.0 | 34370.0 |
| Sentence: fine | 91 | 0.516 | 0.502 | 0.000 | 1.000 | 1.000 |
| Sentence: prison | 91 | 0.429 | 0.498 | 0.000 | 0.000 | 1.000 |
| Sentence: forfeiture | 61 | 0.148 | 0.358 | 0.000 | 0.000 | 1.000 |
| Sentence: supervised release | 61 | 0.344 | 0.479 | 0.000 | 0.000 | 1.000 |
| Sentence: community service | 61 | 0.033 | 0.180 | 0.000 | 0.000 | 0.000 |
| Victim: investor | 91 | 0.659 | 0.477 | 0.000 | 1.000 | 1.000 |
| Victim: government | 91 | 0.187 | 0.392 | 0.000 | 0.000 | 1.000 |
| Victim: public | 91 | 0.220 | 0.416 | 0.000 | 0.000 | 1.000 |
| Victim: other | 91 | 0.110 | 0.314 | 0.000 | 0.000 | 1.000 |

Table 1: Summary Statistics on Fraud Cases (Continued)

Panel D. WestLaw Cases

| | Obs | Mean | Std Dev | P5 | P50 | P95 |
|---|---|---|---|---|---|---|
| Fraud duration (months) | 206 | 27.880 | 31.822 | 0.467 | 14.600 | 94.867 |
| Fraud type: fiduciary duty | 207 | 0.425 | 0.496 | 0.000 | 0.000 | 1.000 |
| Fraud type: wire | 207 | 0.034 | 0.181 | 0.000 | 0.000 | 0.000 |
| Fraud type: corporate | 207 | 0.435 | 0.497 | 0.000 | 0.000 | 1.000 |
| Fraud type: securities | 207 | 0.353 | 0.479 | 0.000 | 0.000 | 1.000 |
| Fraud type: other | 207 | 0.164 | 0.371 | 0.000 | 0.000 | 1.000 |
| Misrepresented: financials | 207 | 0.527 | 0.501 | 0.000 | 1.000 | 1.000 |
| Misrepresented: product | 207 | 0.295 | 0.457 | 0.000 | 0.000 | 1.000 |
| Misrepresented: use of funds | 207 | 0.082 | 0.275 | 0.000 | 0.000 | 1.000 |
| Misrepresented: other | 207 | 0.097 | 0.296 | 0.000 | 0.000 | 1.000 |
| Victim: investor | 207 | 0.589 | 0.493 | 0.000 | 1.000 | 1.000 |
| Victim: public | 207 | 0.159 | 0.367 | 0.000 | 0.000 | 1.000 |
| Victim: corporate | 207 | 0.005 | 0.070 | 0.000 | 0.000 | 0.000 |
| Victim: other | 207 | 0.169 | 0.376 | 0.000 | 0.000 | 1.000 |

Panel E. Class Action Cases

| | Obs | Mean | Std Dev | P5 | P50 | P95 |
|---|---|---|---|---|---|---|
| Case settled | 224 | 0.540 | 0.499 | 0.000 | 1.000 | 1.000 |
| Settled amount ($k) | 224 | 16142.6 | 67429.8 | 0.000 | 917.0 | 55000.0 |
| 1934 Act Sec9: securities trading | 224 | 0.004 | 0.067 | 0.000 | 0.000 | 0.000 |
| 1934 Act Sec10b: antifraud | 224 | 0.804 | 0.398 | 0.000 | 1.000 | 1.000 |
| 1934 Act Sec14a: proxy solitication fraud | 224 | 0.080 | 0.272 | 0.000 | 0.000 | 1.000 |
| 1934 Act Sec14e: tender offer antifraud | 224 | 0.009 | 0.094 | 0.000 | 0.000 | 0.000 |
| 1934 Act Sec20a: control person liability | 224 | 0.830 | 0.376 | 0.000 | 1.000 | 1.000 |
| 1933 Act Sec11: registration statement | 224 | 0.415 | 0.494 | 0.000 | 0.000 | 1.000 |
| 1933 Act Sec12a2: prospectus & oral commu. | 224 | 0.201 | 0.402 | 0.000 | 0.000 | 1.000 |
| 1933 Act Sec15: control person liability | 224 | 0.406 | 0.492 | 0.000 | 0.000 | 1.000 |

This table presents the summary statistics for our sample of venture fraud cases. Panel A reports the number of cases from various sources. We restrict to US VC-backed startups founded since 2000. For startups that achieved IPO, we restrict to SEC/DOJ/news cases where fraud start date is before IPO, as well as class action cases with class period start date within 2 years of IPO. Panels B, C, D, and E presents case attributes for SEC, DOJ, WestLaw, and class action cases, respectively.

Table 2: Summary Statistics on IPOs and Class Action Sample

Panel A: Sample of IPOs (2002–2024)

| | Obs | Mean | Std Dev | P5 | P50 | P95 |
|---|---|---|---|---|---|---|
| VC-Backed | 4094 | 0.354 | 0.478 | 0.000 | 0.000 | 1.000 |
| Has class action | 4094 | 0.085 | 0.278 | 0.000 | 0.000 | 1.000 |
| Class action start within 1y IPO | 4094 | 0.042 | 0.200 | 0.000 | 0.000 | 0.000 |
| Class action start within 2y IPO | 4094 | 0.054 | 0.227 | 0.000 | 0.000 | 1.000 |
| Class action start within 3y IPO | 4094 | 0.060 | 0.237 | 0.000 | 0.000 | 1.000 |
| Log(Total Assets)$\_IPOt-1$ | 1882 | 4.728 | 2.227 | 0.876 | 4.711 | 8.332 |
| Revenue Growth$_IPOt-1;t$ | 1813 | 103.651 | 494.166 | -6.216 | 17.957 | 405.602 |

Panel B: Sample of non-Dismissed Class Actions (2002–2024)

| | Obs | Mean | Std Dev | P5 | P50 | P95 |
|---|---|---|---|---|---|---|
| VC-Backed | 429 | 0.583 | 0.494 | 0.000 | 1.000 | 1.000 |
| Securities fraud | 429 | 0.790 | 0.408 | 0.000 | 1.000 | 1.000 |
| Dismissed case | 429 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Settled case | 429 | 0.594 | 0.492 | 0.000 | 1.000 | 1.000 |
| Ongoing case | 429 | 0.406 | 0.492 | 0.000 | 0.000 | 1.000 |
| Class action start within 1y IPO | 429 | 0.499 | 0.501 | 0.000 | 0.000 | 1.000 |
| Class action start within 2y IPO | 429 | 0.643 | 0.480 | 0.000 | 1.000 | 1.000 |
| Class action start within 3y IPO | 429 | 0.706 | 0.456 | 0.000 | 1.000 | 1.000 |
| Class action filing within 1y IPO | 429 | 0.205 | 0.404 | 0.000 | 0.000 | 1.000 |
| Class action filing within 2y IPO | 429 | 0.436 | 0.496 | 0.000 | 0.000 | 1.000 |
| Class action filing within 3y IPO | 429 | 0.599 | 0.491 | 0.000 | 1.000 | 1.000 |
| IPO year | 429 | 2013.795 | 5.833 | 2004.000 | 2014.000 | 2021.000 |
| Start class period year | 429 | 2016.247 | 5.292 | 2006.000 | 2018.000 | 2022.000 |
| Filing year | 429 | 2017.410 | 5.349 | 2007.000 | 2019.000 | 2024.000 |
| Class action start - IPO year | 429 | 2.348 | 3.244 | -0.003 | 1.005 | 8.726 |
| Class period start - end year | 429 | 2.531 | 2.148 | 0.146 | 1.924 | 6.854 |
| Class action filing - IPO year | 429 | 3.549 | 3.426 | 0.329 | 2.362 | 9.866 |
| Settlement amount (M$) | 255 | 21.535 | 60.786 | 0.000 | 7.700 | 74.000 |
| Log(Total Assets)$\_IPOt-1$ | 273 | 5.135 | 2.131 | 1.586 | 5.025 | 8.446 |
| Revenue Growth$_IPOt-1;t$ | 267 | 204.234 | 841.873 | -4.010 | 42.190 | 665.281 |

Panel C: Sample of Class Actions that Start within 2 years of IPO

| | Obs | Mean | Std Dev | P5 | P50 | P95 |
|---|---|---|---|---|---|---|
| VC-Backed | 276 | 0.591 | 0.493 | 0.000 | 1.000 | 1.000 |
| Securities fraud | 276 | 0.736 | 0.442 | 0.000 | 1.000 | 1.000 |
| Class period start - end year | 276 | 2.376 | 1.911 | 0.319 | 1.847 | 6.069 |
| Class action filing within 2y IPO | 276 | 0.678 | 0.468 | 0.000 | 1.000 | 1.000 |
| Class action filing within 3y IPO | 276 | 0.906 | 0.293 | 0.000 | 1.000 | 1.000 |
| IPO year | 276 | 2015.283 | 5.607 | 2005.000 | 2017.000 | 2021.000 |
| Filing year | 276 | 2016.935 | 5.737 | 2006.000 | 2019.000 | 2024.000 |
| Start class period year | 276 | 2015.822 | 5.626 | 2005.000 | 2018.000 | 2022.000 |
| Class action filing - IPO year | 276 | 1.609 | 1.021 | 0.205 | 1.432 | 3.455 |
| Settlement amount (M$) | 166 | 23.364 | 72.415 | 1.000 | 7.500 | 71.000 |
| Log(Total Assets)$\_IPOt-1$ | 181 | 5.411 | 2.087 | 1.762 | 5.559 | 8.446 |
| Revenue Growth$_IPOt-1;t$ | 177 | 287.265 | 1015.046 | -3.321 | 55.000 | 1112.565 |

This table reports the summary statistics for the IPO and class action sample used in Column 1 of Table 4, Panel A.

Table 3: Summary Statistics on Prediction Panel

| | Obs | Mean | Std Dev | P5 | P50 | P95 |
|---|---|---|---|---|---|---|
| $\mathbb{1}$(Fraud start) $\times$ 100 | 766,834 | 0.058 | 2.414 | 0.000 | 0.000 | 0.000 |
| Firm age | 766,834 | 5.982 | 5.094 | 0.000 | 5.000 | 16.000 |
| Ln(valuation) | 153,517 | 3.117 | 1.752 | 0.336 | 3.049 | 6.106 |
| Ln(raised) | 390,895 | 1.269 | 2.134 | -2.303 | 1.366 | 4.619 |
| Board size | 218,793 | 3.365 | 2.159 | 1.000 | 3.000 | 7.000 |
| Board VC controlled | 210,656 | 0.198 | 0.399 | 0.000 | 0.000 | 1.000 |
| Board shared control | 210,656 | 0.305 | 0.461 | 0.000 | 0.000 | 1.000 |
| Investors' converted ownership | 266,500 | 0.457 | 0.207 | 0.135 | 0.450 | 0.803 |
| Founders' payoff convexity | 270,496 | 0.270 | 0.461 | 0.000 | 0.000 | 1.000 |
| Ln(unique investors) | 477,027 | 1.334 | 0.994 | 0.000 | 1.386 | 2.996 |
| Investor_%non-IVC | 531,919 | 0.551 | 0.354 | 0.000 | 0.550 | 1.000 |
| Lead investors' ln(past exits) | 453,274 | 2.157 | 1.567 | 0.000 | 2.043 | 4.876 |
| Team_solo | 503,959 | 0.467 | 0.499 | 0.000 | 0.000 | 1.000 |
| Team_hasserial | 503,959 | 0.214 | 0.410 | 0.000 | 0.000 | 1.000 |
| Team_topschool% | 411,198 | 0.274 | 0.407 | 0.000 | 0.000 | 1.000 |
| Team_avg. age | 311,810 | 42.617 | 10.777 | 27.000 | 41.500 | 62.000 |
| Missing valuation | 766,834 | 0.800 | 0.400 | 0.000 | 1.000 | 1.000 |
| Missing investor exits | 766,834 | 0.409 | 0.492 | 0.000 | 0.000 | 1.000 |
| Missing investor type | 766,834 | 0.306 | 0.461 | 0.000 | 0.000 | 1.000 |
| Missing team info | 766,834 | 0.343 | 0.475 | 0.000 | 0.000 | 1.000 |
| Missing team edu info | 766,834 | 0.517 | 0.500 | 0.000 | 1.000 | 1.000 |
| Missing board info | 766,834 | 0.725 | 0.446 | 0.000 | 1.000 | 1.000 |
| Missing term sheet info | 766,834 | 0.652 | 0.476 | 0.000 | 1.000 | 1.000 |

This table reports the summary statistics for the hazard-style firm panel used in our predictive analysis in Table 5.

## Table 4. Class Actions Against VC- vs Non-VC-Backed Firms

Panel A: Full period (2002–2023)

| Start Class Period within: | 1 year of IPO<br>(1) | 2 years of IPO<br>(2) | 3 years of IPO<br>(3) | 1 year of IPO<br>(4) | 2 years of IPO<br>(5) | 3 years of IPO<br>(6) |
|---|---|---|---|---|---|---|
| VC-backed | 0.0434*** | 0.0481*** | 0.0522*** | 0.0358*** | 0.0478*** | 0.0533*** |
| | (0.006) | (0.007) | (0.006) | (0.010) | (0.012) | (0.012) |
| Log(Total Assets)_$IPOt-1$ | | | | 0.0092*** | 0.0101*** | 0.0102*** |
| | | | | (0.002) | (0.003) | (0.002) |
| Revenue Growth_$IPOt-1;t$ | | | | 0.0000* | 0.0001** | 0.0001** |
| | | | | (0.000) | (0.000) | (0.000) |
| SIC-3 Sector FE × IPO Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.113 | 0.127 | 0.134 | 0.128 | 0.152 | 0.155 |
| N | 4,094 | 4,094 | 4,094 | 1,705 | 1,705 | 1,705 |
| Mean dep. var. | 0.0418 | 0.0545 | 0.0598 | 0.0669 | 0.0815 | 0.0868 |

Panel B: Post Dot-com Bubble Period (2002–2012)

| Start Class Period within: | 1 year of IPO<br>(1) | 2 years of IPO<br>(2) | 3 years of IPO<br>(3) | 1 year of IPO<br>(4) | 2 years of IPO<br>(5) | 3 years of IPO<br>(6) |
|---|---|---|---|---|---|---|
| VC-backed | 0.0252 | 0.0368 | 0.0423* | 0.0293 | 0.0377 | 0.0418 |
| | (0.017) | (0.023) | (0.023) | (0.023) | (0.029) | (0.027) |
| Log(Total Assets)_$IPOt-1$ | | | | 0.0035 | -0.0034 | -0.0025 |
| | | | | (0.004) | (0.006) | (0.006) |
| Revenue Growth_$IPOt-1;t$ | | | | 0.0000 | 0.0000 | 0.0000 |
| | | | | (0.000) | (0.000) | (0.000) |
| SIC-3 Sector FE × IPO Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.168 | 0.176 | 0.182 | 0.189 | 0.223 | 0.229 |
| N | 1,462 | 1,462 | 1,462 | 500 | 500 | 500 |
| Mean dep. var. | 0.0274 | 0.0369 | 0.0404 | 0.0460 | 0.0640 | 0.0680 |

Table 4. Class Actions Against VC- vs Non-VC-Backed Firms (Continued)

Panel C: Recent Period (2013–2023)

| Start Class Period within: | 1 year of IPO (1) | 2 years of IPO (2) | 3 years of IPO (3) | 1 year of IPO (4) | 2 years of IPO (5) | 3 years of IPO (6) |
|---|---|---|---|---|---|---|
| VC-backed | 0.0490*** | 0.0515*** | 0.0552*** | 0.0370*** | 0.0494*** | 0.0553*** |
| | (0.009) | (0.011) | (0.011) | (0.013) | (0.016) | (0.017) |
| Log(Total Assets)_$IPOt-1$ | | | | 0.0103*** | 0.0128*** | 0.0127*** |
| | | | | (0.003) | (0.003) | (0.003) |
| Revenue Growth_$IPOt-1;t$ | | | | 0.0001** | 0.0001*** | 0.0001*** |
| | | | | (0.000) | (0.000) | (0.000) |
| SIC-3 Sector FE × IPO Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.093 | 0.107 | 0.114 | 0.110 | 0.132 | 0.134 |
| N | 2,632 | 2,632 | 2,632 | 1,205 | 1,205 | 1,205 |
| Mean dep. var. | 0.0498 | 0.0642 | 0.0707 | 0.0755 | 0.0888 | 0.0946 |

This table reports OLS estimates of the relationship between (formerly) VC-backed companies and the likelihood of a securities class action lawsuit within one to three years of the firm's IPO. We define the timing of litigation relative to the IPO using the start date of the class period. The sample includes U.S. publicly listed firms that went public between 2002 and 2023. Class actions are identified using the SCAC database and are limited to those filed between 2002 and 2024; cases that were ultimately dismissed are excluded. VC-backed status is determined using PitchBook deal data and Jay Ritterâs IPO database (for older IPOs). Columns (4)â(6) include controls for firm size (log of total assets) and revenue growth, measured using Compustat data from the year prior to the IPO or the IPO year. Note that merging with Compustat reduces the sample size by approximately half. All regressions include SIC-4 industry fixed effects based on CRSP classifications. Standard errors are clustered at the SIC-4 level and are reported in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 5: Panel Prediction of Fraud Among VC-Backed Firms

| Dependent Variable: | $\mathbb{1}$(Fraud start) × 100 | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Firm age | -0.000 | -0.001* | -0.001** | -0.001 | -0.001 | 0.000 |
| | (0.001) | (0.000) | (0.001) | (0.001) | (0.001) | (.) |
| Ln(valuation) | 0.176*** | 0.174*** | 0.175*** | 0.176*** | 0.175*** | 0.168*** |
| | (0.021) | (0.020) | (0.020) | (0.021) | (0.021) | (0.021) |
| Ln(raised) | 0.013*** | 0.010*** | 0.010*** | 0.012*** | 0.012*** | 0.013*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.004) |
| Board size | 0.019*** | | | 0.018*** | 0.020*** | 0.033*** |
| | (0.005) | | | (0.005) | (0.005) | (0.007) |
| Board VC controlled | -0.097*** | | | -0.102*** | -0.104*** | -0.167*** |
| | (0.026) | | | (0.027) | (0.026) | (0.034) |
| Board shared control | -0.017 | | | -0.018 | -0.018 | -0.013 |
| | (0.016) | | | (0.016) | (0.017) | (0.026) |
| Investors' converted ownership | -0.093*** | | | -0.115*** | -0.117*** | -0.165*** |
| | (0.032) | | | (0.034) | (0.035) | (0.053) |
| Founders' payoff convexity | 0.019* | | | 0.023** | 0.024** | 0.055** |
| | (0.011) | | | (0.011) | (0.012) | (0.028) |
| Ln(unique investors) | | 0.018*** | | 0.030*** | 0.029*** | 0.088*** |
| | | (0.004) | | (0.005) | (0.005) | (0.011) |
| Investor_%non-IVC | | 0.038*** | | 0.031*** | 0.031*** | 0.052** |
| | | (0.010) | | (0.010) | (0.010) | (0.022) |
| Lead investors' ln(past exits) | | -0.008*** | | -0.006** | -0.005** | -0.013** |
| | | (0.003) | | (0.002) | (0.002) | (0.005) |
| Team_solo | | | 0.012* | 0.016** | 0.015* | 0.000 |
| | | | (0.007) | (0.007) | (0.008) | (.) |
| Team_hasserial | | | 0.012 | 0.014 | 0.011 | 0.000 |
| | | | (0.009) | (0.009) | (0.009) | (.) |
| Team_topschool% | | | 0.008 | 0.010 | 0.006 | 0.000 |
| | | | (0.010) | (0.010) | (0.010) | (.) |
| Team_avg. age | | | -0.001*** | -0.001** | -0.001*** | 0.001 |
| | | | (0.000) | (0.000) | (0.000) | (0.001) |
| FE: ind-year | Yes | Yes | Yes | Yes | Yes | Yes |
| FE: ind-1st round year | No | No | No | No | Yes | No |
| FE: ind-year, firm | No | No | No | No | No | Yes |
| Missing value indicators | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 766834 | 766834 | 766834 | 766834 | 766834 | 766834 |
| $R^2$ | 0.007 | 0.007 | 0.007 | 0.007 | 0.011 | 0.101 |
| Outcome Mean | 0.058 | 0.058 | 0.058 | 0.058 | 0.058 | 0.053 |

The sample is a hazard-style panel of firm-years for US VC-backed firms founded since 2000. For non-fraud firms, the panel includes all years from firm founding to the earlier of 2024 or firm exit. For fraud firms the panel goes from firm founding year to the fraud start year. The dependent variable is 100 in the fraud start year for fraud firms and is 0 otherwise. We restrict to SEC/DOJ/WestLaw cases where fraud start year is no later than IPO year, and class action cases with class date within 2 years of IPO. The sample contains 447 fraud cases. Missing value indicators are included but are not reported for brevity (their coefficients are in Table A.4). Standard errors are clustered at the industry level and are reported in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 6: Cross-Sectional Prediction of Class Action Suit Among VC-Backed IPOs

| Dependent Variable: | $\mathbb{1}$(Fraud start) × 100 | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Firm age | -0.222* | -0.236* | -0.175 | -0.232** |
| | (0.114) | (0.122) | (0.119) | (0.116) |
| Ln(valuation) | 4.532*** | 4.729*** | 4.860*** | 4.401*** |
| | (0.688) | (0.680) | (0.695) | (0.653) |
| Ln(raised) | 1.525*** | 0.985** | 1.228*** | 1.147** |
| | (0.398) | (0.474) | (0.343) | (0.476) |
| Board size | 0.219 | | | 0.169 |
| | (0.370) | | | (0.353) |
| Board VC controlled | -5.011* | | | -5.154* |
| | (2.718) | | | (2.701) |
| Board shared control | -2.992 | | | -2.881 |
| | (1.812) | | | (1.756) |
| DualClassIPO | 6.699** | | | 6.689** |
| | (2.946) | | | (2.948) |
| Investors' converted ownership | -5.866 | | | -5.969 |
| | (4.549) | | | (4.539) |
| Founders' payoff convexity | 2.061 | | | 1.985 |
| | (1.322) | | | (1.326) |
| Ln(unique investors) | | 0.800 | | 1.742** |
| | | (0.759) | | (0.877) |
| Investor_%non-IVC | | 3.909* | | 3.267 |
| | | (2.345) | | (2.455) |
| Lead investors' ln(past exits) | | -0.355 | | -0.178 |
| | | (0.389) | | (0.384) |
| Team_solo | | | 1.748 | 2.020* |
| | | | (1.123) | (1.097) |
| Team_hasserial | | | 1.248 | 1.490 |
| | | | (1.690) | (1.691) |
| Team_topschool% | | | 0.389 | 0.564 |
| | | | (1.976) | (1.999) |
| Team_avg. age | | | -0.046 | -0.042 |
| | | | (0.036) | (0.033) |
| FE: ind, yr | Yes | Yes | Yes | Yes |
| Missing value indicators | Yes | Yes | Yes | Yes |
| N | 2475 | 2475 | 2475 | 2475 |
| $R^2$ | 0.088 | 0.084 | 0.083 | 0.089 |
| Outcome Mean | 8.252 | 8.252 | 8.252 | 8.252 |

The sample is a cross-section of US VC-backed firms founded since 2000 that went public. The dependent variable is 100 for fraud firms involved in class actions and is 0 otherwise. We restrict to class action cases with class date within 2 years of IPO. The sample contains 204 fraud cases from class action suits. Missing value indicators are included but are not reported for brevity. Standard errors are clustered at the industry level and are reported in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 7: Panel Prediction of Fraud Among VC-Backed Firms: Including News-Based Cases

| Dependent Variable: | $\mathbb{1}$(Fraud start) × 100 | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Firm age | -0.000 | -0.001 | -0.002* | -0.001 | -0.001 | 0.000 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (.) |
| Ln(valuation) | 0.249*** | 0.248*** | 0.252*** | 0.249*** | 0.247*** | 0.233*** |
| | (0.031) | (0.031) | (0.031) | (0.031) | (0.031) | (0.032) |
| Ln(raised) | 0.019*** | 0.015*** | 0.018*** | 0.016*** | 0.017*** | 0.021*** |
| | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) | (0.005) |
| Board size | 0.044*** | | | 0.043*** | 0.044*** | 0.070*** |
| | (0.009) | | | (0.009) | (0.009) | (0.014) |
| Board VC controlled | -0.166*** | | | -0.176*** | -0.180*** | -0.307*** |
| | (0.038) | | | (0.039) | (0.039) | (0.063) |
| Board shared control | -0.052*** | | | -0.054*** | -0.059*** | -0.084** |
| | (0.019) | | | (0.019) | (0.021) | (0.034) |
| Investors' converted ownership | -0.056 | | | -0.094** | -0.088** | -0.132* |
| | (0.039) | | | (0.040) | (0.042) | (0.073) |
| Founders' payoff convexity | 0.024* | | | 0.030** | 0.033** | 0.090** |
| | (0.014) | | | (0.014) | (0.015) | (0.035) |
| Ln(unique investors) | | 0.032*** | | 0.047*** | 0.046*** | 0.143*** |
| | | (0.006) | | (0.008) | (0.008) | (0.020) |
| Investor_%non-IVC | | 0.041*** | | 0.033*** | 0.035*** | 0.054** |
| | | (0.011) | | (0.011) | (0.011) | (0.023) |
| Lead investors' ln(past exits) | | -0.010*** | | -0.007*** | -0.006** | -0.020*** |
| | | (0.003) | | (0.003) | (0.003) | (0.006) |
| Team_solo | | | 0.028*** | 0.036*** | 0.034*** | 0.000 |
| | | | (0.011) | (0.011) | (0.012) | (.) |
| Team_hasserial | | | 0.012 | 0.014 | 0.012 | 0.000 |
| | | | (0.011) | (0.011) | (0.012) | (.) |
| Team_topschool% | | | 0.010 | 0.012 | 0.006 | 0.000 |
| | | | (0.013) | (0.013) | (0.013) | (.) |
| Team_avg. age | | | -0.001*** | -0.001*** | -0.001*** | 0.000 |
| | | | (0.000) | (0.000) | (0.000) | (0.002) |
| FE: ind-year | Yes | Yes | Yes | Yes | Yes | Yes |
| FE: ind-1st round year | No | No | No | No | Yes | No |
| FE: ind-year, firm | No | No | No | No | No | Yes |
| Missing value indicators | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 768453 | 768453 | 768453 | 768453 | 768453 | 768453 |
| $R^2$ | 0.009 | 0.009 | 0.008 | 0.009 | 0.011 | 0.092 |
| Outcome Mean | 0.083 | 0.083 | 0.083 | 0.083 | 0.082 | 0.077 |

The sample is a hazard-style panel of firm-years for US VC-backed firms founded since 2000. For non-fraud firms, the panel includes all years from firm founding to the earlier of 2024 or firm exit. For fraud firms the panel goes from firm founding year to the fraud start year. The dependent variable is 100 in the fraud start year for fraud firms and is 0 otherwise. We restrict to SEC/DOJ/WestLaw/news cases where fraud start year is no later than IPO year, and class action cases with class date within 2 years of IPO. The sample contains 634 fraud cases. Missing value indicators are included but are not reported for brevity. Standard errors are clustered at the industry level and are reported in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 8: VC Market Condition at Initial Round and Future Fraud

| Dependent Variable: | $\mathbb{1}$(Future Fraud) $\times$ 100 | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Ind-yr avg. valuation multiple | 0.078** | 0.078** | 0.082** |
| | (0.034) | (0.034) | (0.036) |
| Ind-yr frac. founder-controlled board | 0.845** | 0.851** | 0.910** |
| | (0.389) | (0.389) | (0.420) |
| Firm age | | 0.020* | 0.001 |
| | | (0.011) | (0.014) |
| Ln(valuation) | | | 0.332*** |
| | | | (0.071) |
| Ln(raised) | | | 0.138*** |
| | | | (0.042) |
| Board size | | | -0.005 |
| | | | (0.061) |
| Board VC controlled | | | 0.216 |
| | | | (0.226) |
| Board shared control | | | 0.153 |
| | | | (0.146) |
| Missing valuation | | | 0.533*** |
| | | | (0.124) |
| Missing board info | | | 0.089 |
| | | | (0.155) |
| FE: ind | Yes | Yes | Yes |
| FE: state-yr | Yes | Yes | Yes |
| N | 72817 | 72817 | 72817 |
| $R^2$ | 0.005 | 0.005 | 0.008 |
| Outcome Mean | 0.595 | 0.595 | 0.595 |

This table shows that VC market condition at firms' initial VC round predicts their future fraud. The sample is at the firm level, covering initial VC rounds between 2000 and 2023. The dependent variable is a dummy indicating whether the firm is ever involved in fraud, multiplied by 100. We employ two proxies for VC market condition at the industry-initial-funding-year level: average initial round valuation multiple and average fraction of founder-controlled boards at initial round. Firm-level controls are firm age, post-money valuation, total raised amount, board size, indicator for VC-controlled board, indicator for shared-control board, and missing valuation indicators for valuation and board info, all measured at initial VC round. All columns include industry fixed effects and state-year fixed effects. Industry is defined as primary industry group in PitchBook (42 industries). Standard errors are double clustered by industry and year and are reported in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

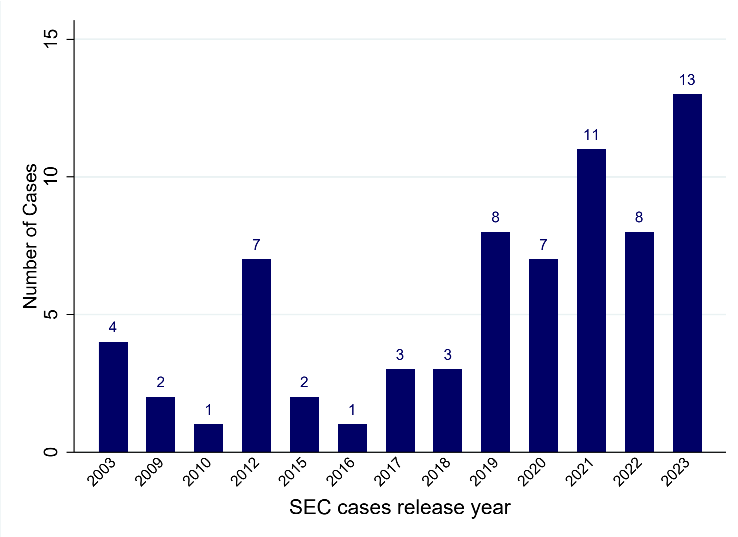
Table 9: Future Founding of VC-Backed Startups by Fraudulent Entrepreneurs

| | (1) Startups founded | (2) Cumulative startups founded | (3) Startups founded | (4) Cumulative startups founded |
|---|---|---|---|---|
| Treat × Post | 0.007** | 0.008 | 0.006 | 0.009 |
| | (0.003) | (0.011) | (0.006) | (0.016) |
| Model | Stacked DID | | Imputation-Based DID | |
| FEs | person, stack-event-year, stack-year | | person, event-year, year | |
| N | 25450 | 25450 | 24048 | 24048 |
| R-sq | 0.659 | 0.935 | - | - |
| Outcome mean | 0.038 | 1.294 | 0.038 | 1.294 |

This table reports the DID estimate for the effect of fraud revelation (i.e., initial charge) on founders' future founding of VC-backed startups. The estimates correspond to the event study plots in Figure 6. The sample is at the person-year level, focusing on a fixed event window from 4 years before to 6 years after initial charge. Columns (1) and (2) estimate a stacked DID following Cengiz et al. (2019), with fixed effects for founder, stack × event year, and stack × calendar year. Columns (3) and (4) use the imputation-based DID from Borusyak et al. (2024). The dependent variable in Columns and (1) and (3) is the number of new startups founded by a person in a year. The dependent variable in Columns (2) and (4) is the cumulative number of startups a person has founded as of a year. For each treated (fraud) founder in the year before charge, we match him/her to 3 control founders, based on the same number of past founded startups, tenure at current startup, gender, immigrant indicator (US or foreign BA degree), top school indicator, sector of current startup, and closest in age. Standard errors are clustered at the person level. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

# Appendix

Figure A.1: SEC Cases



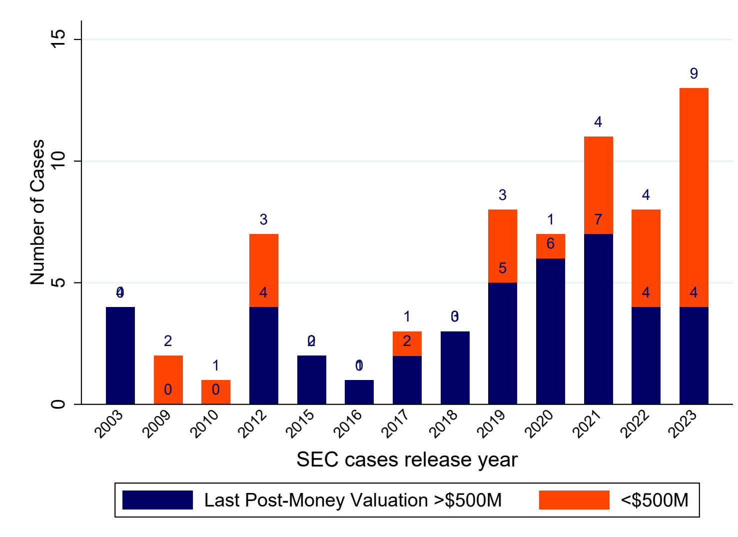
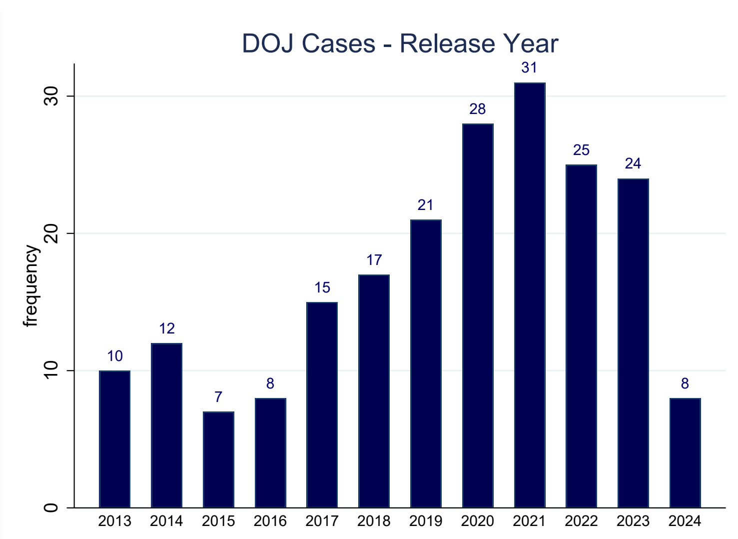
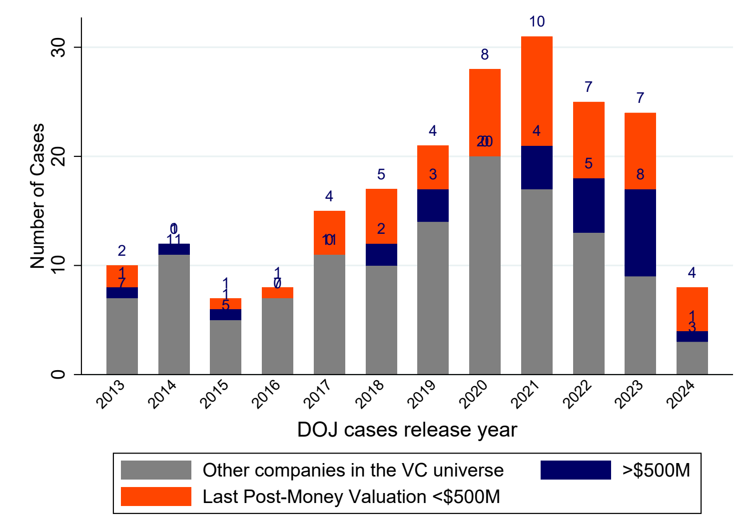




Figure A.2: DOJ Cases



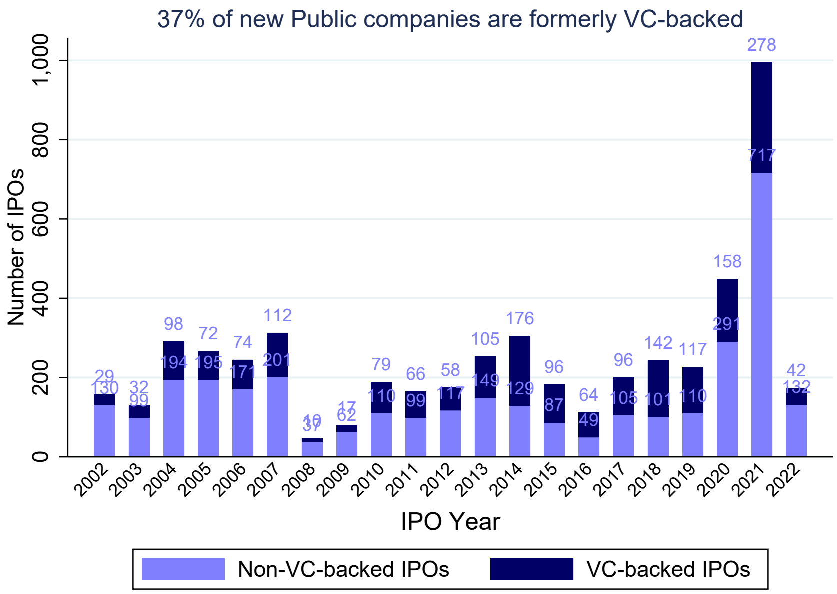

(a) DOJ Cases in the PB VC universe




(b) DOJ Cases involving VC-backed Companies with VC valuation

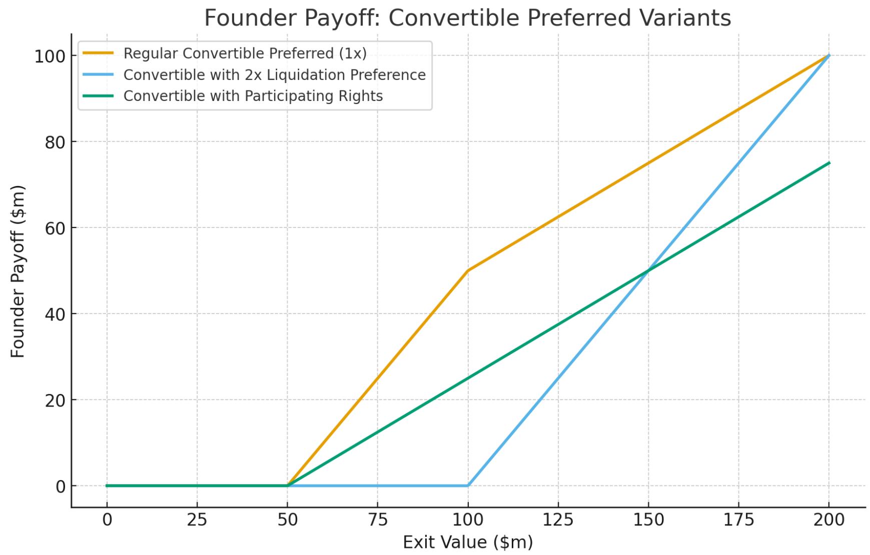
This figure plots the trends in the number of SEC and DOJ cases against VC-backed firms.

Figure A.3: Number of VC-Backed and Non-VC-Backed IPOs




This figure plots the number of VC-backed and non-VC-backed firms by IPO year.

Figure A.4: Founder' Payoff When Investor Has High Liquidation Preference or Participating Rights




This figure plots founder's payoff from common shares when VCs, who hold convertible preferred shares, have high liquidation preference or participation rights. The orange line represents the baseline case of covertible preferred without participating rights but with 1x liquidation preference. The blue line represents high liquidation preference with 2x multiple but without participating rights. The green line represents participating preferred with 1x liquidation preference.

Table A.1: Top Class Actions Against VC-Backed Firms Filed within Two Years of IPO, Ranked by Settlement Amount

| **Company** | **IPO Date** | **Key Allegations (IPO/Pre-IPO)** | **Class Period / Case Filing** | **Settlement Outcome** |
|---|---|---|---|---|
| LendingClub Corp. | Dec 11, 2014 | Misrepresenting loan quality and internal controls | Class period: Dec 11, 2014 – May 2016; Case filed 2016 (*In re LendingClub Sec. Litig.*, N.D. Cal.) | $125 million (2019) |
| Zuora, Inc. | Apr 12, 2018 | Misrepresenting performance of the softwares; Overstated consumer demand | Class period: Apr 12, 2018 – May 30, 2019; Case filed 2019 (*Roberts v. Zuora, Inc.*, N.D. Cal.) | $75.5 million (2023) |
| Groupon, Inc. | Nov 4, 2011 | Overstated revenue by using improper accounting methods (e.g., gross" rather than net" revenue) | Class period: Nov 4, 2011 – Mar 30, 2012; Case filed 2012 (*In re Groupon, Inc. Sec. Litig.*, N.D. Ill.) | $45 million (2016) |
| GoHealth, Inc. | Jul 15, 2020 | IPO registration statement omitted elevated churn, customer retention problems, unfavorable revenue-sharing, and deteriorating product mix | Class period: Jul 14, 2020 – Jan 10, 2021; Case filed Sep 2020 (*In re GoHealth, Inc. Sec. Litig.*, N.D. Ill.) | $29.25 million (2024) |
| Lyft, Inc. | Mar 29, 2019 | Misrepresenting the frequency of sexual assaults involving drivers and passengers. Misrepresenting the risks of treating drivers as independent contractors. | Class period: Mar 29, 2019 – post-IPO disclosures; Case filed Apr 2019 (*In re Lyft, Inc. Sec. Litig.*, N.D. Cal.) | $25 million (2022) |
| Zynga, Inc. | Dec 16, 2011 | Withholding key metrics: declining bookings and user engagement in its core games; Overstated growth, especially on Facebook's platform. | Class period: Dec 16, 2011 – Jul 25, 2012; Case filed 2012 (*In re Zynga, Inc. Sec. Litig.*, N.D. Cal.) | $23 million (2015) |
| Bumble, Inc. | Sep 10, 2021 (SPO) | SPO registration statement failed to disclose declining paying users, negative impact of price increases, and Badoo payment transition issues | Class period: Sep 10, 2021 – Jan 24, 2022; Case filed Jan 2022 (*In re Bumble, Inc. Sec. Litig.*, S.D.N.Y.) | $18 million (2023) |
| Fitbit, Inc. | Jun 18, 2015 | Misrepresenting the accuracy of their heart-rate technology. Fitbit knew the problem based on internal testing and consumer complaints prior IPO; Inflated demand and revenues | Class period: Jun 18, 2015 – Jan 6, 2016; Case filed Jan 2016 (*In re Fitbit, Inc. Sec. Litig.*, N.D. Cal.) | $15 million (2020) |
| Marrone Bio Innovations, Inc. | Aug 2, 2013 | Misstated revenues and product viability; accounting fraud inflated growth | Class period: Aug 2, 2013 – Sep 2, 2014; Case filed 2014 (*In re Marrone Bio Innovations, Inc. Sec. Litig.*, E.D. Cal.) | $12 million (2016) |

Table A.2: Major Enforcement Cases from DOJ and SEC for VC-Backed Private Firms, Ranked by Prison Length and Fine Amount

**Panel A. Top DOJ Cases**

| Company | Year | Key allegations (concise) | Prison (month) | Other Sentence | Major VC backers |
|---|---|---|---|---|---|
| FTX | 2024 | Misappropriation of customer funds; securities/wire fraud | 300 | $11.02 Billion in forfeiture + $12.7 Billion to compensate victims | Sequoia; Paradigm; Temasek; SoftBank; Tiger Global; OTPP |
| Theranos | 2022 | Misled investors and patients; blood-testing technology did not work as claimed | 290 | $54 Million in restitution | High-net-worth/family offices (e.g., Murdoch, Walton, DeVos) |
| Slync.io | 2024 | Defrauded investors; misappropriated $25M+ from company; wire fraud and money laundering | 240 | $65 Million in restitution | Goldman Sachs Growth; ACME Ventures; Blumberg Capital; 235 Capital Partners; Correlation Ventures; Gaingels |
| Sanovas, Inc. | 2020 | Siphoned $2.6M+ from company; wire fraud & money laundering; used funds to buy $2.5M home; false statements/obstruction | 135 | 3 years supervised release | Undisclosed Investor |
| Bitwise Industries | 2024 | Defrauded investors and lenders using fabricated financials; wire-fraud conspiracy; losses >$100M | 132 | $114 Million in restitution | 529 Ventures; Cap Table Coalition; Gaingels; Ingeborg Investments |

**Panel B. Top SEC Cases**

| Company | Year | Key allegations | Fine Amount (Thousands) | Major VC backers |
|---|---|---|---|---|
| Terraform Labs | 2024 | Securities fraud tied to UST/LUNA collapse (civil penalties) | 420000 | Galaxy Digital; Coinbase Ventures; Pantera; Hashed |
| Luckin Coffee | 2020 | Fabricated revenue/expenses; accounting fraud | 180000 | Centurium Capital; Joy Capital; GIC; CICC |
| Nikola | 2021 | Exaggerated technology/products and business prospects (post-SPAC) | 125000 | (Institutional/PIPE heavy) ValueAct; Fidelity |
| Robinhood | 2020 | Misleading customers (payment for order flow disclosures / best execution) | 65000 | Sequoia; NEA; Ribbit; DST; Thrive |
| Telegram Group Inc. | 2020 | Unregistered offering of âGramâ digital tokens (Section 5 violations); court-ordered return of investor funds | 18500 | Manta Ray Ventures, Golden Falcon Capital, 3e Capital Group |

## Table A.3. Robustness Checks: Class Actions of VC- vs Non-VC-Backed Firms

Panel A: Class Action Filings (2002–2023)

| Filing within: | 1 year of IPO (1) | 2 years of IPO (2) | 3 years of IPO (3) | 1 year of IPO (4) | 2 years of IPO (5) | 3 years of IPO (6) |
|---|---|---|---|---|---|---|
| VC-backed | 0.0236*** | 0.0374*** | 0.0445*** | 0.0164** | 0.0378*** | 0.0443*** |
| | (0.005) | (0.006) | (0.006) | (0.007) | (0.009) | (0.012) |
| Log(Total Assets)_$IPOt-1$ | | | | 0.0054*** | 0.0077*** | 0.0103*** |
| | | | | (0.001) | (0.002) | (0.002) |
| Revenue Growth_$IPOt-1;t$ | | | | 0.0000** | 0.0001** | 0.0001** |
| | | | | (0.000) | (0.000) | (0.000) |
| SIC-3 Sector FE × IPO Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.114 | 0.106 | 0.119 | 0.126 | 0.128 | 0.138 |
| N | 4,094 | 4,094 | 4,094 | 1,705 | 1,705 | 1,705 |
| Mean dep. var. | 0.0181 | 0.0366 | 0.0515 | 0.0293 | 0.0540 | 0.0762 |

Panel B: Including Dismissed Class Actions

| Start Class Period within: | 1 year of IPO (1) | 2 years of IPO (2) | 3 years of IPO (3) | 1 year of IPO (4) | 2 years of IPO (5) | 3 years of IPO (6) |
|---|---|---|---|---|---|---|
| VC-backed | 0.0487*** | 0.0675*** | 0.0749*** | 0.0443*** | 0.0727*** | 0.0781*** |
| | (0.008) | (0.009) | (0.009) | (0.011) | (0.016) | (0.017) |
| Log(Total Assets)_$IPOt-1$ | | | | 0.0111*** | 0.0114*** | 0.0112*** |
| | | | | (0.002) | (0.002) | (0.003) |
| Revenue Growth_$IPOt-1;t$ | | | | 0.0000* | 0.0000* | 0.0000** |
| | | | | (0.000) | (0.000) | (0.000) |
| SIC-3 Sector FE × IPO Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.125 | 0.143 | 0.152 | 0.145 | 0.153 | 0.157 |
| N | 4,535 | 4,535 | 4,535 | 2,001 | 2,001 | 2,001 |
| Mean dep. var. | 0.0573 | 0.0820 | 0.0964 | 0.0865 | 0.1209 | 0.1414 |

Panel C: Only non-dismissed class actions with settlement amound ≥$3M

| Start Class Period within: | 1 year of IPO (1) | 2 years of IPO (2) | 3 years of IPO (3) | 1 year of IPO (4) | 2 years of IPO (5) | 3 years of IPO (6) |
|---|---|---|---|---|---|---|
| VC-backed | 0.0132* | 0.0172*** | 0.0188*** | 0.0117 | 0.0202** | 0.0220** |
| | (0.007) | (0.006) | (0.005) | (0.010) | (0.009) | (0.009) |
| Log(Total Assets)_$IPOt-1$ | | | | 0.0057*** | 0.0058*** | 0.0063*** |
| | | | | (0.001) | (0.001) | (0.002) |
| Revenue Growth_$IPOt-1;t$ | | | | 0.0000 | 0.0000 | 0.0000 |
| | | | | (0.000) | (0.000) | (0.000) |
| SIC-3 Sector FE × IPO Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.112 | 0.136 | 0.143 | 0.135 | 0.163 | 0.168 |
| N | 3,876 | 3,876 | 3,876 | 1,572 | 1,572 | 1,572 |
| Mean dep. var. | 0.0175 | 0.0245 | 0.0263 | 0.0293 | 0.0388 | 0.0407 |

Panel D: Non-dismissed class actions filed before 2019

| Start Class Period within: | 1 year of IPO (1) | 2 years of IPO (2) | 3 years of IPO (3) | 1 year of IPO (4) | 2 years of IPO (5) | 3 years of IPO (6) |
|---|---|---|---|---|---|---|
| VC-backed | 0.0265*** | 0.0386*** | 0.0437*** | 0.0281** | 0.0433*** | 0.0477*** |
| | (0.009) | (0.010) | (0.010) | (0.012) | (0.014) | (0.014) |
| Log(Total Assets)_$IPOt-1$ | | | | 0.0085*** | 0.0088*** | 0.0097*** |
| | | | | (0.002) | (0.002) | (0.003) |
| Revenue Growth_$IPOt-1;t$ | | | | 0.0000 | 0.0000 | 0.0001 |
| | | | | (0.000) | (0.000) | (0.000) |
| SIC-3 Sector FE × IPO Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.133 | 0.148 | 0.160 | 0.136 | 0.165 | 0.181 |
| N | 2,496 | 2,496 | 2,496 | 1,100 | 1,100 | 1,100 |
| Mean dep. var. | 0.0349 | 0.0453 | 0.0489 | 0.0527 | 0.0682 | 0.0727 |

Table A.3. Robustness Checks: Class Actions of VC- vs Non-VC-Backed Firms (Continued)

Panel E: Class actions at IPO

| Class period starts | Same day as IPO | 10 days of IPO | 30 days of IPO | Filing within 1 year of IPO |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| VC-backed | 0.0257*** | 0.0310*** | 0.0310*** | 0.0164** |
| | (0.008) | (0.008) | (0.008) | (0.007) |
| Log(Total Assets)_$IPOt-1$ | 0.0066*** | 0.0080*** | 0.0080*** | 0.0054*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Revenue Growth_$IPOt-1;t$ | 0.0000 | 0.0000 | 0.0000 | 0.0000** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| SIC-3 Sector FE × IPO Year FE | Yes | Yes | Yes | Yes |
| $R^2$ | 0.105 | 0.116 | 0.116 | 0.126 |
| N | 1,705 | 1,705 | 1,705 | 1,705 |
| Mean dep. var. | 0.0416 | 0.0493 | 0.0493 | 0.0293 |

This table reports OLS estimates and tests the robustness of the relationship between formerly VC-backed companies and the likelihood of a securities class action lawsuit within one to three years of the firm's IPO (baseline results reported in Table 4). Panel A uses the filing date, rather than the start of the class period, to define the timing of class action to IPO. Panel B reintroduces cases that were dismissed. Panel C restricts the sample to settled class actions with publicly available settlement amounts exceeding $3 million. Panel D restricts the sample to class actions filed between 2002 and 2019. All regressions include SIC-4 industry fixed effects based on CRSP classifications. Panel E restricts to class actions with class period starting on the same day, within 10 days, or 30 days as the IPO date. All regressions include SIC-3 industry fixed effects based on CRSP classifications, interacted with IPO year fixed effects. Standard errors are clustered at the SIC-3 level and are reported in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table A.4: Panel Prediction of Fraud Among VC-Backed Firms: Coefficients on Missing Value Indicators

| Dependent Variable: | $\mathbb{1}$(Fraud start) $\times$ 100 | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Missing valuation | 0.390*** | 0.413*** | 0.417*** | 0.390*** | 0.389*** | 0.402*** |
| | (0.048) | (0.051) | (0.052) | (0.048) | (0.048) | (0.051) |
| Missing board info | 0.046*** | | | 0.044*** | 0.051*** | 0.073*** |
| | (0.013) | | | (0.013) | (0.014) | (0.024) |
| Missing term sheet info | 0.015 | | | 0.010 | 0.013 | 0.026 |
| | (0.011) | | | (0.011) | (0.011) | (0.017) |
| Missing investor exits | | 0.031*** | | 0.034*** | 0.035*** | 0.071*** |
| | | (0.010) | | (0.011) | (0.011) | (0.019) |
| Missing investor type | | 0.003 | | -0.003 | -0.003 | -0.018 |
| | | (0.011) | | (0.011) | (0.011) | (0.022) |
| Missing team info | | | 0.027** | 0.035*** | 0.030** | 0.000 |
| | | | (0.011) | (0.012) | (0.013) | (.) |
| Missing team edu info | | | -0.010 | -0.008 | -0.006 | 0.000 |
| | | | (0.008) | (0.009) | (0.009) | (.) |
| FE: ind-year | Yes | Yes | Yes | Yes | Yes | Yes |
| FE: ind-1st round year | No | No | No | No | Yes | No |
| FE: ind-year, firm | No | No | No | No | No | Yes |
| Other controls | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 766834 | 766834 | 766834 | 766834 | 766834 | 766834 |
| $R^2$ | 0.007 | 0.007 | 0.007 | 0.007 | 0.011 | 0.101 |
| Outcome Mean | 0.058 | 0.058 | 0.058 | 0.058 | 0.058 | 0.053 |

This table reports the coefficients on missing value indicators included in Table 5. Standard errors are clustered at the industry level and are reported in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table A.5: Panel Prediction of Fraud Among VC-Backed Firms: More Team Characteristics

| Dependent Variable: | $\mathbb{1}$(Fraud start) $\times$ 100 | |
|---|---|---|
| | (1) | (2) |
| Firm age | -0.001 | -0.000 |
| | (0.000) | (0.000) |
| Ln(valuation) | 0.142*** | 0.142*** |
| | (0.017) | (0.017) |
| Ln(raised) | 0.009*** | 0.011*** |
| | (0.002) | (0.002) |
| Team_solo | 0.011** | 0.014*** |
| | (0.005) | (0.005) |
| Team_hasserial | 0.011 | 0.013 |
| | (0.008) | (0.008) |
| Team_topschool% | 0.009 | 0.012 |
| | (0.009) | (0.009) |
| Team_avg. age | -0.000** | -0.000** |
| | (0.000) | (0.000) |
| Team_immigrants% | -0.011 | -0.010 |
| | (0.009) | (0.009) |
| Team_female% | 0.007 | 0.005 |
| | (0.010) | (0.010) |
| FE: ind-yr | Yes | Yes |
| Other controls | No | Yes |
| N | 989033 | 989033 |
| $R^2$ | 0.006 | 0.006 |
| Outcome Mean | 0.045 | 0.045 |

This table show robustness of Table 5 to including more team characteristics. Columns 1 excludes all other control variables in Table 5 while Column 2 includes them but omits reporting their coefficients for brevity. The sample is a hazard-style panel of firm-years for US VC-backed firms founded since 2000. For non-fraud firms, the panel includes all years from firm founding to the earlier of 2024 or firm closure. For fraud firms the panel goes from firm founding year to the fraud start year. The dependent variable is 100 in the fraud start year for fraud firms and is 0 otherwise. We restrict to SEC/DOJ/WestLaw cases where fraud start year is no later than IPO year, and class action cases with class date within 2 years of IPO. The sample contains 441 fraud cases from SEC, DOJ, and westlaw, class action suits. Standard errors are clustered at the industry level and are reported in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table A.6: Panel Prediction of Fraud Among VC-Backed Firms: News-Based Cases Only

| Dependent Variable: | $\mathbb{1}$(Fraud start) × 100 | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Firm age | 0.000 | 0.000 | -0.001 | -0.001 | -0.000 | 0.000 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (.) |
| Ln(valuation) | 0.196*** | 0.196*** | 0.199*** | 0.195*** | 0.194*** | 0.181*** |
| | (0.026) | (0.026) | (0.027) | (0.026) | (0.026) | (0.027) |
| Ln(raised) | 0.011*** | 0.009*** | 0.012*** | 0.009*** | 0.009*** | 0.012*** |
| | (0.003) | (0.003) | (0.004) | (0.003) | (0.003) | (0.004) |
| Board size | 0.044*** | | | 0.043*** | 0.044*** | 0.068*** |
| | (0.008) | | | (0.008) | (0.008) | (0.012) |
| Board VC controlled | -0.136*** | | | -0.146*** | -0.153*** | -0.248*** |
| | (0.031) | | | (0.031) | (0.032) | (0.059) |
| Board shared control | -0.040*** | | | -0.043*** | -0.051*** | -0.073*** |
| | (0.015) | | | (0.015) | (0.016) | (0.024) |
| Investors' converted ownership | -0.041 | | | -0.076** | -0.064* | -0.122** |
| | (0.034) | | | (0.034) | (0.036) | (0.061) |
| Founders' payoff convexity | 0.029** | | | 0.034** | 0.037*** | 0.086*** |
| | (0.014) | | | (0.013) | (0.014) | (0.032) |
| Ln(unique investors) | | 0.028*** | | 0.040*** | 0.040*** | 0.113*** |
| | | (0.006) | | (0.007) | (0.008) | (0.019) |
| Investor_%non-IVC | | 0.032*** | | 0.025*** | 0.026*** | 0.032* |
| | | (0.008) | | (0.008) | (0.008) | (0.017) |
| Lead investors' ln(past exits) | | -0.006** | | -0.004 | -0.003 | -0.013*** |
| | | (0.002) | | (0.002) | (0.002) | (0.005) |
| Team_solo | | | 0.030*** | 0.036*** | 0.035*** | 0.000 |
| | | | (0.009) | (0.009) | (0.009) | (.) |
| Team_hasserial | | | 0.006 | 0.008 | 0.008 | 0.000 |
| | | | (0.009) | (0.009) | (0.009) | (.) |
| Team_topschool% | | | 0.000 | 0.002 | -0.002 | 0.000 |
| | | | (0.009) | (0.010) | (0.010) | (.) |
| Team_avg. age | | | -0.001*** | -0.001*** | -0.001*** | -0.001 |
| | | | (0.000) | (0.000) | (0.000) | (0.002) |
| FE: ind-year | Yes | Yes | Yes | Yes | Yes | Yes |
| FE: ind-1st round year | No | No | No | No | Yes | No |
| FE: ind-year, firm | No | No | No | No | No | Yes |
| Missing value indicators | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 766672 | 766672 | 766672 | 766672 | 766672 | 766672 |
| $R^2$ | 0.007 | 0.007 | 0.007 | 0.007 | 0.009 | 0.079 |
| Outcome Mean | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 | 0.050 |

The sample is a hazard-style panel of firm-years for US VC-backed firms founded since 2000, using only news-based fraud cases. For non-fraud firms, the panel includes all years from firm founding to the earlier of 2024 or firm closure. For fraud firms the panel goes from firm founding year to the fraud start year. The dependent variable is 100 in the fraud start year for fraud firms and is 0 otherwise. We restrict to news cases with initial news year is no later than IPO year. The sample contains 398 news-based cases. Standard errors are clustered at the industry level and are reported in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table A.7: Panel Prediction of Fraud Among VC-Backed Firms: Raised More than \$1M, \$5M, or \$10M

| Dependent Variable: | $\mathbb{1}$(Fraud start) $\times$ 100 | | |
|---|---|---|---|
| Sample: | Raised \$1M+ | Raised \$5M+ | Raised \$10M+ |
| | (1) | (2) | (3) |
| Firm age | -0.002** | -0.004*** | -0.005** |
| | (0.001) | (0.002) | (0.002) |
| Ln(valuation) | 0.295*** | 0.365*** | 0.409*** |
| | (0.035) | (0.044) | (0.048) |
| Ln(raised) | 0.024*** | 0.022*** | 0.021*** |
| | (0.006) | (0.006) | (0.006) |
| Board size | 0.042*** | 0.044*** | 0.050*** |
| | (0.008) | (0.009) | (0.010) |
| Board VC controlled | -0.170*** | -0.185*** | -0.211*** |
| | (0.038) | (0.042) | (0.047) |
| Board shared control | -0.063*** | -0.063** | -0.075* |
| | (0.022) | (0.029) | (0.039) |
| Investors' converted ownership | -0.117*** | -0.122** | -0.142** |
| | (0.042) | (0.052) | (0.060) |
| Founders' payoff convexity | 0.037** | 0.042* | 0.046* |
| | (0.017) | (0.023) | (0.027) |
| Ln(unique investors) | 0.060*** | 0.079*** | 0.096*** |
| | (0.010) | (0.013) | (0.016) |
| Investor_%non-IVC | 0.040*** | 0.063*** | 0.081*** |
| | (0.014) | (0.022) | (0.028) |
| Lead investors' ln(past exits) | -0.009** | -0.009* | -0.010* |
| | (0.004) | (0.005) | (0.006) |
| Team_solo | 0.043*** | 0.067*** | 0.085*** |
| | (0.012) | (0.018) | (0.023) |
| Team_hasserial | 0.008 | 0.018 | 0.020 |
| | (0.013) | (0.018) | (0.023) |
| Team_topschool% | 0.013 | 0.022 | 0.012 |
| | (0.018) | (0.027) | (0.033) |
| Team_avg. age | -0.001** | -0.001*** | -0.001*** |
| | (0.000) | (0.000) | (0.001) |
| FE: ind-yr | Yes | Yes | Yes |
| Missing value indicators | Yes | Yes | Yes |
| N | 632238 | 414190 | 315332 |
| $R^2$ | 0.008 | 0.009 | 0.010 |
| Outcome Mean | 0.114 | 0.159 | 0.196 |

The table is similar to Table 5 but restricts to firm-years with at least \$1M, \$5M, or \$10M of cumulated financing. Standard errors are clustered at the industry level and are reported in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table A.8: Panel Prediction of Fraud Among VC-Backed Firms: Dropping Crypto/Blockchain Companies

| Dependent Variable: | $\mathbb{1}$(Fraud start) × 100 | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Firm age | 0.000 | -0.000 | -0.001 | -0.000 | -0.001 | 0.000 |
| | (0.000) | (0.000) | (0.001) | (0.001) | (0.001) | (.) |
| Ln(valuation) | 0.168*** | 0.166*** | 0.167*** | 0.168*** | 0.167*** | 0.160*** |
| | (0.020) | (0.019) | (0.019) | (0.020) | (0.020) | (0.021) |
| Ln(raised) | 0.014*** | 0.011*** | 0.012*** | 0.013*** | 0.013*** | 0.014*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.004) |
| Board size | 0.018*** | | | 0.017*** | 0.020*** | 0.032*** |
| | (0.005) | | | (0.005) | (0.005) | (0.007) |
| Board VC controlled | -0.096*** | | | -0.100*** | -0.103*** | -0.162*** |
| | (0.026) | | | (0.026) | (0.026) | (0.034) |
| Board shared control | -0.016 | | | -0.017 | -0.018 | -0.009 |
| | (0.016) | | | (0.016) | (0.017) | (0.025) |
| Investors' converted ownership | -0.079** | | | -0.098*** | -0.099*** | -0.138** |
| | (0.031) | | | (0.033) | (0.034) | (0.054) |
| Founders' payoff convexity | 0.017 | | | 0.020* | 0.020* | 0.055** |
| | (0.011) | | | (0.011) | (0.011) | (0.026) |
| Ln(unique investors) | | 0.015*** | | 0.026*** | 0.025*** | 0.082*** |
| | | (0.003) | | (0.005) | (0.005) | (0.010) |
| Investor_%non-IVC | | 0.036*** | | 0.029*** | 0.028*** | 0.054** |
| | | (0.010) | | (0.010) | (0.010) | (0.023) |
| Lead investors' ln(past exits) | | -0.008*** | | -0.006** | -0.005** | -0.012** |
| | | (0.002) | | (0.002) | (0.002) | (0.005) |
| Team_solo | | | 0.013* | 0.017** | 0.016** | 0.000 |
| | | | (0.007) | (0.007) | (0.007) | (.) |
| Team_hasserial | | | 0.011 | 0.014 | 0.012 | 0.000 |
| | | | (0.009) | (0.009) | (0.009) | (.) |
| Team_topschool% | | | 0.004 | 0.006 | 0.002 | 0.000 |
| | | | (0.009) | (0.009) | (0.010) | (.) |
| Team_avg. age | | | -0.001*** | -0.000*** | -0.001*** | 0.001 |
| | | | (0.000) | (0.000) | (0.000) | (0.001) |
| FE: ind-year | Yes | Yes | Yes | Yes | Yes | Yes |
| FE: ind-1st round year | No | No | No | No | Yes | No |
| FE: ind-year, firm | No | No | No | No | No | Yes |
| Missing value indicators | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 752263 | 752263 | 752263 | 752263 | 752263 | 752263 |
| $R^2$ | 0.007 | 0.007 | 0.007 | 0.007 | 0.011 | 0.094 |
| Outcome Mean | 0.054 | 0.054 | 0.054 | 0.054 | 0.054 | 0.050 |

The table shows robustness of Table 5 to dropping firms in cryptocurrency or blockchain verticals. The sample is a hazard-style panel of firm-years for US VC-backed firms founded since 2000. For non-fraud firms, the panel includes all years from firm founding to the earlier of 2024 or firm closure. For fraud firms the panel goes from firm founding year to the fraud start year. The dependent variable is 100 in the fraud start year for fraud firms and is 0 otherwise. We restrict to news cases with initial news year is no later than IPO year. The sample contains 407 news-based cases. Standard errors are clustered at the industry level and are reported in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

# A Extending Malenko and Ewens (2024) Startup Board Data

We follow the methods and procedures outlined in EW2024 to construct our startup board dataset, using PitchBook board seat data and Form D filings from Q1 2009 to Q4 2024.

To focus on the startups most relevant to our study, we apply several filters to the PitchBook universe. First, we require that a firm has received at least one financing round backed by a venture capital firm, accelerator, or angel investor. Second, we exclude startups whose first observed financing is at a late stage. Third, the firm must have filed at least one Form D during the sample period.[23]

## A.1 PitchBook Board Seat Data

We construct a comprehensive dataset of board seats by combining PitchBook's firm-level board member data with individual employment histories. We first drop board observers and advisors based on their role descriptions. Board members are then categorized into three groups. Those identified as executives via a curated list of role keywords are labeled as executive directors. Among the remaining members, those affiliated with investors (VCs, accelerators, or angel investors) are labeled as VC directors. The rest are classified as independent directors.

## A.2 Form D Filings

Form D filings play a crucial role in identifying executive directors, as emphasized by EW2024. When cross-validating the PitchBook-based board member composition against those derived in EW2024 using a common sample, we find that our measures align closely for VC and independent directors, but substantially underestimate the number of executive directors. This underscores the importance of incorporating Form D data.

Form D filings also help address missing end-date information in PitchBook. Therefore, we clean date variables following EW2024. We extract the earliest and latest filing dates for each director-firm-type combination, as well as the first filing in which the director is no longer listed. For directors with ongoing appointments as of the latest filing, we impute an end date of December 1, 2025. If a director disappears from a subsequent filing, we set their end date as either six months prior to that filing or, if the interval between filings is shorter than six months, the midpoint between the two.

[23] We do not require a Form D filing immediately following each financing round, as our analysis tracks board composition annually rather than by financing round.

## A.3 Merging PitchBook and Form D

We match Form D records to PitchBook firms using the CIK identifier. In cases of duplicate CIKs (typically due to M&A), we retain the parent or surviving entity. Among approximately 63,000 PitchBook firms with CIKs, about 46,000 filed at least one Form D with director information. For PitchBook firms missing CIKs, we implement a name-based match to Form D company names using exact matching, which adds CIKs (and thus Form D links) for about an additional 1,700 firms.

When individuals are associated with multiple CIKs (across subsidiaries, for example), we aggregate the data at the person-PitchBook company level by taking the earliest start date and latest end date across all CIKs.

We then expand the board dataset in two steps. First, we merge individuals across PitchBook and Form D by matching the first and last name components (to address naming variation) and applying exact matches within CIKs. Second, for unmatched individuals, we use fuzzy matching with an 85% similarity threshold.[24] This process leads to a matched sample of roughly 130,000 person-firm pairs.

Since both PitchBook and Form D are imperfect on their own, we retain unmatched directors as complements. Specifically, we keep unmatched investor and independent directors from PitchBook, dropping unmatched executive directors due to likely misclassification, which adds about 40,000 directors. From unmatched Form D records, we add around 40,000 executive directors. For remaining unmatched individuals, we use PitchBook's person file to infer director types based on their affiliation with known investors in the focal firm. Directors for whom type cannot be determined are dropped (about 18% of the full sample).

Next, we assign director service periods as follows. For the start date, we use the earlier of the two available dates in the matched sample. For non-overlapping observations, we take whichever date is available. If a VC director is missing a start date, we assign the date of their first investment in the company. Independent directors without a start date are excluded from the sample. For the end date, if both Form D and business exit dates are available, we use the earlier of the two.[25] When only a single end date is available, we adopt it. For firms without an end date from any source (PitchBook, Form D, or exit events), we require that their most recent financing occurred in 2023 or later; otherwise, we consider them failed and assign an end date two years after their last financing round. All remaining firms receive a default end date of December 1, 2025.

After these procedures, we obtain approximately 200,000 unique director-firm pairs across roughly 39,000 distinct firms.

[24] This process increases the matched sample by approximately 19%. In the matched dataset, over 70% of PitchBook directors appear in Form D, and about 57% of Form D directors are found in PitchBook.

[25] Business exit is defined as the earliest occurrence of failure, M&A, or IPO.

## A.4 Constructing Firm-Year Panel

To construct the firm-year panel, we first aggregate the director data to a firm-level panel with cumulative counts of each director type from the first to the last year with board information. We define the start year as the later of (1) the first financing year or (2) the first year with board data.[26] The end year is either the last business year or 2025, whichever comes first. We then fill in the full sequence of years and backfill director composition over time.

## A.5 Cross-Validation with EW2024

We validate our board composition measures by replicating EW2024's results using our constructed dataset. EW2024's sample spans 2002 to 2017, while ours covers 2009 to 2024. To ensure comparability, we restrict to overlapping firms whose first observed year is after 2008 and limit the analysis to overlapping firm-years. This yields 17,365 firm-year observations.[27]

Table A.9 compares summary statistics between our measures and those in EW2024. Figure A.5 replicates Figures 1 to 4 of EW2024 using our board composition measures. The close alignment in both statistics and trends supports the validity of our construction approach and suggests that our measures are generalizable to a broader and more recent set of startup firms.

## A.6 Construct the Final Sample

We expand EW2024 along both the time series and cross-sectional dimensions. In the time series, we extend only forward for firms that appear in EW2024. For new firms, we utilize our own series.

Because of measurement errors and data source discrepancies around extension years, observed changes in board composition may not reflect true events. To address this, we implement three extension methods to assess robustness:

1. **Forward Extension by Firm-Year:** Extend each series only forward from the final EW observation, tolerating small discrepancies in the transition year.

2. **Substitution for Overlapping Years:** For firm-years from 2009 onward where EW and our data overlap, substitute EW values with ours; in all other years, apply the forward–extension rule.

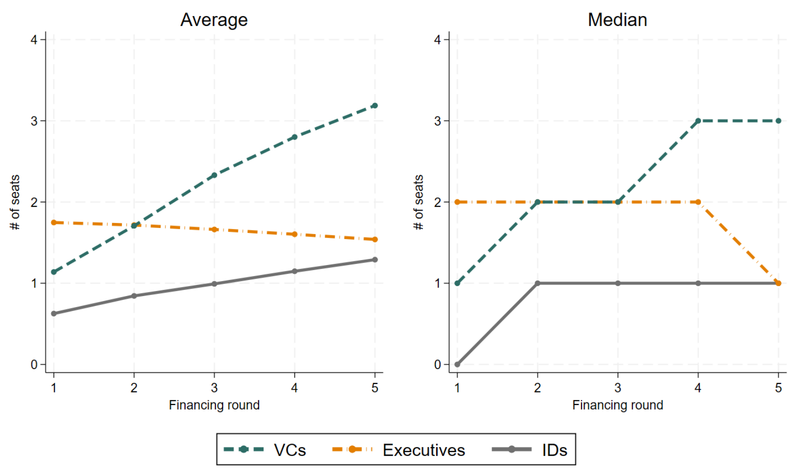
3. **Sample Expansion without Overlap:** Use our data for any firm whose first observation falls after 2008 or which does not appear in the EW; for all others, we retain the original EW

[26] If both financing and board start years are available, we use the latter; if financing data is missing, we default to the board start year.

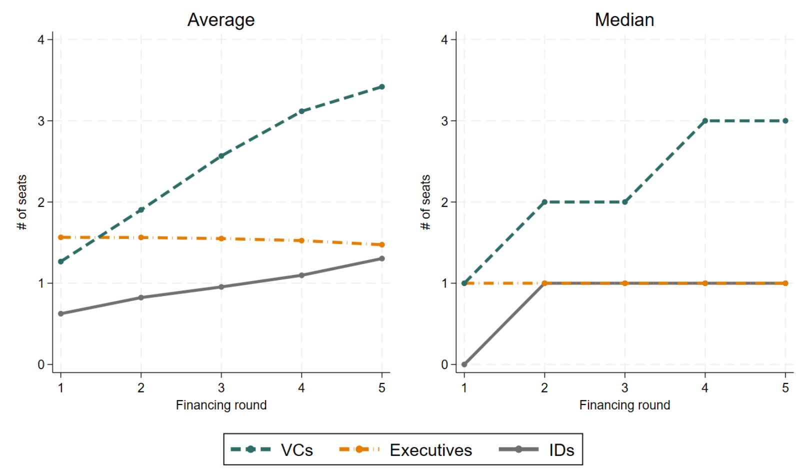
[27] EW2024 includes 18,819 firm-years starting in 2009, suggesting our sample coverage is highly comparable.

records. This strategy prevents mixing data sources within a single firm, thereby maximizing our cross-sectional coverage while minimizing measurement error from data source transitions.

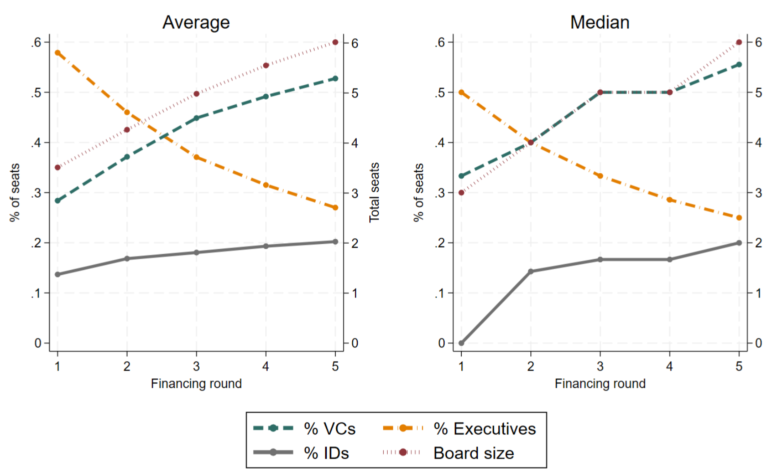
Figure A.5: Validating Extended Startup Board Data: Replicating EW2024 Figures 1 to 4



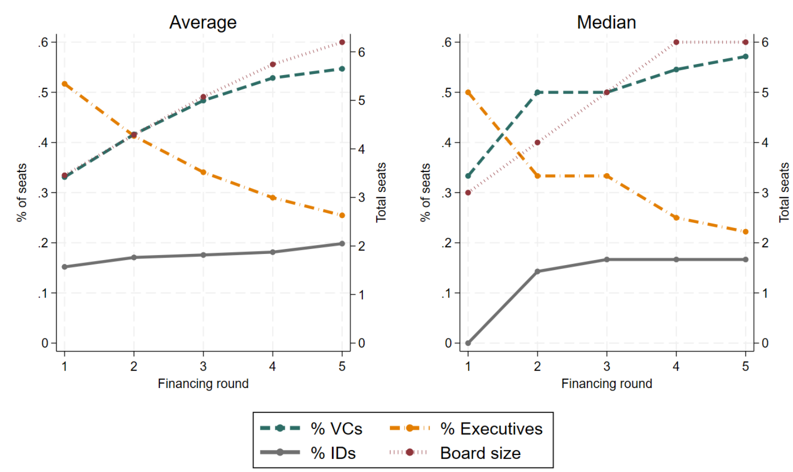

(a) Figure 1: EW2024




(b) Figure 1: Our Sample



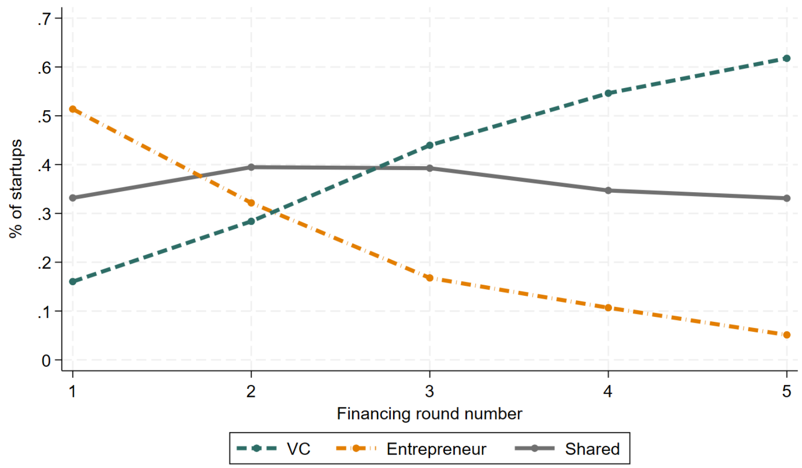

(c) Figure 2: EW2024



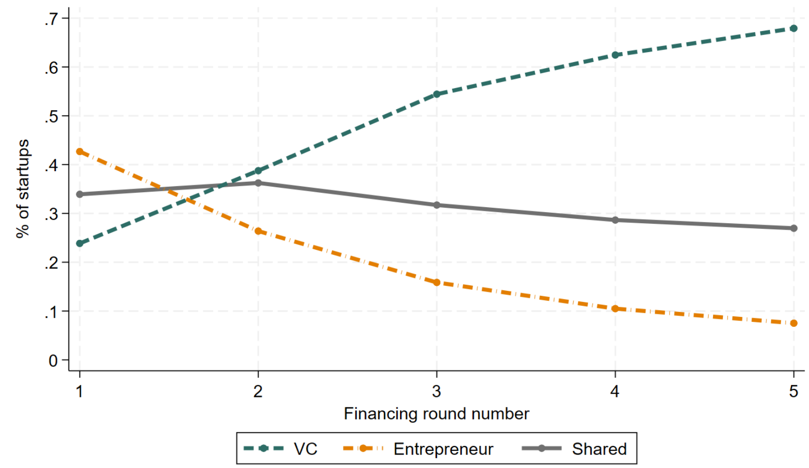

(d) Figure 2: Our Sample

Figures on the left column are generated based on measures in EW2024, and Figures on the right column are generated based on measures in our sample.

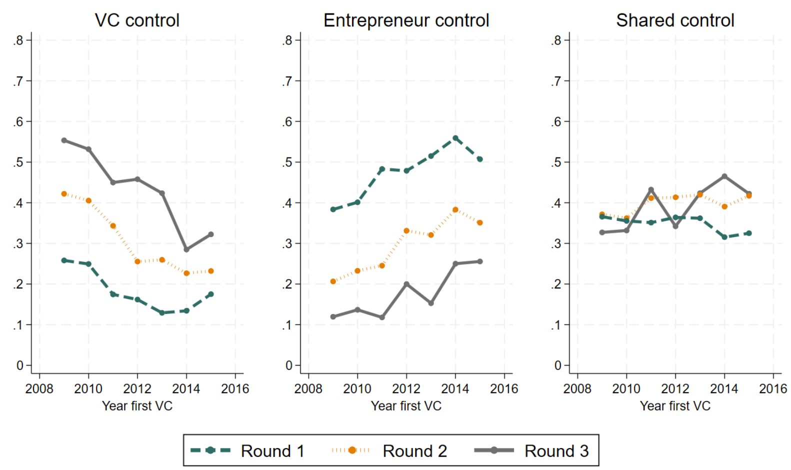
Figure A.5: Validating Extended Startup Board Data: Replicating EW2024 Figures 1 to 4 (Continued)



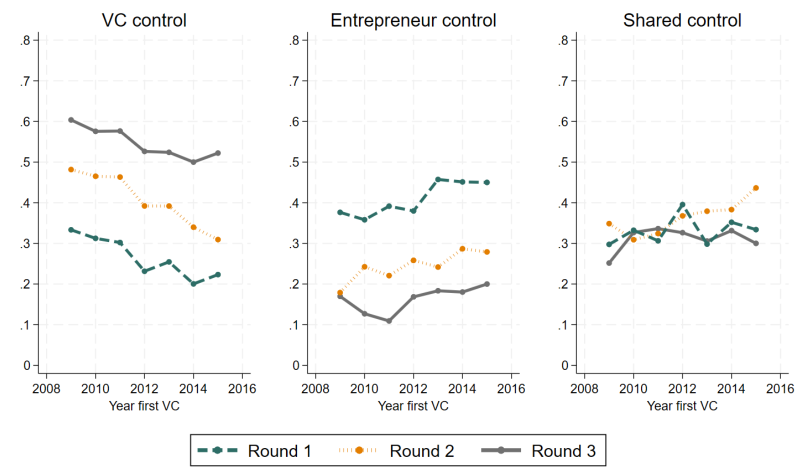

(e) Figure 3: EW2024




(f) Figure 3: Our Sample




(g) Figure 4: EW2024




(h) Figure 4: Our Sample

Figures on the left column are generated based on measures in EW2024, and Figures on the right column are generated based on measures in our sample.

Table A.9: Comparison of Overlapped Sample: Summary Statistics

| | EW2025 | | | | |
|---|---|---|---|---|---|
| Variable | Obs | Mean | Std. Dev. | Min | Max |
| num_independent | 17,186 | 0.797 | 1.052 | 0 | 8 |
| num_executives | 17,186 | 1.672 | 0.747 | 0 | 8 |
| num_vcs | 17,186 | 1.589 | 1.369 | 0 | 11 |
| | Our Sample | | | | |
| num_independent | 17,186 | 0.757 | 1.023 | 0 | 7 |
| num_executives | 17,186 | 1.468 | 0.836 | 0 | 6 |
| num_vcs | 17,186 | 1.712 | 1.525 | 0 | 11 |

The table shows the distribution of three board type measures among the 17,365 overlapped firm-year observations between our sample and EW 2024.

# B Appendix: Identifying Fraud Cases from Various Data Sources

## B.1 SEC Enforcement Actions

To identify SEC enforcement actions involving VC-backed startups and their founders, we begin by scraping three types of SEC releases from 1995 to 2023: litigation releases (11,209), administrative proceedings (17,763), and Accounting and Auditing Enforcement Releases (AAERs) (3,230), which overlap with the first two when related to financial reporting violations. From these releases, we extract 64,830 unique respondents associated with 32,545 releases.

Our first matching approach links SEC respondents to PitchBook companies and individuals using a combination of exact and fuzzy name matching. We retain matches where the release date occurs after the company's founding and first financing, and within five years of either its last financing round or IPO. For individuals, we ensure the release date falls within five years of their departure from a VC-backed company. This procedure yields 206 unique releases involving 108 companies and 214 executives in PitchBook.

In a complementary second approach, we apply keyword-based filters to litigation release texts to flag potential cases involving VC-backed startups or founders (e.g., "venture capital", "startup", "securities fraud", "Ponzi scheme"). We manually verify and match flagged cases to PitchBook. This step identifies an additional 37 SEC releases linked to 31 companies and 88 individuals.

Combining both approaches, we identify 226 unique SEC enforcement actions from 1996 to 2023, associated with 214 companies and 261 founders/CEOs. Most cases occur after 2010, with 42 cases filed in 2023 alone. To refine the final dataset, we exclude cases involving non-relevant violations such as delinquent filings, unregistered broker activity, or market manipulation. Finally, we read and manually validate each case. After imposing filters relevant for our analysis (e.g., founded post 2000, has non-missing fraud start year), 99 SEC cases enter our final sample used in predictive analysis.

## B.2 DOJ Enforcement Actions

We collect information on criminal enforcement actions from the U.S. Department of Justice (DOJ) by scraping all publicly available releases from 2013 to April 2024, resulting in a dataset of 203,226 cases. 65% of these cases include subject-matter tags; among the tagged cases, 18,125 are labeled "financial fraud" and 1,494 are tagged Securities, Commodities, & Investment Fraud".

To identify cases involving startups, we proceed in several steps. First, we filter for cases tagged as financial or securities fraud and containing startup-related keywords (e.g., "founder", "venture capital", "Silicon Valley"). This yields 543 possible cases, of which 243 are manually matched to

162 unique companies in PitchBook.

Second, we apply a machine learning classifier (BERT) to predict the crime type where tags are missing. Among the predicted fraud cases, we manually verify 477 cases as involving startup fraud, identifying 90 additional matches to PitchBook. Third, we expand the training set to flag additional cases. 952 additional cases are predicted as startup-related fraud. We manually review these cases and confirm 477 of these, with 219 linked to PitchBook. Fourth, we review 3,250 additional cases predicted as fraud but not initially linked to VC-backed firms. Manual review yields 599 more matches with PitchBook.

Finally, we restrict the combined sample to U.S.-based firms in the PitchBook VC universe. We manually validate each case to make sure they are linked to the right PitchBook company. This results in 183 unique VC-backed companies involved in 227 DOJ cases. After imposing various filters, the final sample that enters our predictive analysis includes 73 DOJ cases.

## B.3 Securities Class Action Lawsuits

Our class action dataset begins with a comprehensive dataset of 6,601 securities class action lawsuits filed between January 1996 and April 2024, drawn from the Stanford Securities Class Action Clearinghouse (SCAC). We restrict the sample to the 2,275 non-dismissed cases involving U.S. public companies filed between 1999 and 2023. For consistency with our other data sources, we further restrict the sample to companies that went public in 2002 or later.

We determine whether each company was venture-backed using PitchBook's VC deal history, supplemented with Ritter's VC flag, which is particularly useful for identifying VC backing prior to 2010. The final sample includes 917 class action cases, of which 536 involve (previously) VC-backed firms. Restricting to cases within 2 years of IPO yields 282 cases by VC-backed firms, out of which 178 entered our predictive analysis after imposing various filters. To enrich the dataset, we add industry classifications retrieved from CRSP, financial information from Compustat and IPO-related information from Jay Ritter's IPO database.

## B.4 Westlaw Database

Westlaw expands fraud-related litigation coverage in two complementary ways. First, its court cases database broadens the legal scope beyond regulatory and securities class actions by capturing common-law fraud suits, contract and commercial disputes, shareholder derivative actions, and bankruptcy proceedings, thus extending the sample to private, creditor-driven, and state court cases that would otherwise be missed. Second, Westlaw dockets add a time dimension by recording cases that never yield a published opinion, often providing the earliest – and sometimes only –

evidence of a firm being sued for fraud. Together, court cases and dockets allow us to capture both the broader legal settings in which fraud allegations arise and the earlier stages of litigation that might otherwise disappear through dismissal or settlement.

#### Westlaw Case Documents

Westlaw cases record litigation that went to court and yielded public opinions. We apply the following filter to identify cases involving fraudulent activities where the defendant is likely a startup or its founder: words in ["fraud", "misrepresent", "deceit", "misconduct", "Ponzi scheme", "embezzl", "securities fraud", "wire fraud"] **AND** any words in ["startup", "start-up", "early-stage", "venture-backed", "tech comp", "founder", "venture capital", "silicon valley"] **AND** be related to topics in ["fraud", "false pretenses", "securities transactions", "securities regulation"]. This query returns 3,250 cases.

The same process is used to identify fiduciary-duty breach related to the following keywords: ["breach of fiduciary duty", "self-deal", "conflicts of interest", "duty of loyalty", "duty of care"]**AND** be related to topics in ["Corporations and Business Organizations"]. This query returns 1,735 cases, with some about half overlapping with the previous sample, suggesting that fiduciary duty breaches are quite often also associated with fraudulent activities.

From each case, we use text parsing to extract the full list of defendants, typically found on the title page. We then identify all defendants that are likely companies. For a substantial number of cases that list only individual defendants, we conduct a detailed review of the factual background for each such case to identify any companies used in the fraud scheme but omitted as principal defendants.

After completing these two rounds of firm name extraction, we match them to our universe of VC-backed companies in PitchBook. The matching process combines fuzzy name matching with manual verification to ensure accuracy, which yields 240 VC-backed firms across 219 cases.

Next, we extract the specific variables required for the analysis. For this step, we employ the GPT API to perform the initial extraction, followed by research assistant verification to confirm the accuracy of the information for each case. Here is a list of variables we extract: 1) Defendant Type: `"Person"`, `"Company"`, `"Person and Company"`, or `"Other"`; 2) Defendant's role in the company during the fraud period; 3) Fraud Type: `"bank"`, `"wire"`, `"mail"`, `"corporate"`, `"federal"`, `"securities"`, `"tax"`, or `"other"`; 4) What was misrepresented: `"financials"`, `"product"`, `"use of funds"`, or `"other"`. 5) Victim: `"investors"`, `"government"`, `"public"`, or `"other"`; 6) Fine amount; 7) Prison month; 8) Charge date; 9) Fraud start date; 10) Fraud end date.

Following the extraction of case-level information, we implement a series of filters to construct the final sample. First, we exclude cases in which fraud charges were ultimately dismissed by the judge. Second, we manually review the case materials to recover any missing information. Lastly, we verify the final outcome of each case, labelling them as dismissed, settled, ongoing, and reach a decision. This process involves substantial judgment and careful case-by-case validation to ensure completeness and accuracy of the data. A similar process is applied to fiduciary-duty-related cases. The final sample consists of 100 cases involving 111 VC-backed firms.

**Dockets and DOJ News Release**

Dockets are essentially court record identifiers: they confirm that a filing exists and track how the status of the case changes within the court filing system over time, but do not provide substantive details about the proceedings or allegations. They allow us to track allegations that never get a chance to yield any public opinion.

We apply a series of advanced search filters on *Westlaw US Dockets* to identify three categories of fraud cases: settled, ongoing, and those resulting in final judgment. Dockets are chronological and often lack factual detail, making it difficult to identify startup defendants through keyword searches. Therefore, we apply a rather generous filter, focusing on the presence of fraud-related allegations and procedural outcomes. Here is the list of filters we apply:**Settled cases:** dockets must contain any words in ["wire fraud", "mail fraud", "bank fraud", "securities fraud", "other fraud"] **AND** any words in ["settlement", "stipulation", "settled", "active", "judgment", "verdict", "summary judgment"] **AND** exclude cases with words ["dismissed", "class action"] **AND** be classified under any of the following legal topics: ["Business Organizations", "Civil", "Contracts", "Fraud & Misrepresentation", "Other Federal Statutes", "Other Fraud"].

This process yielded approximately 22,498 dockets related to fraud and 13,125 dockets related to fiduciary duty breaches. Due to restrictions imposed by the data provider, we were not permitted to download the full content of each docket for this amount. Instead, we obtained spreadsheets containing only basic metadata for each docket, including the title, filing date, court, and presiding judge (if available). The case title, uniformly formatted as âplaintiff v. defendant,â provides one defendant name for each case, which may refer to either a firm or an individual. We begin by screening the dockets using the defendant name listed in the title. From these records, we extract the list of defendant firms and match them to the PitchBook universe of VC-backed firms using a combination of fuzzy matching and manual review. For defendants that are individuals rather than firms, we match them to board and executive names from PitchBook using exact name matching. Each match is then manually verified by reviewing the docket content, which provides a full list of all defendants, including firm names. A case is retained only when the defendantâs firmâindividual pair corresponds to the same board/executiveâfirm pair identified in PitchBook.

Through this procedure, we identified 327 dockets linked to 290 unique VC-backed firms founded in or after 2000 as defendants. The next step involved collecting the detailed variables required for our analysis. To obtain usable information, we manually searched for associated court filings for each docket. These searches relied on a wide variety of sources, including Westlaw, Justia, official government websites, specialized legal platforms such as Law360, and relevant news or legal articles. Using this process, we were able to locate and download approximately 90% of the filings linked to our docket sample.

Once the filings were assembled, we employed the GPT API to extract structured variables of interest, consistent with the approach described earlier. Automated extraction was followed by rigorous manual verification to ensure the reliability of our dataset. Specifically, we confirmed that: 1) the defendant company matched the PitchBook record; 2) the suit had not been dismissed; 3) the filing contained information on the alleged period of fraudulent activity.

Next, we search the DOJ news release database in Westlaw using the same fraud-related keywords, restricting to those issued before 2014. This process yields 1,709 releases. We then apply the GPT API to extract defendant information, including both firms and individuals. Firms are matched to PitchBook by firm name. When only individual names are disclosed, we identify the associated firms by performing exact name matching between individuals and the PitchBook board and team database, using first and last names as the matching key. This procedure results in three VC-backed startups founded in or after 2000. For these cases, we collect relevant information, including the alleged fraud period, directly from the DOJ news releases.

### B.5 Crunchbase and PitchBook News Feeds

Finally, we use news articles from Crunchbase and PitchBook to construct a dataset that includes alleged fraud involving VC-backed startups. News contain credible allegations and ongoing investigations, which provides a broader set of "likely fraud" or "concerns for fraud", allowing us to mitigate concerns about omitted frauds. A downside of this source is that the news feed data begins primarily in 2017, and for this reason we only include it in robustness checks.

We aggregate all news articles linked to startups or their founders. We then screen articles using a combination of keyword filters (see Appendix for more information), manual checks by research assistants, and a machine learning classifier trained on over 6,000 hand-labeled articles to identify articles about startup fraud.

We manually check all news articles and identify 5,622 fraud-related news articles covering 1,103 unique startups. These include firms linked to SEC or DOJ actions, class action lawsuits, and other cases involving credible fraud allegations. This news-based sample is merged with our legal datasets for further analysis. After imposing various filters, 496 news-based cases enters into our predictive analysis.