Once Bitten, Twice Shy: Evidence From Venture Capital and Scam Startups

Yanying Lyu*

January 13, 2022

Abstract

Scam startups are booming over years. However, little evidence is on the effect of those startups on the venture capital (VC) investment decisions. In this paper I first construct a novel dataset of fraudulent startups using the release news from U.S. Department of Justice (DOJ) and Securities and Exchange Commission (SEC). I then investigate how VCs react after they are cheated by fraudulent startups in a difference-in-differences framework. The main finding is that VCs decline investments in new startups with respect to the amount of investment and the number of deals after they are cheated. The effect is mainly driven by the declined investment activities in industries to which the fraudulent startups belong. The heterogeneity analyses show that the effect is more pronounced for VCs which had more investments or had invested in higher valuated scam startups. My results suggest that learning and trust play critical roles in VCs investment activities.

Keywords: Venture Capital; Scam; Startups; Trust; Learning; Investment

JEL Classification: G24; G30; G41; M13

*Department of Finance, Tilburg University; Email: y.lyu_3@tilburguniversity.edu

1 Introduction

As Elizabeth Holmes, the founder of a famous scam biotech startup Theranos, was convicted of four fraud-related charges in federal court, the public began to realize dark sides of emerging fraudulent start-ups.[1] Scam startups are booming in recent years. Over the past ten years, 70% to 80% of all startups exaggerate rate of return which they do not obtain, while as many as 40% of all startups are a complete failure, losing all of their original investment (Gleason et al. (2021)). Silicon Valley is even called "a land of fake promise".[2] There is a culture of "fake it till you make it" within startups to attract VC investment in Silicon Valley.[3] For example, Theranos, valued at $9 billion in a 2014 funding round, "fakes" the existence of its core product and uses it to raise billions of dollars investment. Considering billions of dollars have poured into venture markets to finance startup business development, no matter whether startups are fake or real, increasing scam startups have caused huge losses in the financial market.[4]

Given the significant amount of money loss, it is important for both academic and industry to gain a better insight into the effect of booming scam startups on VCs' investment decisions. However, there is little evidence in the previous literature. The main reason is that no standard scam startups database exits. In this paper, I fill this gap by constructing a novel database of scam startups from official releases and investigating how involved VCs

[1] https://www.justice.gov/usao-ndca/us-v-elizabeth-holmes-et-al

[2] See https://glanalytics.ca/startup-fraud/

[3] The term describes startups that "fake" having a product or technology in order to attract potential customers and lure investors into funding the sale and marketing of such product. They would do so until they "make" the actual fully functioning and market ready version of the product.

[4] In the US alone, VCs assets under management have more than doubled since 2013, from $267 billion to $548 billion at year-end 2020; in 2020 alone, VC firms raised $74.5 billion in new funds which will soon be deployed into thousands of startups globally.

adjust their investment activities after scam startups are known by the public.

Learning plays an important role in VCs' decisions making. As Multiarmed-Bandit model shows, VCs have dynamic beliefs when they make investments (Berry and Fristedt (1985)). These beliefs are shaped by learning their past investments and outcomes, and they affect VCs' future investment decisions. When VCs know that the startup they invested in is fake and all investments are lost, their beliefs will be adjusted by learning from this failure experience. In addition, trust is positively related to investment (Gurun et al. (2018); Bottazzi et al. (2016); Duffner et al. (2009)). After the revelation of portfolio companies' frauds, the trust between VCs and startups will drop. I assume both aspects affect the relationship between VCs investment and startups. Therefore, the hypothesis for this paper posits that, VCs will learn from past experience of investing in scam startups, the trust between VCs and startups will decrease, then VCs decline investments in new startups after they are cheated.

I construct my novel dataset from U.S. Department of Justice (DOJ) and Securities and Exchange Commission (SEC), which provide detailed digital lawsuits against behaviors violating federal rules or securities laws. Like previous literature collecting frauds sample (Dyck et al. (2010); Choi (2007)), I start by using federal securities class actions to construct the novel sample of scam startups. I use a web crawler to download all cases related to financial frauds from DOJ and SEC. Then I build the words library from the lawsuits contents of ten representative scam startups to identify which is a fraudulent startup case and get scam startups dataset. Finally, I find VCs investment information from VentureXpert by using startups names collected from lawsuits contents.

In the empirical part, I create cohorts for each affected venture capitalist and follow Gormley and Matsa (2011) to use difference-in-differences (DID) method. After controlling the firm-cohort and year-cohort fixed effects, the estimation results show that after they are cheated by scam startups, venture capital would decrease investment a lot, especially in industries where frauds happened. The estimators are potential biased because of the combination of staggered treatment timing and treatment effect heterogeneity, either across groups or over time. Since affected VCs are totally different from unaffected VCs, simply comparing cheated VCs to non-cheated VCs could cause potential possible endogeneity and spurious correlation between the VCs investment activities and impacts of scam startups.

I do robustness tests to solve those concerns. In the first robustness test, I narrow the criteria for selecting control groups. The control groups are VCs which never been cheated but have invested in industries where frauds happened. I also use propensity-score matching (PSM) to find similar peers. The results are all robust to what I find in the baseline.

Overall, my findings indicate that VCs will learn from failure experience and decrease their investment after they are cheated by scam startups, especially in the industries where frauds happen. They reallocate capital into unaffected industries. The effects are more significant for VCs which invested more in scam startups or invested in scam startups with high valuation.

This paper has the following contributions. First, it fills the gap of literature on the effects of booming scam startups on financial investment. To the best of my knowledge, it is the first paper to examine the effects of fraudulent startups on venture capital investment activities. I construct a novel dataset of fraudulent startups and involved venture capitalists to provide empirical evidence that VCs will adjust their investment decisions after being cheated by startups. They will decrease investment in industries where frauds happen.

Second, it provides new evidence on studies of the importance of learning and trust to

VCs investments. Prior literature show that trust plays a critical role in the financial intermediation industry. For example, Bottazzi et al. (2016) use hand-collected data to empirically document that trust among nations positively predicts venture capital firms' investment decisions. Guiso et al. (2008) show that trust affects the willingness to invest money in shares and explains the limited participation in the stock market. I examine VCs' investment activities after they are cheated by scam firms. The trust between VCs and startups decreases after investment failure. Learning is also important to VCs investment decisions, which improves the investors' understanding of the various investment opportunities and improves their future decisions (Sorensen (2008)). This paper provides new evidence from scam startups investments failure.

Third, it contributes to literature which investigates the effect of litigation on VCs' performance. Previous studies examine the effect of litigation on VCs (Atanasov et al. (2012)) and startups luring funding abilities (Cumming et al. (2017)). This paper analyses lawsuits contents to show the effect of portfolio firms' litigations on VCs' investment activities.

The remaining paper is organized as follows. Section 2 describes the data sources and presents the summary statistics. Section 3 discusses the empirical strategy and presents the main results. In section 4, I report two robustness checks. Section 5 provides further heterogeneity analysis on the baseline. I summarize in section 6.

2 Data

In this section, I describe data sources and present the summary statistics of main variables in this paper.

2.1 Startup Frauds Lawsuits

First, I construct a novel dataset of fraudulent startups. The initial startup fraud lawsuits data are from enforcement actions initiated by DOJ and SEC.[5] Lawsuits in DOJ and SEC are reliable data sources for research on financial frauds and are widely used. As Karpoff et al. (2008) say, focusing on SEC and DOJ actions to discipline financial reporting violations can yield a clean sample of cases in which firms violate rules. Heese et al. (2021) also emphasis the effect of DOJ cases on corporate governance. Li and Cohen (2021) use cases from DOJ and SEC to collect firms alleged to be bribing.

DOJ and SEC provide detailed information of financial misconduct cases. DOJ has 28 topics based on lawsuit contents starting from year 2013, ranging from civil rights, drugs to violent crime. The most relevant topic for my research is financial frauds, in which department's fight against sophisticated economic crime. The section, Litigation Releases in SEC, focuses on firms' criminal cases concerning civil lawsuits brought by the commission in federal court. The official releases are typically posted on the same day that the legal case is filed and are immediately available starting from 1995. Therefore, I focus on cases from SEC during the period from 1995 to 2020. I combine these two data sources to construct my final startup fraud lawsuits data and my final sample period is between 1995 and 2020.

I use a web crawler to download a large sample of 11,672 digital cases initialed by DOJ and 10,381 digital lawsuits released by SEC. Figure 3 plots the time distribution of all cases. Most of cases are concentrated between 2016 and 2020. Each case includes the lawsuit title, the dates on which the case was filed and closed, the detailed content describing how the startup lies to the public, the names of defendants and the sentence and fine. Figure 1a and

[5]DOJ: https://www.justice.gov/usao/pressreleases; SEC: https://www.sec.gov/litigation/litreleases.htm

1b provide two examples of DOJ and SEC lawsuits, respectively.

To identify the fraudulent startups cases from all financial frauds cases in DOJ and SEC, I do the following procedures. First, using the 10 representative scam startups' lawsuit contents downloaded from DOJ and SEC, I define the training library of text for startup frauds identification.[6] First, I convert words into lower case. Second, I drop all punctuation marks, numbers, stop words (e.g., a/an, the, they, etc.) and remove all non-English characters. Then I tokenize sentences into single words. Lastly I retrieve each word's root format using the method called lemmatization. I only keep noun and adjective words because they can convert reasonable meanings. After cleaning textual document, I then count the frequency of common words appearing in those lawsuits. I manually check the top 100 most common words to select the key words to build a scam startups dictionary.[7]

Third, I follow Hassan et al. (2019) to use the tracking method to compare every case content to the words frequency table. If 70% of the 100 most common words of scam startups dictionary appear in a lawsuit content, I treat it as a potential startup lawsuit for the next step check. This method yields over 10,000 cases. If I lower 70% to 65%, 750 more cases are selected. I manually check those 750 cases but find they are not related to scam startups financial frauds.

Lastly, I manually read each case to identify the company involved in the lawsuit. I drop the case if it is not related to a startup but to individuals or a public firm. After dropping cases not related to startup companies, I check whether the rests are belonged to one of the following three categories: (i) products/technologies frauds, (ii) financial statement frauds,

[6]The scam startups and their DOJ and SEC release news are reported in Appendix A2.
[7]See Online Appendix A3 and Figure 2 for the frequency of these words.

and (iii) personal characteristic frauds. The above three categories are what VCs consider most when they make investment decisions.

- Products/technologies frauds. A startup lies to the public about its products or technologies to mislead investors to raise money. It defrauds investors that it has developed advanced products or high technologies which do not exist. It cheats investors by promising high return of investment. Kaplan and Strömberg (2004) find technology, product or service are factors VCs consider most when they make investment decisions. A famous example is Theranos, it defrauded doctors and patients by making false claims concerning its technology to provide accurate, fast and reliable blood tests which never exit.

- Financial statement frauds. This kind of fraud is also called accounting frauds. VCs use discounted cash flow (DCF), internal rate of return (IRR) (Graham and Harvey (2001)), or multiple of invested capital (MOIC) (Gompers et al. (2016)) to evaluate investment opportunities. Corporations misrepresent or deceive investors into believing that they are more profitable than they actually are or will be. They get investment by allegedly reporting false financial statement to the public. One example is Benja, a digital advertising company. The account receivables and financial statements CEO provided to investors were misstated and false and a majority of the purported revenue was fabricated.

- Personal characteristic frauds. Startups' founders/CEOs/CFOs lie to the public about their education, working experience or special abilities, which convinces investors that they are reliable and companies are profitable. Previous literature find VCs place the greatest importance on the management/founding team. The team quality is mentioned most frequently as the most important factor (Bernstein et al. (2017); Gompers et al. (2020); Kaplan and Strömberg (2000)). A typical example is Telomolecular Corporation, a biotechnology startup company. It claimed to have developed nanoparticle technology that could eradicate cancer and treat other age-related diseases. It also claimed Telomolecular had a deep management team with experience taking companies public. According to the DOJ release, it raised 6.7 million from around 400 investors.

Once a startup lawsuit is classified into one of above three categories, it can be regarded as a scam startup observation in the final startup fraud lawsuits sample. In summary, 547 lawsuits for 621 unique scam startup companies are in the final sample. From the lawsuit cases contents, I can retrieve the information of startups' names, defendants, publication release dates, sentence and fine, which helps me to link startups dataset with other data sources.

2.2 Investment Data

I download venture capital investment data from VentureXpert, which provides detailed investment information including investment date, amount, round, and some characteristics of VC: age, the number of funding it manages, companies it invests, fundraising it gets.

I then match the collected 621 scam startups against database from VentureXpert. For each startup, I search its name and find venture capitalists which once invested in it. In total, 63 startups are matched with 201 VCs have invested in. In addition, I use Crunchbase as a supplementary to manually collect the venture funding information for the remaining 598 startups once I cannot find data in VentureXpert.[8] Totally, I get 81 startups with 217 venture capitalists involved in the final sample. Figure 3 plots the time distribution of them.

2.3 Main Variables

To examine the effect of being cheated by scam startups on venture capital investment, I construct the following variables. The main explanatory variable I am interested in is *Cheated*, a dummy variable that equals one if a venture capital has invested in one or more fraudulent startups, or zero otherwise.

I use two variables to measure the investment activities of venture capital. The first measurement is *Deal Num*, which is the total number of deals a venture capital makes in a given year. The second measurement is *Deal Amount*. It is calculated as the total amount of investment a venture capital makes in a given year measured by million dollar.

I also construct several VC characteristics variables as control variables. *Fund Year* is defined as the given year minus the last year which venture capital received fundraising. *Company Num* is defined as the cumulative number of companies venture capital has invested in. *VC Age* refers to the age of venture capital. *Fund Num* is the cumulative number of fund venture capital manage. *Total Deal Amount* is the cumulative amount of investment a venture capital has made measured by million dollar. *Total Deal Num* is the cumulative number of investment a venture capital has made.

The definition and sources of all variables used in this paper are presented in Appendix A1.

[8] https://www.crunchbase.com/

2.4 Summary Statistics

Table 1 presents the summary statistics of the main variables used in the paper. Table 2 reports difference tests. Panel A presents the difference between VCs which have invested in fraudulent startups and VCs which have never invested in. It can be seen that cheated VCs are older, larger and make more investments and manage more funds compared to non-cheated VCs. This suggests that controlling for these differences when analyzing the effect of being cheated on VCs performance is crucial for the validity of my inferences. Panel B presents the difference between fake startups and real startups. From the table, there is no significant difference in equity, debt or age between two types of companies.

3 Empirical Strategy and Results

3.1 Empirical Strategy

For each venture capitalist which is cheated, I build a comparison group of unaffected VCs (VCs have never invested in scam startups) that are present in VentureXpert. In my sample, 5% of venture capitals have invested in more than one scam startups, to have a clean sample, I keep the earliest one as cheated event. To estimate the effect of being cheated by scam startups on VC investment, I compare investment changes in the affected and unaffected venture capitalists around the first time of being reported in DOJ or SEC. The empirical strategy generally follows Gormley and Matsa (2011). More specifically, I construct a cohort of cheated and non-cheated firms using firm-year observations for the two years before and the two years after the lawsuits. I then pool the data across 81 cohorts (i.e., across all cheated venture capitals) and estimate the average treatment effect. I use the following regression model as the baseline regression:

$$y_{ict} = \beta Cheated_{ict} \times Post_{ict} + \mathbf{X}'_{ict} + \alpha_{ic} + \delta_{tc} + \epsilon_{ict} \tag{1}$$

where y_{ict} is the measure of venture capital investment. I use two variables: the number of investment deal venture capital i makes in year t (*Deal Num*) and the amount of investment venture capital i makes in year t (*Deal Amount*). $Cheated_{ict}$ is a dummy variable that equals one if venture capital i has invested in one or more fraudulent startups, and zero otherwise. $Post_{ict}$ is the dummy variable that equals one if year $t = 0, 1, 2$, zero if year $t = -2, -1$. $\mathbf{X}'_{ict}$ are a vector of control variables. α_{ic} is the firm-cohort fixed effects. δ_{tc} is the year-cohort fixed effects. ϵ_{ict} is the error term. I allow the firm and year fixed effects to vary by cohort, because this approach is more conservative than including simple fixed effects. Standard errors are clustered at firm level.

3.2 Main Results

The regression estimation results of equation (1) are reported in Table 3. The dependent variable in all column is reported at the top of the table. I include year-cohort and firm-cohort fixed effect in all specifications. All variables are defined in Table A1

I first look at the effect of being cheated on the VCs' following up investment activities in all industries. Column (1) and (2) report the results of VC investment amount. When no control variables are added in the regression specification, the estimated coefficient of *Post* $\times$ *Cheated* is statistically significant and negative (-0.743). However, the estimated coefficient becomes not significant even at 10% level and the magnitude declines a lot from -0.743 to -0.147 when control variables are included in the specification. These results suggest that after being cheated by scam startups, VCs do not decrease their total investment too much compared to other VCs which are not cheated by scam startups. Column (3) and (4) report results on the number of VC investment deals. The estimated coefficients of $Post \times Cheated$ are both statistically significant and negative (-0.167 and -0.064) with and without control variables, suggesting that cheated VCs will decrease their investment deals by more after the event compared to those not being cheated VCs. This clearly supports the hypothesis that learning affects the likelihood of making an investment. In addition to being statistically significant, the estimated coefficient measures an economically important effect. I focus on column (4), it suggests affected VCs will decrease the number of investment deals by 6.4% afterwards.

I then look at VC investments in the specific industries, the industries in which fraudulent startups belong to. The results are reported in columns (5) to (8). The dependent variable in column (5) to (6) is the amount of investment in affected industries and in column (7) to (8) is the number of deals a venture capital makes in them. I include year-cohort and firm-cohort fixed effect in all specifications. Control variables are included in column (6) and (8).

I find VCs will decrease their investment in industries where frauds happened. As column (6) shows, the estimated coefficient of $Post \times Cheated$ is statistically significant and negative (-0.350), suggesting after being cheated by scam startups, VCs will decrease their investment in those indutries by 35%. Column (3) and (4) report the results of VC investment deals. The estimated coefficient of $Post \times Cheated$ is statistically significant and negative (-0.084), suggesting VCs will decrease their investment deals in those industries by 8.4%. Considering all findings in the baseline, affected VCs will reallocate their capital into the industries where frauds do not happen. Affected VCs learn from scam startups investment failure experience, they change their future investment strategies.

4 Robustness Tests

As prior literature have pointed out, staggered DID designs often do not provide valid estimations (Baker et al. (2021); Sun and Abraham (2021); Callaway and Sant'Anna (2021)). Goodman-Bacon (2021) documents that staggered DiD approach is to take weighted average of all possible two-group/two-period DiD estimators, and treatment effect estimates are skewed by comparisons between earlier-treated to later-treated when there are heterogeneity in the average treatment effect on the treated (ATT). This suggests that the baseline estimator is the combination of staggered treatment timing and treatment effect heterogeneity, either across groups or over time (i.e., dynamic treatment effects), leads to biased DiD estimates.

In addition, simply comparing cheated VCs to non-cheated VCs could cause potential possible endogeneity and spurious correlation between the VCs investment activities and impacts of scam startups. Because as Table 2 shows, affected VCs are totally different from unaffected VCs. Cheated VCs take a small part of total VCs, but they are more experienced, larger and tend to invest more money in more deals. This suggests that simply comparing cheated VCs to non-cheated VCs could cause biased estimation because of possible endogeneity. Noted that cheated VCs have more investments, which means it is more possible for them to invest in scam startups compared to not being cheated VCs.

To overcome those concerns, I construct several alternative control groups to identify peer unaffected VCs that are otherwise similar to the affected ones.

4.1 Alternative control group

In the baseline sample, I choose all VCs which have never invested in scam startups as the control group. To avoid potential possible endogeneity problem I discuss above, in the robustness check, I narrow control group selection criteria into those which have never invested in scam startups but have invested at least once in industries where frauds happen. Table 4 reports the results. The sample is smaller than the baseline. I find VCs do not decrease their total investment too much compared to other VCs which are not cheated by scam startups. But they will make less investment in all industries and fraudulent industries. The results are similar to the baseline. As column (6) and (8) show, the estimated coefficients of $Post \times Cheated$ are statistically significant and negative.

4.2 PSM-DID method

For the analysis that follows, I select a sample of peer VCs that have never been cheated. It is important to carefully select peer VCs that are as similar to the affected VCs as possible, as otherwise my tests could be biased by possible endogeneity. To allay concerns of endogeneity, I identify peer VCs that have never been cheated but are as similar to cheated VCs along all four performance proxies—age, the number of deals, funds under management, the number of companies they invest in, they are calculated as cumulative number in the year of the investment. I employ the most commonly used methodologies–propensity score matching (PSM) method for matching. Of 217 VCs, I am able to find 165 cheated VCs with 165 peers. The final sample comprises 330 unique VCs (165 treated + 165 matched). The regression is the following.

$$y_{it} = \beta Cheated_{it} \times Post_{it} + \mathbf{X}'_{it} + \alpha_i + \delta_t + \epsilon_{it} \tag{2}$$

where y_{it} is the measure of venture capital investment. $Cheated_{it}$ is a dummy variable that equals one if venture capital i has invested in one or more fraudulent startups, and zero otherwise. $Post_{it}$ is the dummy variable that equals one if year $t = 0, 1, 2$, zero if year $t = -2, -1$. $\mathbf{X}'_{it}$ are a vector of control variables. α_i is the firm fixed effects. δ_t is the year fixed effects. ϵ_{it} is the error term. The results are presented in Table 5. The results are robust to the baseline.

5 Heterogeneity Tests

5.1 VC investment experience

The decline in the posterior belief is stronger for those who had higher belief, as the participants in the venture become more pessimistic about the likelihood of success (Bergemann and Hege (1998)). If VCs have more positive belief on a startup, they invest more and have higher evaluation.

Therefore I hypothesize that if VCs have invested more in scam startups, they could decline belief and investment afterwards. In addition, if startups have higher valuation, VCs have more confidence and belief in them, then VCs future investments could drop a lot.

To test the first hypothesis, I classify the affected VCs sample into two subgroups according to the amount and number of deals they made in scam startups. Table 6 and Table 7 present the results. Table 6 divides group according to the amount VCs made. The explanation variable is Post $\times$ Cheated $\times$ Amount, which is the interaction among Post, Cheated and Amount. Cheated is a dummy variable that takes the value of one if the venture capital invested in scam startups, and takes the value of zero otherwise. Post is a dummy variable that takes the value of one in the years after being cheated. Amount is a dummy variable that equals one if VC invested more amount in fraudulent startups. The results support my hypothesis that for VCs which made larger investment, they will decline more investment after being cheated. Table 7 divides group according to the number of deal VCs made. The results support my hypothesis.

5.2 Startups valuation

In this section, I divide sample depending on the valuation of scam startups VCs invested in to test the second hypothesis. I use the total amount of investment a startup received as the measurement of its valuation. The results are reported in Table 8. The independent variable is Post $\times$ Cheated $\times$ Valuation, which is the interaction among Post, Cheated and Valuation. Cheated is a dummy variable that takes the value of one if the venture capital invested in scam startups, and takes the value of zero otherwise. Post is a dummy variable that takes the value of one in the years after being cheated. Valuation is a dummy variable that equals one if VC invested in higher valuation startup. The coefficients are negative and significant, which mean if VCs have invested in scam startups which have higher valuation, they will decrease investment more in the future.

6 Conclusion

This paper examines the effect of booming scam startups on venture capital investment. I look at this question by constructing a novel database. Using DiD method and a battery of robustness checks, I find that VCs decrease their investment after being cheated by fake startups with respect to the amount of investment and the number of deals. The effect is more pronounced for VC firms which invested more in scam startups or invested in high valuated scam startups. This finding reveals that VCs investments are affected by past failure experience, learning and trust play important roles in VCs investment decisions.

References

Atanasov, Vladimir, Vladimir Ivanov, and Kate Litvak, 2012, Does reputation limit opportunistic behavior in the vc industry? evidence from litigation against vcs, *Journal of Finance* 67, 2215–2246.

Baker, Andrew, David F Larcker, and Charles CY Wang, 2021, How much should we trust staggered difference-in-differences estimates?, *Available at SSRN 3794018* .

Bergemann, Dirk, and Ulrich Hege, 1998, Venture capital financing, moral hazard, and learning, *Journal of Banking & Finance* 22, 703–735.

Bernstein, Shai, Arthur Korteweg, and Kevin Laws, 2017, Attracting early-stage investors: Evidence from a randomized field experiment, *Journal of Finance* 72, 509–538.

Berry, Donald A, and Bert Fristedt, 1985, Bandit problems: sequential allocation of experiments (monographs on statistics and applied probability), *London: Chapman and Hall* 5, 7–7.

Bottazzi, Laura, Marco Da Rin, and Thomas Hellmann, 2016, The importance of trust for investment: Evidence from venture capital, *Review of Financial Studies* 29, 2283–2318.

Callaway, Brantly, and Pedro HC Sant'Anna, 2021, Difference-in-differences with multiple time periods, *Journal of Econometrics* 225, 200–230.

Choi, Stephen J, 2007, Do the merits matter less after the private securities litigation reform act?, *Journal of Law, Economics, & Organization* 23, 598–626.

Cumming, Douglas, Bruce Haslem, and April Knill, 2017, Entrepreneurial litigation and venture capital finance, *Journal of Financial and Quantitative Analysis* 52, 2217–2250.

Duffner, Stefan, Markus M Schmid, and Heinz Zimmermann, 2009, Trust and success in venture capital financing—an empirical analysis with german survey data, *Kyklos* 62, 15–43.

Dyck, Alexander, Adair Morse, and Luigi Zingales, 2010, Who blows the whistle on corporate fraud?, *Journal of Finance* 65, 2213–2253.

Gleason, Kimberly C, Yezen H Kannan, and Christian Rauch, 2021, Fraud in startups: What stakeholders need to know, *Available at SSRN 3978552* .

Gompers, Paul, Steven N Kaplan, and Vladimir Mukharlyamov, 2016, What do private equity firms say they do?, *Journal of Financial Economics* 121, 449–476.

Gompers, Paul A, Will Gornall, Steven N Kaplan, and Ilya A Strebulaev, 2020, How do venture capitalists make decisions?, *Journal of Financial Economics* 135, 169–190.

Goodman-Bacon, Andrew, 2021, Difference-in-differences with variation in treatment timing, *Journal of Econometrics* .

Gormley, Todd A, and David A Matsa, 2011, Growing out of trouble? Corporate responses to liability risk, *Review of Financial Studies* 24, 2781–2821.

Graham, John R, and Campbell R Harvey, 2001, The theory and practice of corporate finance: Evidence from the field, *Journal of financial economics* 60, 187–243.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, 2008, Trusting the stock market, *Journal of Finance* 63, 2557–2600.

Gurun, Umit G, Noah Stoffman, and Scott E Yonker, 2018, Trust busting: The effect of fraud on investor behavior, *Review of Financial Studies* 31, 1341–1376.

Hassan, Tarek A, Stephan Hollander, Laurence Van Lent, and Ahmed Tahoun, 2019, Firm-level political risk: Measurement and effects, *The Quarterly Journal of Economics* 134, 2135–2202.

Heese, Jonas, Ranjani Krishnan, and Hari Ramasubramanian, 2021, The department of justice as a gatekeeper in whistleblower-initiated corporate fraud enforcement: Drivers and consequences, *Journal of Accounting and Economics* 71, 101357.

Kaplan, Steven N, and Per Strömberg, 2000, How do venture capitalists choose investments, *Workng Paper, University of Chicago* 121, 55–93.

Kaplan, Steven N, and Per ER Strömberg, 2004, Characteristics, contracts, and actions: Evidence from venture capitalist analyses, *Journal of Finance* 59, 2177–2210.

Karpoff, Jonathan M, D Scott Lee, and Gerald S Martin, 2008, The cost to firms of cooking the books, *Journal of Financial and Quantitative Analysis* 43, 581–611.

Li, Bo, and Lauren Cohen, 2021, The political economy of anti-bribery enforcement, *Available at SSRN 3994355* .

Sorensen, Morten, 2008, Learning by investing: Evidence from venture capital, in *AFA 2008 New Orleans Meetings Paper*.

Sun, Liyang, and Sarah Abraham, 2021, Estimating dynamic treatment effects in event studies with heterogeneous treatment effects, *Journal of Econometrics* 225, 175–199.

FOR IMMEDIATE RELEASE Friday, June 15, 2018

Theranos Founder and Former Chief Operating Officer Charged In Alleged Wire Fraud Schemes

Elizabeth Holmes and Ramesh "Sunny" Balwani Are Alleged To Have Perpetrated Multi-million Dollar Schemes To Defraud Investors, Doctors, and Patients.

SAN JOSE - A federal grand jury has indicted Elizabeth A. Holmes and Ramesh "Sunny" Balwani, announced Acting United States Attorney Alex G. Tse, Federal Bureau of Investigation (FBI) Special Agent in Charge John F. Bennett; Food and Drug Administration (FDA) Commissioner Scott Gottlieb; and U.S. Postal Inspection Service (USPIS) Inspector in Charge Rafael Nuñez. The defendants are charged with two counts of conspiracy to commit wire fraud and nine counts of wire fraud. According to the indictment returned yesterday and unsealed today, the charges stem from allegations Holmes and Balwani engaged in a multi-million dollar scheme to defraud investors, and a separate scheme to defraud doctors and patients. Both schemes involved efforts to promote Palo Alto, Calif.-based Theranos.

Holmes, 34, of Los Altos Hills, Calif., founded Theranos in 2003. Theranos is a private health care and life sciences company with the stated mission to revolutionize medical laboratory testing through allegedly innovative methods for drawing blood, testing blood, and interpreting the resulting patient data. Balwani, 53, of Atherton, Calif., was employed at Theranos from September of 2009 through 2016. At times during that period, Balwani worked in several capacities including as a member of the company's board of directors, as its president, and as its chief operating officer.

(a) DOJ Lawsuit Example

SEC Charges E-Commerce Startup and CEO with Defrauding Investors

Litigation Release No. 24968 / November 23, 2020

Securities and Exchange Commission v. Benja Inc, et al., No. 3:20-cv-08238 (N.D. Cal. filed November 23, 2020)

The Securities and Exchange Commission today charged a San Francisco, California-based e-commerce start-up and its chief executive officer with misleading investors about purported contracts with well-known consumer brands.

According to the SEC's complaint, from 2018 to 2020, Andrew J. Chapin, the founder and CEO of Benja Inc., told investors that Benja was a successful online advertising platform that generated millions of dollars in revenue from popular consumer clothing brands and retailers. In reality, as the complaint alleges, Benja never did business with the companies. The complaint further alleges that in order to secure investments, Chapin enlisted one or more associates to help induce investments from venture capital investors by impersonating representatives of Benja's purported customers and the supposed founder of a venture capital fund who falsely claimed to have made a large investment in Benja. According to the complaint, Chapin also provided an investor with forged contracts and doctored bank statements.

(b) SEC Lawsuit Example

Figure 1: DOJ and SEC Lawsuits Examples

This figure gives lawsuit excerpts from DOJ and SEC website.

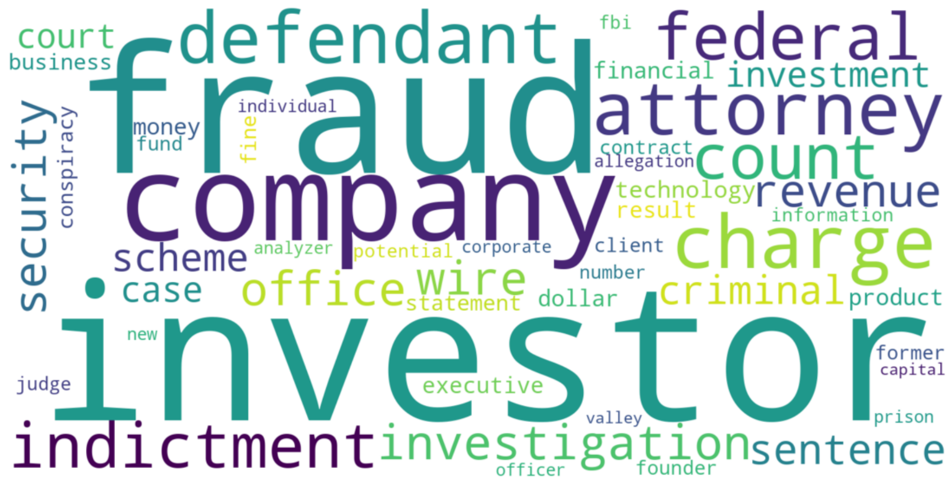


Figure 2: Key Words Word Cloud

This figure plots the word cloud of key words of ten representative scam startups lawsuits contents.

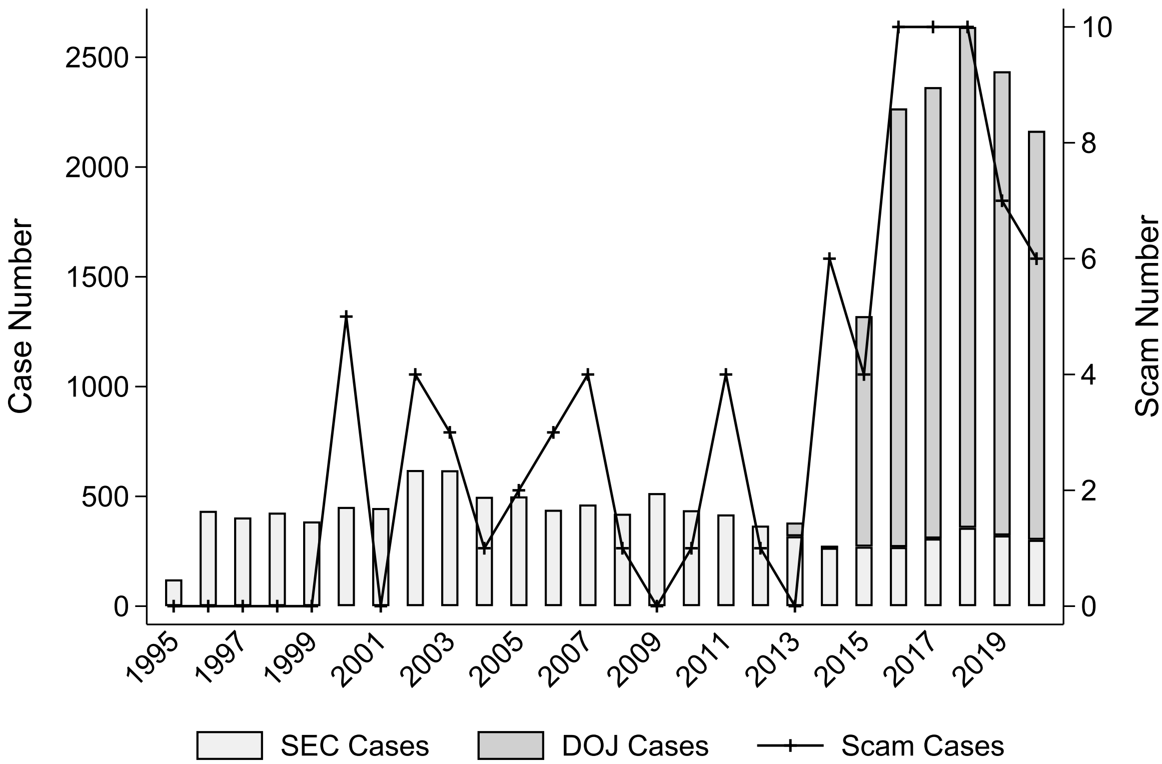
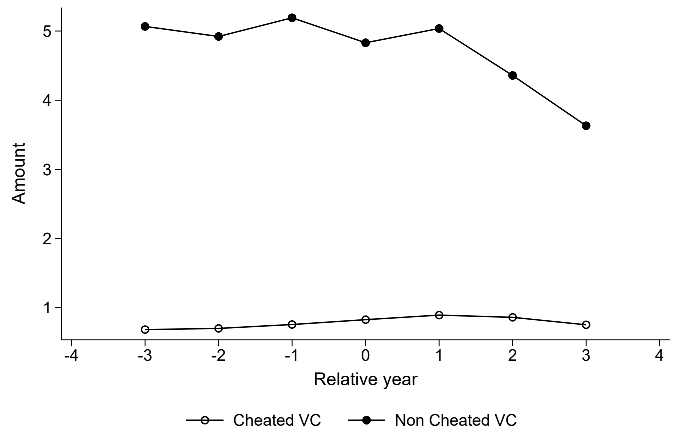
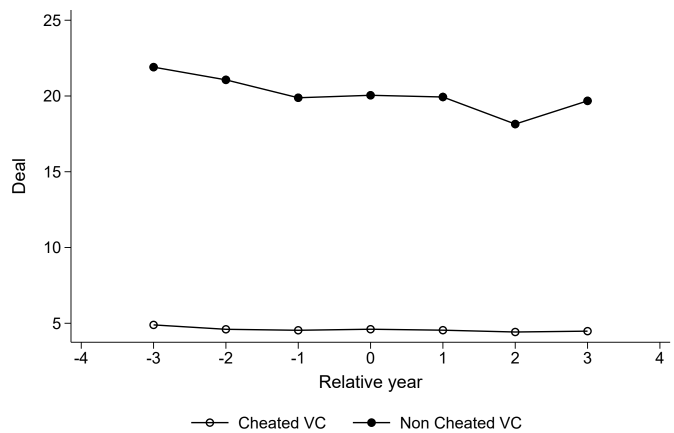


Figure 3: Time Series of the Number of Lawsuits

This figure plots the time distribution of lawsuits over the 1995–2020 period. The bar chat reports distribution of all cases downloaded from DOJ and SEC. The line shows distribution of cases the in final sample.



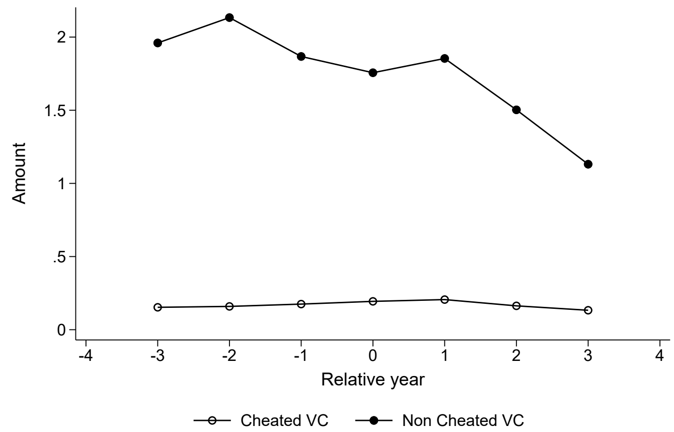
(a) VC Investment Amount



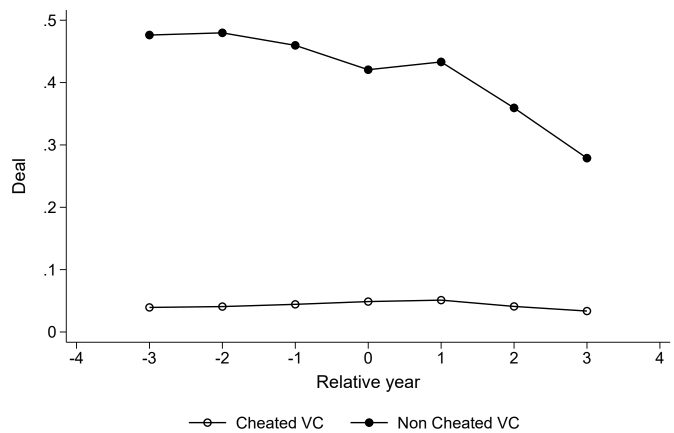
(b) VC Investment Deal

Figure 4: VC Investment

This figure shows the VCs investment in all industries around the event of the announcement of scam startups. Figure(a) reports the amount changes around the event. Figure(b) reports the number of deal changes.



(a) VC Investment Amount



(b) VC Investment Deal

Figure 5: VC Investment in Frauds Industries

This figure shows the VCs investment in industries where frauds happen around the event. Figure(a) reports the amount changes. Figure(b) reports the number of deal changes around the event.

Table 1: Summary Statistics

This table reports the summary statistics of main variables used in this paper. All variables are defined in the Appendix A1. All variables are transformed to natural logarithm. All continuous variables are winsorized at 1% level.

	N	Mean	SD	Min	P25	Median	P75	Max
Deal Amount	3,516,219	4.13	2.83	0.00	0.80	4.71	6.24	9.65
Deal Num	3,516,219	1.33	0.72	0.69	0.69	1.10	1.79	3.69
Deal Amount IND	813,441	4.12	2.54	0.00	2.67	4.60	5.93	12.96
Deal Num IND	813,441	1.14	0.62	0.69	0.69	0.69	1.39	5.03
Fund Year	3,516,219	0.57	0.86	0.00	0.00	0.00	1.10	2.89
Company Num	3,516,219	2.51	1.64	0.00	1.10	2.48	3.71	6.32
VC Age	3,491,273	2.21	1.00	0.00	1.61	2.30	2.94	4.38
Funds Num	3,516,219	1.79	1.27	0.00	0.69	1.79	2.77	4.68
Total Deal Amount	3,142,171	7.14	2.21	2.07	5.57	7.20	8.75	11.92
Total Deal Num	3,516,219	2.58	1.69	0.00	1.10	2.48	3.83	6.54

Table 2: Difference Test

This table reports the difference tests results. Panel A reports the difference between cheated VC and non cheated VC. Panel B reports the difference between scam startups and non scam startups. All variables are defined in the Appendix A1. All continuous variables are winsorized at 1% level.

Panel A: Difference of Non Cheated VC and Cheated VC

	(1)	(2)	(3)	(4)	(5)	(6)
	Non Cheated VCs		Cheated VCs		(2) − (4)	
	N	Mean	N	Mean	Difference	t-value
Deal Num	333,362	383.33	95,190	1,302.29	-918.97	-313.49
Deal Amount	333,362	6,024,632.66	95,190	6,836,163.37	-811,530.71	-4.21
Company Num	333,362	181.45	95,190	519.69	-338.24	-230.71
Found Year	333,362	1,991.32	95,190	1,980.39	10.93	160.11
Fund Num	333,362	11.85	95,190	25.39	-13.55	-177.71

Panel B: Difference of Scam Stratups and Real Startups

	(1)	(2)	(3)	(4)	(5)	(6)
	Real Startups		Scam Startups		(2) − (4)	
	N	Mean	N	Mean	Difference	t-value
Equity	483,385	192.39	512	256.62	-64.22	-0.99
Debt	483,385	1,223,853.61	512	585,923.83	637,929.78	0.26
Company Found Year	211,831	2,005.40	247	2,005.40	-0.00	-0.00

Table 3: Baseline Results

This table reports the baseline results on the effect of being cheated by scam startups on venture capital investment activities. In column (1)–(2), the dependent variable is the amount of investment in all industries. In column (3)–(4), the dependent variable is the number of deals in all industries. In column (5)–(6), the dependent variable is the amount of investment in the fraudulent industries. In column (7)–(8), the dependent variable is the number of deals of investment in the fraudulent industries. The dependent variables are transformed into natural logarithm form. The main independent variable is Post × Cheated. Cheated is a dummy variable that takes the value of one if the venture capital invested in scam startups, and zero otherwise. Post is a dummy variable that takes the value of one in the years after being cheated. All other variables are defined in the Appendix A1. All variables are transformed to natural logarithm. All specifications include year-cohort and firm-cohort fixed effects. All continuous variables are winsorized at 1% level. Robust standard errors are applied. t-statistics are reported within parentheses under the estimates. *, **, and *** indicate that the coefficient is statistically significant at the 10%, 5%, and 1% level,respectively.

	All Industries				Frauds Industries			
	(1) Deal Amount	(2) Deal Amount	(3) Deal Num	(4) Deal Num	(5) Deal Amount	(6) Deal Amount	(7) Deal Num	(8) Deal Num
Post × Cheated	-0.743***	-0.147	-0.167***	-0.064**	-0.437***	-0.254**	-0.101***	-0.055**
	(-7.47)	(-1.60)	(-5.63)	(-2.49)	(-5.16)	(-2.17)	(-5.99)	(-2.44)
Fund Year		-0.076***		-0.066***		-0.020***		-0.010***
		(-23.93)		(-86.95)		(-8.10)		(-18.65)
Company Num		-0.692***		-0.409***		0.211***		0.009**
		(-29.32)		(-74.49)		(10.64)		(2.05)
VC Age		-0.286***		-0.218***		-0.156***		-0.047***
		(-47.71)		(-142.46)		(-25.74)		(-39.22)
Funds Num		-1.746***		-0.941***		-0.648***		-0.183***
		(-246.97)		(-497.92)		(-91.09)		(-123.44)
Total Deal Amount		1.827***		-0.011***		0.242***		-0.005***
		(548.60)		(-19.83)		(104.22)		(-10.18)
Total Deal Num		0.510***		1.595***		0.323***		0.224***
		(22.02)		(298.66)		(17.16)		(52.29)
Year × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
N	18,589,788	2,807,239	3,120,368	2,807,239	18,589,788	2,807,239	18,589,788	2,807,239
Adj. R^2	0.56	0.66	0.67	0.79	0.50	0.67	0.54	0.71

Table 4: Alternative Control Group

This table reports the regression analyses based on a sample of VCs that have never invested in scam startups but have invested in their industries. In column (1)–(2), the dependent variable is the amount of investment in all industries. In column (3)–(4), the dependent variable is the number of deals in all industries. In column (5)–(6), the dependent variable is the amount of investment in the fraudulent industries. In column (7)–(8), the dependent variable is the number of deals of investment in the fraudulent industries. The dependent variables are transformed into natural logarithm form. The main independent variable is Post × Cheated. Cheated is a dummy variable that takes the value of one if the venture capital invested in scam startups, and zero otherwise. Post is a dummy variable that takes the value of one in the years after being cheated. All other variables are defined in the Appendix A1. All variables are transformed to natural logarithm. All specifications include year-cohort and firm-cohort fixed effects. All continuous variables are winsorized at 1% level. Robust standard errors are applied. t-statistics are reported within parentheses under the estimates. *, **, and *** indicate that the coefficient is statistically significant at the 10%, 5%, and 1% level,respectively.

	All Industries				Frauds Industries			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Deal Amount	Deal Amount	Deal Num	Deal Num	Deal Amount	Deal Amount	Deal Num	Deal Num
Post × Cheated	-0.659***	-0.139	-0.247***	-0.049**	-0.304***	-0.174**	-0.065***	-0.040***
	(-6.57)	(-1.50)	(-9.25)	(-2.01)	(-4.87)	(-2.07)	(-5.59)	(-2.58)
Fund Year		-0.074***		-0.063***		-0.007***		-0.003***
		(-17.66)		(-64.03)		(-3.27)		(-5.75)
Company Num		-0.783***		-0.451***		0.396***		0.084***
		(-21.56)		(-53.94)		(21.19)		(21.25)
VC Age		-0.376***		-0.255***		-0.129***		-0.032***
		(-41.33)		(-105.07)		(-20.88)		(-25.76)
Funds Num		-1.605***		-0.892***		-0.310***		-0.064***
		(-157.31)		(-317.19)		(-43.11)		(-43.89)
Total Deal Amount		1.867***		-0.010***		0.037***		-0.007***
		(374.74)		(-11.10)		(17.80)		(-14.76)
Total Deal Num		0.570***		1.629***		-0.009		0.016***
		(15.98)		(200.97)		(-0.48)		(4.34)
Year × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
N	6,432,265	1,549,363	6,432,265	1,549,363	6,432,265	1,549,363	6,432,265	1,549,363
Adj. R^2	0.61	0.67	0.72	0.78	0.26	0.30	0.26	0.28

Table 5: PSM-DID

This table reports the regression analyses based on a sample of VCs selected by propensity score matching method. In column (1)–(2), the dependent variable is the amount of investment in all industries. In column (3)–(4), the dependent variable is the number of deals in all industries. In column (5)–(6), the dependent variable is the amount of investment in the fraudulent industries. In column (7)–(8), the dependent variable is the number of deals of investment in the fraudulent industries. The dependent variables are transformed into natural logarithm form. The main independent variable is Post × Cheated. Cheated is a dummy variable that takes the value of one if the venture capital invested in scam startups, and zero otherwise. Post is a dummy variable that takes the value of one in the years after being cheated. All other variables are defined in the Appendix A1. All variables are transformed to natural logarithm. All specifications include year-cohort and firm-cohort fixed effects. All continuous variables are winsorized at 1% level. Robust standard errors are applied. t-statistics are reported within parentheses under the estimates. *, **, and *** indicate that the coefficient is statistically significant at the 10%, 5%, and 1% level,respectively.

	All Industries				Frauds Industries			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Deal Amount	Deal Amount	Deal Num	Deal Num	Deal Amount	Deal Amount	Deal Num	Deal Num
Post × Cheated	-0.521	-0.397	-0.134***	-0.097**	-0.291**	-0.262**	-0.079**	-0.072**
	(-1.56)	(-1.25)	(-2.88)	(-2.27)	(-2.25)	(-2.08)	(-2.45)	(-2.31)
Fund Year		-1.029**		-0.057		-0.112		-0.028
		(-2.48)		(-1.17)		(-0.70)		(-0.75)
Company Num		4.119		0.638		-0.945		0.030
		(1.46)		(1.57)		(-0.64)		(0.10)
VC Age		-2.004***		-0.293***		-0.158		0.025
		(-3.33)		(-3.80)		(-0.79)		(0.47)
Funds Num		-0.808		-0.714***		-1.512***		-0.652***
		(-0.52)		(-4.82)		(-2.72)		(-4.85)
Total Deal Amount		1.042***		0.105***		0.283***		0.006
		(3.03)		(3.66)		(3.07)		(0.25)
Total Deal Num		-3.721		0.035		1.857		0.482*
		(-1.50)		(0.09)		(1.33)		(1.69)
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y
N	2,068	2,068	2,068	2,068	1,984	1,984	1,984	1,984
Adj. R^2	0.63	0.65	0.85	0.88	0.78	0.80	0.86	0.88

Table 6: VCs Investment Amount

This table reports the regression results for the effects of being cheated for higher and lower venture capital investment. In column (1)–(2), the dependent variable is the amount of investment in all industries venture capital make. In column (3)–(4), the dependent variable is the deal of investment in all industries venture capital make. In column (5)–(6), the dependent variable is the amount of investment in the fraudulent industries venture capital make. In column (7)–(8), the dependent variable is the deal of investment in the fraudulent industries venture capital make. The independent variable is Post $\times$ Cheated $\times$ Amount, which is the interaction among Post, Cheated and Amount. Cheated is a dummy variable that takes the value of one if the venture capital invested in scam startups, and takes the value of zero otherwise. Post is a dummy variable that takes the value of one in the years after being cheated. Amount is a dummy variable that equals one if VCs made higher investment in scam startups. Control variables include last fund year, total known companies, age, funds, total equity and total deal. All variables are defined in the Appendix A1. All variables are transformed to natural logarithm. All specifications include year-cohort and firm-cohort fixed effects. All continuous variables are winsorized at 1% level. Robust standard errors are applied. t-statistics are reported within parentheses under the estimates. *, **, and *** indicate that the coefficient is statistically significant at the 10%, 5%, and 1% level,respectively.

	All Industries				Frauds Industries			
	(1) Deal Amount	(2) Deal Amount	(3) Deal Num	(4) Deal Num	(5) Deal Amount	(6) Deal Amount	(7) Deal Num	(8) Deal Num
Post × Cheated × Amount	-0.615***	-0.115*	-0.117***	-0.057***	-0.386***	-0.189**	-0.083***	-0.037**
	(-7.82)	(-1.85)	(-5.42)	(-3.07)	(-5.29)	(-2.15)	(-6.09)	(-2.34)
Fund Year		-0.076***		-0.066***		-0.020***		-0.010***
		(-23.93)		(-86.95)		(-8.10)		(-18.66)
Company Num		-0.692***		-0.409***		0.211***		0.009**
		(-29.32)		(-74.50)		(10.64)		(2.05)
VC Age		-0.286***		-0.218***		-0.156***		-0.047***
		(-47.71)		(-142.47)		(-25.75)		(-39.22)
Funds Num		-1.746***		-0.941***		-0.648***		-0.183***
		(-246.97)		(-497.92)		(-91.09)		(-123.44)
Total Deal Amount		1.827***		-0.011***		0.242***		-0.005***
		(548.60)		(-19.83)		(104.22)		(-10.18)
Total Deal Num		0.510***		1.595***		0.323***		0.224***
		(22.02)		(298.67)		(17.16)		(52.29)
Year × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
N	18,589,788	2,807,239	3,120,368	2,807,239	18,589,788	2,807,239	18,589,788	2,807,239
Adj. R^2	0.56	0.66	0.67	0.79	0.50	0.67	0.54	0.71

Table 7: VCs Investment Number

This table reports the regression results for the effects of being cheated for more and less venture capital investment. In column (1)–(2), the dependent variable is the amount of investment in all industries venture capital make. In column (3)–(4), the dependent variable is the deal of investment in all industries venture capital make. In column (5)–(6), the dependent variable is the amount of investment in the fraudulent industries venture capital make. In column (7)–(8), the dependent variable is the deal of investment in the fraudulent industries venture capital make. The independent variable is Post × Cheated × Deal, which is the interaction among Post, Cheated and Deal. Cheated is a dummy variable that takes the value of one if the venture capital invested in scam startups, and takes the value of zero otherwise. Post is a dummy variable that takes the value of one in the years after being cheated. Deal is a dummy variable that equals one if VCs made more investment in scam startups. Control variables include last fund year, total known companies, age, funds, total equity and total deal. All variables are defined in the Appendix A1. All variables are transformed to natural logarithm. All specifications include year and firm fixed effects. All continuous variables are winsorized at 1% level. Robust standard errors are applied. t-statistics are reported within parentheses under the estimates. *, **, and *** indicate that the coefficient is statistically significant at the 10%, 5%, and 1% level,respectively.

	All Industries				Frauds Industries			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Deal Amount	Deal Amount	Deal Num	Deal Num	Deal Amount	Deal Amount	Deal Num	Deal Num
Post × Cheated × Deal	-0.680***	-0.112	-0.164***	-0.070***	-0.412***	-0.212**	-0.093***	-0.045**
	(-7.64)	(-1.43)	(-6.19)	(-2.96)	(-5.13)	(-1.99)	(-5.88)	(-2.24)
Fund Year		-0.076***		-0.066***		-0.020***		-0.010***
		(-23.93)		(-86.95)		(-8.10)		(-18.65)
Company Num		-0.692***		-0.409***		0.211***		0.009**
		(-29.32)		(-74.49)		(10.64)		(2.05)
VC Age		-0.286***		-0.218***		-0.156***		-0.047***
		(-47.71)		(-142.47)		(-25.75)		(-39.22)
Funds Num		-1.746***		-0.941***		-0.648***		-0.183***
		(-246.97)		(-497.92)		(-91.09)		(-123.44)
Total Deal Amount		1.827***		-0.011***		0.242***		-0.005***
		(548.60)		(-19.83)		(104.22)		(-10.18)
Total Deal Num		0.510***		1.595***		0.323***		0.224***
		(22.02)		(298.66)		(17.16)		(52.29)
Year × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
N	18,589,788	2,807,239	3,120,368	2,807,239	18,589,788	2,807,239	18,589,788	2,807,239
Adj. R^2	0.56	0.66	0.67	0.79	0.50	0.67	0.54	0.71

Table 8: Scam Startups Valuation

This table reports the regression results for the effects of being cheated for higher and lower startups valuation. In column (1)–(2), the dependent variable is the amount of investment in all industries venture capital make. In column (3)–(4), the dependent variable is the deal of investment in all industries venture capital make. In column (5)–(6), the dependent variable is the amount of investment in the fraudulent industries venture capital make. In column (7)–(8), the dependent variable is the deal of investment in the fraudulent industries venture capital make. The independent variable is High Valuation, which is the interaction among Post, Cheated and Valuation. Cheated is a dummy variable that takes the value of one if the venture capital invested in scam startups, and takes the value of zero otherwise. Post is a dummy variable that takes the value of one in the years after being cheated. Valuation is a dummy variable that equals one if VC invested in higher valuation startup. Control variables include last fund year, total known companies, age, funds, total equity and total deal. All variables are defined in the Appendix A1. All variables are transformed to natural logarithm. All specifications include year and firm fixed effects. All continuous variables are winsorized at 1% level. Robust standard errors are applied. t-statistics are reported within parentheses under the estimates. *, **, and *** indicate that the coefficient is statistically significant at the 10%, 5%, and 1% level,respectively.

	All Industries				Frauds Industries			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Deal Amount	Deal Amount	Deal Num	Deal Num	Deal Amount	Deal Amount	Deal Num	Deal Num
Post × Cheated × Valuation	-0.569***	-0.093*	-0.099***	-0.017	-0.381***	-0.200**	-0.083***	-0.040**
	(-7.72)	(-1.65)	(-4.81)	(-1.60)	(-5.43)	(-2.28)	(-6.15)	(-2.45)
Fund Year		-0.076***		-0.043***		-0.020***		-0.010***
		(-23.93)		(-79.45)		(-8.10)		(-18.66)
Company Num		-0.692***		-0.267***		0.211***		0.009**
		(-29.32)		(-69.16)		(10.64)		(2.05)
VC Age		-0.286***		-0.155***		-0.156***		-0.047***
		(-47.71)		(-133.74)		(-25.75)		(-39.22)
Funds Num		-1.746***		-0.719***		-0.648***		-0.183***
		(-246.97)		(-500.38)		(-91.09)		(-123.44)
Total Deal Amount		1.827***		-0.002***		0.242***		-0.005***
		(548.60)		(-5.46)		(104.22)		(-10.18)
Total Deal Num		0.510***		1.142***		0.323***		0.224***
		(22.02)		(304.82)		(17.16)		(52.29)
Year × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
N	18,589,788	2,807,239	3,120,368	2,807,239	18,589,788	2,807,239	18,589,788	2,807,239
Adj. R^2	0.56	0.66	0.67	0.74	0.50	0.67	0.54	0.71

Appendix

Table A1: Definition and Source

This table reports the definition of variables and their data sources.

Variable	Definition	Source
Cheated	Dummy variable which is one if venture captial has invested in scam startups, otherwise is zero	DOJ/SEC
Deal Amount	The amount of investment venture capital makes in a given year	VenturExpert
Deal Num	The number of investment venture capital makes in a given year	VenturExpert
Deal Amount IND	The amount of investment venture capital makes in the frauds industries in a given year	VenturExpert
Deal Num IND	The number of investment deal venture capital makes in the frauds industries in a given year	VenturExpert
Fund Year	The given year minus the last year which venture capital received fundraising	VenturExpert
Company Num	The cumulative number of companies venture capital invested	VenturExpert
VC Age	The age of venture capital	VenturExpert
Fund Num	The cumulative number of fund venture capital manage	VenturExpert
Total Deal Amount	The cumulative amount of investment a venture capital have made	VenturExpert
Total Deal Num	The cumulative number of investment a venture capital have made	VenturExpert

Table A2: Ten Representative Scam Startups

This table reports 10 representative scam startups I use to build words dictionary. The DOJ and SEC links are shown below:

Startups	SEC	DOJ
Theranos	https://www.sec.gov/litigation/litreleases/2018/lr24069.htm	https://www.justice.gov/usao-ndca/pr/theranos-founder-and-former-chief-operating-officer-charged-alleged-wire-fraud-schemes
Telomolecular	https://www.sec.gov/litigation/litreleases/2008/lr20745.htm	https://www.justice.gov/usao-edca/pr/former-rancho-cordova-executive-pleads-guilty-securities-fraud
Savtira	https://www.sec.gov/litigation/litreleases/2015/lr23339.htm	https://www.justice.gov/usao-mdfl/pr/former-executives-defunct-tampa-technology-company-indicted-investment-fraud
Mozido	https://www.sec.gov/litigation/litreleases/2018/lr24092.htm	https://www.justice.gov/usao-me/pr/defendants-charged-multimillion-dollar-investment-fraud
Osiris Therapeutics	https://www.sec.gov/litigation/litreleases/2017/lr23978.htm	https://www.justice.gov/usao-sdny/pr/former-chief-financial-officer-osiris-therapeutics-inc-pleads-guilty-lying-auditors
Outcome Health	https://www.sec.gov/litigation/litreleases/2019/lr24675.htm	https://www.justice.gov/opa/pr/former-executives-and-employees-health-technology-start-charged-1-billion-scheme-defraud
Starship	https://www.sec.gov/litigation/litreleases/2020/lr24812.htm	https://www.justice.gov/usao-sdny/pr/acting-manhattan-us-attorney-and-fbi-assistant-director-announce-securities-and-wire
InfrAegis	https://www.sec.gov/litigation/litreleases/2019/lr24525.htm	https://www.justice.gov/usao-ndil/pr/founder-suburban-tech-company-sentenced-9-years-defrauding-investors-out-more-9-million
Youplus	https://www.sec.gov/litigation/litreleases/2020/lr24854.htm	https://www.justice.gov/usao-ndca/pr/santa-clara-man-charged-running-bogus-artificial-intelligence-investment-fraud-scheme
Benja	https://www.sec.gov/litigation/litreleases/2020/lr24968.htm	https://www.justice.gov/usao-ndca/pr/ceo-charged-securities-and-bank-fraud-alleged-scheme-raise-funds-digital-advertising

Table A3: Startups Dictionary

This table reports the frequency of the top 100 most common words appearing in the ten representative lawsuits contents.

Word	Frequency	Word	Frequency	Word	Frequency
investor	76	client	10	agarwal	6
fraud	65	former	10	liberty	6
company	53	number	9	law	6
charge	37	judge	9	maine	6
attorney	37	information	9	payment	6
defendant	36	potential	9	service	6
federal	29	fbi	9	partner	6
count	28	allegation	8	alleges	6
indictment	28	officer	8	example	6
investigation	24	analyzer	8	venture	6
office	23	prison	8	snack	5
wire	23	corporate	8	david	5
security	22	individual	8	stock	5
revenue	22	new	8	gain	5
sentence	21	commit	7	president	5
criminal	19	restitution	7	field	5
scheme	18	victim	7	director	5
case	16	capital	7	imprisonment	5
investment	16	hubbard	7	inspector	5
court	15	advertising	7	purchase	5
technology	15	valley	7	addition	5
product	14	crime	7	thing	5
financial	14	employee	7	postal	5
money	14	division	7	jersey	5
dollar	14	monster	7	calif	5
executive	13	holmes	7	software	5
result	13	solar	6	appearance	5
business	12	innocent	6	personal	5
statement	12	exchange	6	panel	5
fine	11	guilty	6	chocolate	5
conspiracy	11	share	6	public	5
contract	10	credit	6	balwani	5
founder	10	commission	6		
fund	10	account	6		