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

Yanying Lyu* Lingbo Shen†

August 6, 2023

Abstract

Scam startups are on the rise in recent years. However, little is known about the impact of those startups on venture capitals' (VC) investment activities. In this paper, we first construct a novel dataset of scam startups using the release of news from the Securities and Exchange Commission (SEC) and the U.S. Department of Justice (DOJ). We then investigate how venture capitalists react after they are cheated by scam startups using a difference-in-differences framework. The main finding is that VCs update their beliefs on new startups after the scams outbreak because the bad signals have been released. VCs reduce investments in new startups and the number of deals compared to a control group of VCs that have not been cheated. The effect is mainly driven by a decline in investment in the industries of scam startups. VCs strengthen screening and monitoring process afterwards. We also find that limited partners (LPs) provide less capital to VCs after they financed a scam startup.

Keywords: Venture Capital; Scam; Startups; Investment; Fundraising

JEL Classification: G24; G30; G41; M13

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

†HBE Department, University of Twente; Email: l.shen-1@utwente.nl

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 the dark sides of emerging scam startups.[1] Theranos, valued at 9 billion dollars in a 2014 funding round, "fakes" the existence of its core product and uses it to raise billions of dollars in investment. This kind of scam startup has been on the rise in recent years. Examples including Theranos, Zenefits, Mozido, and Hampton Creek have all hit the headlines recently. Silicon Valley is even called "a land of fake promise". Startups "fake" having a product or technology to attract potential customers and lure investors into funding the sale and marketing of a such product, perhaps they will "make" ideas come true in the future. Considering billions of dollars have been poured into scam startups, as well as the reputation of investors and trust among venture capital partners are destroyed, it is crucial for both the academic and industry to gain a better insight into the effects of scam startups on the financial market.

However, there is little evidence in the previous literature. The main reason is that no standard scam startups database exits. In this paper, we fill this gap by constructing a novel dataset of scam startups from official releases to investigate how involved VCs adjust their investment activities when startups they invested in turn out to be scams.

The two main data sources we use in this paper are Securities and Exchange Commission (SEC) and U.S. Department of Justice (DOJ). SEC and DOJ provide detailed lawsuits information against behaviors violating federal rules or securities laws. 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 our 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. We use the web crawling to get 10,381 cases released by SEC during 1995 to 2020 and 11,672 cases from DOJ during 2013 to 2020. The processes of one case refer to the following steps in the legal proceedings: initial appearance, arraignment, the guilt stage, and the sentencing stage. While some enforcement actions may be resolved quickly and simultaneously upon the initial release of information about the case, the majority of actions typically unfold over multiple regulatory proceedings that can take several years to fully resolve. To alleviate the potential concern that the observed effect may be driven by the lawsuits themselves, rather than the underlying scam, we conduct additional empirical tests and robustness checks. Specifically, we separate the cases in the dataset into two groups: those that resulted in a guilty verdict and those that resulted in a not guilty verdict. By comparing the results of these two groups, we can help ensure that any observed effects are indeed driven by the underlying scam, rather than simply by the occurrence of a lawsuit or legal action.

Theoretically, each investment brings not only a monetary payoff for VCs but also more information. By

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

learning from their past investment returns, investors can gain insights into which industries and startups are worth investing in, and which ones are not. Sorensen (2008) estimates his learning model by using U.S. data from venture capital investments. He finds that VCs' investment decisions are affected not only by immediate returns but also by the option value of learning. Furthermore, he demonstrates that the success of VCs is predicated on their past learning experiences, indicating that VCs who engage in more learning tend to be more successful. Goldfarb et al. (2007) find learning behavior is derived from the investment histories across VCs. Using a model that links past investments and outcomes to subsequent investments within the same industry, they find that investors only learn from their own past investments. The effects of this learning are more significant in the industries where they have previously invested. Similarly, VCs learn from their past failure experience after financing scam startups. As a result, they improve their understanding of various investment opportunities, particularly in industries where scams have occurred.

In addition, according to modern information-based theories, markets discipline financial intermediaries that fail as delegated monitors (Dehejia and Wahba (2002); Boyd and Prescott (1986)). Specifically, a VC firm's reputation is affected when it fails to stop preventing fraud in its portfolio companies. VCs with damaged reputations are less likely to attract new startups in the future because they are perceived as ineffective monitors (Tian et al. (2016)). In other words, after companies are exposed as scams, VCs' reputations are destroyed, which affects their future attractiveness to startups.

Considering the above theories, we hypothesis that VCs learn from their experiences of failure after being deceived by scam startups. They adjust their beliefs about new startups and improve their future decisions regarding various investment opportunities, particularly from industries where scams happen.

To test these hypotheses, we first construct the novel dataset from SEC and DOJ. Following previous literature collecting scam samples (Dyck et al. (2010); Choi (2007); Ahern (2017)), we start by using federal securities class actions to construct the novel sample of scam startups. We use a web crawler to download all cases related to financial frauds from DOJ and SEC. The resulting large sample consists of 11,672 digital cases initialed by DOJ and 10,381 digital lawsuits released by SEC. To identify the scam startup cases from all frauds cases, we then follow Hassan et al. (2019) to build the words library from the lawsuits contents of ten representative scam startups. If 70% of the 100 most common words of scam startups dictionary appear in a lawsuit content, we treat it as a potential startup lawsuit for the next step check. Then we totally collect 621 scam startups. Finally, we find VCs investment information from VentureXpert by searching scam startups names collected from the legal documents. our final sample includes 81 scam startups with 217 venture capitalists involved.

In the empirical part, we follow Gormley and Matsa (2011) to use the stacked difference-in-differences (DID) method. After controlling the firm-cohort and year-cohort fixed effects, the estimation results show that after startups that VCs invested in turning out to be scam, venture capitals decrease subsequent investment. The investment decreases by 4.5% especially in industries where frauds happened. If We use the staggered DiD to just compare the control and treated group, the estimators are potentially biased because of the combination of staggered treatment timing and treatment effect heterogeneity, either across groups or over time (Goodman-Bacon (2021)). In our sample, more cases happened in the later period, which means there are many later cases comparing to the earlier cases under the staggered design. Since each event has a different treatment effect, more weights putting into later–earlier pairs makes estimator biased. But by stacking and aligning events in the event time[2], the approach We use in the baseline is like a setting where the events happen contemporaneously, it can prevent using past treated units as effective comparison units, which may occur with a staggered design.

Overall, our baseline findings indicate that VCs decrease their investments after the occurrence of scams. Previous investment experiences in a given industry allow VCs to make inferences about future ventures in that industry. The impact is more significant for VCs who have had more serious experiences of failure. These findings are consistent with previous empirical and theoretical research, which has documented a positive relationship between VCs' successful experiences and their future investments (Tian et al. (2016); Bergemann and Hege (1998); Hochberg et al. (2007)).

In the further test, we study the effects on VCs following fundraising activities. Past performance is the key determinant of new fund flows for venture capitalists (Kaplan and Schoar (2005); Balboa and Martí (2007)). Updating assessments of a partnership's ability by limited partners (LPs) affects VCs' ability to obtain new fundraising. Poor performance leads to a decline in LPs seeking stakes in subsequent funds. Investing in scam startups damages VCs' reputation, which in turn affects their ability to attract follow-on funding. In empirical tests, we found that affected VCs take more time to attract less funding in the seven years after scams occurred.

This paper has the following contributions. First, it fills a gap in the literature by examining the impact of scam startups on financial investment, specifically on venture capital investment activities. To the best of our knowledge, this is the first paper to investigate the effects of scam startups on venture capital investments. Mahendiran and Conti (2022), a concurrent paper, find that startups producing similar technologies as the scam perpetrators become less likely to obtain financing and raise smaller amounts after the fraud is reported in the news. Although both of Mahendiran and Conti (2022) and this paper focus on scam startups, we significantly differ in the research question, empirical focus, and methodology. Mahendiran and Conti (2022) investigate the impact of misconduct on the financing and exit opportunities of entrepreneurial ventures that are technologically related to the misconduct perpetrators, while our paper examines the effects of scam startups on venture capital investment activities. The definitions of scam startups and the data collection processes are different. Mahendiran and Conti (2022) collect scam startups information by searching for all the articles mentioning the words silicon valley/startup and harassment, startup and allegation/scandal, and startup and lawsuit/infringement/fraud/economic espionage through LexisNexis. However, we use DOJ and SEC lawsuits releases to identify three types of scam startups including products/technologies scams, financial statement scams, and teams' characteristic scams.

[2]The event time refers to the first time case appeared in SEC/DOJ.

Second, this paper contributes to literature which investigates the litigation effects in VC industry. Previous studies in this stream include Atanasov et al. (2012) which find that litigated VCs suffer declines in future business and Cumming et al. (2017) investigating the interaction between entrepreneurial plaintiff firm litigation and venture capital. In contrast to previous studies, this paper focuses specifically on lawsuits related to startup fraud and their impact on VCs' future investment activities. While previous research has explored the litigation effects in the venture capital industry, this paper contributes to the literature by examining the specific type of lawsuits related to startup fraud.

Third, it contributes to the literature on the learning process of VCs when making investment decisions. While previous research, such as Sorensen (2008), has developed a model of learning through investing in the venture capital industry, this paper provides novel empirical evidence that VCs learn new information and update their preferences when they encounter scam startups. Specifically, this paper demonstrates how scam startups can serve as a valuable source of information for VCs, informing their future investment decisions.

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, we report some robustness checks results. We discuss the underlying channels in section 5. In section 6, we conduct some additional analyses. The section 7 contains several tests to disentangle underlying reasons. We conclude in section 8.

2 Data

2.1 Scam Startups Dataset Construction

In this section, we describe the construction process of the novel scam startups dataset.

2.1.1 Startup Scams Lawsuits

The building block of the scam startups dataset is the enforcement actions initiated by SEC and DOJ.[3] Lawsuits in SEC and DOJ are reliable data sources for research on financial frauds. As noted by Karpoff et al. (2008), focusing on SEC and DOJ actions to discipline financial reporting violations can yield a clean sample of cases in which firms violate rules. SEC and DOJ lawsuits data are also widely used in the previous studies. For example, Heese et al. (2021) emphasize the effect of DOJ cases on corporate governance. Li and Cohen (2021) use cases from SEC and DOJ to collect firms alleged to be bribing.

The SEC and DOJ provide detailed information on cases of financial misconduct. The SEC's "Litigation Releases" section focuses on criminal cases involving civil lawsuits brought by the commission in federal court. Official releases are typically posted on the same day that legal cases are filed and have been available since 1995. The DOJ categorizes its releases into 28 topics based on lawsuit contents, including civil rights, drugs, and violent crime, with financial fraud being the most relevant topic for this research, beginning from

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

2013. Figure 1b and 1a provide two examples of SEC and DOJ lawsuits, respectively. Each case includes a detailed narrative history of the allegations, comprising the lawsuit title, the dates on which the case was filed and closed, a description of how the startup lies to the public, the names of defendants, and the sentence and fine. We combine these two data sources to construct our final startup fraud lawsuits data. The final sample period covers 1995 to 2020. To obtain the data, we use a web crawler to download a large sample of 10,381 digital lawsuits released by SEC and 11,672 digital cases initialed by DOJ.

2.1.2 Scam Startups Identification

To identify the scam startups cases from all financial frauds cases in DOJ and SEC, we do the following procedures. First, following Hassan et al. (2019), we define the training library of text for startup frauds identification. The training library contents are 10 representative scam startups' lawsuit contents downloaded from DOJ and SEC.[4] To construct the scam startups lawsuit training library, first, we convert all these 10 representative scam startups lawsuits texts into lower cases. Second, we drop all punctuation marks, numbers, and stop words (e.g., a/an, the, they, etc.) and remove all non-English characters. Third, we tokenize sentences into single words. Lastly, we retrieve each word's root format using the method called lemmatization. We only keep the noun and adjective words because they can convert reasonable meanings. After cleaning the textual documents, we then count the frequency of common words appearing in those lawsuits. We manually check the top 100 most common words to select the keywords to build a scam startups dictionary.[5]

After constructing the training library, we follow Hassan et al. (2019) to use the tracking method to compare every cleaned case content to the top 100 words.[6] If 70% of the 100 most common words of scam startups dictionary appear in a lawsuit content, we treat it as a potential startup lawsuit for the next step check. This method yields over 10,000 cases. If we lower the standard from 70% to 65%, 750 more cases are selected. We manually check those 750 cases but find they are not related to scam startups financial frauds.

The last step is manually reading each case to sort out the companies involved in the lawsuit. We drop the case if it is not related to a startup but an individual or a public firm. After dropping cases not related to startup companies, we check whether the rests are belonged to one of the following three categories: (i) products/technologies scams, (ii) financial statement scams, and (iii) teams' characteristic scams. The above three categories are what VCs consider most when they make investment decisions.

- Products/technologies scams. 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 a high return on 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

[4]The scam startups and their DOJ and SEC release news are reported in Appendix A2.
[5]See Online Appendix A3 and Figure 2 for the frequency of these words.
[6]Each lawsuit text is cleaned as what We do for the training library texts.

false claims concerning its technology to provide accurate, fast and reliable blood tests which never exit.

- Financial statement scams. 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 statements 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.

- Teams' characteristic scams. 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 finds 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 the 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. Figure 3 displays the time distribution of these cases. Using information from primary source documents, we can retrieve details such as startup names, defendants, publication release dates, sentences, and fines. This information allows me to link the startups dataset with other data sources.

2.2 Investment Data

The data on VC investment activities, including investment date, amount, and round, were primarily downloaded from VentureXpert, one of the most comprehensive data sources available for research on venture capital. In addition, we downloaded data on VC characteristics, such as age, the number of funding it manages, the companies it invests in, and the amount of fundraising it has received.

We match the 621 collected scam startups to VentureXpert by searching for each startup's name and identifying the venture capitalists that had previously invested in it. A total of 63 startups are matched with 201 VCs that have invested in them. For the remaining 598 startups, we use Crunchbase as a supplementary source to manually collect the venture funding information when it is not available in VentureXpert.[7] our

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

final sample consists of 81 startups with 217 venture capitals involved. Figure 4 plots the time distribution of scam startups of the final sample. The blue bar represents the scam startup cases from SEC, the grey bar represents the scam startups case released by DOJ, and the red bar indicates cases released by both sources. At the beginning, only SEC cases enter into the final scam startups sample because DOJ's releasement sample starts from 2013. However, from 2014 onwards, most cases appeared in both sources, with only a small fraction of cases coming from either source alone. This suggests that our data collection strategy, which combines both sources, does not introduce selection bias compared to relying on just one source.

Figure 5 shows the distribution of scam startups over time and across industries. During the sample period, the majority of scam startups are in the technical services and business services industries.

2.3 Main Variables

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

We use two variables to measure the investment activities of venture capitalists. The first measure is *Deal Num*, which is the total number of deals a venture capital makes in a given year. The second measure is *Deal Amount*. It is calculated as the total amount of investment a venture capital makes in a given year measured by million dollars. We also use two variables to gauge the fundraising activities of venture capitalists. The first one is *Fundraising Amount*, which is the amount of funding a venture capitalist gets in a given year measured by million dollar. The second one is *Fundraising Time*. It is defined as the duration of getting the next fundraising.

Following Atanasov et al. (2012) and Cumming et al. (2017), we construct several VC characteristics variables as control variables. *Fund Year* is defined as the given year minus the last year in 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 funds 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 investments a venture capital has made. *Industry Deal Amount* is calculated as the amount of investment a venture capitalist makes in all frauds industries in a given year. *Industry Deal Num* is the number of investment deals a venture capitalist makes in all frauds industries in a given year. *Performance* is calculated as the ratio of successful exits under one fundraising. *Industry Performance* is the ratio of successful exits in the given industry under one fundraising.

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

2.4 Summary Statistics

Table 1 presents the summary statistics of the main variables used in the paper. Table 2 reports difference tests results. Panel A illustrates the observed differences between VCs who have invested in scam startups and VCs who have never invested. It is evident that cheated VCs tend to be older, larger in terms of the number of funds they manage, and have higher levels of investment and fundraising activities than non-cheated VCs. These findings underscore the importance of accounting for these differences while examining the impact of being cheated on VC performance to ensure the validity of our conclusions. This also suggests that We capture the important VCs even the number of treated VCs is smaller compared to the control groups. Panel B shows the differences in venture capital between scam startups and real startups. It indicates that, on average, scam startups attract more capital than real startups, despite the fact that they represent a smaller percentage of the total sample.

3 Empirical Strategy and Results

3.1 Empirical Strategy

To estimate the effect of being cheated by scam startups on VC investment, we compare investment changes in the affected and unaffected venture capitalists around the first time of being reported in SEC or DOJ. Figure 6 and Figure 7 plot the average changes in investment amount and deal surrounding the event for treated and control VCs, respectively. These figures verify the parallel trends assumption by showing that there is no treatment effect prior to the revelation of the scams.

There are two key challenges that We face under this setting. Firstly, a small proportion (5%) of VCs in our sample have invested in multiple scam startups, resulting in timing overlaps among events for each affected VC. In order to ensure a clean sample, we only consider the first scam startup in which they invested. Secondly, the standard staggered DiD method comparing the control and treated groups causes the estimators biased because of the combination of staggered treatment timing and treatment effect heterogeneity. Most events in our sample occur in the later period, which biases the estimators as they put more weight on "later" treated to "earlier" treated comparisons. To address this issue, we stack and align events in the event time, which is similar to a setting where events occur contemporaneously. This approach prevents the use of past treated units as effective comparison units, which could occur with a standard staggered design.

The empirical strategy generally follows Gormley and Matsa (2011). More specifically, We 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. We then pool the data across 15 cohorts (i.e., across all event years) and estimate the average treatment effect. More specifically, one cohort consists of the treated VCs in the same year and all non treated VCs as the control group, since there are events in different 15 years, then We create 15 cohorts in our regression. We use the following regression model as the baseline regression:

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

where y_{ict} is the measure of vc investment. We use two variables as proxies: the number of investment deal VC firm i makes in year t (*Deal Num*) and the amount of investment VC firm i makes in year t (*Deal Amount*). $Cheated_{ic}$ is a dummy variable that equals one if VC firm i has invested in one or more scam startups, and zero otherwise. $Post_{ct}$ 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. We 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. We include year-cohort and firm-cohort fixed effect in all specifications. All variables are defined in Table A1.

We first look at the effect of being cheated on the VCs' following up investment activities in all industries. In Table 3, column (1) and (2) report the results of VC investment amount. Without controlling for any additional variables, the estimated coefficient of *Post* × *Cheated* is statistically insignificant and negative (-0.104). The estimated coefficient becomes significant at 1% level and its magnitude declines substantially from -0.104 to -0.023 when control variables are included in the specification. These results suggest that after being cheated by scam startups, VCs do not significantly reduce their total investment relative to VCs who have not been cheated by scam startups. The results in column (3) and (4) indicate that cheated VCs decrease their investment deals after being cheated by scam startups. Both estimated coefficients of *Post* × *Cheated* are statistically significant and negative (-0.185 and -0.082) with and without control variables, suggesting that cheated VCs decrease their investment deals more after the event compared to those not being cheated VCs. This clearly supports the hypothesis that past failure experience affects the likelihood of making future investments. In addition to being statistically significant, the estimated coefficient measures an economically important effect. As column (4) shows, the affected VCs decrease the number of investment deals by 8.2% after the scam revelation.

Table 3 also indicates the effects on VC investments in the specific industries where scams happen. The results are reported in columns (5) to (8). The dependent variable in column (5) to (6) is the amount of investment in scam industries and in column (7) to (8) is the number of deals in them. Control variables are included in column (6) and (8). We find VCs decrease their investments in industries where scam happened. Column (6) shows that the estimated coefficient of *Post* × *Cheated* is statistically significant and negative (-0.045) suggesting after being cheated by scam startups, VCs decrease their investment in scam industries by 4.5%. Column (7) and (8) report the results of VC investment deals. The estimated coefficient of *Post* × *Cheated* is statistically significant and negative (-0.085), suggesting VCs decrease their investment deals in scam industries by 8.5%. This finding is consistent with the idea that VCs learn from their own past investments in a specific industry and adjust their investment behavior accordingly, which is known as industry-level learning Goldfarb et al. (2007).

These findings suggest that VCs are able to learn from their past experiences and adjust their future investment decisions accordingly. Specifically, the results indicate that VCs who have been cheated by scam startups in the past are more cautious and tend to decrease their investment amounts and number of deals in industries where scams have occurred. This behavior is consistent with the idea that VCs update their beliefs about the likelihood of future success based on their previous experiences. Moreover, the findings are robust to using a subsample of first round investments. As Table A4 shows treated VCs decrease their first round investment by 2% in the amounts and 8% in the number of deals.

4 Robustness Tests

In this section, we provide several DiD methods to test the baseline results.

4.1 Standard DID

Normal DID involves comparing timing groups to each other, also known as the two-way fixed effects DD estimator (TWFEDD). As mentioned previously, this method may result in biased estimators, as treatment effects may be heterogeneous for each event. In our sample, there are many later-earlier comparisons, which places more weight on them. In this part, we use the following TWFE regression as the robustness test.

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

where y_{it} is the measure of venture capital investment. $Cheated_{it}$ is a dummy variable that equals one after venture capital i has invested in one or more scam startups, and zero otherwise. $Post_{it}$ is the dummy variable that equals one after the startups turning out to be scam and zero otherwise. $\mathbf{X}'_{it}$ 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. Standard errors are clustered at firm level. The results are reported in Table 4.

In Table 4, column (1) and (2) report the effects on VC investment amount. The estimated coefficient keeps significant when control variables are included in the specification. These results suggest that after being cheated by scam startups, VCs decrease their total investment more 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 (-0.138 and -0.044) with and without control variables, suggesting that cheated VCs decrease their investment deals more after the event compared to those not being cheated VCs. This clearly supports the hypothesis that past investment experience affects the likelihood of making future investments. In addition to being statistically significant, the estimated coefficient measures an economically important effect. As column (4) reveals, the affected VCs decrease the number of investment deals by 4.4% afterwards. The effects of being cheated on the specific industries are reported in column (5) to (8). The estimators are significantly negative after adding control variables, indicating that being cheated by a scam startup has a negative impact on VC investment in the same industry. All results are robust to the findings in the baseline.

4.2 PSM-DID method

According to Panel A of Table 2, the treated group is relatively smaller than the control group. In order to address the concern that the estimated treatment effects are driven by selection bias, we have narrowed down the control group by selecting VCs that are similar to the treated group. For the analysis that follows, we select a sample of peer VCs that have never been cheated but have similar characteristics to the cheated VCs. It is crucial to carefully select such peer VCs to minimize any possible endogeneity concerns that may bias our tests. To allay concerns of endogeneity, we identify peer VCs that have never been cheated but are as similar to cheated VCs across all four performance dimensions, including age, the number of deals, funds under management, the number of companies they invest in. These dimensions are calculated as cumulative number in the year of the investment. We employ the most commonly used methodology–propensity score matching (PSM) method for matching. Out of 217 VCs, we am able to find 165 cheated VCs with 165 peers. The final sample comprises 330 unique VCs (165 treated + 165 matched). The regression model is the following.

$$y_{it} = \beta Cheated_{it} \times Post_{it} + \mathbf{X}'_{it} + \alpha_i + \delta_t + \epsilon_{it} \tag{3}$$

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 scam 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. Overall, the findings are generally robust to comparing VCs that have invested in scam startups with those that have not based on a propensity-score matching algorithm.

4.3 Cengiz et al. (2019) Approach

To address concerns that news reports of scams may have preceded the SEC/DOJ publication or that VCs may have known about the scams earlier, we extend the estimation window to cover a longer period before the events. As noted by Cengiz et al. (2019), using an asymmetric estimation window does not effect the result. Specifically, we construct an eight-year estimation window ranging from t = -4 to t = 2 and stack events in each event time. Table 6 provides evidence that the main results are robust to different event windows. The estimated coefficients for the independent variable remain significantly negative across all event windows.

These findings suggest that even when extending the previous time period under consideration, the negative effects of being cheated by scam startups are still statistically significant.

5 Channel

The empirical results so far show that being cheated by scam startups can cause VCs decrease investment activities in the following years. In this section, we propose several potential reasons for this finding.

5.1 Industry Signal

The industry releases bad signals when scam happens. According to Bayes' rule, rational agents update their beliefs after receiving new information. VCs are more careful for the future investment opportunities in the industry where scam happens (Sharma et al. (2015); Sorenson and Stuart (2001)). If the baseline results are indeed attributable to the learning mechanism, one would expect the effect to be stronger in the industries where the scams happened, since VCs have adjusted their beliefs on those industries.

To test this hypothesis, we interact Post × Cheated with Frauds, Frauds is a dummy variable that takes the value of one if the industry has frauds and zero otherwise. The results, presented in Table 7, indicate that the negative relationship between being cheated by scam startups and VCs' future investment activities is significantly stronger in industries with a history of scam activities.

5.2 Investment Experience

Empirical evidence suggests that VCs engage in more learning if they suffer a greater loss (Sorensen, 2008). Therefore, it can be expected that the reduction in investment would be more severe for VCs that have experienced more significant failures. To test this hypothesis, the sample is split based on the VCs' level of investment experience, into low and high investment groups. We replicate the analysis in Table 3.

The results, as presented in Table 8, provide support for the conjecture that VCs adjust their strategies by learning from past experience. Specifically, in column (1)-(4), the independent variable is Post × Cheated × Amount. Amount is the dummy variable that takes the value of one if the venture capitalist invested more money than the average in scam startups, and zero otherwise. In column (5)-(8), the independent variable is Post × Cheated × Deal. Deal is the dummy variable that takes the value of one if the venture capitalist made more investment deals than the average in scam startups, and zero otherwise. Similarly, we find that VCs who have invested in a larger number of scam startups tend to reduce their investments more in the aftermath of being cheated, as evidenced by the statistically significant interaction effect among Post, Cheated, and Deal.

6 Further Tests

In the additional analysis We conduct supplementary tests to have a more comprehensive understanding of the impacts of scam startups on VCs investment activities. First, we conduct tests to see whether VCs fundraising is also affected by such scams.. Second, we look at VCs exit performance. Third, we revisit the concern regarding the results of SEC/DOJ investigation by examining the guilty cases directly.

6.1 Fundraising Activities

The performance of an existing fund is an important factor for venture capitalists when raising capital for a follow-on fund. Previous performance can significantly influence the outcome of future fundraising efforts. In light of this, we posit that VCs' reputation is negatively impacted following unsuccessful investments in scam startups, which could impede their ability to attract further fundraising.

To measure VCs' fundraising abilities, this study considers two key factors: the amount of capital raised from LPs, referred to as "Amount", and the time taken to raise the funds, referred to as "Time". The duration of launching a subsequent fund can also provide important insights into a VC firm's fundraising capacity.

6.1.1 Fundraising Amount

The first hypothesis suggests that venture capitalists who have experienced investment failures in scam startups receive less follow-on fundraising. To test this hypothesis, we examine a 10-year time window around the event, considering the average holding period for the current funding to be 4-5 years. The results, as shown in Table 9, suggest that VCs receive significantly less follow-on funding after a failure experience in scam startups, with the amount of fundraising decreasing by three times compared to the past. These findings are consistent with previous research that suggests VCs who have performed better in the past tend to raise more capital (Barber and Yasuda (2017); Chung et al. (2012)).

6.1.2 Fundraising Speed

Another measure of fundraising ability is the time it takes for VCs to raise follow-on funds. Therefore, we hypothesize that VCs who have experienced failures in investing in scam startups take longer to raise subsequent funds. As the average holding period for current funds is typically around 4-5 years, we analyze a 10-year time window around the failure event. The results, presented in Table 10, suggest that the frequency of fund launches decreases for VCs who have been affected by scams, even after controlling for firm-cohort and year-cohort fixed effects.

6.2 Guilty Verdict

As described in Appendix 8, the trial process typically spans several years. Therefore, using the first appearance of a startup in the SEC/DOJ investigation as the sole criterion for determining scam status may lead to misclassifications, as some startups may ultimately be found innocent. To address this concern, we exclude the 26 startups that have not yet reached a final verdict, thus focusing solely on cases where the verdict is guilty. We re-estimate the baseline regression with this revised sample, and the results are presented in Table 11. This approach helps to mitigate the possibility that the observed effects are driven by litigation rather than actual scam activities.

6.3 Exit Performance

In this section, we analyze the impact of being cheated by scam startups on the performance of venture capitalists. As prior studies have done, we measure venture success through successful exits such as IPOs and acquisition exits (Das et al. (2011); Sorensen (2004); Hochberg et al. (2007)). We extend the observation period to five years following the scam incident. The results in Table 12 indicate that the past experience of failure investment in scam startups does not have a significant association with exit performance.

7 Beliefs or Less Funding?

In previous sections, we found that VCs learn from their failure experiences and adjust their investment strategies accordingly. Additionally, it becomes more difficult for VCs to secure follow-on funding after being cheated by a scam startup. In this section, we aim to distinguish whether the negative effects on investment activities are due to VCs updating their beliefs on the industry or due to a decrease in fundraising. If VCs learn from their past failure experiences, they should improve their screening and monitoring procedures in future investment activities. Furthermore, reputable VCs find it easier to get fundraising (Gompers and Lerner (1999); Barber and Yasuda (2017); Cumming et al. (2005)), therefore, if reputable VCs decrease their investments more significantly after being cheated, the reason cannot be due to a lack of funding. In the following tests, we distinguish between these two reasons.

7.1 Screening and Monitoring

In this section, we conduct tests to examine whether VCs have strengthened their screening and monitoring procedures in response to the failure experience of investing in scam startups. To do so, we analyze the investment distance and the investment stage.

7.1.1 The Investment Distance

In this section, we examine whether the distance between VCs and their invested companies changes after experiencing scam investments, as some VCs prefer to monitor their investments by making frequent visits to the businesses they invest in. To test this, we analyze whether treated VCs prefer to invest closer to their location in order to better monitor their investments.

The results, presented in Table 13, show that VCs decrease the geographic distance between themselves and their invested companies after the scam event. After controlling for relevant variables and cohort fixed effects, column (2) indicates that treated VCs decrease investment distance by 3.4% in the following two-year investments. Column (4) further demonstrates that treated VCs decrease investment distance in scam industries by 2.4%. These results support the hypothesis that VCs strengthen their monitoring procedures on new investments following a negative experience with scam investments.

7.1.2 The Investment Stage

In this section, we examine whether treated VCs decrease their investments in early-stage ventures after strengthening their screening procedure. Since early-stage ventures are considered to be riskier, we hypothesize that the VCs will be more cautious in investing in these ventures. The results are presented in Table 14, where we focus on the first investments made by VCs. Column (2) and (4) show that treated VCs significantly reduce their investments in early-stage ventures, particularly in the scam industries, which supports the hypothesis that VCs strengthen their screening of new investment opportunities.

7.2 Reputable VCs

Previous studies have established that reputable VCs have greater access to fundraising compared to their counterparts (Gompers (1996); Barber and Yasuda (2017)). If reputable VCs decrease their investments significantly, despite having access to ample funds, it would imply that their decreased investments are due to the loss of belief in a specific industry rather than a lack of fundraising. To test this, we use two measures of reputation: "better past performance" and "old". A VC with superior past performance, better than the average level, is regarded as reputable. Similarly, a VC that is older than the average is also considered reputable.

Table 15 shows that VCs with better past performance decrease more investment in the future, particularly in the scam industries. This suggests that VCs' decreased investment is not driven by less funding but rather by a belief lost in the specific industry. Similarly, Table 16 shows that older VCs decrease more investment in the scam industries, again suggesting that belief updating rather than less funding is the driving force behind the decrease in investment. Overall, the results of the analysis of VC reputation are consistent with the previous findings and further support the hypothesis that VCs' decreased investment is due to belief updating rather than less funding.

8 Conclusion

This paper examines the effect of growing scam startups on venture capital investment. We look at this question by constructing a novel dataset. Through the use of the DID method and various robustness checks, the results indicate that VCs tend to decrease their investments in response to being deceived by scam startups. The decrease is more pronounced in industries with a higher frequency of scams. To further explore this finding, we examine potential channels that may explain this phenomenon. One of the key factors identified is the VC's past performance, which plays an important role in their follow-on investment activities. VCs learn from past failures and adjust their beliefs about new investment opportunities accordingly. The effect is even more pronounced for VCs who have experienced more serious failures. Additionally, we examine the VCs' fundraising activities, screening and monitoring processes, and exit performance after being cheated by scam startups. Overall, these findings shed light on the serious effects of scam startups on the venture capital industry and provide insights into how VCs respond to such incidents.

References

Ahern, Kenneth R, 2017, Information networks: Evidence from illegal insider trading tips, *Journal of Financial Economics* 125, 26–47.

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.

Balboa, Marina, and José Martí, 2007, Factors that determine the reputation of private equity managers in developing markets, *Journal of Business Venturing* 22, 453–480.

Barber, Brad M, and Ayako Yasuda, 2017, Interim fund performance and fundraising in private equity, *Journal of Financial Economics* 124, 172–194.

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.

Boyd, John H, and Edward C Prescott, 1986, Financial intermediary-coalitions, *Journal of Economic theory* 38, 211–232.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, 2019, The effect of minimum wages on low-wage jobs, *Quarterly Journal of Economics* 134, 1405–1454.

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

Chung, Ji-Woong, Berk A Sensoy, Lea Stern, and Michael S Weisbach, 2012, Pay for performance from future fund flows: the case of private equity, *The Review of Financial Studies* 25, 3259–3304.

Cumming, Douglas, Grant Fleming, and Jo-Ann Suchard, 2005, Venture capitalist value-added activities, fundraising and drawdowns, *Journal of Banking & Finance* 29, 295–331.

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

Das, Sanjiv R, Hoje Jo, and Yongtae Kim, 2011, Polishing diamonds in the rough: The sources of syndicated venture performance, *Journal of Financial Intermediation* 20, 199–230.

Dehejia, Rajeev H, and Sadek Wahba, 2002, Propensity score-matching methods for nonexperimental causal studies, *Review of Economics and statistics* 84, 151–161.

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

Goldfarb, Brent, David Kirsch, and David A Miller, 2007, Was there too little entry during the dot com era?, *Journal of Financial Economics* 86, 100–144.

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, and Josh Lerner, 1999, What drives venture capital fundraising?

Gompers, Paul A, 1996, Grandstanding in the venture capital industry, *Journal of Financial economics* 42, 133–156.

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.

Hassan, Tarek A, Stephan Hollander, Laurence Van Lent, and Ahmed Tahoun, 2019, Firm-level political risk: Measurement and effects, *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.

Hochberg, Yael V, Alexander Ljungqvist, and Yang Lu, 2007, Whom you know matters: Venture capital networks and investment performance, *The Journal of Finance* 62, 251–301.

Kaplan, Steven N, and Antoinette Schoar, 2005, Private equity performance: Returns, persistence, and capital flows, *Journal of Finance* 60, 1791–1823.

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* .

Mahendiran, Shreekanth, and Annamaria Conti, 2022, Too hot to handle: The effects of misconduct on entrepreneurial venture outcomes, in *Workng Paper, University of Lausanne*.

Sharma, AK, et al., 2015, Venture capitalists' investment decision criteria for new ventures: a review, *Procedia-Social and Behavioral Sciences* 189, 465–470.

Sorensen, Morten, 2004, How smart is smart money? an empirical two-sided matching model of venture capital, *Unpublished working paper. University of Chicago* .

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

Sorenson, Olav, and Toby E Stuart, 2001, Syndication networks and the spatial distribution of venture capital investments, *American journal of sociology* 106, 1546–1588.

Tian, Xuan, Gregory F Udell, and Xiaoyun Yu, 2016, Disciplining delegated monitors: When venture capitalists fail to prevent fraud by their ipo firms, *Journal of Accounting and Economics* 61, 526–544.

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.

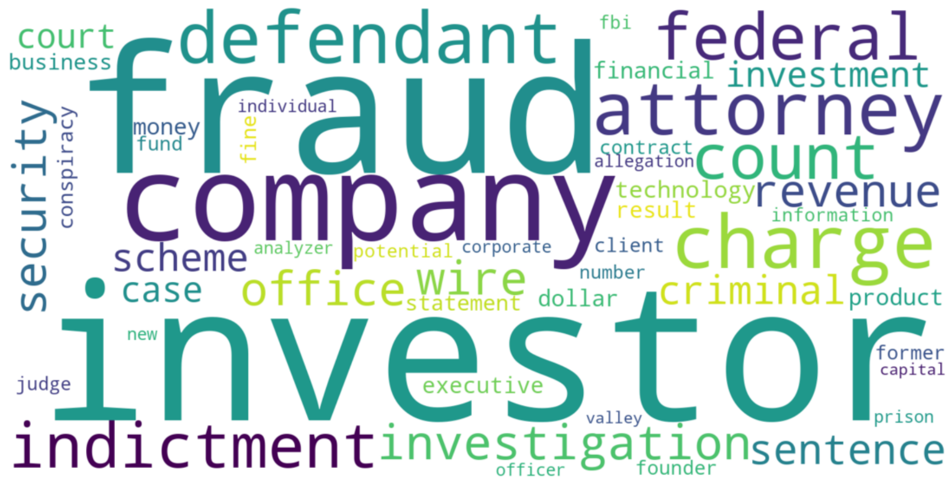


Figure 2: Key Words Word Cloud

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

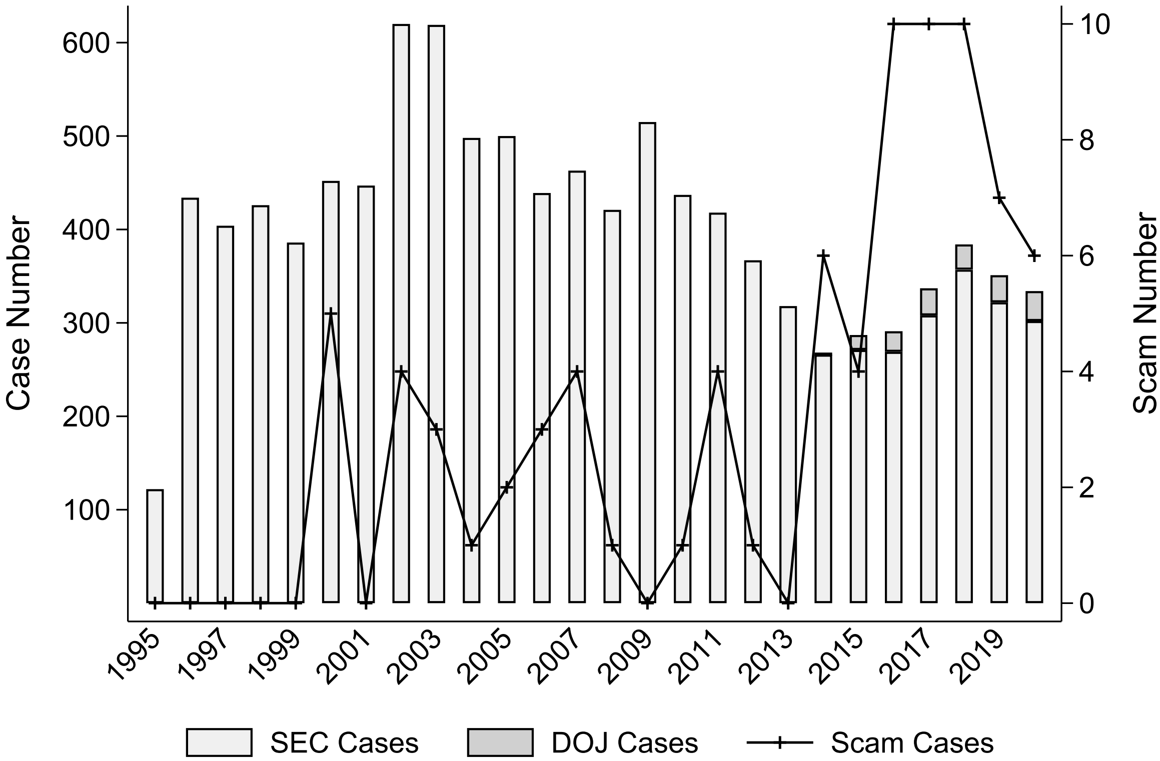


Figure 3: Time Series of the Number of Frauds Lawsuits

This figure plots distribution of frauds cases shown in SEC and DOJ during the sample period. The line "Scam Cases" includes the cases that are in the final sample and contain VCs information.

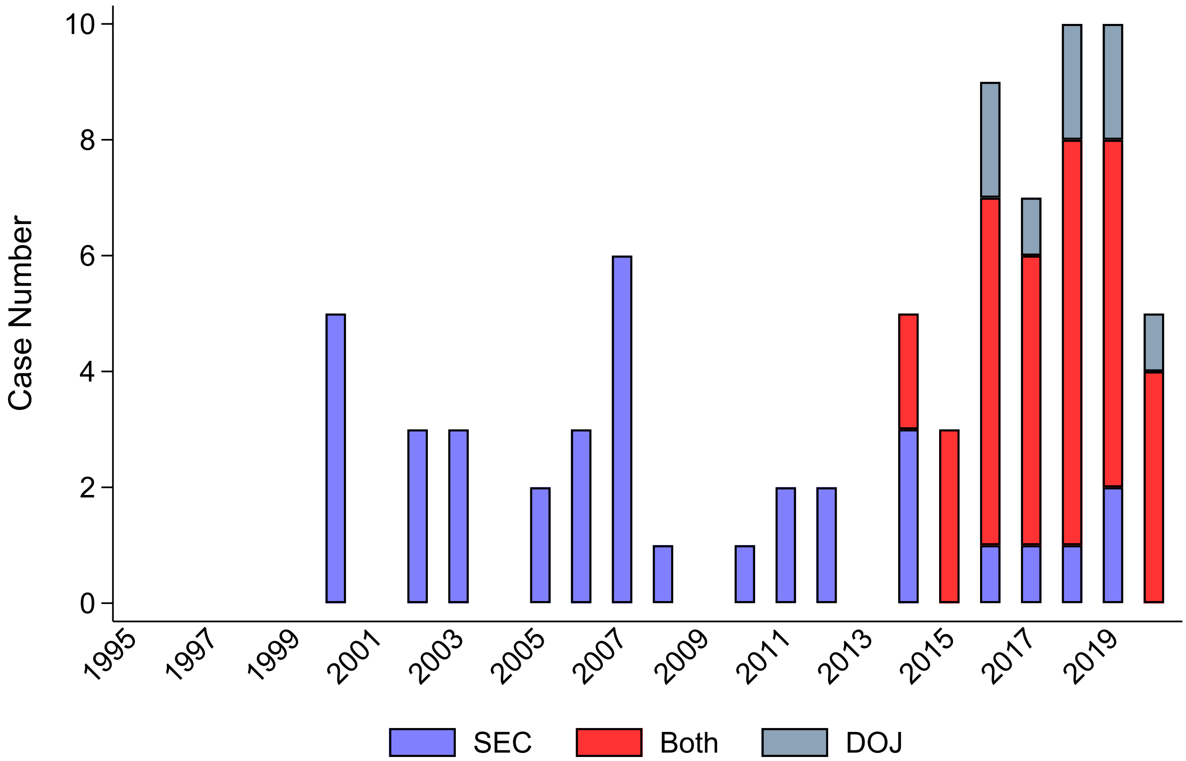


Figure 4: Time Series of Cases in the Final Sample

This figure plots distribution of cases the in final sample.

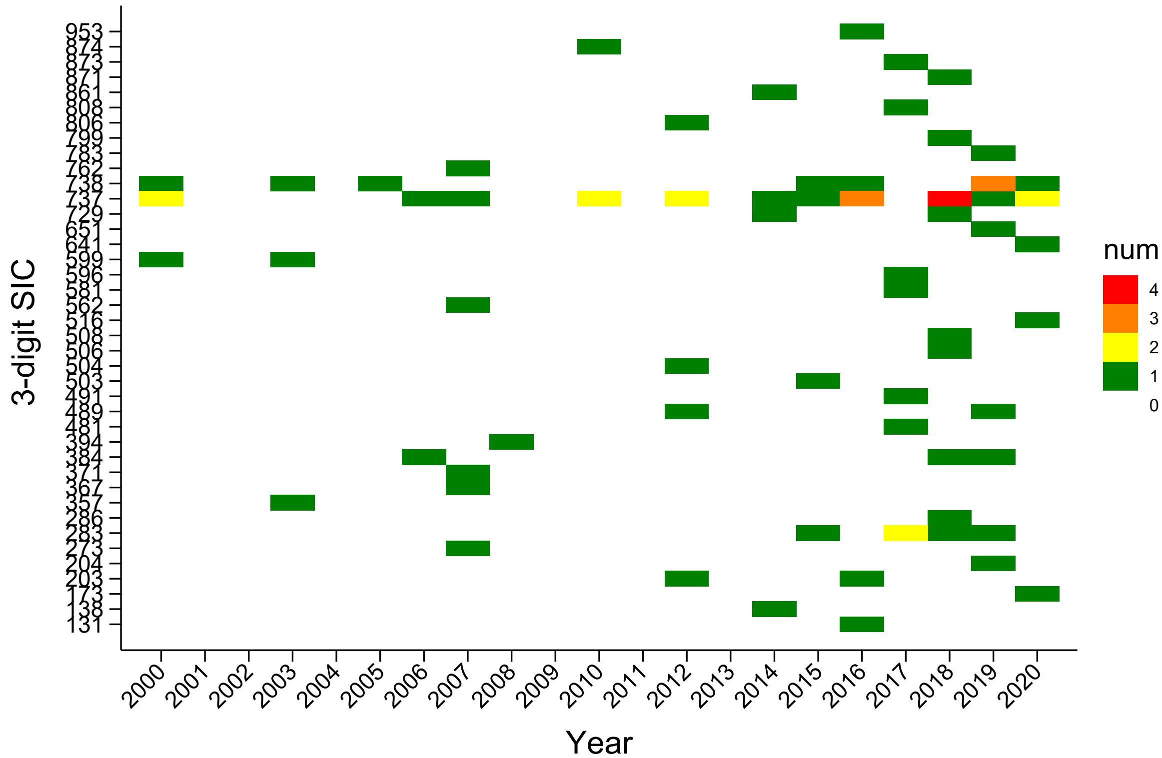
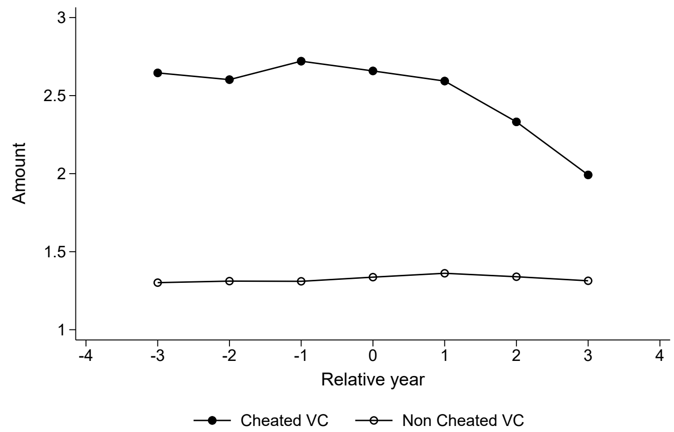
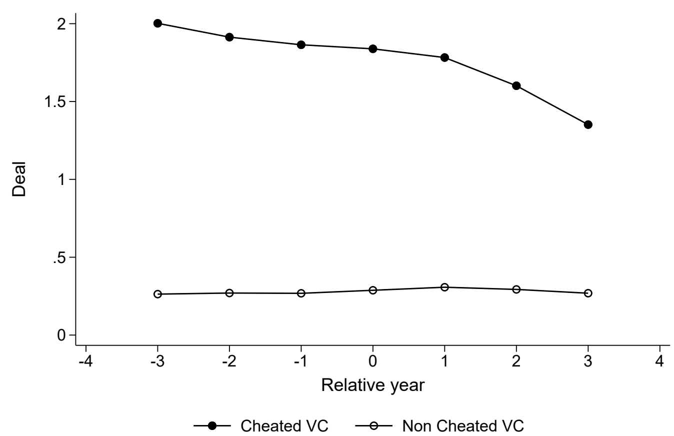


Figure 5: The time and industry distribution

This figure plots time and industry distribution of cases in the final sample.



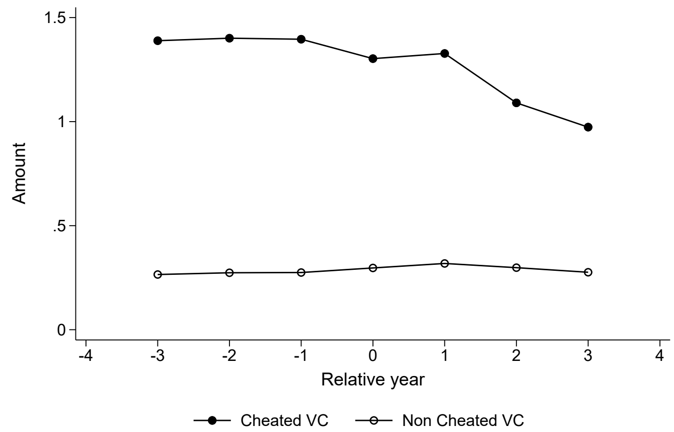
(a) VC Investment Amount



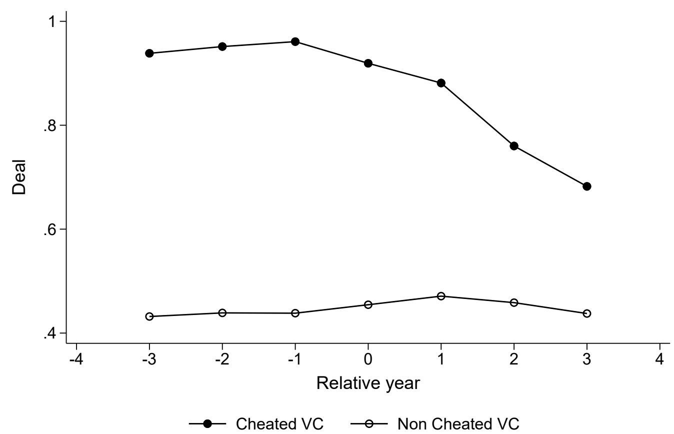
(b) VC Investment Deal

Figure 6: 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 7: 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	28,658	2.01	1.60	0.00	0.69	1.79	3.01	9.84
Deal Num	28,658	1.68	0.94	0.69	1.10	1.39	2.20	5.79
Deal Amount IND	11,724	1.91	2.33	0.00	0.44	1.37	2.60	8.18
Deal Num IND	11,724	1.03	0.52	0.69	0.69	0.69	1.10	4.30
Industry Performance	25,480	0.46	0.37	0.00	0.00	0.50	0.75	1.00
Performance	25,480	0.45	0.24	0.00	0.29	0.45	0.60	1.00
Industry Deal Num	21,045	2.81	1.49	0.69	1.61	2.64	3.83	7.93
Industry Deal Amount	21,043	3.28	1.90	0.00	1.87	3.20	4.52	10.08
Company Num	28,658	2.94	1.53	0.69	1.61	2.83	3.99	7.99
VC Age	28,168	2.33	0.85	0.00	1.79	2.40	2.89	5.28
Funds Num	28,658	2.17	1.06	0.69	1.39	2.08	2.89	5.57
Total Deal Amount	28,656	3.63	2.05	0.00	2.14	3.61	4.98	10.73
Total Deal Num	28,658	3.13	1.61	0.69	1.79	3.00	4.25	8.52

Table 2: Balancedness Test

This table reports the difference tests results. Panel A reports the difference between cheated VC and non-cheated VC. 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	11,239	1.00	217	1.18	-0.19	-5.86
Deal Amount	11,239	1.07	217	1.19	-0.12	-1.34
Company Num	11,229	2.02	217	4.40	-2.39	-27.19
Found Year	11,239	2,000	217	1,991	9	7.46
Fund Num	7,351	1.15	171	2.11	-0.96	-20.15

Panel B: Difference of Scam Startups and Real Startups

	(1)	(2)	(3)	(4)	(5)	(6)
	Real Startups		Scam Startups		(2) − (4)	
	N	Mean	N	Mean	Difference	t-value
Deal Amount	483,663	0.86	557	1.06	-0.19	-4.85
Deal Num	483,665	2.70	557	2.96	-0.27	-6.26
Company Found Year	483,665	1,990	557	1,988	2	3.52
Round	483,665	3.61	557	3.96	-0.34	-2.40
Debt	483,665	0.08	557	0.08	-0.00	-0.17

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 scam industries. In column (7)–(8), the dependent variable is the number of deals of investment in the scam industries. 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. Standard errors are clustered at firm level. *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	-0.104	-0.023**	-0.185***	-0.082**	-0.119	-0.045**	-0.182***	-0.085**
	(-1.00)	(-2.02)	(-3.87)	(-1.99)	(-1.16)	(-2.00)	(-3.83)	(-2.07)
Industry Deal		-0.164**		-0.100***		-0.142**		0.038
		(-2.47)		(-3.50)		(-2.13)		(1.25)
Industry Amount		0.343***		-0.008		0.445***		0.007
		(9.56)		(-0.71)		(12.68)		(0.56)
VC Age		-0.119**		-0.302***		-0.156***		-0.245***
		(-2.33)		(-10.74)		(-2.96)		(-8.59)
Funds Num		-0.994***		-0.966***		-0.852***		-0.917***
		(-15.32)		(-26.28)		(-13.12)		(-24.69)
Total Deal Amount		0.736***		0.020*		0.556***		0.006
		(21.36)		(1.80)		(16.66)		(0.48)
Total Deal Num		0.286***		1.247***		0.248***		1.046***
		(4.29)		(40.11)		(3.69)		(31.42)
Year × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
N	221,284	168,003	221,284	168,003	195,783	159,027	195,788	159,027
Adj. R^2	0.56	0.71	0.64	0.78	0.57	0.70	0.65	0.76

Table 4: TWFE DID

This table reports the regression analyses based on normal staggered DiD 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 scam industries. In column (7)–(8), the dependent variable is the number of deals of investment in the scam industries. 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 and firm fixed effects. All continuous variables are winsorized at 1% level. Standard errors are clustered at firm level. *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	0.436***	-0.264***	-0.138**	-0.044***	0.699	-0.074**	0.344	-0.267**
	(3.47)	(-3.42)	(-2.00)	(-4.77)	(0.43)	(-1.98)	(1.14)	(-2.08)
Industry Deal		0.128		0.146***		-0.816*		-0.568***
		(1.41)		(3.30)		(-1.74)		(-5.19)
Industry Amount		0.320***		0.050***		0.465*		0.035
		(5.65)		(3.07)		(1.75)		(0.58)
VC Age		-0.356***		-0.353***		0.536*		0.066
		(-5.02)		(-6.74)		(1.87)		(0.83)
Funds Num		-0.280***		-0.520***		-1.046***		-0.402***
		(-3.56)		(-9.91)		(-3.37)		(-3.38)
Total Deal Amount		0.935***		-0.011		0.665*		-0.015
		(15.30)		(-0.49)		(1.91)		(-0.21)
Total Deal Num		-0.616***		0.652***		0.461		0.946***
		(-6.48)		(11.64)		(0.84)		(6.78)
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y
N	482,007	441,727	482,007	441,727	108,462	100,420	108,462	100,420
Adj. R^2	0.75	0.84	0.81	0.91	0.59	0.61	0.59	0.62

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 scam industries. In column (7)–(8), the dependent variable is the number of deals of investment in the scam industries. 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 and firm fixed effects. All continuous variables are winsorized at 1% level. Standard errors are clustered at firm level. 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	-0.087	-0.035*	-0.075*	-0.031*	0.183	-0.005**	0.052	-0.004***
	(-1.01)	(-1.75)	(-1.67)	(-1.69)	(0.72)	(-2.02)	(0.54)	(-3.03)
Industry Amount		-0.183		-0.008		-0.119		0.104
		(-0.48)		(-0.07)		(-0.13)		(0.29)
Industry Deal		1.015*		0.559**		2.975*		-0.297
		(1.65)		(2.30)		(1.74)		(-0.59)
Company Num		0.839		0.759**		-0.799		0.012
		(1.26)		(2.33)		(-0.53)		(0.02)
VC Age		-0.157		-0.197		-0.188		0.001
		(-0.36)		(-0.86)		(-0.16)		(0.00)
Funds Num		-1.411***		-0.946***		0.088		-0.584
		(-3.09)		(-4.18)		(0.08)		(-1.41)
Total Deal Amount		1.976***		0.105		1.578		-0.044
		(5.15)		(0.93)		(1.57)		(-0.13)
Total Deal Num		-1.808**		0.156		-3.349		1.060
		(-2.07)		(0.47)		(-1.61)		(1.39)
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y
N	2,226	1,180	2,226	1,180	324	301	324	301
Adj. R^2	0.77	0.77	0.85	0.85	0.62	0.63	0.75	0.76

Table 6: Cengiz et al. (2019) Approach

This table reports the regression analyses based on a new stacked DiD method (Cengiz et al. (2019)). 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 scam industries. In column (7)–(8), the dependent variable is the number of deals of investment in the scam industries. 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. Standard errors are clustered at firm level. 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	-0.070	-0.010	-0.184***	-0.077*	-0.121	-0.046	-0.184***	-0.085**
	(-0.67)	(-0.11)	(-3.58)	(-1.72)	(-1.18)	(-0.53)	(-3.85)	(-2.06)
Industry Deal		-0.155**		-0.100***		-0.140**		0.036
		(-2.43)		(-3.65)		(-2.08)		(1.18)
Industry Amount		0.336***		-0.007		0.444***		0.008
		(9.54)		(-0.64)		(12.57)		(0.66)
VC Age		-0.140***		-0.310***		-0.156***		-0.251***
		(-2.77)		(-11.18)		(-2.89)		(-8.60)
Funds Num		-0.950***		-0.948***		-0.839***		-0.909***
		(-14.88)		(-26.49)		(-12.77)		(-24.21)
Total Deal Amount		0.745***		0.024**		0.552***		0.005
		(22.02)		(2.19)		(16.47)		(0.43)
Total Deal Num		0.265***		1.236***		0.246***		1.045***
		(4.12)		(41.16)		(3.64)		(31.17)
Year × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
N	262,947	199,181	262,947	199,181	188,120	153,516	188,125	153,516
Adj. R^2	0.55	0.70	0.64	0.77	0.57	0.70	0.64	0.76

Table 7: Frauds Industry

This table reports the regression results for the investment in specific scam 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. The main independent variable is Post × Cheated × Frauds. 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. Frauds is a dummy variable that takes the value of one if the industry has scam and zero otherwise. 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. Standard errors are clustered at firm level. 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			
	(1) Deal Amount	(2) Deal Amount	(3) Deal Num	(4) Deal Num
Post × Cheated × Frauds	0.000	-0.181***	-0.033**	-0.053*
	(0.00)	(-5.96)	(-2.06)	(-1.74)
Post × Cheated	-0.053	-0.142	-0.203**	-0.121
	(-0.30)	(-0.91)	(-2.24)	(-1.46)
Post × Frauds	-0.061***	-0.058***	-0.070***	-0.061***
	(-8.97)	(-9.50)	(-21.09)	(-20.78)
Cheated × Frauds	-0.441***	-0.380***	-0.332***	-0.287***
	(-2.87)	(-2.71)	(-4.61)	(-4.00)
Industry Deal		-0.150**		-0.085***
		(-2.20)		(-2.91)
Industry Amount		0.342***		-0.006
		(9.04)		(-0.52)
VC Age		-0.105**		-0.283***
		(-2.08)		(-10.20)
Funds Num		-0.993***		-0.972***
		(-15.61)		(-27.18)
Total Deal Amount		0.753***		0.027**
		(21.40)		(2.52)
Total Deal Num		0.256***		1.230***
		(3.88)		(40.15)
Year × Cohort FE	Y	Y	Y	Y
Firm × Cohort FE	Y	Y	Y	Y
N	265,986	202,528	265,986	202,528
Adj. R^2	0.55	0.70	0.65	0.78

Table 8: Investment Experience

This table reports the regression results for the channel test based on different investment in scam startups. In column (1)–(4), the dependent variable is the investment in all industries. In column (5)–(8), the dependent variable is the investment in the scam industries. The independent variables are Post × Cheated × Amount and Post × Cheated × Deal. Amount is the dummy variable that takes the value of one if the venture capitalist invested more money in scam startups, and zero otherwise. Deal is the dummy variable that takes the value of one if the venture capitalist made more investment deals in scam startups, and zero otherwise. 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. Standard errors are clustered at firm level. *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 Num	(3) Deal Amount	(4) Deal Num	(5) Deal Amount	(6) Deal Num	(7) Deal Amount	(8) Deal Num
Post × Cheated × Amount	-0.098	-0.049***			-0.093***	-0.043*		
	(-0.59)	(-3.55)			(-4.55)	(-1.84)		
Post × Cheated × Deal			-0.140*	-0.143*			-0.133	-0.132**
			(-1.85)	(-1.82)			(-0.61)	(-2.23)
Post × Cheated	0.048	-0.020	0.040	-0.013	0.032	-0.023	0.025	-0.016
	(0.39)	(-0.37)	(0.38)	(-0.28)	(0.26)	(-0.43)	(0.23)	(-0.34)
Industry Deal	-0.313***	0.057**	-0.313***	0.057**	-0.259***	0.242***	-0.259***	0.242***
	(-5.06)	(2.33)	(-5.06)	(2.33)	(-4.17)	(9.42)	(-4.17)	(9.42)
Industry Amount	0.373***	-0.023**	0.373***	-0.023**	0.529***	0.001	0.529***	0.001
	(9.94)	(-2.32)	(9.94)	(-2.32)	(14.04)	(0.06)	(14.04)	(0.06)
VC Age	-0.322***	-0.410***	-0.322***	-0.410***	-0.358***	-0.397***	-0.358***	-0.397***
	(-7.33)	(-17.15)	(-7.33)	(-17.15)	(-8.22)	(-15.92)	(-8.22)	(-15.92)
Funds Num	-1.011***	-0.957***	-1.011***	-0.958***	-0.846***	-0.873***	-0.846***	-0.873***
	(-19.11)	(-33.60)	(-19.11)	(-33.61)	(-16.49)	(-30.07)	(-16.49)	(-30.08)
Total Deal Amount	0.915***	0.024**	0.915***	0.024**	0.677***	-0.002	0.677***	-0.002
	(24.34)	(2.53)	(24.34)	(2.53)	(17.69)	(-0.20)	(17.69)	(-0.20)
Total Deal Num	0.320***	1.260***	0.321***	1.260***	0.254***	1.012***	0.254***	1.012***
	(5.20)	(51.13)	(5.20)	(51.13)	(4.10)	(37.49)	(4.10)	(37.49)
Year × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
N	146,771	146,771	146,771	146,771	138,641	138,641	138,641	138,641
Adj. R^2	0.71	0.81	0.71	0.81	0.69	0.80	0.69	0.80

Table 9: Fundraising Amount

This table reports the regression results for the treated VCs new fundraising activities. In column (1)-(2), the dependent variable is the amount of fundraising VCs get every year. The independent variable is Post × Cheated. Control variables include the amount of fundraising before frauds happened, 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. Standard errors are clustered at firm level. 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.

	(1) Amount	(2) Amount
Post × Cheated	-2.813***	-3.160***
	(-4.47)	(-3.97)
Pre Amount	0.777***	0.791***
	(25.89)	(22.07)
Last Fund Year		0.632***
		(8.54)
Total Known Company		1.836***
		(3.59)
Age		-0.028
		(-0.23)
Funds		-0.101
		(-0.39)
Total Equity		0.002
		(0.02)
Total Deals		-2.022***
		(-3.97)
Year × Cohort FE	Y	Y
Firm × Cohort FE	Y	Y
N	97,041	96,844
Adj. R^2	0.48	0.50

Table 10: Fundraising Speed

This table reports the regression results for the timing of treated VCs new fundraising activities. In column (1)-(2), the dependent variable is a dummy variable which equals to one if VC gets the fundraising. The independent variable is Post × Cheated. Control variables include the amount of fundraising before frauds happened, 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. Standard errors are clustered at firm level. 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.

	(1) Time	(2) Time
Post × Cheated	0.007	0.007*
	(0.29)	(1.70)
Fund Year		-0.015***
		(-6.33)
Company Num		0.088***
		(2.87)
VC Age		-0.039***
		(-4.94)
Funds Num		-0.041***
		(-3.84)
Total Deal Amount		-0.001
		(-0.31)
Total Deal Num		-0.036
		(-1.22)
Year × Cohort FE	Y	Y
Firm × Cohort FE	Y	Y
N	97,041	96,844
Adj. R^2	0.28	0.28

Table 11: Guilty Verdict

This table reports the regression analyses based on the cases where the verdicts are guilty. 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 main independent variable is Post × Cheated. 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. Standard errors are clustered at firm level. 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	-0.163	-0.027***	-0.234***	-0.114***	-0.179	-0.050***	-0.232***	-0.118***
	(-1.41)	(-3.27)	(-4.55)	(-2.61)	(-1.57)	(-3.52)	(-4.53)	(-2.70)
Industry Deal		-0.164**		-0.100***		-0.142**		0.038
		(-2.47)		(-3.50)		(-2.13)		(1.24)
Industry Amount		0.343***		-0.008		0.444***		0.007
		(9.55)		(-0.72)		(12.67)		(0.55)
VC Age		-0.119**		-0.303***		-0.156***		-0.245***
		(-2.34)		(-10.74)		(-2.96)		(-8.59)
Funds Num		-0.994***		-0.967***		-0.853***		-0.917***
		(-15.32)		(-26.27)		(-13.12)		(-24.68)
Total Deal Amount		0.736***		0.020*		0.556***		0.006
		(21.34)		(1.78)		(16.63)		(0.45)
Total Deal Num		0.286***		1.248***		0.249***		1.046***
		(4.30)		(40.10)		(3.70)		(31.41)
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y
N	221,055	167,788	221,055	167,788	195,554	158,812	195,559	158,812
Adj. R^2	0.56	0.71	0.64	0.78	0.57	0.70	0.65	0.76

Table 12: Exit Performance

This table reports the regression analyses based on the investment performance. 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 main independent variable is Post × Cheated. 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. Standard errors are clustered at firm level. 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.

	(1) Exit	(2) Exit
Post × Cheated	-1.001***	-0.825***
	(-4.78)	(-3.78)
Industry Deal		0.378***
		(5.33)
Industry Amount		0.041*
		(1.65)
VC Age		-0.223**
		(-2.56)
Funds Num		-1.007***
		(-9.27)
Total Deal Amount		-0.005
		(-0.20)
Total Deal Num		0.977***
		(12.69)
Year × Cohort FE	Y	Y
Firm × Cohort FE	Y	Y
N	165,897	126,375
Adj. R^2	0.50	0.53

Table 13: Investment Distance

This table reports the regression analyses based on the investment distance. In column (1)–(2), the dependent variable is the distance of investment in all industries. In column (3)–(4), the dependent variable is the distance of investment in scam industries. The main independent variable is Post × Cheated. 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. Standard errors are clustered at firm level. *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)
	Deal Distance	Deal Distance	Deal Distance Ind	Deal Distance Ind
Post × Cheated	-0.076	-0.034***	-0.033**	-0.024**
	(-1.20)	(-3.59)	(-2.46)	(-2.36)
Industry Deal		-0.034		0.159**
		(-0.48)		(1.99)
Industry Amount		-0.018		-0.039
		(-0.65)		(-1.29)
VC Age		-0.043		-0.089
		(-0.49)		(-0.94)
Funds Num		-0.211**		-0.227**
		(-2.36)		(-2.50)
Total Deal Amount		0.009		0.047
		(0.29)		(1.35)
Total Deal Num		0.459***		0.253***
		(5.32)		(2.76)
Year × Cohort FE	Y	Y	Y	Y
Firm × Cohort FE	Y	Y	Y	Y
N	142,911	109,976	132,469	106,670
Adj. R^2	0.26	0.28	0.25	0.27

Table 14: Investment Stage

This table reports the regression analyses based on the investment stage. In column (1)–(2), the dependent variable is the number of investment in the first stage in all industries. In column (3)–(4), the dependent variable is the number of investment in the first stage in scam industries. The main independent variable is Post × Cheated. 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. Standard errors are clustered at firm level. 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)
	Early Stage	Early Stage	Early Stage	Early Stage
Post × Cheated	0.033	-0.327*	-0.008	-0.231**
	(0.10)	(-1.72)	(-0.03)	(-2.38)
Industry Deal		0.295		0.492***
		(1.41)		(2.84)
Industry Amount		0.218**		0.198**
		(2.05)		(2.09)
VC Age		-1.207***		-0.872***
		(-6.11)		(-5.27)
Funds Num		-3.416***		-2.771***
		(-14.49)		(-14.68)
Total Deal Amount		-0.311***		-0.221**
		(-2.79)		(-2.28)
Total Deal Num		3.460***		2.427***
		(14.80)		(12.92)
Year × Cohort FE	Y	Y	Y	Y
Firm × Cohort FE	Y	Y	Y	Y
N	72,201	64,535	69,809	63,182
Adj. R^2	0.57	0.63	0.52	0.58

Table 15: Past Performance

This table reports the regression analyses based on the past performance. The main independent variable is Post × Cheated × Better. 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. 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. Standard errors are clustered at firm level. *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 × Better	-0.568	-0.088	-0.205	-0.002**	-0.587	-0.099***	-5.166	-3.716**
	(-1.19)	(-0.30)	(-1.52)	(-2.32)	(-1.22)	(-4.33)	(-1.45)	(-2.12)
Post × Cheated	-0.215	-0.100	-0.167*	-0.109*	-0.206	-0.065	-3.406	-3.167
	(-1.28)	(-0.69)	(-1.79)	(-1.66)	(-1.30)	(-0.46)	(-1.45)	(-1.28)
Industry Deal		0.181		0.520*		-0.249		21.284***
		(0.20)		(1.77)		(-0.26)		(2.66)
Industry Amount		0.157		0.113		0.492		4.707
		(0.35)		(1.14)		(1.05)		(1.43)
VC Age		-0.683		-0.136		-0.581		-2.866
		(-0.95)		(-0.40)		(-0.80)		(-0.28)
Funds Num		-0.391		-0.592**		-0.394		2.781
		(-0.83)		(-2.48)		(-0.84)		(0.42)
Total Deal Amount		1.516***		-0.109		1.103*		-3.776
		(2.68)		(-0.87)		(1.77)		(-1.03)
Total Deal Num		-0.952		0.525*		-0.447		-13.059
		(-0.90)		(1.70)		(-0.39)		(-1.65)
Year × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
N	1,743	1,580	1,743	1,580	1,711	1,580	1,711	1,580
Adj. R^2	0.72	0.80	0.82	0.87	0.72	0.79	0.75	0.76

Table 16: Age

This table reports the regression analyses based on the VCs age. 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 scam industries. In column (7)–(8), the dependent variable is the number of deals of investment in the scam industries. The main independent variable is Post $\times$ Cheated $\times$ Age. 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. Standard errors are clustered at firm level. 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 × Age	-0.335	-0.184***	-0.206	-0.011	0.259	-0.161**	2.039	-0.736***
	(-1.01)	(-3.84)	(-1.37)	(-0.09)	(0.82)	(-2.51)	(0.51)	(-3.18)
Post × Cheated	-0.513*	-0.025	-0.281***	-0.125	-0.474*	-0.007	-5.134*	-3.801
	(-1.84)	(-0.15)	(-2.93)	(-1.62)	(-1.67)	(-0.04)	(-1.76)	(-1.30)
Post × Age	-0.559**	0.119	-0.317***	0.020	-0.502**	0.133	-3.830**	2.109
	(-2.44)	(0.81)	(-3.23)	(0.35)	(-2.45)	(0.95)	(-2.46)	(0.99)
Industry Deal		0.885		0.564**		0.580		25.260***
		(0.76)		(2.28)		(0.47)		(3.06)
Industry Amount		0.507		0.162		0.777**		5.857
		(1.46)		(1.55)		(2.25)		(1.62)
VC Age		-0.057		0.022		0.088		4.919
		(-0.09)		(0.06)		(0.14)		(0.48)
Funds Num		-0.905*		-0.597**		-0.812*		3.255
		(-1.77)		(-2.57)		(-1.67)		(0.48)
Total Deal Amount		0.612**		-0.063		0.416*		-3.327*
		(2.45)		(-0.98)		(1.70)		(-1.68)
Total Deal Num		-0.637		0.399		-0.478		-18.286**
		(-0.66)		(1.51)		(-0.45)		(-2.28)
Year × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
N	1,762	1,589	1,762	1,589	1,730	1,589	1,730	1,589
Adj. R^2	0.69	0.78	0.83	0.87	0.69	0.78	0.75	0.77

Appendix

A1 SEC Litigation Releases and DOJ Press Releases

In this section, I introduce the institutional background of my two data sources: Securities and Exchange Commission (SEC) litigation releases and Department of Justice (DOJ) press releases.[8]

In terms of the difference in the function of enforcement agencies, the SEC takes enforcement actions and brings civil penalties, while the DOJ is responsible for civil suits and all criminal prosecutions. However, both the SEC and DOJ often enforce through joint investigations and settlement negotiations. Criminal enforcement of the federal securities laws is done through the U.S. Department of Justice and the individual U.S. Attorney's offices throughout the country. The Division of Enforcement of SEC provides assistance to United States Attorneys throughout the country by, among other things, providing access to the SEC investigative files and assigning the SEC staff to assist those offices as Special Assistant U.S. Attorneys.

A1.1 SEC Enforcement and Litigation Releases

The Division of Enforcement of SEC, created in August 1972, is to consolidate enforcement activities that previously had been handled by the various operating divisions at the SEC's headquarters in Washington. The responsibility of the Division of Enforcement is to detect and investigate a wide range of potential violations of the federal securities laws and regulations. It is also responsible for litigating the SEC's civil enforcement proceedings in the federal courts and in administrative proceedings. Although the securities laws prohibit fraudulent conduct both criminally and civilly, the SEC is responsible only for civil enforcement and administrative actions.

Many different events and sources of information can trigger a SEC investigation. An investigation, however, is not the same as a prosecution. Investigations involve fact finding by the SEC staff and are usually not public. In this way, the mere existence of an investigation does not harm an individual or entity. During an investigation, neither the staff nor the SEC makes any determination of wrongdoing. If, however, the staff ultimately believes that there has been a violation of the securities laws, it generally will make a recommendation to the SEC to take further action. The SEC then determines whether to file a public civil lawsuit in court or to institute a public administrative proceeding and whether to accept offers of settlement, if there are any.

In civil suits, the SEC seeks injunctions, which are orders that prohibit future violations; a person who violates an injunction is subject to fines or imprisonment for contempt. In addition, the SEC often seeks civil money penalties and the disgorgement of illegal profits. In certain circumstances, the SEC also may seek, among other things, a court order barring or suspending individuals from acting as corporate officers or directors. The Division of Enforcement posts releases describing the SEC's litigation in federal district court.[9] These releases are one of my data sources in this paper. I use a web crawler to download all litigation cases released by SEC.

[8] The information is mainly from SEC's and DOJ's website. For more information, please refer to https://www.sec.gov/enforce/Article/enforce-about, https://www.sec.gov/about/offices/oia/oia_enforce/overviewenfor.pdf, https://www.justice.gov/criminal-fraud, and https://www.justice.gov/criminal-fraud/file/1472076/download

[9] https://www.sec.gov/litigation/litreleases.htm

A1.2 DOJ and Press Releases

An investigation by the SEC could lead to a civil enforcement proceeding by the SEC or a criminal prosecution by the DOJ. Section 24 of the Securities Act of 1933 and Section 32 of the Exchange Act of 1934 provide that any person or entity may be criminally prosecuted for willful violations of the federal securities laws. Whether an investigation for violations of the federal securities laws will proceed civilly or criminally depends on the circumstances of each case. However, criminal prosecutions are typically more severe since they entail significant criminal penalties, injunctions, disgorgement orders, and possibly jail time. Along with the assistance of U.S. attorneys throughout the nation, criminal violations of the securities laws are prosecuted by the Fraud Section of the Criminal Division within the DOJ.

The Fraud Section plays a unique and essential role in the Department's fight against sophisticated economic crime. The Fraud Section has three litigating units: The Foreign Corrupt Practices Act (FCPA) Unit, The Health Care Fraud (HCF) Unit, and The Market Integrity and Major Frauds (MIMF) Unit. The most relevant unit to this paper is the MIMF Unit which focuses on the prosecution of complex and sophisticated securities, commodities, corporate, investment, and cryptocurrency-related fraud cases. The MIMF Unit works in parallel with regulatory partners at the SEC, Commodity Futures Trading Commission (CFTC), and other agencies to tackle major national and international fraud schemes.

The DOJ periodically publishes press releases of ongoing cases on their website.[10] I use a web crawler to download all cases in the category of Financial Fraud.

[10] https://www.justice.gov/news

Table A1: Definition and Source

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

Variable	Definition	Source
Baseline Analysis		
Cheated	Dummy variable which is one if a VC has invested in scam startups, otherwise is zero	DOJ/SEC
Post	Dummy variable which is one after the startups turn out to be scam, otherwise is zero	DOJ/SEC
Deal Amount	The amount of investment a VC makes in a given year	VentureXpert
Deal Num	The number of investment a VC makes in a given year	VentureXpert
Deal Amount IND	The amount of investment a VC makes in the specific frauds industries in a given year	VentureXpert
Deal Num IND	The number of investment deal a VC makes in the specific frauds industries in a given year	VentureXpert
Industry Deal Amount	The amount of investment a VC makes in all frauds industries in a given year	VentureXpert
Industry Deal Num	The number of investment deal a VC makes in all frauds industries in a given year	VentureXpert
Fund Year	The given year minus the last year which a VC received fundraising	VentureXpert
Company Num	The cumulative number of companies a VC has invested	VentureXpert
VC Age	The age of a VC	VentureXpert
Fund Num	The cumulative number of fund a VC manages	VentureXpert
Total Deal Amount	The cumulative amount of investment a VC has made	VentureXpert
Total Deal Num	The cumulative number of investment a VC has made	VentureXpert
Fundraising Analysis		
Amount	The amount of funds a VC gets in a given year	VentureXpert
Time	Duration of a VC gets the next fundraising	VentureXpert
Pre Amount	The cumulative amount of funds a VC has got	VentureXpert
Industry Performance	The ratio of successful exits in the given industry under one fundraising	VentureXpert
Performance	The ratio of successful exits under one fundraising	VentureXpert
Better	Dummy variable which is one if a VC has better investment performance than the average, otherwise is zero	VentureXpert
Age	Dummy variable which is one if a VC is older than the average, otherwise is zero	VentureXpert

Table A2: Ten Representative Scam Startups

This table reports 10 representative scam startups I use to build words dictionary. The amount of funding they raised are shown.

Startups	VC (US$ millions)
Theranos	1,300
Zenefits	584
Outcome Health	500
Mozido	314
Hampton Creek	240
NS8	158
UBiome	110
Nanotech Engineering	68
Raze Therapeutics	24
Trustify	19

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		

Table A4: First Round Investment

This table reports the regression analyses based on the first round investment. 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 main independent variable is Post × Cheated. 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	-0.106	-0.023	-0.187***	-0.083**	-0.121	-0.021***	-0.184***	-0.085**
	(-1.01)	(-0.26)	(-3.90)	(-2.00)	(-1.18)	(-3.53)	(-3.85)	(-2.06)
Industry Deal		-0.161**		-0.101***		-0.140**		0.036
		(-2.40)		(-3.50)		(-2.08)		(1.18)
Industry Amount		0.342***		-0.007		0.444***		0.008
		(9.45)		(-0.63)		(12.57)		(0.66)
VC Age		-0.121**		-0.311***		-0.156***		-0.251***
		(-2.32)		(-10.79)		(-2.89)		(-8.60)
Funds Num		-0.980***		-0.957***		-0.839***		-0.909***
		(-14.95)		(-25.72)		(-12.77)		(-24.21)
Total Deal Amount		0.735***		0.019*		0.552***		0.005
		(21.21)		(1.76)		(16.47)		(0.43)
Total Deal Num		0.281***		1.247***		0.246***		1.045***
		(4.19)		(39.75)		(3.64)		(31.17)
Year × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm × Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y
N	212,701	162,239	212,701	162,239	188,120	153,516	188,125	153,516
Adj. R^2	0.55	0.70	0.64	0.77	0.57	0.70	0.64	0.76