Disentangling Crowdfunding from Fraudfunding**

Douglas Cumming*, Lars Hornuf†, Moein Karami‡, and Denis Schweizer§

Abstract

Crowdfunding fraud is an oft-repeated pronounced concern of many regulators. Using Kickstarter and Indiegogo, the two largest crowdfunding platforms, we conduct an exhaustive search of all fraud cases from 2010 through 2015, spanning nine countries. We present the first ever evidence that fraudsters in crowdfunding markets have specific characteristics: they are less likely to have engaged in prior crowdfunding activities, they are less likely to have a social media presence, and they are more likely to provide poorly worded and confusing campaign pitches with a greater number of enticements through pledge categories.

JEL Classification: G21, G24, G32, K22, L26

Keywords: Crowdfunding, Entrepreneurial Finance, Fraud, Internet

* York University - Schulich School of Business, Professor and Ontario Research Chair, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada, Web: http://ssrn.com/author=75390, Phone: +1-416-736-2100 ext. 77942, Fax: +1-416-736-5687, e-mail: dcumming@schulich.yorku.ca.

† University of Trier, Assistant Professor of Law and Economics, University of Trier, Behringstraße 21, Trier, 54296, Germany, and Affiliated Research Fellow, Max Planck Institute for Innovation and Competition, Web: http://hornuf.com, Phone +49 651 201 4744, Fax +49 651 201 4742, e-mail: hornuf@uni-trier.de.

‡ Concordia University, PhD Candidate of Finance, John Molson School of Business Building, 1450 Guy, Montreal, Quebec, Canada H3H 0A1, e-mail: Moein.Karami@concordia.ca.

§ Concordia University, Associate Professor of Finance, John Molson School of Business Building, 1450 Guy, Montreal, Quebec, Canada H3H 0A1, Phone: +1 514-848-2424, ext. 2926, Fax: +1-514-848-4500, e-mail: Denis.Schweizer@concordia.ca.

**Acknowledgments: We are grateful to Harjeet S. Bhabra, Steven Bradford, Jonathan M. Karpoff, Sofia Johan, Juliane Proelss, Rahul Ravi, Armin Schwienbacher, Thomas Walker, and Tingyu Zhou, as well as participants of the seminar at the John Molson School of Business, and the conference "Entrepreneurship, the Internet, and Fraud: Managerial and Policy Implications" (Montreal, Canada) for helpful comments and suggestions. This project has been supported by the Social Sciences & Humanities Research Council of Canada. Denis Schweizer gratefully acknowledges the financial support provided through the Manulife Professorship.

Disentangling Crowdfunding from Fraudfunding

ABSTRACT

Crowdfunding fraud is an oft-repeated pronounced concern of many regulators. Using Kickstarter and Indiegogo, the two largest crowdfunding platforms, we conduct an exhaustive search of all fraud cases from 2010 through 2015, spanning nine countries. We present the first ever evidence that fraudsters in crowdfunding markets have specific characteristics: they are less likely to have engaged in prior crowdfunding activities, they are less likely to have a social media presence, and they are more likely to provide poorly worded and confusing campaign pitches with a greater number of enticements through pledge categories.

JEL Classification: G21, G24, G32, K22, L26

Keywords: Crowdfunding, Entrepreneurial Finance, Fraud, Internet

It's a credit to Kickstarter and the collective power of the crowd to identify fraud...
- CNN Money, June 17, 2013[1]

If you utter the word "crowdfunding" in front of a dusty old-fashioned securities lawyer, make sure you have a fully charged defibrillator on hand. Perhaps a fully equipped contingent of ER doctors and nurses. It won't be pretty.
- Financial Post, July 31, 2013[2]

1. Introduction

A central tenet of securities laws is the disclosure of fair and complete information, which in turn encompasses the prohibition of fraud, misrepresentation, and market misconduct (La Porta et al., 2006).[3] Over the years, a great deal of research has focused on the causes and consequences of fraud in the context of disclosure laws and enforcement (Fang, Huang, and Karpoff, 2016; Giannetti and Wang, 2016; Khanna, Kim, and Lu, 2015; Wang, Winton, and Yu, 2010). Research has shown that up to 14% of U.S. publicly listed firms engage in fraud each year, and fraud can cost from 20%-38% of firm value. This equates to hundreds of billions of dollars in lost value in the U.S. each year (Dyck et al., 2010, 2014; 2016; Karpoff et al., 2008a,b, 2012).

The crowdfunding phenomenon represents the antithesis of what securities laws were meant to accomplish. In the spirit of the colorful *Financial Post* quote above, and in view of the less stringent disclosure rules and dearth of enforcement surrounding crowdfunding, we may expect to find that fraud is a massive problem in crowdfunding markets. Moreover, because crowdfunding portals act as gatekeepers (Coffee, 2006) for project creators that seek funding from the crowd, the questions arise as to whether fraud can be detected *ex ante*, and whether portals should be required to implement more appropriate prescreening mechanisms.

[1] See http://money.cnn.com/2013/06/17/technology/kickstarter-scam-kobe-jerky.
[2] See http://business.financialpost.com/fp-comment/extraordinary-popular-delusions-and-the-madness-of-crowd funding.
[3] See also https://www.sec.gov/about/laws.shtml.

In this paper, we conduct a thorough and methodical search of media reports from 2010 through 2015 across nine countries (Australia, Canada, China, Germany, Hong Kong, Israel, Spain, the U.K., and the U.S.) to identify all crowdfunding fraud cases for all projects on Kickstarter and Indiegogo (the two most commonly used reward- and donation-based platforms).[4] We consider "detected" fraud in cases of outright misrepresentation, such as in the "Kobe Red" case, which involved the production of Japanese beer-fed Kobe beef jerky.[5] Kickstarter ultimately suspended this project a few minutes before the scheduled end date of the campaign's funding period. We define "suspected" fraud in cases where 1) the rewards are significantly delayed (more than one year), 2) campaign initiators cease communicating with their backers for more than six months after an unmet delivery date, or 3) the promised product is never delivered and the backers are not fully refunded (see section 2 for more details about our classification method). In either case, when these detected or suspected fraud cases are reported in the news media and on consumer advocacy websites, they are picked up in our dataset.

In this paper, we explore the factors predicting detected or suspected crowdfunding fraud. As with other fraud research, we face the issue of detection. The probability of a fraud being detected is a function of the probability of it having been committed and the rate of detection. This issue is relevant for our work in two respects. First, it affects the overall frequency of reported fraud discussed above. Second, it affects the interpretation of our data and our multivariate tests.

While fraud in crowdfunding has been the concern of many regulators (Hornuf and Schwienbacher, 2016), it is potentially of greater importance for the industry itself. If crowdfunding

[4] We do not study fraud in equity crowdfunding markets. As of April 30, 2016, only two fraudulent campaigns in equity crowdfunding have been reported: 1) Ascenergy, which raised USD $5 million from approximately 90 investors on Crowdfunder, Fundable, and EquityNet (see http://ncfacanada.org/the-first-investment-crowdfunding-fraud-what-does-this-mean-for-the-industry), and 2) vibewrite, which raised EUR 560,000 on Seedmatch, and was charged with a delayed filing of insolvency (see http://www.gruenderszene.de/allgemein/vibewrite-insolvenzverschleppung-anzeige).
[5] The "crowd" detected the fraud because it noticed several suspicious campaign characteristics, such as little personal information about project creators and discrepancies between the high cost of production and the low goal amount requested from backers; see footnote 1.

exhibits high levels of fraud, this new way of financing could rapidly disappear again. Our central theoretical proposition is that fraud can be predicted based on economic and behavioral theories. Our analysis finds a major link with the likelihood that fraudulent crowdfunders use social media, such as Facebook, where there is a stronger connection between crowdfunders and the crowd.

On the one hand, we acknowledge that social media could facilitate fraud detection, because a fraudulent act is more likely to appear in our dataset if the project creator uses social media. Higher social media presence provides the crowd with more information about the creator and the project, and thus leads to a higher probability of detection if the creator is a fraudster. On the other hand, we theorize that social media presence is a predictor of crowdfunding fraud. Fraudsters may be less likely to appear on social media in an effort to avoid public scrutiny of their crimes. In that case, we should observe a negative correlation between social media presence and crowdfunding fraud in our dataset.

Our data ultimately indicate the latter: There is a strong and robust negative correlation between a crowdfunding fraudster being on social media in our dataset and the likelihood of a fraudulent activity. This implies that either 1) our data does not suffer from systematic biases of missing fraud cases, or 2) if a bias exists, the negative relationship between social media presence and crowdfunding fraud is even stronger than the observed negative effect. In either case, our conclusion of a negative relationship between social media presence and the probability of commission of a crowdfunding fraud remains unchanged. Moreover, this evidence is not merely statistically significant but also economically important. For example, if all other things remain equal, a Facebook presence reduces the likelihood of fraud commission by over 50%.

Our paper is related to a growing literature on crowdfunding that has, to date, focused primarily on the determinants of funding success (e.g., Agrawal et al., 2015; Ahlers et al., 2015; Bayus, 2013; Belleflamme et al., 2013, 2014; Colombo et al., 2015; Mollick, 2014; Vismara, 2016). Other papers have examined the role of securities regulation in crowdfunding markets (Bradford, 2012; Hornuf and Schwienbacher, 2016). However, prior work did not empirically study the frequency of fraud in crowdfunding markets or its determinants. To this end, we make an empirical contribution by documenting the frequency of crowdfunding fraud and its empirical determinants.

The remainder of this paper is organized as follows. Section 2 provides an overview of the legal treatment of fraud in crowdfunding markets, while section 3 summarizes our hypotheses. The data are introduced in section 4, and the methodology is discussed in section 5. Section 6 provides univariate and multivariate analyses of the data, followed by a discussion of the results and several robustness checks. Section 7 provides univariate evidence on the differences between fraud and non-fraud campaigns after they end, and section 8 concludes.

2. Legal Applications to Fraud in Crowdfunding Markets

Securities laws in the U.S. have several antifraud provisions that allow investors and the SEC to bring legal actions. These provisions apply in the context of a purchase or sale of a security. While equity crowdfunding and peer-to-peer lending issuers almost inevitably offer securities (Bradford, 2012), neither donation- nor reward-based crowdfunding includes securities as defined under the Securities Act § 2(a)(1) or the Exchange Act § 3(a)(10). Thus, backers cannot recover damages from fraudulent campaign creators under U.S. securities laws. Moreover, the SEC has no jurisdiction over these matters, and consequently cannot impose fines or achieve injunctive relief, as would be possible for fraudulent security offerings on traditional capital markets.

However, many jurisdictions provide common law or general civil law code fraud actions, even if no securities are involved. In the U.S., for example, backers can take action under state law if the following five elements are present: 1) the creator makes a false statement related to a material fact, 2) the creator knows that the statement is untrue, 3) the creator intends to deceive the backer, 4) the backer reasonably relied on the statements of the creator when making a decision to invest, and 5) the backer was injured, which in a crowdfunding context is likely if funds are lost and no product was delivered. In order to recover money pledged by crowdfunding, a backer would therefore have to show a court that the campaign creator committed a fraud, *and* that the backer relied on false statements in choosing to invest.

One problem with private remedies is that the amount of the claims often does not justify the costs of litigation. Class actions may be potentially suitable in cases where many backers deceived by the same creator can consolidate their claims. Given that the pledges of most crowdfunding contributions are extremely small, however, even class actions may not be feasible, because legal cases are too expensive, time consuming, and emotionally exhausting relative to the expected refund. Thus, the most effective remedies need to come through government agencies.

Finally, there are criminal provisions prohibiting fraud in a crowdfunding context. The Federal Trade Commission (FTC) has jurisdiction when crowdfunding involves the sale of a good (which is typically true with pre-purchases, and potentially in cases when rewards are offered). Importantly, the FTC has the authority to impose monetary penalties on fraudulent campaign creators. Moreover, it may also obtain civil penalties if fraudulent entrepreneurs persistently violate its standards.

Currently, we are aware of only a single case where the FTC acted on a crowdfunding fraud: a case involving a campaign set up by Erik Chevalier, which was known as *The Doom That Came To Atlantic City!*, and was created under the business synonym *The Forking Path, Co.* In June 2012, 1,246 backers had pledged a total of USD $122,874 for Chevalier to develop a new board game. As part of the campaign, he promised backers that they could pre-purchase a copy of the game as well as specially designed action figures. However, after fourteen months, Chevalier declared that he had terminated the project and intended to refund the backers. According to the FTC, instead of creating the game, Chevalier had spent most of the money on his own expenses, such as rent, a move to Oregon, personal equipment, and licenses for an unrelated project (FTC 2015). As a result, the FTC filed a complaint for a permanent injunction, followed by an order of judgment for USD $111,793.71 (*FTC v. Chevalier*, No. 3:15-cv-01029-AC [D. Or. June 10, 2015]). The judgment was suspended, however, due to Chevalier's inability to pay.

In another Kickstarter campaign called *Asylum Playing Cards*, Edward J. Polchlopek III, the president of Altius Management, LLC, attracted 810 backers pledging a total of USD $25,146 in October 2012. In this case, the campaign creator promised backers he would print and market a deck of playing cards created by a Serbian artist. After failing to deliver the promised rewards, and ending communication with the crowd in July 2013, the King County Superior Court ordered a total of USD $668 in restitution be made to 31 backers living in Washington State. Furthermore, court commissioner Henry Judson ordered another USD $1,000 per violation (USD $31,000 in total) in civil penalties for violating the state Consumer Protection Act, as well as USD $23,183 to cover the costs and fees of bringing the case (*State of Washington v. Polchlopek*, No. 14-2-12425-SEA [Wash. Super. Ct. April 30, 2014]).

To summarize, fraudsters in a pre-purchase crowdfunding campaign might anticipate being detected as the campaign progresses and the delivery date approaches. Despite the sometimes weak incentives of backers who may have pledged only small amounts to bring legal actions, fraudsters are still subject to prosecutions by FTC or state attorneys general. However, the inactivity of government agencies such as the FTC until 2015, and the lack of private actions, may have provided fraudsters with sufficient incentives to engage in deceptive activities.

3. Theory and Hypotheses

Economists and psychologists have provided various explanations for why individuals engage in fraudulent activities. Currently, the most prominent comes from Becker's (1968, 1993) theory of crime and punishment, which explains criminal activities using the "rational actor" model. In this model, the probability and severity of the anticipated punishment decrease the expected utility from deception, while the personal gains of engaging in a fraud increase the likelihood that a crowdfunding campaign creator will commit a crime. Thus, a campaign creator should be more likely to deceive backers when the expected penalties for deception are weak.

In contrast, if detection is almost certain, and the expected punishment exceeds any personal gains from fraud, campaign creators are likely to refrain from deceptive tactics. Becker's (1968, 1993) theory is supported by numerous experimental studies showing, for example, that detection probability has a larger impact on deterring fraud than the magnitude of punishment (Nagin and Pogarsky, 2003). In another experiment, where subjects could misrepresent the financial statements of a firm, increasing the utility of fraud by 30% increased actual fraudulent behavior by 17% (Gibson et al., 2013).

In a crowdfunding context, detection is often a matter of timing. Consider the case where a crowdfunding campaign has reached the delivery date but the creator has ceased credible communication. It might become obvious to many backers that the campaign has failed. If there is no good explanation for the failure, and if the campaign creator has, e.g., spent the money on personal expenses or another project, the fraud is recognized quickly once backers vent their anger on the project website and Internet blogs. Arguable, creators might anticipate such actions and try to carefully hide the scam. Concealing the fraud might be easy if few backers support the campaign. Generally, the crowd responds quickly though when the creator fails to deliver the long-awaited product.

The second variable in Becker's (1968, 1993) model, as we mention above, is the magnitude of punishment. The analysis in the previous subsection shows that the *expected* fines do not necessarily exceed the funds collected. The difference may be a close proxy for the creator's personal gains from the fraud. In the case of Erik Chevalier, the FTC also mandated that he deliver a copy of the order to his business partners, which may be an additional non-monetary punishment. However, the financial penalty is still likely to be too lenient to effectively deter fraudulent behavior.

More recently, social psychologists have argued that, when people are acting in a dishonest manner, they nevertheless remain concerned with maintaining a positive self-concept (Gino et al., 2009; Jiang, 2013; Mann et al. 2016a, Mazar et al., 2008). For example, Mann et al. (2016b) focus on non-violent crimes, and find that internal sanctions provided the strongest deterrent to such crimes. The effect of legal sanctions was weaker, and varied across countries.

As a result, fraud in crowdfunding campaigns may not follow a solely economic calculation by the project creator, but may also reflect his or her personal attitudes. Using the predictions from the theory of crime and punishment, as well as self-concept theory, portals and backers may be able to identify fraudsters *ex ante* by evaluating certain campaign features.

In the realm of crowdfunding, we have identified four broad themes where backers could theoretically identify fraud based on available information: 1) *creator(s)' characteristics/ background*, 2) *social media affinity*, 3) *campaign funding and reward structure*, and 4) *campaign description details*.

First, in line with the theory of crime and punishment, we expect fraudsters not to provide their real names, so as to deter effective detection. Moreover, fraudsters may not use their real name because an alias does not put any name-specific reputation at risk. Setting up a new crowdfunding campaign is also easier if you were previously caught committing a fraud under a different name. Further, it has been shown that signing a document can increase moral saliency (Shu et al., 2011a; Shu et al., 2011b), and consequently decrease cheating behavior. In this vein, providing one's name may remind creators of their own moral standards, and deter them from entering into a deception.

Note further that creating multiple campaigns constitutes a cost to the creator in terms of both time and money, and would thus decrease the utility from committing a fraud. Additionally, in line with Diamond (1989), creators build a reputation by engaging more frequently in the market and would consequently suffer a larger loss if they engage in fraud. Creating multiple supposedly successful campaigns in order to commit one fraud may also conflict later with a creator's self-perception. If a creator does run multiple campaigns, then evaluating his own success stories and ability may make it mentally difficult to perceive himself as fully dishonest (Colombo and Shafi, 2016; see also Bertoni et al., 2011; and Colombo and Piva, 2012). Similarly, creators who have previously backed other projects are likely to believe in the democratic and supportive idea of crowdfunding (Kim and Hann, 2015). This can make it difficult for them to reconcile the idea of leading a scam later on. In effect, we predict a negative relationship between crowdfunding fraud and the intensity with which a creator uses crowdfunding as a backer or a creator, which is illustrated in Hypothesis 1.

> **Hypothesis 1 (Creator(s)' Characteristics and Background):** *Crowdfunding fraudsters do not provide their formal names and do not engage in other crowdfunding activities.*

Second, backers can easily screen creators' social media activities on the Internet. Fraudsters may try to avoid this scrutiny by, e.g., not having any social media presence because social media facilitates fraud detection. Furthermore, a social media presence is an indicator that the creator has more to lose from cheating in terms of social connections and is potentially also subject to closer monitoring via social media contacts as compared to a creator without a media presence. On the other hand, fraudsters may also manipulate personal or professional social media information, such as a Facebook page that falsely lists the number of friends or likes of a project (Wessel et al., 2015).

Hence, consistent with a purely rationality based conjecture it is *per se* not clear whether fraudsters have more or less social media contacts.

However, the more professional the scam, the costlier it becomes in terms of time and money, and the lower the personal gain from fraud. The same holds for external links on a campaign website that lead to other fake websites that supposedly support the trustworthiness of the campaign. Moreover, fraudsters need to consider that being connected to actual friends on Facebook, and providing many external links to business partners or people who endorse the project, can be emotionally costly to the creator once the fraud is uncovered. Supporters of the project may question the creator intensely, which can make it more uncomfortable to come up with plausible justifications (Shalvi et al., 2015).

Thus, we posit that fraudsters concerned about maintaining positive self-images will be less present on social media. We should thus observe a negative correlation between social media use and fraud.

> **Hypothesis 2 (Social Media Affinity):** *Crowdfunding fraudsters are less well-connected in the social media arena, and provide fewer external links.*

Third, the funding and reward structure of the campaign can provide credible signals, in the spirit of Spence (1973). For example, more confident creators may *ex ante* restrict the duration of the funding period because they strongly believe their project will be fully funded very rapidly. In contrast, we may observe a different rationale with fraudsters, because they very likely cannot send credible signals to the crowd. Therefore, fraudsters may believe it is optimal to keep the funding period ongoing to raise as much capital as possible. However, longer funding periods may also make detection more likely, thus also increasing the risk of not receiving funds. Consequently, it remains an empirical question whether a longer funding period reduces or increases the probability of fraud.

Furthermore, because fraudsters do not intend to ship product or continue communication with backers, they tend to be more open to raising small amounts by as many backers as possible. While fraud in a crowdfunding campaign is almost inevitably detected once the creator fails to deliver the product, the ultimate prosecution of the scam may be a more important factor to the fraudster. As noted above, the smaller the amount invested by backers, the less likely the amount of the claims will justify the costs of litigation. In line with this conjecture, we believe fraudsters will target as many different backers as possible, ideally spending only smaller amounts of money. One way to achieve this is by creating many different pledge categories, so that backers can easily provide many levels of small size contributions.

Research on the manipulation of stock markets has long explored so-called “pump and dump” schemes. These schemes involve fraudsters acquiring long positions in stocks before heavily promoting them through online chat forums or by spoof trading (deleting orders before execution to keep up appearances of a very active order book). Fraudsters thereby encourage other investors to purchase these stocks at successively higher prices, and then sell their own shares in large quantities at the higher prices. In a similar way, crowdfunding fraudsters can more heavily promote a campaign by offering many project enticements with various reward levels (Belleflamme et al., 2014; Mollick, 2014).

> **Hypothesis 3 (Campaign Funding and Reward Structure):** *Longer funding periods, smaller minimum pledges, and a large number of pledge categories are positively correlated with the likelihood of crowdfunding fraud.*

Fourth, it is commonly accepted that perpetrating securities fraud in publicly traded firms is easier when confusion exists among investors (Fischel, 1982; Perino, 1998; Simmonds et al., 1992). In crowdfunding markets, the main place for a crowdfunder to learn about a project is through the description, which is normally a few thousand words (Cumming et al., 2014). Crowdfunding fraudsters are therefore less likely to provide a clearly worded description in order to foster confusion and ideally perpetrate the fraud without detection.

Just as in academia, writing concise and convincing descriptive text for a project also requires effort and skill. In line with Becker's (1968, 1993) theory of crime and punishment, it constitutes a cost to the fraudster to create such a text, which reduces his or her personal gains from the scam. Furthermore, it is likely to be difficult to accurately and perfectly describe a product that does not exist, and was never intended to exist in the first place. Ultimately, fraudsters may target a less educated and broader crowd, which would need to be addressed in simpler, easier to read terminology. We therefore derive Hypothesis 4:

> **Hypothesis 4 (Campaign Description Details):** *Crowdfunding fraudsters use simple wording (i.e., easier to read descriptions) and are less likely to present transparent pitches.*

The next sections of this paper empirically examine these four main hypotheses, while accounting for other relevant factors that may influence the probability of a campaign being fraudulent.

4. Sample Construction

The legal definition of fraud, as outlined in section 2, is not easy to operationalize for an empirical study on crowdfunding because only a few cases were decided by an ordinary judge so far. Therefore, we focus on what is considered industrywide as *detected fraud*, and as *suspected fraud* (see, for example, *Crowdfund Insider*[6] and Table 1, panel B, for an overview).

[6] See http://www.crowdfundinsider.com/2014/03/34255-crowdfunding-fraud-big-threat.

The first category, *detected fraud*, includes 1) *pre-emptive fraud*, which occurs when a suspected crowdfunding campaign is either suspended by the portal or cancelled by the creator before money is transferred to the creator's account after funding has ended. Both are typically a consequence of a significant number of backer complaints to the platform provider, or of numerous postings in forums or on blogs that the campaign carries a risk of fraud, and 2) *attempted fraud*, which occurs if the fraud was not originally detected during the campaign's funding period, and the amount raised is transferred to the campaign initiators. After the funding is completed, backers may still find out that creators, for example, attempted to resell pre-existing products as part of their campaign, or that they misrepresented material facts, have used intellectual property they do not hold legal rights to, or that the project is a fake altogether. The fraud can be confirmed through news articles about the campaigns (e.g., an actual lawsuit against the creator may have been brought), or there may be news reports that the project is fraudulent.

The second category, *suspected fraud*, occurs when rewards are substantially delayed by more than one year (condition 1a), creators have ceased communication with their backers for more than six months after the promised delivery date (condition 1b), and rewards are not delivered until our data collection date and backers are not fully refunded (condition 1c), or the rewards are changed, to the disadvantage of the backers (condition 2). Detection of campaigns where rewards have been changed is straightforward by studying news articles on a particular campaign, or by reading comments posted by backers after the rewards are delivered. For example, consider the case of Indiegogo's *Kreyos: The Only Smartwatch with Voice and Gesture Controls*. The creators raised more than USD $1.5 million, which was 1,500% of their goal amount, but subsequently failed to deliver the quality and features promised (violating condition 2), with news articles claiming the creators spent a significant amount of the money raised on personal expenses. The project was described on the Kickscammed website as "Over $1 million in funding, and looks like the company has folded after delivering a very sub-standard product that did not meet specifications and has an inoperable App". On the campaign's "comments-section" many backers let off steam by complaining about product quality and that they would never pledge on a crowdfunding campaign again.[7]

However, if the delivery of the rewards is delayed, it can be more difficult to distinguish between projects that have failed, and those experiencing normal setbacks, such as unforeseen challenges or technological issues. To overcome this problem, we consider a campaign with delayed rewards as being in the suspected fraud category, but *only* if 1) rewards are significantly delayed (for at least one year), 2) the creator has not communicated with backers for at least six months after the originally promised delivery date, and 3) the promised reward is not delivered until the end of our observation period, and backers were not fully refunded.[8]

To better understand which campaigns are flagged as *suspected fraud* (case of late delivery), we explore the Kickstarter campaign of FINsix, which aimed to build a phone-sized charger for a laptop. The latest promised delivery date for the charger was March 2015, but the charger was not delivered until March 2016, thereby meeting condition 1a. However, the campaign creators were in constant communication with their backers (fifteen times over the course of the year). Therefore, condition 1b was not fulfilled, and the campaign would not be classified as *suspected fraud*.

We followed the FINsix campaign to see if the rewards were finally delivered, and they were.[9] So, even if the delivery was clearly delayed for more than one year, and even if the creators had

[7] In the "Kreyos" case, some comments include: "I recieved end of last year sometime. Never worked or synced up to the phone. Sent emails for support, no reply. This was daylight robbery and nothing can be done about it apparently. Its ugly and it been in drawer ever since. Useless…"; "How many of us are out there? Class Action possibility!"; "This is why I stopped using Indiegogo."

[8] Our observation period ranges from 2010 through 2015, and our data collection ends on April 30, 2016. With our identification strategy, we are unable to distinguish "fraud" from "late delivery" for all delivery dates from May 2015 onward. However, this is only relevant for two campaigns. In both of these cases, the rewards were delayed by more than ten months, and the creators had not been in communication with the backers after the promised delivery date for more than six months. Therefore, we also consider these two cases to be fraudulent.

[9] We cease checking after April 30, 2016, which is the end of our observation period.

ceased communicating with their backers for more than six months but finally delivered the reward, we would not code this as *suspected fraud* because the promised reward was ultimately delivered without any disadvantageous changes to product quality.

Note that there are other potential forms of fraud in crowdfunding that we do not focus on here because they are difficult or impossible to detect in a consistent and comprehensive manner. These include so-called *stillborn fraud*, where a potential fraud campaign is rejected by the crowdfunding portal before it is launched. Fraud is also not necessarily limited to project creators; there have been cases of reported fraud by crowdfunding backers and even by some portals themselves.

One notorious backer fraud case involves a Kickstarter donor, Encik Farhan, who targeted over 100 projects with fraudulent pledges. He then withdrew his cash once the rewards were received by using his credit card's chargeback policy[10] (chargebacks can be requested by buyers who claim, e.g., that they purchased something that was not delivered, or that the item was materially misrepresented). The credit card company investigates and often rules in favor of the cardholder, as it can be difficult for merchants to prove the fraudulent intent of a buyer. These actions can have a substantial effect on creators' finances, because of the losses as well as the chargeback fees they incur.[11]

There is no commercial database available for fraud cases in crowdfunding, so we hand-collect data from the two leading crowdfunding platforms—Kickstarter and Indiegogo. Our sample covers all actual and potential fraud campaigns identified through a website called Kickscammed (http://kickscammed.com). Kickscammed's purpose is to offer the crowd an opportunity to report suspicious or fraudulent activities on both Kickstarter and Indiegogo.

[10] https://ignitiondeck.com/id/crowdfunding-platform-backer-fraud.
[11] http://www.theverge.com/2013/11/8/5081806/kickstarter-alleged-chargeback-fraud-hits-over-100-campaigns.

We include all reported scams in crowdfunding campaigns on both portals, and, as of April 30, 2016, we were able to identify and confirm 196 fraud cases for the 2010-2015 period that were reported on Kickscammed and met our criteria to be categorized as detected fraud or suspected fraud. However, Kickscammed's website does not necessarily cover all instances of fraudulent activity on both portals, so we complement our dataset with a news search using Google, Factiva, and LexisNexis.

We first searched for the term "fraud," in combination with "crowdfunding" in general. Second, we used different synonyms for fraud, such as "cheat," "untruth," "deception," "scam," "dishonesty," "trickery," "double-dealing," and "unlawful." Since there are many types of conceivable fraud, we also combined the fraud synonyms with the words "accounting," "payroll," "double check," "over-ordering," "friendship," "benefit," "monetary gain," "pre-empted," "stillborn," "attempted," "suspected," "perceived," "backer," "backer-creator," "broker," and "portal."

Third, we added a country dimension to the search strategy to capture the four countries with the largest amount of backer contributions: "the U.S.," "the U.K.," "Germany," and "Australia." Twenty new fraud cases that were not previously identified on Kickscammed were obtained through the news search. Our initial dataset is therefore comprised of 216 fraudulent campaigns (204 from Kickstarter and 12 from Indiegogo) (see Table 1, panel A).

In order to identify a non-fraud control group with similar characteristics, we apply a propensity score matching (PSM) algorithm. The vector of control variables includes the goal amount, portal dummy variable, country of the creator, year the campaign was posted (start dates of the matched non-fraud campaigns are within six-month intervals before and after the start date of the fraud campaign), and business category. When a fraud campaign had multiple matches with identical propensity scores, we used a random uniform function to choose one match among all the non-fraud matches with identical propensity scores. In this way, we aim to avoid sorting the data, which could potentially affect our results.

To derive our final dataset, we exclude 1) seven fraud campaigns for which no non-fraud campaigns could be found through propensity score matching because there was no control campaign launched within the six-month intervals before and after the start date of the fraud campaign from the same country in the same business category, and 2) two fraud campaigns that were missing information on matching criteria. We therefore consider 207 one-to-one pairs of matched non-fraud campaigns in our analysis. Table 1, panel A, summarizes the fraud and PSM non-fraud samples. The final dataset contains 414 crowdfunding campaigns: 207 fraud cases, and 207 PSM non-fraud matches. We also checked the campaign webpages of all the non-fraud matches to ensure that none of them were suspected of engaging in any fraudulent behavior.

Panel B of Table 1 also illustrates the differences in the number of fraud cases across Indiegogo and Kickstarter. Specifically, on Indiegogo, any campaign can be posted without a review process if the creators abide by certain basic rules, such as, e.g., not launching projects related to firearms or narcotics. Furthermore, on Indiegogo, a creator can choose to keep all the money raised in a campaign even if it does not meet its capital goal. 95% choose this option (see Cumming et al., 2014).

In contrast, on Kickstarter, the rules are stricter, and creators can only keep the money they raise if they reach at least the stated goal amount. However, presumably because of its better branding and marketing, Kickstarter has far more completed campaigns and has raised much higher volume amounts than Indiegogo. According to the Kickstarter website, more than USD $2.5 billion has been pledged to more than 110,000 successful projects, compared to USD $0.8 billion raised for 275,000 projects on Indiegogo.[12]

Regarding the coverage of fraudulent campaigns, our data shows that Kickstarter had 197 cases, dramatically more than the 10 found on Indiegogo. One reason for this large difference may be because Kickstarter has occasionally suspended potentially fraudulent cases, which are deterministically classified by us as "detected" fraud cases, while Indiegogo does not suspend projects after they have launched. Moreover, all proposed campaigns undergo a prescreening by Kickstarter before being posted and made available for funding, and Kickstarter rejects about 25% of all submitted campaigns (see Lakhani et al., 2014). This review process tends to filter out potential fraudulent campaigns in the first place, but also increases backers' expectations that they will not be (or are unlikely to be) subject to scams. Given this expectation, backers are more likely to report suspected or fraudulent campaigns. In our "Robustness Checks" section, we rerun our results using a Kickstarter sample only.

The chronological sequence of the initiation date, categories, and raised volumes in USD of detected fraudulent campaigns are shown in panel C of Table 1. We find that backers have spent approximately USD $30 million thus far on detected fraudulent campaigns, and the number of cases has increased steadily since 2010. After excluding suspended, cancelled, and failed projects (in terms of reaching goal amounts), our data show that more than $27 million was successfully raised by fraudulent campaigns. The reduction in the number of fraudulent campaigns over the past two years is presumably attributable to potentially fraudulent campaigns that have delivery dates later than the second half of 2015, which have not been identified as fraudulent yet. Fraud

[12] See https://www.kickstarter.com/help/stats and https://go.indiegogo.com/blog/2015/12/2015-crowdfunding-infographic-statistics-tech-film-social.html (both accessed on August 4, 2016).

campaigns are most common in the product design business category (53 cases), and have raised the most money within the technology business category (almost $10 million).

Fraud campaigns by country for each respective year are shown in panel D of Table 1. In our sample, fraud cases occurred most frequently in campaigns launched by creators in the U.S. (183 cases), Canada (8 cases), and the U.K. (8 cases), followed by creators in Israel (2), Spain (2), and Australia, China, Germany, and Hong Kong (1 each).

—Please insert Table 1 about here—

We hand-collect information from Kickstarter and Indiegogo on 26 explanatory campaign variables, calculated based on the information from the campaign's webpage or from social media web pages associated with the campaign/creator. We group the variables into four blocks: 1) creator(s)' characteristics/background, 2) social media affinity, 3) campaign funding and reward structure, and 4) campaign description details. See Table 2 for variable descriptions and calculation methods.[13]

—Please insert Table 2 about here—

Table 3 shows the quality of our PSM using probit estimates for the probability of a fraudulent campaign. We find that all the variables (*Goal Amount*, *Portal Dummy*, *Country*, *Year*, and *Category*) included in the PSM are perfectly balanced between fraud and non-fraud campaigns, as we find no statistical differences between them. Consequently, our results are not driven by differences in goal amount, portal, country, category, or year of campaign launch.

—Please insert Table 3 about here—

[13] For three campaigns, there was no information available on the main control variables in the campaign's funding structure on the campaign's web page (moreover, two campaigns were missing data on *Duration*, and one campaign was missing data on *No. of Pledge Categories* and *Min. Pledge Dummy*). Also, the project description text was unavailable for three campaigns, resulting in no data on text indices for them. See also Table 1 for more details.

Table 4 shows the descriptive statistics for the twenty-nine explanatory variables. Due to data availability (some variables, such as *Ln* (*No. of FB Likes)*, reduce the sample size considerably—see Table 4), and the fact that some variables cannot be considered simultaneously (e.g., a Facebook page is a precondition for *Ln* (*No. of FB Likes*)) because of severe multicollinearity (e.g., variance inflation factors (VIF) larger than five when including all possible explanatory variables), we use only twelve variables for our main analyses. However, we consider the others in subsequent robustness checks (see Table A4 in the online appendix for the correlations between all explanatory variables, and Table 5 for those used in the main analyses).

—Please insert Tables 4 and 5 about here—

5. Methodology

In the empirical analysis, we first specify a baseline regression model, and then incrementally append and operationalize our core theoretical concepts, *social media affinity*, *campaign funding and reward structure*, and *campaign description details*, with more refined variables. For the baseline regression, we apply a logistic regression model to comprehensively analyze the determinants of our dependent variable *Fraud*, which equals 1 if the campaign is in our fraud sample, and 0 otherwise. The basic structure of our baseline regression model in Table 7 is as follows:

$$Fraud\ (0/1) = \alpha + \sum_i \gamma_i \cdot Creator(s)'\ Characteristics/Background_i + \sum_j \xi_j \cdot Social\ Media\ Affinity_j + \sum_k \varphi_k \cdot Campaign\ Funding\ and\ Reward\ Structure_k + \sum_l \phi_l \cdot Campaign\ Description\ Details_l + \varepsilon. \quad (1)$$

The main explanatory variables in the *creator(s)' characteristics/background* block are *Natural Person*, *Formal Name*, *No. of Creator-Backed Projects*, and *No. of Creator-Created Projects*. The *social media affinity* block includes *Facebook_Personal*, *Facebook_Page*, and *No. of External*

Links. The *campaign funding and reward structure* block includes *Duration*, *No. of Pledge Categories*, and *Min. Pledge Amount*. Finally, the *campaign description details* block includes *ARI* and *Video Pitch*. We do not include year, country, portal, or business category fixed effects because our samples have been initially matched and are perfectly balanced on those variables (see also Bertoni et al., 2011; Grilli and Murtinu, 2014; and Lee et al., 2015 on time variation and access to finance). However, we do use robust standard errors, which are one-way-clustered by thirteen business categories in all regressions.

The main reason for not using two-way clustering (based on business category and year) is that we do not have observations for all category-year combinations. This restricts us from calculating all standard errors. Furthermore, even with observations for all category-year combinations, the six-year time span would not be sufficient to conduct two-way-clustered standard errors. That would involve estimating autocovariances between residuals, which can be downward-biased and is a widely discussed procedure in the econometrics literature (see e.g., Hurwicz, 1950). According to Thompson (2011), larger numbers of time periods are needed (at least twenty-five) for the estimates of two-way standard errors to be accurate.

In our next steps, we run several robustness checks, where we operationalize our theoretical concepts with different variables that may have an impact on fraud. Note that we do not include them in our baseline equation in order to avoid multicollinearity and data availability problems.

We first retain the variables for our theoretical concept *social media affinity*, and further operationalize it by using the *Facebook* and *LinkedIn* variables. In Table 8, we include them separately, and, in specification (6), we include them simultaneously (except for *Facebook*, which is a linear combination of *Facebook_Personal* and *Facebook_Page*):

$$Fraud\ (0/1) = \alpha + \xi_1 \cdot Facebook + \xi_2 \cdot Facebook_{Personal} + \xi_3 \cdot Facebook_{Page} + \xi_4 \cdot$$

$$No.of\ External\ Links + \xi_5 \cdot LinkedIn + \sum_i \gamma_i \cdot Creator(s)'\ Characteristics/$$

$$Background_i + \sum_k \varphi_k \cdot Campaign\ Funding\ and\ Reward\ Structure_k + \sum_l \phi_l \cdot Campaign\ Description\ Details_l + \varepsilon. \quad (2)$$

An alternative to measuring social media affinity is to consider the connections creators have on Facebook. We consider them in Table 9, and we use *Ln (No. of FB Connections)*, *Ln (No. of FB Friends)*, and *Ln (No. of FB Likes)* as proxies (specifications (4)-(6)):

$$Fraud\ (0/1) = \alpha + \xi_1 \cdot Ln(No.of\ FB\ Connections) + \xi_2 \cdot Ln(No.of\ FB\ Friends) + \xi_3 \cdot Ln(No.of\ FB\ Likes) + \sum_i \gamma_i \cdot Creator(s)'\ Characteristics/Background_i + \sum_k \varphi_k \cdot Campaign\ Funding\ and\ Reward\ Structure_k + \sum_l \phi_l \cdot Campaign\ Description\ Details_l + \sum_t \delta_t \cdot Year_t + \sum_c \lambda_c \cdot Category_c + \varepsilon_t. \quad (3)$$

We use the natural logarithm of variables to reduce the potential influence of outliers on the results.[14] We control for year and category fixed effects in Equation (3). This is because, due to data limitations, our fraud and non-fraud campaign samples that have data items available on their social media connections are not necessarily the propensity score matched pairs, and may not be balanced on year or on campaign category.

Next, we are interested in identifying the potential influence of differences in funding and reward structure, especially involving the various durations and pledge categories:

$$Fraud\ (0/1) = \alpha + \varphi_1 \cdot Min.PledgeDummy + \varphi_2 \cdot Avg.Small - 2\ PledgeDummy + \varphi_3 \cdot Avg.Small - 3\ PledgeDummy + \varphi_4 \cdot Min.PledgeAmount + \varphi_5 \cdot Avg.Small -$$

[14] Note that the number of connections can change over time (e.g., people can unfriend a campaign creator, or "unlike" a campaign page). Thus, for some campaigns, the social media page will become unavailable, and, as a result, we may not be able to obtain information on connections. This is one explanation for the lower number of observations (see Mollick, 2014) for the same complex of problems. According to Mollick (2014), project success may lead to an increase in the number of "likes" on a *Facebook* page associated with an individual campaign. But this is not necessarily the case for a founder's personal *Facebook* page, as many creators tend to keep their social media accounts separate. In our setting, we focus on personal pages that are linked to campaign pages (those with available data). However, we have already used information about campaign creators' social media presence on the start date of campaigns in Equations (1) and (2), which we expect to capture a larger percentage of the variation. For robustness, we check whether there is any explanatory power in the number of connections (ex post), which serves as a measure for the ex ante number of social media connections.

$$2\ Pledge\ Amount + \varphi_6 \cdot Avg.Small - 3\ Pledge\ Amount + \sum_i \gamma_i \cdot Creator(s)'\ Characteristics/Background_i + \sum_j \xi_j \cdot Social\ Media\ Affinity_j + \sum_l \phi_l \cdot Campaign\ Description\ Details_l + \varepsilon. \tag{4}$$

The variables we consider in Table 10 include: *Min. Pledge Dummy*, *Avg. Small-2 Pledge Dummy*, *Avg. Small-3 Pledge Dummy*, *Min. Pledge Amount*, *Avg. Small-2 Pledge Amount*, and *Avg. Small-3 Pledge Amount* for different specifications. This helps us detect any explanatory power stemming from a campaign's funding and reward structure that is not included in our main explanatory variables.

Finally, we check the robustness of our results on campaign description details by analyzing five different measures of text readability: Automated Readability Index (*ARI*), Coleman-Liau index (*CL*), Gunning Fog index, Flesch-Kincaid grade level (*FKG*), and Flesch Reading Ease score (*FRE*). The model is as follows (the *campaign description details* block only includes *Video Pitch*, because the text measures are examined separately):

$$Fraud\ (0/1) = \alpha + \phi_1 \cdot ARI + \phi_2 \cdot CL + \phi_3 \cdot Gunning\ Fog + \phi_4 \cdot FKG + \phi_5 \cdot FRE + \sum_i \gamma_i \cdot Creator(s)'\ Characteristics/Background_i + \sum_j \xi_j \cdot Social\ Media\ Affinity_j + \sum_k \varphi_k \cdot Campaign\ Funding\ and\ Reward\ Structure_k + \sum_l \phi_l \cdot Campaign\ Description\ Details_l + \varepsilon. \tag{5}$$

6. Empirical Results

6.1. Univariate Results

We begin by discussing our results in a univariate setting. In subsequent analyses, we focus on multivariate settings in order to include multiple possible determinants of fraud simultaneously. Table 6 gives the first broad results for a difference in means t-test about how fraudulent campaigns differ from non-fraud campaigns for all twenty-nine of our explanatory variables. Most importantly, Table 6 shows that, on average, fraudulent campaigns tend to have fewer numbers of creator-created projects (row 4), and are less present or active on Facebook.

For example, initiators of fraudulent campaigns use Facebook less often (rows 5-7), and they tend to have fewer connections in logarithmic terms (row 10). We find no statistically significant differences among the numbers of “friends” or “likes,” however, between fraudulent and non-fraudulent campaigns (rows 11-12). This is somewhat different from findings in the area of crowdlending, where borrowers with more social ties are consistently more likely to have their loans funded and to receive lower interest rates, but the loans *ex post* perform worse (Freedman and Ginger, 2014). This may be because fraud campaigns are using fake profiles to increase their numbers of “likes.” This phenomenon is well recognized, and portals always warn backers to thoroughly vet a creator’s social media connections.

We also find that the number of external links is negatively correlated with fraud in crowdfunding campaigns (row 8). It seems that external links may serve a kind of certification role. Thus, the more external links provided, the higher the reputational capital that can be lost in the case of a fraudulent campaign.

In terms of campaign funding and reward structure, it seems that differences in, e.g., the numbers of pledge categories and pledge amounts do not differ between fraud and non-fraud campaigns (rows 14-20). However, campaign durations tend to be significantly longer for fraudulent campaigns (row 13).

Finally, we find that the descriptions of fraudulent campaigns are easier to read. The descriptions can also be interpreted as less sophisticated, because most readability measures correspond to the number of years of formal education needed to understand the text in the first reading (rows 21-22 and 24-25). The rationale behind this finding is either that fraudsters are targeting a wider and presumably less educated crowd, or that they have no real intention of using the funds raised for the stated purpose, and therefore put less effort into the campaign descriptions.

We find no differences between fraud and non-fraud campaigns' use of video pitches (row 29). This may be because creators are well aware that video pitching can strongly impact the probability of successful fundraising, and is strongly encouraged by portals. Previous research has documented a clearly positive correlation between videos and funding success (see Mollick, 2014).

—Please insert Table 6 about here—

6.2. Multivariate Results

We now turn to our baseline model, which uses multivariate regressions to evaluate the correlations among the four blocks of explanatory variables—*creator(s)' characteristics/ background*, *social media affinity*, *campaign funding and reward structure*, and *campaign description details*—with fraud. Table 7 summarizes our results from multivariate logistic regressions for the determinants of fraud in Equation (1). We first consider the blocks separately, and specification (5) considers all simultaneously.

We find that *No. of Creator-Backed Projects* and *No. of Creator-Created Projects* are both significantly negatively correlated with fraud (rows 3-4). Thus, we cannot reject Hypothesis 1 that project creators who have engaged more heavily in prior crowdfunding activities (as either creators or backers) are less likely to carry out fraud campaigns. However, we find no statistically significant relationship between a natural person profile or formal profile name and fraudulent campaigns (rows 1-2). This is attributable to the fact that on, e.g., Kickstarter, project creators must verify their identity through an automated process, and that information appears on their profile (although not necessarily as their "profile name") regardless of whether they use a formal profile name.

As shown in Table 7, all three of our main explanatory variables in the s*ocial media affinity* block have a strongly negative relationship with fraud (rows 5-7). Therefore, having a personal Facebook page associated with a campaign decreases the probability of fraudulent activity by 50% (significant at a 1% level). Having a Facebook page associated with a campaign also has a negative effect on the probability of fraudulent activity (significant at a 10% level). And the number of external links provided on the campaign website (e.g., a link to a YouTube video associated with the campaign, a LinkedIn profile, a startup's web page, etc.) has a strongly negative effect on the probability of a campaign being fraudulent. Overall, we cannot reject Hypothesis 2 that fraudsters tend to be less present on social media.

Moreover, we find that fraudulent campaigns tend to *ex ante* choose longer durations for their funding periods. This is in accordance with Hypothesis 3, which states that a crowdfunding scam is correlated with the funding and reward structure of a campaign, and is particularly positively correlated with the duration of the funding period. A campaign creator can choose a duration of between one and ninety days as his funding period. Our results indicate that Kickstarter, e.g., strongly encourages creators to choose shorter durations by showing the differences in success rates of shorter- and longer-duration campaigns. As shown by Table 7 (row 8), a ten-day increase in a funding period increases the probability of fraud by 30% (significant at a 1% level). This is also in accordance with the signaling argument that high-quality campaigns choose shorter campaign durations to signal their quality as well as their confidence in getting successfully funded.

Our results also show that the number of pledge categories has a significantly positive relationship with fraud (row 9). This also supports Hypothesis 3, that crowdfunding fraudsters are more likely to offer a larger number of rewards at differing levels. We find no statistical significance for the *Min. Pledge Dummy* (row 10). This may be because most reward-based crowdfunding campaigns offer small amounts as minimum contributions for monetary payoffs, and the campaigns do not substantially differ on this variable.

Finally, Table 7 shows that the project descriptions of fraudulent campaigns tend to have lower automated readability indexes (ARI). In other words, they tend to be easier to read. ARI is an approximate representation of the number of formal years of education needed to understand the text on a first reading. A one-level ARI increase decreases the probability of fraud by 7% (row 11, significant at a 5% level). More specifically, a change in readability level from a high school graduate level to a college graduate level decreases the probability of fraud by 28%. Hence, we cannot reject Hypothesis 4, that fraudsters target a less educated crowd by using less sophisticated and easier to read language.

We find no statistically significant effect of *Video Pitch* on fraud. This may be because more than 92% of our 414 cases use a video pitch to describe their projects (Table 4, row 29). As previously discussed in the univariate analysis, video pitches can significantly affect funding success, and thus most campaigns *a priori* have one.

—Please insert Table 7 about here—

6.3. Robustness Checks

As discussed in the "Methodology" section, we further check the robustness of our results by operationalizing our four theoretical concepts with different, more refined, explanatory variables. First, as shown in Equation (2), we operationalize *social media affinity* by using the variables previously mentioned, and we add *Facebook* and *LinkedIn* as additional explanatory variables. Table 8 shows that our results remain robust if we measure *social media affinity* by *Facebook*, *Facebook_Personal*, *Facebook_Page*, or *No. of External Links* (rows 1-4). We were unable to detect any statistically significant effect of *LinkedIn* on fraudulent activity. This may be attributable to LinkedIn's less frequent use as a social media affiliation in the crowdfunding arena (only 3.1% of our observations had a LinkedIn page associated with the campaign creator; see Table 4, row 9).

—Please insert Table 8 about here—

In the next set of regressions, presented in Equation (3), we focus on social media connections. To avoid the problem that outliers may be driving our results, we take the natural logarithm of number of connections, which is defined as number of friends of a personal Facebook page associated with campaign creator(s), plus the total likes of a Facebook page associated with a campaign. Because we estimate Equation (3) for campaigns with data items available on social media connections, we lose the balancing of our fraud and non-fraud samples on years and campaign business categories. Therefore, we include year fixed effects and category fixed effects in the model.

As Table 9 shows, there is a negative relationship between *Ln (No. of FB Connections)* and the probability of fraud (Table 9, specification (1)). This further supports Hypothesis 2, that fraudsters are less well-connected on social media. We found no statistically significant separate impact for the number of Facebook friends or the number of Facebook likes on fraudulent activity in subsequent specifications of Table 9.

Note there are two major drawbacks to using number of connections as an explanatory variable. First, the information is hand-collected as of the time of data collection, and can therefore vary over time. Second, fraudulent campaigns may use fake profiles to increase the number of "friends" or "likes" on their pages, in order to mislead potential backers. However, we formally examine the data on number of connections to test whether they have any explanatory power in our models.

—Please insert Table 9 about here—

Table 10 shows our results for the multivariate regressions presented in Equation (4), focusing on all explanatory variables in *campaign funding and reward structure*. Although the coefficients on two of our main explanatory variables in this block remain significant (*Duration* and *No. of Pledge Categories*), we were unable to detect any statistically significant explanatory power for the six variables that aim to capture the differences among other variables related to funding structure (specifically, the minimum pledge effect (specifications (1)-(6)).

As Table 4 (row 15) shows, overall, more than 64% of our observations had a minimum pledge of less than the median (62% of non-fraud and 67% of fraudulent campaigns-Table 6, row 15). However, we examined *Avg. Small-2 Pledge Dummy*, *Avg. Small-3 Pledge Dummy*, *Avg. Small-2 Pledge Amount*, and *Avg. Small-3 Pledge Amount* in order to capture any differences between our fraud and non-fraud samples that stem from minimum pledge requirements. We were unable to detect any statistically significant differences. As previously discussed, this may be due to the fact that reward-based crowdfunding campaigns frequently set low minimum requirements for funding for backers to participate and obtain monetary payoffs.

—Please insert Table 10 about here—

Finally, Table 11 shows our results for *campaign description details*, as presented in Equation (5). Previously, we identified a significantly negative relationship between ARI and fraud, i.e., the probability of fraud increases when the project description is easier to read (and presumably less sophisticated and poorly worded). We further check the robustness of our results by using four other measures of text readability (specifications (2)-(5)).

As Table 11 shows, the Coleman-Liau index (*CL*) and the Flesch-Kincaid Grade level (*FKG*) both exhibit significantly negative correlations with fraudulent activity, which further validate our inferences. While the coefficient on the Gunning Fog index is also negative, it is not significant at a 10% level. We attribute this to the fact that Gunning Fog aims primarily to capture the fraction of complex words by syllable count, which may not be a suitable measure for complex words and subsequently for text readability. It thus may not be in line with our other proxies that focus primarily on the number of characters and average sentence length. Moreover, the positive coefficient on the Flesch Reading Ease score (*FRE*) is in accordance with our inferences. For this variable, easier to read texts will have higher scores.

—Please insert Table 11 about here—

In unreported regressions, we use several possible proxies for our main variables of interest in Table 7 to check the robustness of our results to different variable definitions and calculation methods. For example, *No. of Creator-Created Projects* has now been defined as a dummy variable that equals 1 if the project creator has previously created a crowdfunding campaign, and 0 otherwise.[15] *Formal Name* has been redefined as a dummy variable that equals 1 if the profile name of the creator is one (or more than one) natural person with a formal name(s), and 0 otherwise. Furthermore, we test the robustness of our results for the influence of outliers by winsorizing. We winsorize all non-dummy variables at the 2.5% level on both sides, and then repeat all analyses for our main Table 7 using only winsorized variables. We find that the results are qualitatively unchanged (see Table A1 in the online appendix).

Also, considering the differences between the Kickstarter and Indiegogo platforms, and the low number of identified fraud cases on Indiegogo, we excluded all Indiegogo fraud cases and repeat all analyses for our main Table 7 for a Kickstarter subsample. We again find that the results are qualitatively unchanged (see Table A2 in the online appendix).

Another possible concern relates to the robustness of the inferences when examining the text readability indices. We have matched the fraud and non-fraud samples based on countries, and, therefore (as shown in Table 3), our results are not affected by countries in which campaigns are

[15] All project creators on Kickstarter are required to provide official identification documentation. Each project is attributed to at least one natural person, and the name is publicly available on the campaign webpage. The creator's profile name can be the formal name or a fantasy name, but all the information on the person associated with the campaign (first and family name) is readily available by clicking on the profile.

initiated. However, one could argue that the project description's readability indices should be tested based on an English-speaking country subsample to further validate our results. Table A3 in the online appendix re-estimates Table 11 for an English-speaking subsample (Australia, Canada, the U.K., and the U.S.). We find that the results for the ARI and CL index remain similar. However, the statistical significance of the FKG and FRE indices is slightly higher than the conventional 10% level, although the coefficients are nearly identical.

To summarize, we find that our results remain robust to using different possible proxies and subsamples, and to winsorizing.

7. What Happens When a Campaign Ends?

All of our previous analyses only included information available at a campaign's start. However, we want to complement the picture by focusing on information available at a campaign's end, such as the amount raised or the number of backers. We posit that, for at least some project creators, the motivation for fraud was not in place when beginning the campaign. Instead, it grew over the course of the campaign, or even after it ended.

For example, some creators may get overwhelmed by the excessive demand for their product, and later become unwilling or unable to manage the logistics of a large number of requests. In other cases, projects may have been too elaborate in the first place. This may have triggered excessive demand for a fancy product, but at the same time caused a problem for the creator, who was too optimistic about his or her capabilities to actually see the product development through to the end.

To shed light on the differences between fraud and non-fraud campaigns after they end, Table 12 shows the univariate results of the difference in means tests of only *ex post* available characteristics (see Table 2, panel B, for variable descriptions). We aim to provide some insights into future research on the characteristics that lead to crowdfunding fraud at points during and after a campaign's launch, and into variables that could potentially mitigate fraud risk by facilitating detection mechanisms.

The results in Table 12 show that the fraud campaigns raised statistically significantly more money than the non-fraud campaigns. Moreover, as shown by the success ratio, they raised significantly more money than their *ex ante* funding goals. Fraud campaigns also have statistically significantly higher numbers of backers. On the one hand, this can be viewed as opposition to the "wisdom of the crowd" argument; on the other hand, we would need more rigorous analyses to determine whether the higher number of total backers for fraud campaigns helped the detection of fraud process.

Table 12 also reveals that non-fraud creators have on average been members of their associated portals for longer periods of time. Their campaign web pages are also updated less frequently, with both lower numbers of updates, and lower numbers of comments posted by the creator and by backers.

—Please insert Table 12 about here—

8. Conclusion

Our paper is the first to provide an in-depth examination of which factors influence a higher probability of crowdfunding fraud detection. We find that the fraud rate in the crowdfunding market is still very low and that fraud is not a random phenomenon, which can be explained by empirical models. Furthermore, we find that penalties for fraud largely come from government agencies such as the Federal Trade Commission or regional courts, but are hardly spurred by the crowd itself. As the penalties are also still small, this again puts the focus on the prescreening procedures and liability of the crowdfunding portals.

The evidence shows that not all scams have *ex ante* been detected by the two largest U.S. crowdfunding portals. While the lack of fraud detection might justify private or public regulation that requires portals to offer standardized prescreening procedures, such regulations could rapidly become obsolete as fraudsters adapt and learn new methods to avoid detection.

In contrast, regulators may require portals to implement some form of prescreening procedure that fulfills a more abstract catalog of quality requirements. One option is a risk model that triggers background checks of project creators without actually prohibiting questionable creators from applying for funding. Once such dynamically adapting fraud detection models are implemented, it may become safe to discuss the phenomenon of crowdfunding with old-fashioned securities lawyers *without* the need for a defibrillator!

We stress again that, just as entrepreneurs learn over time how to run successful campaigns, fraudsters also tend to adapt their behavior based, for example, on results such as those provided here. For portals and backers, it may thus become even more important to fully verify the social media contacts of campaign creators. It would be instructive to track our results over time to learn whether and how fraudsters are honing their craft.

All joking aside, we believe regulators around the world are correct in their attempts to protect less sophisticated crowd members. Until recently, most crowdfunding laws largely targeted specific branches of crowdfunding (primarily equity). Reward-based crowdfunding has not been regulated under specific laws except in a few jurisdictions, such as Germany (see Klöhn et al, 2016).

A natural question to ponder is whether we believe there is a great deal of undetected fraud that has not been picked up in our sample. For example, the drop in fraud cases from 2013 to 2014, and again from 2014 to 2015 (Table 1, panel C), may imply a certain amount of undiscovered fraud in recent years. In our empirical analyses, we explore whether there is a relationship to the determinants of fraud, or merely differences in the likelihood of detection. We expect future research to elaborate on this point, particularly if undetected cases from 2010-2015 reveal themselves, or as more data on the topic become available.

In practice, crowdfunding has largely taken place under the donation- and reward-based models, which we explore in-depth here. In the U.S., these activities are not regulated under specific crowdfunding rules that are tailored toward this new form of entrepreneurial finance, which occurs on the Internet. And, as we have seen, backers who are victimized by fraudsters often do not bring any legal action because the emotional and legal costs often outweigh the rewards. Therefore, fully effective enforcement must come from government agencies.

Furthermore, once equity crowdfunding emerges more fully in the U.S. (given the SEC's formal ruling on it in May 2016), we may observe that fraud in crowdfunding takes on a completely different twist. This is because equity crowdfunding campaigns are much more complex, involve higher investment amounts, and usually comprise an entire venture, not just one small part. As the complexity of crowdfunding grows, we expect the nature of fraud to also evolve, and become more difficult to both detect and punish. Note that, under a reward-based model, fraud generally occurs because founders do not develop or deliver promised products. However, under equity crowdfunding, founders may engage in a whole realm of unethical and illegal activities, such as running several different startups at a time, violating their fiduciary duties, or engaging in asset substitution, risk shifting, or similar tactics, which can be much harder to detect. Whether these dire predictions will ultimately emerge, however, should be investigated empirically once the new market develops.

References

Agrawal, Ajay, Catalini, Christian, and Goldfarb, Avi, 2015. Crowdfunding: Geography, Social Networks, and the Timing of Investment Decisions, *Journal of Economics and Management Strategy* 24, 253–274.

Ahlers, Gerrit K.C., Cumming, Douglas J., Guenther, Christina, and Schweizer, Denis, 2015. Signaling in Equity Crowdfunding, *Entrepreneurship Theory and Practice* 39, 955–980.

Bayus, Barry, 2013. Crowdsourcing New Product Ideas over Time: An Analysis of the Dell IdeaStorm Community, *Management Science* 59, 226–244.

Becker, Gary, 1968. Crime and Punishment: An Economic Approach, *Journal of Political Economy* 76, 493–517.

Becker, Gary, 1993. Nobel Lecture: The Economic Way of Looking at Behavior, *Journal of Political Economy* 101, 385–409.

Belleflamme, Paul, Lambert, Thomas, and Schwienbacher, Armin, 2013. Individual Crowdfunding Practices, *Venture Capital: An International Journal of Entrepreneurial Finance* 15, 313–333.

Belleflamme, Paul, Lambert, Thomas, and Schwienbacher, Armin, 2014. Crowdfunding: Tapping the Right Crowd, *Journal of Business Venturing* 29, 585–609.

Bertoni, Fabio, Colombo, Massimo G., and Grilli, Luca, 2011. Venture Capital Financing and the Growth of High-Tech Start-Ups: Disentangling Treatment from Selection Effects, *Research Policy* 40, 1028–1043.

Bradford, Steven, 2012. Crowdfunding and the Federal Securities Laws, *Columbia Business Law Review* 2012, 1–150.

Coffee, John, 2006. Gatekeepers: The Professions and Corporate Governance, Oxford University Press.

Colombo, Massimo G., Franzoni, Chiara, and Rossi-Lamastra, Cristina, 2015. Internal Social Capital and the Attraction of Early Contributions in Crowdfunding Projects, *Entrepreneurship Theory and Practice* 39, 75–100.

Colombo, Massimo G., and Piva, Evila, 2012. Firms' Genetic Characteristics and Competence-Enlarging Strategies: A Comparison between Academic and Non-Academic High-Tech Start-ups, *Research Policy* 41, 79–92.

Colombo, Massimo, and Shafi, Kourosh, 2016. Does Reward-Based Crowdfunding Help Firms Obtain Venture Capital and Angel Finance? *Available at SSRN:* http://ssrn.com/abstract=2785538.

Cumming, Douglas J., Leboeuf, Gael, and Schwienbacher, Armin, 2014. Crowdfunding Models: Keep-it-All versus All-or-Nothing, Working Paper, University of Lille and York University.

Diamond, Douglas, 1989. Reputation Acquisition in Debt Markets, *Journal of Political Economy* 97(4), 828-862.

Dyck, Alexander, Morse, Adair, and Zingales, Luigi, 2010. Who Blows the Whistle on Corporate Fraud? *Journal of Finance* 65, 2063–2253.

Dyck, Alexander, Morse, Adair, and Zingales, Luigi, 2014. How Pervasive is Corporate Fraud? Working Paper, University of Chicago.

Fang, Vivian W., Huang, Allen H., and Karpoff, Jonathan M., 2016. Short Selling and Earnings Management: A Controlled Experiment, *Journal of Finance* 71, 1251-1294.

Fischel, Daniel R., 1982. Use of Modern Finance Theory in Securities Fraud Cases Involving Actively Traded Securities, *Business Lawyer* 38, 1–20.

Freedman, Seth, and Ginger, Zhe Jin, 2014. The Signaling Value of Online Social Networks: Lessons from Peer to Peer Lending, NBER Working Paper 19820, available at: http://www.nber.org/papers/w19820.

FTC (2015). Crowdfunding Project Creator Settles FTC Charges of Deception: Defendant Spent Backers' Money on Personal Expenses, available at: https://www.ftc.gov/news-events/press-releases/2015/06/crowdfunding-project-creator-settles-ftc-charges-deception (accessed 2 May 2016).

Giannetti, Mariassunta, and Wang, Tracy Yue, 2016. Corporate Scandals and Household Stock Market Participation, Journal of Finance 71, 2591–2636.

Gibson, Rajna, Tanner, Carmen, and Wagner, Alexander, 2013. Preferences for Truthfulness: Heterogeneity Among and Within Individuals, *American Economic Review* 103, 532–548.

Gino, Francesca, Ayal, Shahar, and Ariely, Dan, 2009. Contagion and Differentiation in Unethical Behavior: The Effect of One Bad Apple on the Barrel, *Psychological Science* 20, 393–398.

Grilli, Luca, and Murtinu, Samuele, 2014. Government, Venture Capital and the Growth of European High-Tech Entrepreneurial Firms, *Research Policy* 43, 1523–1543.

Hornuf, Lars, and Schwienbacher, Armin, 2016. Should Securities Regulation Promote Equity Crowdfunding? *Small Business Economics*, forthcoming.

Hurwicz, Leonid, 1950. Least squares bias in time series, *Statistical Inference in Dynamic Economic Models* 10, 365–383.

Jiang, Ting, 2013. Cheating in Mind Games: The Subtlety of Rules Matters, *Journal of Economic Behaviour and Organization* 93, 328–336.

Karpoff, Jonathan M., Koester, Allison, Lee, D. Scott, and Martin, Gerald S., 2012. A Critical Analysis of Databases Used in Financial Misconduct Research, Mays Business School Research Paper No. 2012_73.

Karpoff, Jonathan M., Lee, D. Scott, and Martin, Gerald S., 2008a. The Consequences to Managers for Cooking the Books, *Journal of Financial Economics* 88, 193–215.

Karpoff, Jonathan M., Lee, D. Scott, and Martin, Gerald S., 2008b. The Consequences to Managers for Financial Misrepresentation, *Journal of Financial Economics* 85, 66–101.

Khanna, Vikramadity A., Kim, E. Han, and Lu, Yao, 2015. CEO Connectedness and Corporate Fraud, *Journal of Finance*, 70, 1203–1252.

Kim, Keongtae, and Hann, Il-Horn, 2015. Does Crowdfunding Democratize Access to Finance? A Geographical Analysis of Technology Projects, available at SSRN: http://ssrn.com/abstract=2334590.

Klöhn, Lars, Hornuf, Lars, and Schilling, Tobias, 2016. The Regulation of Crowdfunding in the German Small Investor Protection Act: Content, Consequences, Critique, Suggestions, *European Company Law* 13, 56–66.

Kuppuswamy, Venkat, and Bayus, Barry, 2014. Crowdfunding Creative Ideas: The Dynamics of Project Backers in Kickstarter, UNC Kenan-Flagler Research Paper No. 2013-15.

Kutner, Michael H., Nachtsheim, Christopher J., Neter, John, and Li, William, 2005. *Applied Linear Statistical Models*, 5th edition. New York: McGraw-Hill.

Lakhani, Karim R., Hutter, Katja, and Friar, Greta, 2014. Prodigy Network: Democratizing Real Estate Design and Financing, *Harvard Business School Case* 614-064.

Lee, Neil, Sameen, Hiba, and Cowling, Marc, 2015. Access to Finance for Innovative SMEs Since the Financial Crisis, *Research Policy* 44, 370–380.

Lin, Mingfeng, and Viswanathan, Siva, 2015. Home Bias in Online Investments: An Empirical Study of an Online Crowdfunding Market, *Management Science* 62, 1393–1414.

Mann, Heather, Garcia-Rada, Ximena, Hornuf, Lars, Tafurt, Juan, and Ariely, Dan, 2016a. Cut from the Same Cloth: Similarly Dishonest Individuals Across Countries, *Journal of Cross-Cultural Psychology* 47, 858-874.

Mann, Heather, Garcia-Rada, Ximena, Hornuf, Lars, and Tafurt, Juan, 2016b. What Deters Crime? Comparing the Effectiveness of Legal, Social, and Internal Sanctions Across Countries, *Frontiers in Psychology (Cognitive Science)* 7, doi: 10.3389/fpsyg.2016.00085.

Mazar, Nina, Amir, On, and Ariely, Dan, 2008. The Dishonesty of Honest People: A Theory of Self- Concept Maintenance, *Journal of Marketing Research* 45, 633–644.

Mollick, Ethan, 2014. The Dynamics of Crowdfunding: An Exploratory Study, *Journal of Business Venturing* 29, 1–16.

Nagin, Daniel, and Pogarsky, Greg, 2003. An Experimental Investigation of Deterrence: Cheating, Self-Serving Bias, and Impulsivity, *Criminology* 41, 167–194.

Perino, Michael A., 1998. Fraud and Federalism: Preempting Private State Securities Fraud Causes of Action, *Stanford Law Review* 50, 273-338.

Shalvi, Shaul, Gino, Francesca, Barkan, Rachel, and Ayal, Shahar, 2015. Self-Serving Justifications. Doing Wrong and Feeling Moral, *Current Directions in Psychological Science* 24, 125–130, doi: 10.1177/0963721414553264.

Shu, Lisa, Gino, Francesca, and Bazerman, Max, 2011a. Dishonest Deed, Clear Conscience: When Cheating Leads to Moral Disengagement and Motivated Forgetting, *Personality & Social Psychology Bulletin* 37, 330–349.

Shu, Lisa, Mazar, Nina, Gino, Francesca, and Ariely, Dan, 2011b. When to Sign on the Dotted Line? Signing First Makes Ethics Salient and Decreases Dishonest Self-Reports, Working Paper, Harvard Business School, 11-117.

Simmonds, Andrew R., Sagat, Kenneth A., and Ronen, Joshua, 1992. Dealing with Anomalies, Confusion and Contradiction in Fraud on the Market Securities Class Actions, *Kentucky Law Journal* 81, 123–186.

Spence, Michael, 1973. Job Market Signaling, *Quarterly Journal of Economics* 87, 355–374.

Thompson, Samuel B., 2011. Simple Formulas for Standard Errors that Cluster by Both Firm and Time, *Journal of Financial Economics* 99, 1–10.

Vismara, Silvio, 2016. Equity Retention and Social Network Theory in Equity Crowdfunding, *Small Business Economics* 46, 579–590.

Wang, Tracy Yue, Winton, Andrew, and Yu, Xiaoyun, 2010. Corporate Fraud and Business Conditions: Evidence from IPOs, *Journal of Finance*, 65, 2255–2292.

Wessel, Michael, Thies, Ferdinand, and Benlian, Alexander, 2015. A Lie Never Lives to be Old: The Effects of Fake Social Information on Consumer Decision-Making in Crowdfunding, ECIS 2015 Completed Research Papers (spring 5-29-2015), available at: http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1200&context=ecis2015_cr.

Table 1: Sample Selection

This table shows the derivation of the sample (panel A), campaign status (panel B), and campaign categories, as well as number of campaigns (No.) and amounts raised in USD (Vol.) for each respective year for the fraud campaigns (panel C). Panel D shows the distribution of fraud campaigns across different countries. Panel A shows in detail the number of identified fraud cases using the Kickscammed website and the news search, as well as the number of identified non-fraud campaigns using propensity score matching (PSM). We dropped nine campaigns from the initial fraud sample because 1) we found no PSM campaign, or 2) campaign information was missing (e.g., the campaign web page was no longer available). Variables in the four categories include only those used in the main multivariate regression table (Table 7, and see also Table 2 for variable descriptions and calculation methods). In panel B, "unfunded" is defined as the goal amount not being met by the end date of the campaign; successful is defined as the goal amount being achieved (and neither suspended nor cancelled). In panel C, the amounts raised in currencies other than USD are converted into USD using the respective average exchange rate in the year the campaign was initiated.

Panel A

Detection	##	#	Kickstarter	Indiegogo	#	PSM Kickstarter	PSM Indiegogo
Kickscammed	-	196	185	11	189	179	10
News Search	-	20	19	1	18	18	0
Total (initial cases)	-	216	204	12	-	-	-
Unmatched	-	9	7	2	-	-	-
Total	**414**	**207**	197	10	**207**	197	10
Main Explanatory Variables							
Creator(s)' Characteristics/Background	414	207	197	10	207	197	10
Subtotal	414	207	197	10	207	197	10
Social Media Affinity	414	207	197	10	207	197	10
Subtotal	414	207	197	10	207	197	10
Campaign Funding Structure	411	206	197	9	205	197	8
Subtotal	411	206	197	9	205	197	8
Campaign Description Details	411	206	196	10	205	195	10
Total	**408**	**205**	**196**	**9**	**203**	**195**	**8**

Panel B

Fraud Category		Status	#	Kickstarter	Indiegogo
Detected Fraud	Pre-Empted	Suspended	14	14	0
		Cancelled by Creator	5	5	0
	Attempted	Fake Project	21	18	3
		Reselling Existing Product	7	7	0
Suspected Fraud		Rewards Changed	8	6	2
		Rewards Not Delivered	152	147	5
		Total	**207**	**197**	**10**

Table 1: Sample Selection—*continued*

Panel C

Business Category	2010	Vol.	2011	Vol.	2012	Vol.	2013	Vol.	2014	Vol.	2015	Vol.	Total:	Total:
3D Printing	0	-	0	-	0	-	3	2,032,371	1	456,953	1	4,681	5	2,494,005
Camera & Photography	0	-	0	-	0	-	1	8,046	2	153,907	1	58,734	4	220,687
Design & Art	0	-	1	14,984	1	127,900	6	1,464,819	2	321,190	0	-	10	1,928,893
Education	1	28,701	0	-	0	-	1	248,116	1	380,747	0	-	3	657,564
Fashion	0	-	0	-	1	94,279	2	25,648	2	114,317	4	711,655	9	945,899
Film/Video/Music	0	-	2	95,994	4	331,594	5	164,875	5	321,576	0	-	16	914,039
Food	0	-	0	-	0	-	4	208,084	1	13,355	0	-	5	221,439
Gadgets & Accessories	0	-	0	-	0	-	0	-	3	1,073,552	4	476,347	7	1,549,899
Hardware	0	-	0	-	0	-	10	1,241,425	7	437,181	2	318,396	19	1,997,002
Product Design	1	87,407	3	616,310	18	2,023,809	21	3,716,613	7	515,060	3	47,311	53	7,006,510
Tabletop & Card Games	0	-	3	73,991	8	661,688	8	162,579	8	287,613	1	13,796	28	1,199,667
Technology	0	-	0	-	5	571,826	5	3,331,197	10	5,613,012	5	742,395	25	10,258,430
Video Games & Comics	0	-	3	192,829	7	93,696	9	229,452	4	311,784	0	-	23	827,761
Total	**2**	**116,108**	**12**	**994,108**	**44**	**3,904792**	**75**	**12,833,225**	**53**	**10,000,247**	**21**	**2,373,315**	**207**	**30,221,795**

Total Amount Raised	**30,221,795**
- Cancelled by Creator	430,312
- Suspended	2,348,866
- Failed	69,294
Total Amount Transferred to Creators	**27,373,323**

Table 1: Sample Selection—*continued*

Panel D

Country	2010	2011	2012	2013	2014	2015	Total
Australia	0	0	0	0	0	1	1
Canada	0	1	1	3	1	2	8
China	0	0	0	0	1	0	1
Germany	0	0	0	0	0	1	1
Hong Kong	0	0	0	0	0	1	1
Israel	0	0	0	1	1	0	2
Spain	0	0	0	0	2	0	2
United Kingdom	0	0	0	4	2	2	8
United States	2	11	43	67	46	14	183
Total	**2**	**12**	**44**	**75**	**53**	**21**	**207**

Table 2: Variable Definitions

This table gives a detailed description of the data-gathering process and calculation methods for all variables (panel A includes only *ex ante* known variables; panel B includes variables only at campaign end (*ex post*).

Variable Name	Description and Calculation
	Dependent Variable
Fraud	Dummy variable indicating whether a campaign is associated with fraudulent activities that equals 1 if a fraudulent activity is detected for a campaign, and 0 otherwise.

Panel A (*ex ante* variables):

	Creator(s)' Characteristics/Background
Natural Person	Dummy variable that equals 1 if the project creator is one/more than one natural person(s) as shown by the profile, and 0 otherwise.
Formal Name	Dummy variable that equals 1 if the project creator is one natural person and uses a formal profile name (i.e., [first name] [last name]), and 0 otherwise.
No. of Creator-Backed Projects	Total number of projects backed by the creator since joining the portal.
No. of Creator-Created Projects	Total number of projects created by the creator since joining the portal.
	Social Media Affinity
Facebook	Dummy variable that equals 1 if either a link to a personal Facebook page associated with the campaign creator(s) or a Facebook page associated with the campaign is provided, and 0 otherwise.
Facebook_Personal	Dummy variable that equals 1 if a link to a personal Facebook page associated with the campaign creator(s) is provided, and 0 otherwise.
Facebook_Page	Dummy variable that equals 1 if a link to a Facebook page associated with the campaign is provided, and 0 otherwise.
No. of External Links	Total number of external links provided on campaign's page.
LinkedIn	Dummy variable that equals 1 if a link to a LinkedIn page of the creator(s) is provided, and 0 otherwise.
Ln (No. of FB Connections)	Natural logarithm of the total friends of personal Facebook page associated with the campaign creator(s), plus the total likes of Facebook page associated with the campaign.
Ln (No. of FB Friends)	Natural logarithm of the total friends of personal Facebook page associated with the campaign creator(s).
Ln (No. of FB Likes)	Natural logarithm of the total likes of Facebook page associated with the campaign.
	Campaign Funding and Reward Structure
Goal Amount	Funding goal (in USD thousands) set by the creator of the project before the start date of the campaign. For campaigns with amounts in other currencies, the USD equivalent is calculated based on the annual average exchange rate (corresponding to the year the campaign was launched). This variable is used for propensity score matching.

Duration	Number of days between the campaign's end date and start date.
No. of Pledge Categories	Total number of pledge categories. Each individual backer can pledge an amount associated with one of the categories and receive a specific reward/benefit.
Min. Pledge Dummy	Dummy variable that equals 1 if *Min. Pledge Amount* is smaller than or equal to 5 USD, and 0 otherwise. The value of 5 USD corresponds to the median of the reference variable, *Min. Pledge Amount*.
Avg. Small-2 Pledge Dummy	Dummy variable that equals 1 if *Avg. Small-2 Pledge Amount* is smaller than or equal to 11.73 USD, and 0 otherwise. The value of 11.73 USD corresponds to the median of the reference variable, *Avg. Small-2 Pledge Amount.*
Avg. Small-3 Pledge Dummy	Dummy variable that equals 1 if *Avg. Small-3 Pledge Amount* is smaller than or equal to 18.77 USD, and 0 otherwise. The value of 18.77 USD corresponds to the median of the reference variable, *Avg. Small-3 Pledge Amount*.
Min. Pledge Amount	Minimum amount (in USD) that any backer needs to pledge to be allowed to participate and receive a certain reward/benefit (associated with the minimum pledge categories).
Avg. Small-2 Pledge Amount	Average of the first two minimum pledge amounts (in USD).
Avg. Small-3 Pledge Amount	Average of the first three minimum pledge amounts (in USD).
	Campaign Description Details
ARI	The Automated Readability Index of the project description text. ARI equals $4.71\left(\frac{Number\ of\ Characters}{Number\ of\ words}\right) + 0.5\times ASL - 21.43$, where ASL is average sentence length (i.e., the number of words divided by the number of sentences). ARI corresponds to a U.S. grade level; the lower the number, the easier the text is to read.
CL	The Coleman-Liau index of the project description text. CL equals $5.88\left(\frac{Number\ of\ Characters}{Number\ of\ words}\right) - 29.6\times ASL$, where ASL is average sentence length (i.e., the number of words divided by the number of sentences). CL corresponds to a U.S. grade level; the lower the number, the easier the text is to read.
Gunning Fog	Gunning Fog index of the project description text. The index equals $0.4\left[ASL + 100\left(\frac{Number\ of\ complex\ words}{Total\ Number\ of\ words}\right)\right]$, where ASL is average sentence length (i.e., the number of words divided by the number of sentences), and $complex\ words$ are words with three or more syllables. The index estimates the years of formal education needed to understand the text on a first reading; the lower the number, the easier the text is to read.
FKG	Flesch-Kincaid grade level of the project description text. FKG equals $0.39\times ASL + 11.8 * ASW - 15.59$, where ASL is average sentence length (i.e., the number of words divided by the number of sentences), and ASW is average number of syllables per word. FKG corresponds to a U.S. grade level; the lower the number, the easier the text is to read.
FRE	Flesch reading ease score of the project description text. FRE equals $206.835 - (1.015\times ASL) - (86.4\times ASW)$, where ASL is average sentence length (i.e., the number of words divided by the number of sentences), and

	ASW is average number of syllables per word. FRE ranges from 0-100; the higher the number, the easier the text is to read.
Video Pitch	Dummy variable that equals 1 if a video pitch is provided on the campaign's page to describe the project, and 0 otherwise.

Panel B (*ex post* variables):

Campaign Success	
Ln (Raised Amount)	Natural logarithm of the total amount raised by the campaign.
Success Ratio	Total amount raised divided by the goal amount of the project creator.
No. of Total Backers	Total number of backers of the crowdfunding project.
Ln (No. of Total Backers)	Natural logarithm of the total number of backers of the crowdfunding project.
% Min. Pledge Backers	Percentage of backers pledging the minimum amount.
% Amount Raised Minimum Backers	Percentage of amount raised by minimum pledge backers.
Creator/Backer Activity	
Creator Portal Familiarity	Number of months between the day creator joined the platform and start date of the campaign.
No. of Updates	Number of total updates posted by creator on the campaign webpage.
No. of Creator Comments	Number of total comments posted by creator on the campaign webpage.
Frequency of Updates	Number of updates divided by number of days between the post date of the first and last updates.
Last Round Absence	Number of days between the date of the last update and the date of the previous update.
Avg. Comments per Update	Number of total comments posted by backers divided by total number of updates.
Avg. Comments per Backer	Number of total comments posted by backers divided by total number of backers.

Table 3: Propensity Score Matching Quality

This table presents probit estimates for the probability of a fraud campaign and the propensity score matched (PSM) non-fraud subsample, where one close neighbor is matched to each fraud campaign. Potential matches are detected from the respective crowdfunding portal based on country, "year" (the start dates of matched non-fraud campaigns are within six months of the start date of the fraud campaign), and category/subcategory criteria. To conduct a one-to-one matching, we calculate propensity scores of the fraud campaign and its potential matches based on goal amount. The dependent variable equals 1 if the campaign was fraudulent, and 0 otherwise. *Country*, *Year*, category dummy variables, and *Portal Dummy* are also included (the country reference group is the U.S., the year reference group is 2015, and the category reference group is technology. *Portal Dummy* equals 1 if the crowdfunding platform is Kickstarter, and 0 if it is Indiegogo). Unreported category dummy coefficients are not all statistically significant in specification (4). When a fraud campaign has multiple matches with identical propensity scores, we use a random uniform function to choose one match among all non-fraud campaigns with identical propensity scores (to avoid sorting of the data, which could potentially affect our results). *t*-statistics using robust standard errors are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
Goal Amount	0.001	0.001	0.001	0.001
	(0.81)	(0.83)	(0.80)	(0.85)
Portal Dummy		0.015	0.014	-0.007
		(0.05)	(0.05)	(-0.02)
Australia		0.031	0.038	0.010
		(0.04)	(0.04)	(0.01)
Canada		-0.000	0.003	0.006
		(-0.00)	(0.01)	(0.02)
China		0.029	0.020	0.027
		(0.03)	(0.02)	(0.03)
Germany		0.030	0.036	0.037
		(0.03)	(0.04)	(0.04)
Hong Kong		0.001	0.008	0.012
		(0.00)	(0.01)	(0.01)
Israel		0.028	0.019	0.026
		(0.04)	(0.03)	(0.04)
Spain		0.006	-0.036	-0.043
		(0.01)	(-0.05)	(-0.07)
United Kingdom		-0.001	0.003	-0.013
		(-0.00)	(0.01)	(-0.04)
Year 2010			-0.214	-0.160
			(-0.35)	(-0.24)
Year 2011			0.020	0.004
			(0.06)	(0.01)
Year 2012			0.015	0.004
			(0.06)	(0.01)
Year 2013			-0.017	-0.027
			(-0.07)	(-0.11)
Year 2014			0.048	0.045
			(0.19)	(0.17)
Category	No	No	No	Yes
Constant	-0.039	-0.054	-0.059	-0.091
	(-0.50)	(-0.18)	(-0.16)	(-0.16)
Observations	414	414	414	414
Pseudo R^2	0.001	0.001	0.002	0.002

Table 4: Summary Statistics

This table gives descriptive statistics (mean, standard deviation, min, and max) for the full sample (207 fraud and 207 non-fraud campaigns) shown in Table 1, if the data items were available (see Table 1 for data availability). All variables are considered in subsequent analyses (see Table 2 for variable descriptions and calculation methods).

Variable	# Obs.	Mean	Std. Dev.	Min	Max
		Creator(s)' Characteristics/Background			
(1) Natural Person	414	0.504	0.500	0	1
(2) Formal Name	414	0.415	0.493	0	1
(3) No. of Creator-Backed Projects	414	16.190	53.974	0	890
(4) No. of Creator-Created Projects	414	1.079	2.712	0	21
		Social Media Affinity			
(5) Facebook	414	0.594	0.491	0	1
(6) Facebook_Personal	414	0.463	0.499	0	1
(7) Facebook_Page	414	0.272	0.446	0	1
(8) No. of External Links	414	1.886	1.424	0	8
(9) LinkedIn	414	0.031	0.174	0	1
(10) Ln (No. of FB Connections)	232	6.801	1.606	1.098	12.572
(11) Ln (No. of FB Friends)	188	6.122	1.326	1.098	8.507
(12) Ln (No. of FB Likes)	95	7.434	1.710	3.218	12.569
		Campaign Funding and Reward Structure			
(13) Duration	412	35.279	11.263	3	90
(14) No. of Pledge Categories	413	12.859	7.596	2	53
(15) Min. Pledge Dummy	413	0.648	0.477	0	1
(16) Avg. Small-2 Pledge Dummy	413	0.535	0.499	0	1
(17) Avg. Small-3 Pledge Dummy	413	0.501	0.500	0	1
(18) Min. Pledge Amount	413	12.787	31.389	0.783	349
(19) Avg. Small-2 Pledge Amount	413	23.252	40.621	1.5	399
(20) Avg. Small-3 Pledge Amount	413	35.853	54.176	2.333	599
		Campaign Description Details			
(21) ARI	411	11.472	2.390	4.8	23.1
(22) CL	411	12.377	2.111	5.9	24.94
(23) Gunning Fog	411	8.672	1.371	5.6	18.8
(24) FKG	411	9.273	1.882	4.4	18.5
(25) FRE	411	58.312	9.871	25.8	84.68
(26) Video Pitch	414	0.927	0.259	0	1

Table 5: Correlation Matrix

This table shows the Pearson correlation coefficients for the main variables (control 1, control 2, control 3, and control 4) if data items are available (see Table 1 for data availability). All variables are considered in subsequent analyses (see Table 2 for variable descriptions and calculation methods). * indicates statistical significance at least at a 5% level. See Table A4 in the online appendix for correlation coefficients for all variables.

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1)	Natural Person	1											
(2)	Formal Name	0.81*	1										
(3)	No. of Creator-Backed Projects	0.10*	0.10*	1									
(4)	No. of Creator-Created Projects	0.03	0.03	0.20*	1								
(5)	Facebook_Personal	0.29*	0.21*	0.02	0.06	1							
(6)	Facebook_Page	-0.09*	-0.09*	-0.04	-0.02	0.07	1						
(7)	External Links	-0.02	-0.02	0.07	0.04	0.07	0.43*	1					
(8)	Duration	0.02	-0.04	0.09	-0.11*	0.02	0.11*	0.03	1				
(9)	No. of Pledge Categories	0.03	0.07	0.03	-0.04	0.05	0.08	0.06	0.00	1			
(10)	Min. Pledge Dummy	0.09	0.07	0.03	-0.09*	0.06	0.00	0.04	-0.09	0.21*	1		
(11)	ARI	-0.09	-0.11*	-0.01	-0.03	-0.02	0.04	0.12*	0.07	-0.05	-0.03	1	
(12)	Video Pitch	-0.07	-0.05	0.05	0.00	0.09	0.10*	0.08	0.05	0.02	-0.01	0.07	1

Table 6: Mean Differences Between Fraud and Non-Fraud Campaigns

This table gives the comparison of means test for fraud (207) and non-fraud campaigns (207), if data items are available (see Table 1 for data availability). All variables are considered in subsequent analyses (see Table 2 for variable descriptions and calculation methods). ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Non-Fraud		**Fraud**		
Variable	***# Obs.***	***Mean***	***# Obs.***	***Mean***	***Difference Test***
		Creator(s)' Characteristics/Background			
(1) Natural Person	207	0.52	207	0.49	0.02
(2) Formal Name	207	0.43	207	0.40	0.04
(3) No. of Creator-Backed Projects	207	21.30	207	11.08	10.21
(4) No. of Creator-Created Projects	207	1.55	207	0.61	0.94***
		Social Media Affinity			
(5) Facebook	207	0.67	207	0.52	0.15***
(6) Facebook_Personal	207	0.53	207	0.40	0.12**
(7) Facebook_Page	207	0.34	207	0.20	0.14***
(8) # External Links	207	2.26	207	1.52	0.73***
(9) LinkedIn	207	0.03	207	0.03	-0.004
(10) Ln (No. of FB Connections)	134	6.94	98	6.61	0.32*
(11) Ln (No. of FB Friends)	106	6.20	82	6.02	0.18
(12) Ln (No. of FB Likes)	64	7.33	31	7.64	-0.31
		Campaign Funding and Reward Structure			
(13) Duration	207	34.048	205	36.521	-2.47**
(14) No. of Pledge Categories	206	12.053	207	13.661	-1.60**
(15) Min. Pledge Dummy	206	0.621	207	0.676	-0.054
(16) Avg. Small-2 Pledge Dummy	206	0.519	207	0.550	-0.031
(17) Avg. Small-3 Pledge Dummy	206	0.495	207	0.507	-0.012
(18) Min. Pledge Amount	206	12.069	207	13.503	-1.43
(19) Avg. Small-2 Pledge Amount	206	22.075	207	24.423	-2.34
(20) Avg. Small-3 Pledge Amount	206	35.265	207	36.439	-1.17
		Campaign Description Details			
(21) ARI	206	11.687	205	11.255	0.43**
(22) CL	206	12.54	205	12.206	0.34*
(23) Gunning Fog	206	8.741	205	8.602	0.13
(24) FKG	206	9.400	205	9.145	0.25*
(25) FRE	206	57.696	205	58.931	-1.23*
(26) VideoPitch	207	0.937	207	0.917	0.019

Table 7: Multivariate Analysis of Fraud Determinants

In this table, we apply logistic regressions to analyze the determinants of fraud, where the dependent variable equals 1 if the campaign is fraudulent, and 0 otherwise. The sample includes all fraud cases for which a one-to-one propensity score non-fraud match campaign can be found. The total sample includes 414 campaigns (207 fraud + 207 non-fraud) (see Table 1). Investigating the variance inflation factors (VIFs) reveals no evidence of multicollinearity, given the mean VIF range from 1.01 to 1.45, and that all individual values are well below the critical value of 5 in specification (5) including all variables (see Kutner et al., 2005). Robust standard errors are one-way-clustered by campaign category. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
			Creator(s)' Characteristics/Background		
(1) Natural Person	0.099				0.172
	(0.29)				(0.49)
(2) Formal Name	-0.191				-0.278
	(-0.67)				(-0.91)
(3) No. of Creator-Backed Projects	-0.003				-0.004*
	(-1.22)				(-1.80)
(4) No. of Creator-Created Projects	-0.163***				-0.143***
	(-4.13)				(-3.41)
			Social Media Affinity		
(5) Facebook_Personal		-0.447**			-0.507***
		(-2.00)			(-2.91)
(6) Facebook_Page		-0.223			-0.501*
		(-1.24)			(-1.94)
(7) No. of External Links		-0.359***			-0.340***
		(-6.27)			(-5.15)
			Campaign Funding and Reward Structure		
(8) Duration			0.022***		0.029***
			(2.84)		(3.47)
(9) No. of Pledge Categories			0.027		0.040**
			(1.58)		(2.34)
(10) Min. Pledge Dummy			0.209		0.185
			(0.98)		(0.86)
			Campaign Description Details		
(11) ARI				-0.076**	-0.074**
				(-2.57)	(-2.52)
(12) Video Pitch				-0.178	0.025
				(-0.54)	(0.13)
(13) Constant	0.229*	0.934***	-1.270***	1.026**	0.404
	(1.74)	(3.66)	(-3.38)	(2.22)	(0.83)
Mean VIF	2.02	1.16	1.04	1.01	1.45
Maximum VIF	2.99	1.24	1.06	1.01	3.14
Observations	414	414	411	411	408
Pseudo R^2	0.029	0.061	0.020	0.006	0.121

Table 8: Focusing on Social Media Affinity

In this table, we apply logistic regressions to analyze the determinants of fraud, where the dependent variable equals 1 if the campaign is fraudulent, and 0 otherwise. The sample is the same as in Table 1. Control 1 (*creator(s)' characteristics/background*) includes *Natural Person*, *Formal Name*, *No. of Creator-Backed Projects*, and *No. of Creator-Created Projects*; control 2 (*social media affinity*) includes *Facebook_Personal*, *Facebook_Page*, and *No. of External Links*; control 3 (*campaign funding and reward structure*) includes *Duration*, *No. of Pledge Categories*, and *Min. Pledge Dummy*; control 4 (*campaign description details*) includes *ARI* and *Video Pitch*. Investigating the variance inflation factors (VIFs) reveals no evidence of multicollinearity, given the mean VIF range from 1.44 to 1.47, and that all individual values are well below the critical value of 5 in specification (6), including all variables (see Kutner et al., 2005). Robust standard errors are one-way-clustered by campaign category. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Facebook	-0.791***					
	(-3.45)					
Facebook_Personal		-0.562***				-0.496***
		(-3.14)				(-2.84)
Facebook_Page			-0.972***			-0.474*
			(-4.05)			(-1.68)
No. of External Links				-0.415***		-0.373***
				(-5.26)		(-4.95)
LinkedIn					0.028	0.823
					(0.08)	(1.59)
Control 1	Yes	Yes	Yes	Yes	Yes	Yes
Control 2	No	No	No	No	No	No
Control 3	Yes	Yes	Yes	Yes	Yes	Yes
Control 4	Yes	Yes	Yes	Yes	Yes	Yes
Mean VIF	1.46	1.47	1.45	1.45	1.45	1.44
Maximum VIF	3.03	3.14	3.02	3.02	3.05	3.14
Observations	408	408	408	408	408	408
Pseudo R^2	0.079	0.068	0.085	0.106	0.056	0.124

Table 9: Focusing on Social Media Affinity (Connections)

In this table, we apply logistic regressions to analyze the determinants of fraud, where the dependent variable equals 1 if the campaign is fraudulent, and 0 otherwise. The sample is the same as in Table 1. Control 1 (*creator(s)' characteristics/background*) includes *Natural Person*, *Formal Name*, *No. of Creator-Backed Projects*, and *No. of Creator-Created Projects*; control 2 (*social media affinity*) includes *Facebook_Personal*, *Facebook_Page*, and *No. of External Links*; control 3 (*campaign funding and reward structure*) includes *Duration*, *No. of Pledge Categories*, and *Min. Pledge Dummy*; control 4 (*campaign description details*) includes *ARI* and *Video Pitch*. Investigating the variance inflation factors (VIF) reveals no evidence of multicollinearity, given the mean VIF range from 1.38 to 1.45, and that individual values are well below the critical value of 5 among all specifications (see Kutner et al., 2005). Year fixed effects and category fixed effects are considered in all specifications due to data limitations, and the unbalanced fraud and non-fraud subsamples are based on years and campaign categories. Robust standard errors are one-way-clustered by campaign category. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)
Ln (No. of FB Connections)	-0.144*		
	(-1.89)		
Ln (No. of FB Friends)		-0.149	
		(-1.12)	
Ln (No. of FB Likes)			0.068
			(0.40)
Control 1	Yes	Yes	Yes
Control 2	No	No	No
Control 3	Yes	Yes	Yes
Control 4	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Category FE	Yes	Yes	Yes
Mean VIF	1.38	1.38	1.45
Maximum VIF	2.52	2.51	2.35
Observations	224	182	88
Pseudo R^2	0.142	0.173	0.151

Table 10: Focusing on Campaign Funding and Reward Structure

In this table, we apply logistic regressions to analyze the determinants of fraud, where the dependent variable equals 1 if the campaign is fraudulent, and 0 otherwise. The sample is the same as in Table 1. Control 1 (*creator(s)' characteristics/background*) includes *Natural Person*, *Formal Name*, *No. of Creator-Backed Projects*, and *No. of Creator-Created Projects*; control 2 (*social media affinity*) includes *Facebook_Personal*, *Facebook_Page*, and *No. of External Links*; control 3 (*campaign funding and reward structure*) includes *Duration*, *No. of Pledge Categories*, and *Min. Pledge Dummy*; control 4 (*campaign's description details*) includes *ARI* and *Video Pitch*. Investigating the variance inflation factors (VIF) reveals no evidence of multicollinearity, given the mean VIF range from 1.44 to 1.46, and that individual values are well below the critical value of 5 among all specifications (see Kutner et al., 2005). Robust standard errors are one-way-clustered by campaign category. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Min. Pledge Dummy	0.185					
	(0.86)					
Avg. Small-2 Pledge Dummy		0.023				
		(0.10)				
Avg. Small-3 Pledge Dummy			0.018			
			(0.11)			
Min. Pledge Amount				0.003		
				(0.59)		
Avg. Small-2 Pledge Amount					0.003	
					(0.56)	
Avg. Small-3 Pledge Amount						0.001
						(0.02)
Control 1	Yes	Yes	Yes	Yes	Yes	Yes
Control 2	Yes	Yes	Yes	Yes	Yes	Yes
Control 3	Yes	Yes	Yes	Yes	Yes	Yes
Control 4	Yes	Yes	Yes	Yes	Yes	Yes
Mean VIF	1.44	1.46	1.46	1.44	1.45	1.45
Maximum VIF	3.14	3.18	3.16	3.13	3.14	3.15
Observations	408	408	408	408	408	408
Pseudo R^2	0.121	0.120	0.120	0.121	0.122	0.121

Table 11: Focusing on Campaign Description Details

In this table, we apply logistic regressions to analyze the determinants of fraud, where the dependent variable equals 1 if the campaign is fraudulent, and 0 otherwise. The sample is the same as in Table 1. Control 1 (*creator(s)' characteristics/background*) includes *Natural Person*, *Formal Name*, *No. of Creator-Backed Projects*, and *No. of Creator-Created Projects*; control 2 (*social media affinity*) includes *Facebook_Personal*, *Facebook_Page*, and *No. of External Links*; control 3 (*campaign funding and reward structure*) includes *Duration*, *No. of Pledge Categories*, and *Min. Pledge Dummy*; control 4 (*campaign's description details*) includes *ARI* and *Video Pitch*. Investigating the variance inflation factors (VIF) reveals no evidence of multicollinearity, given that the mean VIF for all specifications is 1.45, and that individual values are well below the critical value of 5 among all specifications (see Kutner et al., 2005). Robust standard errors are one-way-clustered by campaign category. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
ARI	-0.074**				
	(-2.52)				
CL		-0.085*			
		(-1.96)			
Gunning Fog			-0.059		
			(-1.07)		
FKG				-0.076*	
				(-1.93)	
FRE					0.016*
					(1.66)
Control 1	Yes	Yes	Yes	Yes	Yes
Control 2	Yes	Yes	Yes	Yes	Yes
Control 3	Yes	Yes	Yes	Yes	Yes
Control 4	Yes	Yes	Yes	Yes	Yes
Mean VIF	1.45	1.45	1.45	1.45	1.45
Maximum VIF	3.14	3.16	3.17	3.14	3.15
Observations	408	408	408	408	408
Pseudo R^2	0.121	0.121	0.118	0.120	0.120

Table 12: Mean Differences Between Fraud and Non-Fraud Campaigns (*ex post* variables)

This table shows the comparison of the mean tests for fraud (207) and non-fraud campaigns (207), if data items are available (see Table 2 appendix for variable descriptions and calculation methods). ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Non-Fraud		**Fraud**		
Variable	***# Obs.***	***Mean***	***# Obs.***	***Mean***	***Difference Test***
			Campaign Success		
(1) Ln (Raised Amount)	207	10.57	207	10.85	-0.27**
(2) Success Ratio	207	4.81	207	6.69	-1.88**
(3) No. of Total Backers	207	983.08	207	1647.44	-664.35***
(4) Ln (No. of Total Backers)	207	6.08	207	6.55	-0.47***
(5) % Min. Pledge Backers	206	11.35	207	10.49	0.86
(6) % Amount Raised Minimum Backers	206	3.92	207	4.28	-0.36
			Creator/Backer Activity		
(7) Creator Portal Familiarity	196	12.29	191	8.37	3.91***
(8) No. of Updates	202	27.5	207	28.44	-0.94
(9) No. of Creator Comments	206	141.19	201	175	-33.80
(10) Frequency of Updates	197	0.07	191	0.12	-0.05*
(11) Last Round Absence	196	87.75	189	79.35	8.39
(12) Avg. Comments per Update	202	10.80	203	42.10	-31.30***
(13) Avg. Comments per Backer	202	0.36	207	1.03	-0.67***

Online Appendix

Table A1: Multivariate Analysis of Fraud Determinants (winsorizing)

In this table, we apply logistic regressions to analyze the determinants of fraud, where the dependent variable equals 1 if the campaign is fraudulent, and 0 otherwise. The sample includes all fraud cases for which a one-to-one propensity score non-fraud match campaign can be found. The total sample includes 414 campaigns (207 fraud + 207 non-fraud) (see Table 1). Investigating the variance inflation factors (VIF) reveals no evidence of multicollinearity. Robust standard errors are one-way-clustered by campaign category, and all non-dummy variables are winsorized at the 2.5% level on both sides. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
			Creator(s)' Characteristics/Background		
(1) Natural Person	0.101				0.151
	(0.30)				(0.44)
(2) Formal Name	-0.184				-0.245
	(-0.65)				(-0.80)
(3) No. of Creator Backed Projects	-0.006				-0.012**
	(-1.08)				(-2.25)
(4) No. of Creator Created Projects	-0.178***				-0.135**
	(-3.76)				(-2.41)
			Social Media Affinity		
(5) Facebook_Personal		-0.443**			-0.494***
		(-1.98)			(-2.62)
(6) Facebook_Page		-0.200			-0.497*
		(-1.09)			(-1.92)
(7) No. of External Links		-0.385***			-0.365***
		(-6.34)			(-5.30)
			Campaign Funding and Reward Structure		
(8) Duration			0.025***		0.034***
			(3.07)		(3.75)
(9) No. of Pledge Categories			0.026		0.044**
			(1.42)		(2.41)
(10) Min. Pledge Dummy			0.216		0.193
			(1.02)		(0.92)
			Campaign Description Details		
(11) ARI				-0.091***	-0.096***
				(-3.10)	(-2.83)
(12) Video Pitch				-0.173	0.071
				(-0.53)	(0.33)
(13) Constant	0.255*	0.966***	-1.344***	1.198**	0.455
	(1.86)	(3.75)	(-3.51)	(2.26)	(0.77)
Observations	414	414	411	411	408
Pseudo R^2	0.025	0.062	0.020	0.008	0.122

Table A2: Multivariate Analysis of Fraud Determinants (Kickstarter sample)

In this table, we apply logistic regressions to analyze the determinants of fraud, where the dependent variable equals 1 if the campaign is fraudulent, and 0 otherwise. The sample includes all Kickstarter fraud cases for which a one-to-one propensity score non-fraud match campaign can be found. The total sample includes 394 campaigns (197 fraud + 197 non-fraud) (see Table 1). Robust sstandard errors are one-way-clustered by campaign category. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
	Creator(s)' Characteristics/Background				
(1) Natural Person	0.070				0.151
	(0.21)				(0.45)
(2) Formal Name	-0.106				-0.204
	(-0.38)				(-0.73)
(3) No. of Creator-Backed Projects	-0.003				-0.004*
	(-1.15)				(-1.69)
(4) No. of Creator-Created Projects	-0.185***				-0.150***
	(-4.58)				(-3.63)
	Social Media Affinity				
(5) Facebook_Personal		-0.485**			-0.566***
		(-2.03)			(-2.87)
(6) Facebook_Page		-0.183			-0.382*
		(-0.99)			(-1.72)
(7) No. of External Links		-0.338***			-0.312***
		(-6.08)			(-4.83)
	Campaign Funding and Reward Structure				
(8) Duration			0.025**		0.030***
			(2.32)		(2.84)
(9) No. of Pledge Categories			0.025		0.036**
			(1.42)		(2.16)
(10) Min. Pledge Dummy			0.204		0.179
			(0.88)		(0.76)
	Campaign Description Details				
(11) ARI				-0.064*	-0.065**
				(-1.96)	(-2.06)
(12) Video Pitch				-0.036	0.124
				(-0.10)	(0.59)
(13) Constant	0.233	0.884***	-1.294***	0.761	0.157
	(1.46)	(3.81)	(-3.01)	(1.53)	(0.30)
Observations	394	394	394	391	391
Pseudo R^2	0.034	0.056	0.019	0.004	0.113

Table A3: Focusing on Campaign Description Details (English-speaking sample)

In this table, we apply logistic regressions to analyze the determinants of fraud, where the dependent variable equals 1 if the campaign is fraudulent, and 0 otherwise. The sample includes all Australian, Canadian, U.K., and U.S. fraud cases for which a one-to-one propensity score non-fraud match campaign can be found. The total sample includes 400 campaigns (200 fraud + 200 non-fraud) (see Table 1). However, text description readability indices are only available for 394 cases. Control 1 (*creator(s)' characteristics/background*) includes *Natural Person*, *Formal Name*, *No. of Creator-Backed Projects*, and *No. of Creator-Created Projects*; control 2 (*social media affinity*) includes *Facebook_Personal*, *Facebook_Page*, and *No. of External Links*; control 3 (*campaign funding and reward structure*) includes *Duration*, *No. of Pledge Categories*, and *Min. Pledge Dummy*; control 4 (*campaign's description details*) includes *ARI* and *Video Pitch*. Robust standard errors are one-way-clustered by campaign category. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
ARI	-0.069**				
	(-2.13)				
CL		-0.089*			
		(-1.81)			
Gunning Fog			-0.039		
			(-0.67)		
FKG				-0.072	
				(-1.62)	
FRE					0.017
					(1.51)
Control 1	Yes	Yes	Yes	Yes	Yes
Control 2	Yes	Yes	Yes	Yes	Yes
Control 3	Yes	Yes	Yes	Yes	Yes
Control 4	Yes	Yes	Yes	Yes	Yes
Observations	394	394	394	394	394
Pseudo R^2	0.120	0.121	0.116	0.118	0.119

Table A4: Correlation Matrix

This table shows the Pearson correlation coefficients for all variables if data items are available. All variables are considered in the analyses as either main variables, or as robustness checks (see Table 2 for variable descriptions and calculation methods). The four missing correlation coefficients (.) can be obtained because of dependence, e.g., a creator needs to have a Facebook page in order to have likes on the Facebook page. * indicates statistical significance at least at a 5% level.

		1	2	3	4	5	6	7	8	9	10	11	12	136	14	15	16	170	18	19	20	21	22	23	24	25	26
1	Natural Person	1.00																									
2	Formal Name	0.82*	1.00																								
3	No. of Creator-Backed Projects	0.10*	0.10*	1.00																							
4	No. of Creator-Created Projects	0.03	0.03	0.21*	1.00																						
5	Facebook	0.16*	0.12*	-0.01	0.02	1.00																					
6	Facebook_Personal	0.29*	0.23*	0.02	0.06	0.77*	1.00																				
7	FacebookPage	-0.10*	-0.10*	-0.04	-0.02	0.50*	0.07	1.00																			
8	No. of External Links	-0.02	-0.02	0.07	0.04	0.27*	0.07	0.45*	1.00																		
9	LinkedIn	0.15*	0.19*	-0.02	-0.04	0.04	0.03	0.01	0.23*	1.00																	
10	Ln (No. of FB Connections)	-0.01	-0.02	0.08	0.01	0.14*	-0.21*	0.47*	0.35*	0.02	1.00																
11	Ln (No. of FB Friends)	0.15	0.14	0.10	-0.01	0.15*	0.15*	-0.03	0.13	-0.09	0.68*	1.00															
12	Ln (No. of FB Likes)	-0.03	-0.09	0.12	-0.01	(.)	-0.09	(.)	0.11	0.09	0.96*	0.10	1.00														
13	Duration	0.02	-0.04	0.09	-0.11*	0.12*	0.02	0.12*	0.03	0.03	0.04	-0.05	0.06	1.00													
14	No. of Pledge Categories	0.03	0.07	0.03	-0.04	0.12*	0.05	0.08	0.06	0.07	-0.06	-0.04	-0.12	0.00	1.00												
15	Min. Pledge Dummy	0.09	0.07	0.03	-0.10*	0.00	0.06	0.00	0.04	-0.04	0.02	0.02	-0.03	-0.09	0.22*	1.00											
16	Avg. Small-2 Pledge Dummy	0.13*	0.08	0.04	-0.06	0.03	0.06	-0.02	0.04	-0.05	0.08	0.12	0.10	-0.08	0.29*	0.58*	1.00										
17	Avg. Small-3 Pledge Dummy	0.13*	0.11*	0.09	-0.01	-0.01	0.04	-0.01	0.06	-0.07	0.08	0.09	0.09	-0.11*	0.29*	0.51*	0.85*	1.00									
18	Min. Pledge Amount	-0.11*	-0.13*	-0.05	0.01	-0.03	-0.04	0.01	-0.05	0.01	0.00	0.02	-0.13	0.09	-0.20*	-0.46*	-0.34*	-0.32*	1.00								
19	Avg. Small-2 Pledge Amount	-0.15*	-0.15*	-0.07	-0.02	-0.02	-0.02	0.04	-0.04	0.01	0.01	-0.08	-0.03	0.11*	-0.23*	-0.39*	-0.45*	-0.41*	0.90*	1.00							
20	Avg. Small-3 Pledge Amount	-0.16*	-0.15*	-0.07	-0.04	-0.03	-0.02	0.01	0.00	0.02	0.00	-0.08	-0.02	0.12*	-0.25*	-0.34*	-0.45*	-0.46*	0.81*	0.93*	1.00						
21	ARI	-0.09	-0.12*	-0.01	-0.03	0.04	-0.02	0.04	0.12*	-0.05	0.05	0.10	0.05	0.07	-0.05	-0.03	-0.07	-0.07	0.08	0.07	0.09	1.00					
22	CL	-0.16*	-0.16*	-0.06	-0.09	0.05	0.02	0.03	0.08	0.00	0.04	0.05	0.08	0.05	-0.09	-0.04	-0.17*	-0.14*	0.09	0.11*	0.14*	0.80*	1.00				
23	Gunning Fog	0.03	-0.03	0.05	0.06	0.02	-0.05	0.04	0.11*	-0.09	0.04	0.12	0.00	0.05	0.02	-0.01	0.07	0.04	0.03	-0.01	0.00	0.70*	0.15*	1.00			
24	FKG	-0.10*	-0.14*	-0.05	-0.06	0.05	-0.02	0.05	0.12*	-0.06	0.02	0.08	-0.03	0.09	-0.08	-0.05	-0.09	-0.10	0.09	0.08	0.01*	0.95*	0.70*	0.74*	1.00		
25	FRE	0.16*	0.17*	0.10*	0.12*	-0.05	0.00	-0.05	-0.10*	0.02	0.00	-0.03	0.04	-0.10*	0.12*	0.07	0.17*	0.16*	-0.10*	-0.11*	-0.14*	-0.86*	-0.87*	-0.40*	-0.9107*	1.00	
26	Video Pitch	-0.07	-0.05	0.05	0.00	0.11*	0.09	0.11*	0.08	0.05	0.12	0.08	-0.03	0.05	0.02	-0.01	-0.11*	-0.11*	0.04	0.08	0.10*	0.07	0.11*	0.00	0.09	-0.13*	1.00