Three Essays on Entrepreneurial Finance

Moein Karami

A Thesis

in John Molson School of Business

Presented in Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy (Business Administration-Finance) at

Concordia University

Montreal, Quebec, Canada

August 2021

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Moein Karami

Entitled: Three Essays on Entrepreneurial Finance

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy Business Administration (Finance)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Chair
Professor Mahesh Sharma

External Examiner
Dr. Alexander Groh

External to Program
Dr. Ketra Schmitt

Examiner
Dr. Maher Kooli

Examiner
Dr. Thomas Walker

Thesis Supervisor
Dr. Denis Schweizer

Approved by

Dr. Cedric Lesage, Graduate Program Director

August 27, 2021
Date of Defence

Dr. Anne-Marie Croteau, Dean John Molson School of Business

ABSTRACT

Three Essays on Entrepreneurial Finance

Moein Karami, Ph.D.

Concordia University, 2021

This dissertation aims to shed light on dynamics of new forms of entrepreneurial finance, in general, and crowdfunding, in particular, from three different following aspects.

First, we conduct an exhaustive search of all media reports on Kickstarter campaign fraud allegations from 2010 through 2015, and determine campaign features that are associated with a higher probability of observing fraud, using multiple matched samples of non-fraudulent campaigns. We also document the short-term negative consequence of possible breaches of trust in the market, using a sample of more than 270,000 crowdfunding campaigns posted from 2010 through 2018 on Kickstarter. Our results show that crowdfunding projects launched around a significant misconduct detection on Kickstarter tend to have a lower probability of success, raise less funds, and attract fewer backers.

Second, using a sample of 230,255 crowdfunding campaigns (2013-2018) on Kickstarter and drawing upon previous empirical evidence, the statistically significant effect of five variables on campaign success is documented. To date, numerous studies have focused on determining factors affecting crowdfunding success, however, it is extremely difficult to compare results across papers as each use incompatible specifications, and different control variables. The identified variables aim to measure the intensity of competition, creator's crowdfunding experience, project quality & creator confidence, portal recognition, and project size. Furthermore, the effect of campaign creator's citizenship, as well as project location, on funding success is investigated.

Third, and drawing upon previous findings on the effect of biological factors on investment behavior and entrepreneurship, a significant positive relationship between fWHR (facial Width-to-Height Ratio) of the hedge fund managers in the sample (1994-2016) and fund's risk is documented. The association between facial masculinity of male entrepreneurs and their fund-raising outcome is also investigated using a sample of ABC channel's "Shark Tank" show (2009-2014). The results are in line with previous findings on the positive correlation between fWHR and testosterone; a hormone which its role in describing behavioral patterns such as competitiveness and risk-taking is well-established. The study sheds light on the factors that are not incorporated in economic models, but may significantly affect financial risk-taking and performance, as well as entrepreneurial outcomes.

Acknowledgements

I would like to express sincere gratitude to my supervisor, Dr. Denis Schweizer, for his invaluable support during my PhD studies. It was a great pleasure to study and work under his supervision.

I would also like to thank my thesis committee members, Dr. Maher Kooli and Dr. Thomas Walker, for their priceless encouragement, and the thesis examiners, Dr. Alexander Groh and Dr. Ketra Schmitt for many helpful comments and feedback.

I am extremely grateful for the financial support provided by Concordia University during my PhD studies.

Dedication

To my wonderful family for their unconditional love,

and,

To my respected teachers, colleagues and friends for their invaluable support.

Contribution of Authors

Part of this thesis is working paper co-authored with Douglas Cumming, Lars Hornuf, and Denis Schweizer.

Cumming, D., Hornuf, L., Karami, M., & Schweizer, D. (2020). Disentangling crowdfunding from fraudfunding. *Max Planck Institute for Innovation & competition research paper 16-09*.

The manuscript is in the revise and resubmit stage at *Journal of Business Ethics*, and has been reformatted and reorganized according to the requirements set out in the guideline of the School of Graduate Studies.

Table of content

List of Figures

List of Tables

Chapter 1: Introduction

Small firms and new ventures have become an increasingly important component of economic development. Entrepreneurial finance aims at addressing key questions which challenge all entrepreneurs: 1) the amount of funding that can and should be raised; 2) when should it be raised and from whom; and 3) how should funding contracts and exit decisions be structured. Inspired by its recent success and popularity, this dissertation, in particular, investigates the dynamics of a new means of entrepreneurial finance, namely, reward-based crowdfunding. In the broadest sense, crowdfunding is the use of small amounts of capital from a large number of individuals/investors to finance a new business venture. Furthermore, we aim at shedding light on the biological underpinnings of entrepreneurship, based on a new school of thought that incorporates human biology into the study of managerial behavior.

In Chapter 2, we explore the occurrence, determinants and consequences of fraud in crowdfunding market, and its implications for business ethics literature. We emphasize the importance and fragility of trust in this nascent industry, and determine campaign features that are associated with a higher probability of observing fraudulent behavior in crowdfunding market.

In Chapter 3, we use a comprehensive sample of crowdfunding campaigns posted through 2013-2018 on Kickstarter, the largest global crowdfunding platform, and document the statistically significant effect of five independent variables on campaign success. We also shed light on the effect of campaign creator’s nationality, as well as project location, on campaign dynamics and funding outcome.

In Chapter 4, we provide an inter-disciplinary literature review on the effect of physiological and hormonal factors on investment behavior as well as entrepreneurial activities in order to introduce avenues for future research in finance and entrepreneurship, on similar topics. Our results shed more light on recent works linking facial metrics and physiological factors to economic behavior, and extend our understanding on biological reasons affecting financial risk-taking, performance, and entrepreneurship.

Chapter 2: Disentangling Crowdfunding from Fraudfunding

Douglas Cumming[1], Lars Hornuf[2], Moein Karami[3], Denis Schweizer[3]

Abstract

Fraud in reward-based crowdfunding market has been of concern to regulators, but it is arguably of greater importance to the nascent industry itself. Despite the importance of this growing industry for entrepreneurial finance, our knowledge of occurrence, determinants and consequences of fraud in crowdfunding market, and its implications for business ethics literature remain limited. In this study, we conduct an exhaustive search of all media reports on Kickstarter campaign fraud allegations from 2010 through 2015, then follow up until 2018 to check the ultimate outcome of the allegedly fraudulent campaigns. First, we construct a sample of 193 fraud cases and categorize them into detected vs suspected fraud based on a set of well-defined criteria. Using multiple matched samples of non-fraudulent campaigns, we determine campaign features that are associated with a higher probability of observing fraudulent behavior. Second, we document the short-term negative consequence of possible breaches of trust in the market, using a sample of more than 270,000 crowdfunding campaigns posted from 2010 through 2018 on Kickstarter platform. Our results show that crowdfunding projects launched around a public announcement of a late and significant misconduct detection on Kickstarter (resulting in suspension of a successful campaign), tend to have a lower probability of success, raise less funds, and attract fewer backers.

[1] College of Business, Florida Atlantic University, 777 Glades Road, Boca Raton, Florida 33431, US, e-mail: cummingd@fau.edu

[2] Faculty of Business Studies and Economics, University of Bremen, Max-von-Laue-Str. 1, 28334 Bremen, Germany, e-mail: hornuf@uni-bremen.de

[3] John Molson School of Business, Concordia University, 1450 Rue Guy, Montreal, Quebec H3H 0A1, Canada, e-mails: moein.karami@concordia.ca, denis.schweizer@concordia.ca

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]

2.1. Introduction

Reward-based crowdfunding (hereafter, crowdfunding) has emerged in recent years as a catalyst for entrepreneurship, an important new means of financing early-stage ventures, and a door opener for successful future financing. As an alternative solution to the capital gap problem for start-ups, crowdfunding can complement or even substitute for other sources of early stage financing, such as venture capital or angel investors. Early-stage ventures have benefited enormously from the availability of crowdfunding, and the positive impact of crowdfunding for new firm creation and future venture capital investments has become evident in recent years (Assenova et al., 2016; Sorenson et al., 2016). This signifies the importance of investigating issues that could negatively affect the crowdfunding market and endanger its long-term existence.

Trust between counterparties is vital in the process of economic exchange, in general, and venture financing, in particular (Brockman et al., 2020; Hain et al., 2016). Therefore, crowdfunding adoption depends significantly on establishing trust in the market. Equity markets have already demonstrated the fragility of trust, and how a breach of trust not only negatively affects specific firms (Davidson and Worrel, 1988), but also can result in the collapse of an entire market segment (Hainz, 2018). The concept of *Trust Triangle* was recently adapted for financial markets and financial fraud (Dupont and Karpoff, 2019). According to the Trust Triangle framework, firms can *ex ante* invest in accountability and build trust through three main channels: first-party, related-party, and third-party enforcement (first leg, second leg, and third leg of the Trust Triangle). These three legs are not equally effective in crowdfunding context. The crowdfunding market is still in its infancy, and campaign creators have no legal obligation, for example, to file income statements or to provide profit and loss accounts to the platform or regulatory bodies suggesting a weak third-party

[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-crowdfunding.

enforcement in the market. Backers must trust campaign creators to use the funds obtained to deliver on the promises made in the campaign description (first-party enforcement) and also trust the platform to conduct a thorough pre-screening of the projects before they are posted (related-party enforcement). Thus, a core element of a functional crowdfunding market is trust between backers, campaign creators, and the platform.

Incidents of fraudulent behavior and misconduct by campaign creators, and the inactivity of the platform in preventing them, can negatively affect the open-mindedness of crowdfunding backers. Therefore, it is important to categorize and document fraudulent cases to 1) assess which factors could be perceived as signals of weak first-party enforcement that can help in predicting subsequent fraudulent behavior, and 2) identify incidents that could have led to a breach of trust associated with weak related-party enforcement, and to analyze their consequences for the crowdfunding market.

In view of the less stringent disclosure rules and the dearth of enforcement surrounding crowdfunding, we categorize fraudulent behavior based on 2010 to 2015 Kickstarter campaign fraud allegation reports (and following up on the reported cases until 2018 to check the ultimate outcome). We conduct a thorough and methodical search of media reports, and include the reported campaigns in our dataset if certain criteria are met, to finalize a sample of Kickstarter campaigns associated with fraudulent behavior from 2010 through 2015. We define "detected" fraud in cases of outright misrepresentation, e.g., in the "Kobe Red" case, a Kickstarter campaign from 2013, which involved the production of Japanese beer-fed Kobe beef jerky.[1] 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 either the promised rewards are significantly changed to the disadvantage of the backers, or the following three conditions are met simultaneously (besides the mention in media reports): 1) the rewards are significantly delayed (more than one year past the promised delivery date); 2) the creator has ceased *credible* communication with backers (through, for example, posting updates on the campaign web page) for at least six months after the promised delivery date; and 3) the promised product is never delivered, and the backers are not refunded. In either case, when these campaigns are reported in the news media or on consumer advocacy

[1] 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.

websites, they are picked up in our dataset and are included in the fraud sample if the abovementioned conditions are also met.

In the first part of our empirical analyses (i.e., *Determinants of Fraud*), using the constructed fraud sample and multiple matched samples of non-fraudulent campaigns, we find that fraudsters are less likely to have engaged in prior crowdfunding activities and use social media, such as Facebook. We also find that fraudsters tend to offer a higher number of enticements through pledge categories and choose longer campaign durations. Finally, results show that fraudsters are more likely to provide easier-to-read campaign pitches based on the readability indices of campaign description. In sum, we identify factors that can be perceived as signals of first-party enforcement, as well as project quality, and our results illustrate their relevance in predicting subsequent fraudulent behavior.

In the second part of our empirical analyses (i.e., *Platform-Wide Consequences of Fraud*), we document that occurrence of a large public crowdfunding scam (campaign reaching goal amount but suspended by the platform in late stages of funding period), has an economically significant negative impact on other concurrent projects. Therefore, a few incidences over a short period of time may cause a tremendously negative spillover effect on the entrepreneurial environment. We collected data on more than 270,000 Kickstarter campaigns posted from 2010 through 2018, and found that as a consequence of Kickstarter "late" suspensions that can be potentially perceived as signal of weak related-party enforcement and inefficient platform pre-screening, the probability of reaching the goal amount for campaigns that are launched around the same date is lower by about 6.38%, and on average the pledged amount decreases by about 9.6%, all else being equal. Crowdfunding market participants (potential backers) tend to be relatively shocked by large, public scams that are suspended by the platform much later than normally expected, questioning the effectiveness of the second leg (related-party enforcement) of the Trust Triangle. Backer's trust in platform integrity is vital especially since the platform revenue is a percentage of the campaigns' raised amounts, leading to a potential agency problem. Backers may react negatively because they perceive these suspended campaigns as first-hand evidence that not only legal enforcement is weak, but also the platform (and their own) scrutiny was not efficient.

Finally, and as robustness check of our results, we show that the identified signals of strong first-party enforcement also positively affect campaign success, which can be thought of as backers' trust level in terms of contributed amounts, and this effect is larger when a "late" platform enforcement (perceived by backers) occurs. We highlight the importance of related-party enforcement and platform scrutiny before projects are posted, especially since crowdfunding platforms do not have mechanisms in place for enforcing accountability (e.g., by charging an insurance fee proportional to the overcontribution to the campaign) once the funds are transferred to the campaign creator.

Our paper is related to a growing literature on crowdfunding that has, to date, focused primarily on the determinants of funding success (see, e.g., Agrawal et al., 2015; Ahlers et al., 2015; Belleflamme et al., 2013; Colombo et al., 2015; Mollick, 2014; Vismara, 2016). Prior research in the context of our study has focused on late deliveries (Mollick, 2014), project or firm failures (Hornuf et al., 2018; Signori and Vismara, 2018), factors that affect backer trust (Liang et al., 2019), mechanisms that can potentially deter misconduct in crowdfunding market (Belavina et al., 2020), and the impact of pro-social framing, altruism and self-interest on crowdfunding success (André et al., 2017; Berns et al., 2020; Defazio et al., 2020). Other papers have examined the role of securities regulation in equity crowdfunding markets (Bradford, 2012; Hornuf and Schwienbacher, 2017), the return on investment in equity crowdfunding (Hornuf et al., 2018; Signori and Vismara, 2018), and the dynamics of crowdfunding project support over time (Hornuf and Schwienbacher, 2018). We contribute to the entrepreneurial finance literature by identifying specific campaign- and creator-related factors that correlate with fraudulent behavior in crowdfunding market. We also document the negative effect of perceived weak platform scrutiny on the success of concurrent crowdfunding campaigns. Our study opens avenues for future research on crowdfunding fraud and its effects on entrepreneurship by developing and integrating new fraud detection models in an entrepreneurial finance setting (see, e.g., Perez et al., 2020).

The remainder of this paper is organized as follows: next section provides an overview of the legal treatment of fraud in crowdfunding markets, which is followed by the formation of our hypotheses. Then we introduce the data and outline our methodology. Thereafter, we present univariate and multivariate empirical analyses, followed by a discussion of the results and several robustness checks. The last section concludes and discusses implications for research, practice, and policy.

2.2. Legal Sanctions on Fraud in Crowdfunding Markets

Law enforcement through third parties—the third leg of the Trust Triangle—is an essential element to deter fraud in markets (Ehrlich, 1973). Securities laws in the U.S. have several antifraud provisions that allow investors and the SEC to bring legal actions and enforce legal rules. 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, 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. 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[1]. 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 thirty-one 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]).

The inactivity of the FTC and the lack of private legal actions does not necessarily imply that fraudulent behavior is absent in crowdfunding markets. The FTC's inactivity can be partially attributed to the high costs of verifying contracts (Lacker and Weinberg, 1989; Townsend, 1979) in crowdfunding context. This is because, in many cases, it is extremely difficult and costly to prove that the creator intended to deceive backers. As a result, many backers and government agencies

[1] See: https://www.ftc.gov/news-events/press-releases/2015/06/crowdfunding-project-creator-settles-ftc-charges-deception

may be disincentivized from bringing presumably fraudulent cases in front of a court. Consequently, it is also important to investigate suspicious crowdfunding activities in order to methodically construct a comprehensive and reliable sample of fraudulent campaigns.

To summarize, fraudsters in a reward-based crowdfunding campaign may anticipate being detected as the campaign progresses and the delivery date approaches. Despite the weak incentives of backers, who may have pledged only small amounts, to bring legal action, fraudsters are still subject to prosecution by the FTC or by state attorneys general. However, the inactivity of government agencies, such as the FTC, until 2019 has signaled the overall absence of the third leg of the Trust Triangle—legal liability. Thus, the lack of private and government actions may provide fraudsters with sufficient incentives to engage in deceptive activities.

2.3. Theory and Hypotheses

Dupont and Karpoff (2019) explain the importance and fragility of trust in the process of economic exchanges. They introduce a framework that involves three primary mechanisms to provide discipline, deter opportunistic behavior, and build sufficient trust in economic relationships to foster healthy economic exchanges.

We have learned from the equity markets that fraudulent activities can result not only in sharp declines in firm performance and share prices (Karpoff et al., 2008; Rezaee, 2005), but also to the collapse of an entire market segment. In 1997, the market segment *Neuer Markt* was established on the German stock exchange, with the goal of financing innovative small and medium-sized growth companies. After a strong start, the segment reached a market capitalization of about $234 billion (Hainz, 2018). However, several incidents of corporate fraud and misconduct eroded the segment's reputation, and it was closed only six years after its launch, with a loss of about 90% from market peak. Similarly, given that crowdfunding is a relatively new funding source, fraud cases can exert very strong destructive power on the segment, and lead to spillover effects on future campaigns.

The three legs of this Trust Triangle are 1) first-party enforcement (i.e., personal ethics, integrity, and culture); 2) related-party enforcement (i.e., market forces and reputational capital); and 3) third-party enforcement (i.e., laws, regulations, and regulators). The legal enforcement by government agencies within the crowdfunding market has been lax until today and regulators have had limited capacity for enforcement, therefore, project creator's integrity and platform enforcement play important roles in determining backers' trust level. This is extremely important because platform revenue is directly related to the campaigns' raised amounts (usually it is a fixed percentage of the raised amount) at the end of the campaign, and the remaining amount will be transferred to the campaign creator. After the campaign ends and the project creator receives the funds, there remains a risk that the creators are not committing the promised level of effort to the venture, or the creators may use the funds to extract private benefits that were not agreed upon, creating a moral hazard problem (Hainz, 2018). This problem can be somewhat reduced by writing complete contracts, which does not seem feasible in a crowdfunding context, but it can also be reduced by strengthening the first- and related-party enforcement in the market.

First, we focus on the first leg of the Trust Triangle (i.e., first-party enforcement) and signals of project quality, in order to develop Hypothesis 1 to 3. We aim to identify creator and campaign characteristics that can be perceived as credible (i.e., difficult to mimic) signals of first-party enforcement.

Economists and psychologists have provided various explanations for why individuals engage in fraudulent activities. In the context of crowdfunding, backers can check the campaign page on the crowdfunding platform, and form their own expectations about the quality of the venture and any probability of fraud. They process information by, for example, reading the campaign description, and watching the campaign video. All the information provided clearly helps reduce the asymmetric information, but it does not eliminate it. Fraudulent campaign creators, on the other hand, have a clear incentive to increase information asymmetries, so that backers cannot *ex ante* differentiate fraudulent from non-fraudulent projects. Therefore, it seems necessary to identify creator and campaign features that can *ex ante* serve as signals of first-party enforcement and are difficult (or costly) to mimic for fraudulent campaigns, as one might argue fraudsters can potentially implement a series of symbolic actions to build trust and increase their chance of success (see, e.g., Zott and Huy, 2007).

In the realm of crowdfunding, we have identified three broad themes where backers could theoretically identify and examine signals of stronger first-party enforcement based on available information: 1) *creator(s)' characteristics/ background*, 2) *creator(s)' social media affinity*, and 3) *campaign characteristics*.

Social psychologists have argued that, when people are acting in a dishonest manner, they nevertheless remain concerned about maintaining a positive self-concept (Gino et al., 2009; Jiang, 2013; Mazar et al., 2008). This idea brings the focus back to the first leg of the Trust Triangle, which suggests that personal ethics play an important role when campaign creators commit fraud. Mann et al. (2016) focus on non-violent crimes, and find that internal sanctions provide the strongest deterrent to such crimes. The effect of legal sanctions was weaker and varied across countries. As a result, fraud in crowdfunding context may not only follow a solely economic calculation by the project creator, but may also reflect his or her personal attitudes and reputation.

We do not expect creators that have already built a rich history of successful campaigns to launch a subsequent fraudulent project. As pointed out by Diamond (1989), creators build their reputation by engaging more frequently in the market, and would consequently suffer a larger loss if they engage in misconduct. In other words, if a creator does run multiple honest campaigns, it not only provides a signal of experience, but may decrease the probability that the person later act dishonestly. Similarly, creators who have previously backed other crowdfunding projects are likely to believe in the democratic and supportive idea of crowdfunding (Cumming et al., 2019b) and this can make it difficult for them to reconcile the idea of leading a scam later on, however, it is noteworthy that backing multiple projects is easier to mimic and less costly (compared to creating projects) for fraudsters as they can contribute small amounts to multiple campaigns to signal their prior crowdfunding activity. In sum, we predict a negative relationship between crowdfunding fraud and the intensity with which a creator uses crowdfunding as a backer or a creator as presented in Hypothesis 1.

Hypothesis 1 (Creator(s)' Characteristics and Background): *Crowdfunding fraudsters are less likely to have engaged in prior crowdfunding activities.*

Moreover, backers can easily screen creators' social media activities on the Internet. If personal ethics and a positive self-image are important to fraudsters, they may try to avoid this scrutiny, by, for example, not having any social media presence, not least because social media also facilitates fraud detection. Furthermore, a social media presence (observable by backers of the campaign) 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. Similar to the earlier work on the effect of media on corporate social responsibility (El Ghoul et al., 2019), we theorize that social media presence can lower the risk of crowdfunding fraud. Moreover, early backers are often friends and family, which is a specific feature of non-equity crowdfunding (Agrawal et al., 2015; Colombo et al., 2015). Arguably, committing an outright scam might be harder if investors consist of friends and family.

Lin et al. (2013) show that, in peer-to-peer lending, online friendships of borrowers act as signals of credit quality and lead to a higher probability of a successful funding. However, 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. Hence, it is not clear *per se* whether elaborate fraudsters have fewer or more social media contacts and whether this is a difficult to mimic feature. The same holds for external links on a campaign website that leads to other fake websites that supposedly support the trustworthiness of the campaign. 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. This would again jeopardize the positive self-image of the campaign creator, and highlights the importance of the first leg of the Trust Triangle. 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 predict a negative correlation between social media use and fraud.

> **Hypothesis 2 (Social Media Affinity):** *Crowdfunding fraudsters are less likely to have social media presence, and provide fewer external links.*

Finally, the *Campaign Funding and Reward Structure* and the *Campaign Description Details*, which we group under the umbrella term *Campaign Characteristics*, can, in the spirit of Spence (1973), provide credible signals of first-party enforcement and project quality. In the context of entrepreneurship, Shailer (1999) develops a theoretical model showing that signals that entrepreneurs provide to prospective lenders (by either providing direct information or taking observable actions) may assist lender in allocating *ex ante* default probabilities based on the lender's prior knowledge of group characteristics. We aim to identify and determine the value of such signals in the context of crowdfunding, and their relationship with fraudulent behavior.

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 are less likely to send credible signals of quality to the crowd. Therefore, fraudsters may believe it is optimal to keep the funding period ongoing to raise as much capital as possible. Longer funding periods may also make detection more likely, and thus also increase the risk of not receiving funds. Consequently, it remains an empirical question as to whether a longer funding period reduces or increases the probability of fraud but considering the low probability of detection in the course of campaign (resulting in campaign suspension) we believe that short duration is a credible signal of the project quality. We, therefore, derive Hypothesis 3.A as follows:

Hypothesis 3.A: *Crowdfunding fraudsters are more likely to choose longer funding periods.*

While fraud in a crowdfunding campaign might be detected by backers once e.g., the creator fails to deliver the product without any explanation or follow up, the ultimate prosecution of the scam may be the most 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 to individual backers. In order to make third-party enforcement even less likely, an optimal strategy for fraudulent campaign creators might be to target backers who are willing to contribute only small pledges. In line with this conjecture, we believe fraudsters will target as many different backers as possible, ideally those who will support the project with smaller pledges. One way to achieve this is by creating many different pledge categories, so that backers can easily provide various small-size contributions. We, therefore, derive Hypothesis 3.B as follows:

Hypothesis 3.B: *Crowdfunding fraudsters are more likely to offer smaller minimum pledge allowance.*

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). Research on the manipulation of stock markets has also long explored so-called "pump and dump" schemes. These schemes involve fraudsters acquiring long positions in stocks, and then heavily promoting them on online chat forums or by spoof trading (deleting orders before execution to keep up appearances of an active order book). Fraudsters thus 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 heavily promote a campaign by offering many project enticements with various reward levels (Belleflamme et al., 2014; Mollick, 2014). Moreover, because they do not intend to ship any product or continue communicating with backers, they can easily raise small amounts by as many backers as possible without being overstrained by excess demand and shipping costs later on. We, therefore, derive Hypothesis 3.C as follows:

> **Hypothesis 3.C:** *Crowdfunding fraudsters are more likely to offer larger number of reward/pledge categories.*

Finally, in crowdfunding markets, fraudulent campaign creators may try to increase information asymmetries to make it more difficult for backers to differentiate between scams and valuable projects. The main place for a backer to learn about a project is through the description, which is normally a few thousand words (Cumming et al., 2019a). Crowdfunding fraudsters are, therefore, less likely to provide a professionally worded description in order to foster confusion and ideally perpetrate the fraud without detection. In contrast, professional entrepreneurs use campaign descriptions to signal their quality. Moreover, it is likely complicated to accurately and professionally describe a product that does not exist and was never intended to exist in the first place. This is in line with findings by Siering et al. (2016), who provide evidence that linguistic and content-based cues in static and dynamic contexts can help predict fraudulent behavior in crowdfunding. Parhankangas and Renko (2017) show that certain linguistic styles increase the probability of success of social campaigns, such as, e.g., those that make the campaign and creator(s) more understandable and more relatable to the crowd. Alternatively, a simple wording of the campaign description (e.g., without a need for many years of formal education in order to understand it in a first read) might help fraudsters in targeting a less educated and broader crowd, which would result in simpler, easier-to-read terminology. We, therefore, derive Hypothesis 3.D as follows:

> **Hypothesis 3.D:** *Crowdfunding fraudsters are more likely to use simply worded campaign descriptions (i.e., lower formal education required to understand the description in a first read).*

Next, we focus on the second leg of the Trust Triangle (i.e., related-party enforcement) in order to develop Hypothesis 4. In general, Reward-based crowdfunding platforms do not conduct sophisticated background checks or due diligence (compared to e.g., equity crowdfunding platforms), but if Kickstarter's "Trust & Safety" team finds a campaign in violation of Kickstarter's rules, the campaign is suspended. Although these suspended cases are not necessarily outright fraud and there is no legal conviction, the platform-wide consequences of observed incidences of misconduct detection, proxied by campaign suspensions, is worth investigating empirically and *a* *priori* not clear. Backers who observe campaigns being suspended for violating platform rules or raising serious suspicion, could rationally infer that related-party enforcement works. On the other hand, backers may learn that fraudulent campaign creators have already scammed many backers prior to suspension, and not only the platform cannot ensure future accountability, but the pre-screening conducted by the platform has not worked efficiently. Hence, large scale campaign suspensions that have already attracted a lot of backers, raised large amount of funds, and are close to the scheduled deadline, might not only substantially weaken backers' confidence in their own fraud detection skills, but also weaken the trust in related-party enforcement. As an upshot of this weakened trust, concurrent crowdfunding campaigns may face difficulties raising capital and achieving their funding goals. We, therefore, derive Hypothesis 4 as follows:

> **Hypothesis 4 (Platform-wide Consequences of Fraud):** *Campaigns posted around a late and visible suspension of a successful crowdfunding project have a lower probability of success, raise less funds, and attract fewer backers.*

2.4. Data

We divide the data collection into two parts. First, we describe how we categorize fraudulent campaigns, derive the respective fraud sample and matched sample of non-fraudulent campaigns, and examine factors that are associated with higher likelihood of observing fraudulent behavior using the constructed sample. Second, we show how we constructed our sample for studying platform-wide consequences of possible breaches of trust. All variable definitions used in empirical analyses can be found in Table 2.1.

—Please insert Table 2.1 about here—

2.4.1. Categorizing Fraudulent Behavior in Crowdfunding

The legal definition of fraud in crowdfunding, as outlined previously, is not simple to operationalize for an empirical study, because only a few cases have been decided by an ordinary judge thus far. In a theoretical context, Belavina et al. (2020) point out that the platform (i.e., Kickstarter) can leave backers exposed to two risks: 1) entrepreneurs may run away with backers' money (funds misappropriation), and 2) product specifications may be misrepresented (performance opacity), however, we aim to operationalize this definition in an empirical setting. Therefore, and using the same basic logic, we focus on what is considered industrywide as *detected fraud* and *suspected fraud* (see, for example, *Crowdfund Insider*[1] for an overview). In what follows, we describe our categorization of fraud, in more detail based on media reported cases resulting in construction of a sample of 193 fraudulent campaigns.

The first category, *detected fraud*, includes 1) *pre-empted fraud*, which occurs when a supposedly fraudulent crowdfunding campaign was reported on media but it either got suspended by the platform 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 creators. However, after the funding is completed, backers may find out that creators attempted to resell pre-existing products as part of their campaign, or that they misrepresented material facts, used intellectual property they do not hold legal rights to, or that the project is a fake altogether. The fraud may 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 has never started.

The second category, *suspected fraud*, occurs when supposedly fraudulent crowdfunding campaign was reported on media, and either 1) the following three conditions (1a, 1b, and 1c) are met simultaneously, or 2) the rewards are changed to the disadvantage of the backers (condition 2). The three conditions are as follows: Rewards are delayed by more than one year from the promised delivery date (condition 1a); the creators have ceased credible communication with backers (through posting updates on the campaign web page) for at least six months after a promised delivery date (condition 1b); and rewards are not delivered, and backers are not partially or fully refunded until December 31, 2018 (condition 1c).

Detection of campaigns where rewards have been significantly changed is straightforwardly accomplished by studying news articles on a particular campaign, or by reading comments posted by backers after the rewards have been delivered. However, if the delivery of the rewards is overdue, it can be more difficult to distinguish between fraudulent projects and those that failed or experienced normal setbacks, such as unforeseen challenges or technological issues. To overcome this problem,

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

we consider a campaign where rewards are delayed for at least one year after the delivery date as being in the suspected fraud category, but *only* if 1) the creator has also not posted meaningful updates for at least six months after the originally promised delivery date, 2) the promised reward is not delivered until the end of our observation period, and 3) backers were not at least partially refunded.[1] To classify projects as suspected fraud, we manually checked and followed all campaigns up to December 31, 2018, to determine whether rewards were finally delivered, whether campaign creators provided credible explanations[2] for the late delivery or failure, or whether backers were at least partially refunded. In cases where a final reward, credible communication, or a refund was delivered, the project was excluded from our suspected fraud sample.[3] We acknowledge that extreme incompetence of project creator can provide alternative explanation for campaign being marked as fraudulent, but even in those cases, not providing explanations for failure through updates can be perceived as a form of serious misconduct.

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 platform 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 platforms themselves.[4]

There is no commercial database available for fraud cases in crowdfunding, and our base media reports sample covers all actual and potential fraud campaigns reported on a website called Kickscammed (http://kickscammed.com). Kickscammed's purpose is to offer the crowd an opportunity to report suspicious or fraudulent activities in crowdfunding campaigns. This website is not linked to Kickstarter and is an independent website where backers can post their opinions and signal a potential scam.

[1] Our observation period for identifying suspected fraudulent campaigns spanned 2010 through 2015, and we classified the campaigns in April 2016. We re-checked all suspected fraud campaigns on December 31, 2018, and excluded those where the rewards were finally delivered, the reason for late delivery/failure was explained, or backers were at least partially refunded.

[2] See Kickstarter's guidelines to define what credible communication means in case of a failed project: https://www.kickstarter.com/fulfillment.

[3] This resulted in a further four exclusions from our base media reports fraud sample.

[4] See http://www.theverge.com/2013/11/8/5081806/kickstarter-alleged-chargeback-fraud-hits-over-100-campaigns.

Table 2.2 shows the steps in constructing the fraudulent campaigns' sample. As of April 30, 2016, we were able to identify and confirm 181 fraud cases for the 2010-2015[1] period that were reported on Kickscammed and met our criteria to be categorized as *detected* or *suspected fraud*. However, Kickscammed's website does not necessarily cover all instances of fraudulent activity on Kickstarter, so we complement our dataset with a news search using Google, Factiva, and LexisNexis. Our initial fraud dataset is, therefore, comprised of 200 fraudulent campaigns and after excluding 7 campaigns for which there is no data available, we finalize a sample of 193 fraudulent cases[2] (see Table 2.2, Panel A).

Panel B of Table 2.2 illustrates the differences in the number of identified fraud cases across different fraud categories. Within our identified fraud cases, 44 campaigns are marked as *detected fraud* (19 "Pre-empted" and 25 "Attempted"), and 149 are marked as *suspected fraud* (5 "Rewards Changed" and 144 "Rewards Not Delivered").[3] Our identified fraudulent campaigns (within the 2010-2015 sample period) seem low in comparison to the total number of projects on Kickstarter. This raises the question of whether we are only observing the tip of the iceberg, or whether perhaps fraud in crowdfunding is difficult to confirm.

Following the discussion by Hainz (2018), we find there are multiple major reasons why fraud in crowdfunding is less observed. Hainz (2018) underscores that 1) the efficiency of the crowd in detecting fraudulent campaigns is relatively high (most backers are frequent backers who have gained experience from previous campaigns); 2) the effectiveness of platforms such as Kickstarter at filtering out fraudulent projects before they are posted is also relatively high; 3) non-reporting of fraudulent campaigns is highly likely, especially when the campaign is unsuccessful and no money has changed hands, in which case neither backers nor platform providers have a high incentive to report it; and 4) backers of successful but fraudulent campaigns may not bother to report fraud if they contributed only a small amount.

---

[1] We use 2010-2015 as the sample in order to ensure sufficient time (until 2018) to identify "suspected fraud" campaigns, especially in the case of rewards not delivered.

[2] In unreported tests, we examined the differences in means across all independent variables used in "determinant of fraud" analyses between fraudulent campaigns identified via *Kickscammed* vs. those identified via *News Search*. Result reveal no evidence of substantial difference in means across the two groups.

[3] The chronological sequence of the initiation date, campaign categories, and raised volumes in USD of fraudulent campaigns are shown in Panel C of Table 2.2. Fraud campaigns are most common in the "Technology" category (56 cases), and have also raised the largest amount within the "Technology" category (more than $11 million). Fraud campaigns by country for each respective year are shown in Panel D of Table 2.2. In our sample, fraud cases occurred most frequently in campaigns launched by creators in the U.S. (171 cases); the U.K. (8 cases); and Canada (7 cases); followed by creators in Israel (2); and Australia, China, Germany, Hong Kong, and Spain (1 each).

—Please insert Table 2.2 about here—

2.4.2. Determinants of Fraud

In order to identify a non-fraud control group with similar characteristics, we apply a propensity score matching (PSM) algorithm. We match our fraudulent campaigns only on campaign-related *demographic* characteristics (year, country, campaign category) and goal amount to ensure we do not select for other factors that could potentially explain fraudulent behavior[1].

We implement the nearest neighbor one-to-one fraud, and the non-fraud matching without the replacement option, to ensure the random component of the sample and to construct our sample for the main analyses. As a robustness check, we also provide results based on one-to-one (with replacement option) and one-to-two matches (with and without replacement options). We consider 386 crowdfunding campaigns (193 one-to-one pairs of matched fraud and non-fraud campaigns) in our main analysis. We also checked the campaign web pages of all the non-fraud matches to ensure that none were suspected of engaging in fraudulent behavior. We hand-collected information from Kickstarter on nineteen explanatory campaign variables, calculated based on the information from the campaign's web page, or from social media web pages associated with the campaign/creator.

2.4.3. Platform-wide Consequences of Fraud

In order to study platform-wide consequences of breaches of trust, we use an event study like setting to demonstrate whether late suspensions of campaigns by Kickstarter, which we classify based on four criteria as public large scams, have a negative effect on the success of other crowdfunding campaigns launched at about the same time. One challenge is to identify the "announcement date" that the fraud became visible to the crowdfunding community (i.e., potential backers). In order to determine the "announcement dates," we use Kickstarter's suspension dates for large successful campaigns associated with misconduct. It is noteworthy that there is no legal proof that these suspended cases were outright fraud. If Kickstarter's "Trust & Safety" team uncovers evidence that a campaign is in violation of Kickstarter's rules, the campaign is, according to Kickstarter's

[1] In an unreported test, we checked the quality of our PSM algorithm for our main analysis by using logit estimates for the probability of a campaign being fraudulent. We find that all the variables (*Goal Amount*, *Country*, *Year*, and *Category*) included in the PSM are well balanced between fraud and non-fraud campaigns, as we find no statistically significant differences between them. Consequently, our results are not driven by differences in goal amount, country, category, or year of campaign launch.

procedures, suspended.[1] We scraped data on all suspended campaigns using the "Explore" function of Kickstarter, which resulted in 1,760 suspended campaigns with suspension dates between January 1, 2010, and September 30, 2018.[2] Table 2.3 provides an overview of the number of suspended campaigns within each main category for the respective year (panel A) and the pledged dollar volumes (panel B). This is the population of suspensions we use to determine the most severe and visible scam campaigns that attracted many backers, as well as their "announcement dates", as we describe below.

The first challenge is to identify "late" suspensions, because the act of suspending a campaign by Kickstarter, if done in early stages of funding period and before the project successfully raises large amounts of funds, could be considered a positive signal to the crowd—a sign of related-party enforcement—and should therefore not have a negative effect on the willingness of backers to provide funding to projects on the platform or the market as a whole. The second challenge is to ensure that such an announcement was visible to a large community of potential crowdfunding backers. We follow a two-step procedure to identify the suspended campaigns (ensuring late suspension and visibility) that have arguably had the highest negative platform-wide consequences, and can be regarded as large, public scam campaigns.

Late suspension criteria: We require, as a first criterion, that at least 20% of the allegedly fraudulent campaign's duration has passed. As a second criterion, we require that less than one week remains until the campaign ends. Both criteria aim to ensure that the suspension was perceived as "late" in the crowdfunding community, and could in fact impact the funding success of other non-fraudulent campaigns. The first criterion reduced the total number of 1,760 suspended campaigns by 859, and the second by 689, leaving us with 212 (see Table 2.3, panel C).

Visibility criteria: Unfortunately, there is no direct measure of campaign visibility available, but we argue that it correlates highly with the number of backers in a campaign, pre-suspension. The third criterion (suspended campaign attracted at least 1,000 backers) is important, because 580 campaigns were suspended before a single backer contributed. If campaigns were suspended by Kickstarter before, e.g., anyone could contribute, backers might believe related-party enforcement has worked, and we should not expect any negative impact on platform-wide funding activities. We

---

[1] See https://help.kickstarter.com/hc/en-us/articles/115005139813-Why-would-a-project-be-suspended.

[2] In order to ensure sufficient time for the last potential suspension to affect campaigns posted at around the same time period, we set this date as three months before our last funded/failed campaign has ended (i.e., December 31, 2018), considering that maximum campaign length is ninety days.

use another proxy for campaign visibility, namely, pledged amount before the campaign was suspended. Therefore, we require, as a fourth criterion, that at least USD $10,000 was contributed to the campaign before suspension. The criterion for the number of backers reduced the number of suspended campaigns by another 198, while the contribution requirement did not result in any further exclusions (see again Table 2.3, panel C). To summarize, based on the four criteria, we identified fourteen suspended campaigns that may have had a sizable negative platform-wide effect (see Table 2.3, panel D).[1]

—Please insert Table 2.3 about here—

We then collected comprehensive data from the Kickstarter website for all campaigns with a goal amount of at least USD $100 (excluding very small donation-like campaigns), which were launched on or after January 1, 2010, ended on or before December 31, 2018, and were either successful/funded (i.e., reached goal amount) or unsuccessful/failed (i.e., pledged amount was less than the campaign's goal amount).[2] Our scraping procedure identified 271,971 unique campaigns within 15 main categories of Kickstarter. Table 2.4 provides an overview of the Kickstarter sample, showing the number of launched campaigns for each year within the main categories (panel A), their respective success rates (panel B), and the summary statistics (all values are converted to USD using static USD rate that is used by Kickstarter to show the amounts in local currencies), as well as the correlation matrix for all variables considered in the analyses of platform-wide consequences of fraud (panel C).

—Please insert Table 2.4 about here—

[1] Note further that the thresholds we use for the four applied criteria did not have a strong effect on the fourteen identified cases. Relaxing the thresholds, within certain margins, would still result in the same fourteen suspended campaigns. For example, changing the first criterion to "at least 50% of the campaign duration has passed," and changing the second criterion to "the campaign was suspended within two weeks of its scheduled deadline," and keeping the visibility criteria the same results in the same fourteen cases.

[2] We do not include any "cancelled" or "suspended" projects in the main sample because their success/failure do not depend on backers' decisions.

2.5. Methods

First, we specify a baseline regression model for the determinants of fraud analyses, using three blocks of characteristics: *creator's characteristics/background*, *social media affinity*, and *campaign characteristics* (*campaign funding and reward structure*, as well as *campaign description details*). For the baseline regression, we apply a logistic regression model to examine the determinants of our dependent variable *Fraud*, which equals 1 if the campaign is in our fraud sample, and 0 otherwise.

The non-fraud campaigns are based on a PSM approach based on available *demographic* variables. In this way, we can ensure that our control sample of non-fraud campaigns is not affected in a systematically different way by national regulations, culture, project category, project size, or time period in which they were seeking crowdfunding (Aggarwal et al., 2016; Attig et al., 2016; El Ghoul et al., 2016). The structure of our baseline logistic regression model is as follows:

$$\begin{aligned} Fraud\ (0/1)_i = {} & \alpha + \textstyle\sum_j \gamma_j \cdot Creator(s)'\ Characteristics/Background_j + \sum_k \xi_k \cdot \\ & Social\ Media\ Affinity_k + \textstyle\sum_l \varphi_l \cdot \\ & Campaign\ Funding\ and\ Reward\ Structure_l + \textstyle\sum_m \phi_m \cdot \\ & Campaign\ Description\ Details_m + \varepsilon_i. \end{aligned} \tag{1}$$

For each campaign i, the main explanatory variables are the j variables in the *creator(s)' characteristics/background* block (*Creator-Backed Projects* and *Creator-Created Projects*). The k variables in the *social media affinity* block include # *External Links* and *Facebook*. The l variables in the *campaign funding and reward structure* block include *Duration*, *Min. Pledge Amount*, and *No. of Pledge Categories*. Finally, the *campaign description details* block includes m variables, the *ARI*, and *Video Pitch*. We do not include year, country, or campaign category fixed effects because our samples have been initially matched and are balanced on those variables (see Bertoni et al., 2011; Grilli and Murtinu, 2014; Lee et al., 2015, on time variation and access to finance). However, we do use robust standard errors, which are one-way-clustered by campaign categories in all regressions, because residuals can be correlated within certain categories (Thompson, 2011).

We run several robustness checks, where we 1) use different nearest neighbor matching procedures (one-to-one and one-to-two, with and without replacement options) for our main analysis, and 2) operationalize our theoretical concepts with different variables and alternative proxies for *creator(s)' characteristics/background* (*Waiting Time (months)*, *Formal Name*, and

Natural Person), *social media affinity* (*Facebook_Page*, *Facebook_Personal*, *LinkedIn*, *Log (FB Connections)*), and project description readability indices (*CL*, *FKG*, and *GF*).

Our aim is not to develop a fraud prediction model to specify the forecasted probability of a campaign being fraudulent for a given set of explanatory variables. This would be extremely difficult to achieve. King and Zeng (2001b) explain that in a case-control design, where the fraction of failure in the data is different from the population, the *estimated probabilities* (i.e., forecasts) are biased and need *prior correction*. King and Zeng (2001a) posit that, for logit models with unknown sampling probability, which is given in our set-up, the *constant term* is biased, however, the parameter estimates remain largely unbiased. Therefore, prior correction is only applied to the constant term. However, the calculation of the correction term, which is to be subtracted from the estimated constant term, requires knowledge about the underlying probability of fraud in the population. This probability is not known to us, because there may be false negatives in the population preventing us from calculating the correction. Thus, we are only interested in the coefficients of the independent variables that have been shown to be unaffected, and that are generalizable to the population (King and Zeng, 2001a, 2001b).

Second, we present the methodology related to our platform-wide consequences of fraud analyses. We require a goal amount of at least USD $100 to avoid micro campaigns. To determine whether dynamics are different for campaigns that are presumably more likely to be related to entrepreneurial activities, we require a goal amount of at least USD $10,000, and we repeat the analyses (see Mollick and Nanda, 2015, for a similar argument for setting this threshold). The structure of our logistic (and OLS) regression model is as follows:

$$\begin{aligned} Success_i = {} & \beta_0 + \beta_{1,a} \cdot Fraud\ Period_i + \beta_{1,b} \cdot Post\ Fraud_i + \beta_2 \cdot Duration_i + \beta_3 \cdot \\ & Waiting\ Time_i + \beta_4 \cdot Featured_i + \beta_5 \cdot Log\ Goal_i + \beta_6 \cdot Daily\ Activity_{i,a} + \\ & \phi_{a,b} + \varphi_{a,b} + \lambda_a + \theta_a + \xi_a + \varepsilon_i, \qquad (2) \end{aligned}$$

for each campaign *i*, *Success* represents the dummy variable *Funded* (Logistic), the variable *Log Pledged* (OLS), or the variable *Log Backers* (OLS). Our main variable of interest is 1) the dummy variable *Fraud Period*, which equals 1 if the campaign's start date is within 14 days ($\mp$14) of the late suspension announcement, and 0 otherwise, or (as an alternative proxy) 2) the dummy variable *Post Fraud*, which equals 1 if the campaign's start date is within the 14 days after the late suspension announcement, and 0 if ended within the 14 days before the announcement (campaigns with other start/end dates are omitted).

If our Hypothesis 4 is supported, we expect to find negative coefficients for $\beta_{1,a}$ and $\beta_{1,b}$ for all three success measures. We control for the three main variables (i.e., *Duration*, *Featured*, and *Log Goal*), which are also used in Mollick (2014) and have a significant influence on campaign success, plus *Waiting Time* to proxy for creator's experience on the platform. We also introduce a new control variable *Daily Activity* to proxy for the level of competition while the project is live.

Classifying a campaign as posted within fraud period is not as straightforward as for an ordinary event study. Campaign suspensions should not be treated as a "one-day" event, because, e.g., campaigns launched before the suspension date that have a deadline scheduled for after the suspension date are affected by the suspension, as are campaigns launched closely after the suspension date. We define a dummy variable "Fraud Period" for each of the 271,971 campaigns, which equals 1 if the campaign is launched within 14 days before/after any of the identified suspension dates.[1] We choose 14 days, because the majority of campaigns have a duration of about 30 days. We also change this definition from $\mp$7 to $\mp$29 days, instead of $\mp$14 days, for the robustness checks.

When using the classification *Fraud Period* to identify the campaigns most likely to be affected by a suspension announcement, we include a series of fixed effects: campaign category (ϕ), year (2010 to 2018) (φ), month of year (January to December) (λ), day of month (first day to last day of respective month) (θ), day of week (Monday to Sunday) (ξ) to capture dynamics in different categories, as well as any time effect that may influence crowdfunding (and platforms) in certain years, certain months within years, and certain days within months. We also include the variable *Daily Activity* (average daily number of projects that were "live" during the campaign's lifetime). This variable captures the effects on campaign success that are directly related to platform activity but have not been picked up by the series on fixed effects. This is highly important, because intuitively one expects that the competition intensity (measured, e.g., by the number of live

[1] For example, if Kickstarter suspends a campaign on March 15, 2015, the "Fraud Period" dummy takes the value of 1 for all campaigns (either funded or failed) that were launched between March 1, 2015, and March 29, 2015. If we have any overlap between two suspension dates, our logic remains the same. For example, if suspension 1 was on March 15, 2015, and suspension 2 on March 25, 2015, the "Fraud Period" takes the value of 1 for all campaigns that were launched between March 1, 2015, and April 8, 2015.

campaigns on the platform competing for funding) is inversely correlated with campaign success, and there is empirical evidence suggesting so (Chen, 2021).

For the alternative classification *Post Fraud*, we determine a direct *pre* vs. *post* fraud comparison in success levels of a subsample of projects that are posted around the identified dates. We also include a series of fixed effects in these analyses: campaign category (ϕ), and year (2010 to 2018) (φ) and use clustered robust standard errors based on campaign categories in all regressions. The alternative classification *Post Fraud* allows a more direct comparison, because it has fewer observations and substantially reduces the need to control for periodic fixed effects.

Finally, we check the robustness of our findings by examining the impact of identified signals of first-party enforcement on the success of crowdfunding campaigns in the 2010 to 2018 Kickstarter sample (controlling for main determinants of success). We aim to show the relevance of these signals in predicting success, especially when there exists a perceived signal of weak related-party enforcement in the market.

2.6. Empirical Results

We use two different samples to study 1) determinants of fraud (credible signals of first-party enforcement), and 2) platform-wide consequences of perceived weak related-party enforcement. We then check the robustness of our results by examining the impact of signals of first-party enforcement (as well as project quality) on project success, especially when related-party enforcement (platform scrutiny) is perceived to be weak.

For studying "determinants of fraud," and in the absence of a verdict, it is important to have a high level of certainty that the identified campaigns are fraudulent, or are at least perceived largely as such. This is why we do not include all campaigns reported on Kickscammed or media in our dataset. Instead, we check whether, e.g., the promised product was finally delivered or any communication was attempted, in order to distinguish "failed" from "fraudulent" projects. To study measurable platform-wide consequences, it is important to identify suspended campaigns that are relatively suspended much later than expected, large in size, have higher number of backers, and higher pledged amounts in order to ensure visibility that, e.g., other backers (besides those directly affected by campaign suspensions) react to a suspension announcement. Therefore, we conducted the filtering process, described previously, to identify those campaigns that presumably had the most damaging effects on the market.

2.6.1. Determinants of Fraud

We begin by discussing our results in a univariate setting, and then we focus on multivariate analyses in order to include multiple possible determinants of fraud simultaneously. Table 2.5 (Table 2.A1 in the Online Appendix) shows the descriptive statistics (correlation matrix) for explanatory variables used in determinant of fraud analyses.

—Please insert Table 2.5 about here—

Table 2.6 presents the results for a difference in means *t*-test about how fraud sample differ from non-fraud matched campaigns on our main explanatory variables. In line with our Hypothesis 1, Table 2.6 shows that, on average, fraudsters tend to have fewer backed projects (about five projects), and create fewer projects (about one project), and a shorter waiting period between the date they opened the account on Kickstarter and the date the project is launched (about three to four months) (Rows 1-3). This univariate comparison provides initial evidence in line with Hypothesis 1 that fraudsters are less likely to have engaged in prior crowdfunding activity.

In accordance with Hypothesis 2, we find that the number of external links is negatively related to fraud (Row 6). It seems that external links may serve a kind of certification role. Thus, the more external links are provided, the higher the reputational capital that can be lost in the case of a fraudulent campaign. We also find that fraudsters are less present or active on Facebook (66% of non-fraud campaigns link either a Facebook page or a personal Facebook account to the campaign web page, compared to only 50% of fraudulent campaigns) (Row 7). The results remain consistent if we examine personal Facebook accounts and Facebook pages separately (Rows 8-9).

In terms of *campaign characteristics*, and in accordance with Hypothesis 3, we found the following results. Campaign durations tend to be longer for fraudulent campaigns, but the difference, on average, is about two days (Row 12). One reason for the small variation in duration is Kickstarter generally recommends a duration of thirty days or less,[1] and most projects follow that advice. We note that fraudulent campaigns provide more pledge categories (Row 14), and that the descriptions of fraudulent campaigns are easier to read (e.g., Row 15). The descriptions can also be interpreted

[1] See: https://www.kickstarter.com/help/faq/creator+questions.

as less sophisticated, because most of the readability measures correspond to the number of years of formal education needed to understand the text upon the first reading. The rationale behind this finding is that fraudsters are either 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, however, find no differences between fraud and non-fraud campaigns' use of video pitches (Row 19). This may be because creators are well aware that video pitching can strongly impact the probability of successful fundraising, and is strongly encouraged by platforms. Previous research has documented a positive correlation between videos and funding success (Mollick, 2014).

—Please insert Table 2.6 about here—

We now turn to our baseline model, which uses multivariate regressions to evaluate the correlations among the three blocks of explanatory variables—*creator(s)' characteristics/ background*, *social media affinity*, and *campaign characteristics*—with fraud. Table 2.7 summarizes our results from multivariate logistic regressions for the determinants of fraud in Equation (1). We consider all the main explanatory variables simultaneously, as the Variance Inflation Factors (VIFs) do not indicate a multicollinearity problem given that mean VIF range from 1.10 to 1.12 and that all individual values are well below the critical value of 5 (see Kutner et al., 2005). Our main analysis is presented in Specification (1), for which the matched non-fraud campaigns are determined by using a one-to-one PSM nearest-neighbor matching method without replacement. For robustness checks, we also show the results with replacement (Specification (2)), and for a one-to-two PSM nearest-neighbor matching method with and without replacement options (Specifications (3) and (4)).

—Please insert Table 2.7 about here—

No. of Creator-Backed Projects is negatively correlated with fraud, and, while the coefficient remains stable throughout the different specifications, it is only statistically significant in Specification (3). We also find that *No. of Creator-Created Projects* is negatively related to fraud, and the coefficient is statistically significant throughout all specifications. This supports Hypothesis 1, that project creators who have higher level of prior crowdfunding activities are less likely to carry out fraudulent campaigns and also confirms that backing multiple projects is an easier to mimic signal for fraudsters compared to previously created projects.

As shown in Table 2.7, our main explanatory variables—*# External Links* and *Facebook*—in the s*ocial media affinity* block have a strongly negative relationship with fraud. Therefore, campaigns that have either a Facebook page or a personal Facebook account associated with the project are about 45% (= EXP (-0.606) - 1) less likely to be fraudulent when compared to their matches (significant at a 5% level—Specification (1)). 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 strong negative correlation with the probability of a campaign being fraudulent. Overall, the results support Hypothesis 2 that fraudsters tend to be less present on social media and provide fewer external links.

Furthermore, in accordance with Hypothesis 3, we find that many *campaign characteristics* are related to the probability of observing fraudulent behavior. In detail, fraudulent campaigns tend to *ex ante* choose longer durations for their funding periods (Hypothesis 3.A). This is also in accordance with the signaling argument that high-quality campaigns choose shorter campaign durations to signal their quality and their confidence in getting successfully funded. We find no statistical significance for the *Min. Pledge Amount* (Hypothesis 3.B). This may be due to the fact that most reward-based crowdfunding campaigns offer small amounts as minimum contributions for non-monetary payoffs, and the campaigns do not substantially differ on this variable. Our results also show that the number of pledge categories has a significant positive relationship with fraud. This provides further evidence in line with Hypothesis 3.C, that crowdfunding fraudsters are more likely to offer a larger number of rewards at differing levels. Finally, Table 2.7 shows that the project descriptions of fraudulent campaigns tend to have lower automated readability indexes (ARI). ARI is an approximate representation of the number of formal years of education needed to comprehend the text on a first reading. A one-level ARI increase from the average score of eleventh grade (U.S. grade level) needed to comprehend the text to twelfth grade decreases the probability of the campaign being in the fraudulent subsample by about 10.5% (= EXP (-0.116) - 1). This supports Hypothesis 3.D, that fraudsters may target a less educated crowd by using less sophisticated and easier to understand language, or that they are not making enough effort to fine-tune their campaign descriptions. We find no statistically significant effect of *Video Pitch* on fraud. This may be because more than 93% of our 386 cases use a video pitch to describe their projects.

We check the robustness of our "determinants of fraud" results by using alternative proxies or complementary explanatory variables in Table 2.8. In order to avoid multicollinearity problem, or inter-dependent definitions across variables, models including all variables simultaneously are not estimated. Therefore, we examine each variable separately, but we retain all the main explanatory variables from the other blocks as "controls."

—Please insert Table 2.8 about here—

First, within the *creator(s)' characteristics/background* block of explanatory variables, we test whether there is a relationship between providing a formal name or being a natural person, and the likelihood of a fraudulent campaign. We find no statistically significant relationship between a natural person profile or a formal profile name and fraudulent campaigns. This is attributable to the fact that, on Kickstarter, for example, project creators must verify their identities through an automated process, and this information appears on their profiles (although not necessarily as their "profile name"), regardless of whether they use formal profile names[1]. However, we find that, similar to backing and creating crowdfunding campaigns, non-fraud sample creators have, on average, been members of the platform for longer periods of time.

We also test for the influence of 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 the 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. Despite finding a negative relationship between *Log (FB Connections)* and the probability of observing fraud, we found no statistically significant separate impact for the number of Facebook friends or the number of Facebook likes on fraudulent activity. One possible explanation for this is that fraud campaigns are using fake profiles to increase their number of "friends" or "likes" in order to mislead potential backers.

Furthermore, within the *campaign description details*, we used the readability index ARI, and identify a significantly negative relationship between ARI and fraud. That is, the probability that the campaign is in our fraudulent sample is higher when the project description is easier to understand. We further check the robustness of our results by using three alternative measures of text readability (see Table 2.8, panel C, Rows 5-7). As Table 2.8, panel C, shows, the Coleman-Liau index (*CL*), the

[1] 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 web page. 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.

Gunning Fog index (*GF*), and the Flesch-Kincaid Grade level index (*FKG*) all exhibit significantly negative correlations with fraudulent activity, which further validate our inferences.

To summarize, we find that our results remain robust to using alternative proxies for prior crowdfunding activity, social media affinity, and readability indices.

2.6.2. Platform-wide Consequences of Fraud

In Table 2.9, we present the results of multivariate logistic and OLS regressions—for our different measures of success—from Equation (2) to test for platform-wide consequences of suspended large, public scam campaigns. In both panels A, and B, of Table 2.9, Specifications (1)-(3) include Kickstarter campaigns with a goal amount of at least USD \$100, and Specifications (4)-(6) show the results for campaigns with a goal amount of at least USD \$10,000. We analyze the determinants of success measured by *Funded* (logistic regression; coefficients are the logs of the odds ratios), *Log Pledged* (OLS regressions), and *Log Backers* (OLS regressions). Campaigns being affected by suspension announcements are classified with the dummy variable, Fraud Period (panel A) or Post Fraud (panel B)[1].

—Please insert Table 2.9 about here—

In Table 2.9, panel A, we find that the coefficient of *Fraud Period* is negative and highly statistically significant for the entire sample, including all campaigns with a goal amount of more than \$100 (see Specifications (1)-(3)). In panel B, we follow a stricter approach, and compare campaigns that ended within 14 days before the announcement (*Post Fraud* = 0) with those begun within 14 days after the announcement (*Post Fraud* = 1). This allows for a more direct comparison, while requiring fewer observations. It also substantially reduces the need to control for the *Daily Activity* variable and the sets of "periodic fixed effects" used in panel A, because the pre- and post-fraud campaigns were launched around the same period and the concern for any potential procyclicality affecting the results is mitigated.

[1] Panel A includes main category, year, month of year (January to December), day of month (first day to last day of respective month), and day of week (Monday to Sunday) fixed effects. Moreover, in panel A, we control for a proxy of platform activity by calculating the average number of daily "live" campaigns during a project's lifetime, namely, *Daily Activity*. Panel B includes main category, and year fixed effects. The time fixed effects are all based on the campaign launch date.

Overall, the results in Table 2.9 provide strong empirical support for our Hypothesis 4, that the occurrence of fraudulent campaigns and their visibility to potential backers have far-reaching consequences for the success (success probability, number of backers, and funds raised) of concurrent crowdfunding campaigns that begin around suspension dates. As panel A, Specification (1) shows, if all else is equal, campaigns posted within fourteen days before/after one of our fourteen identified Kickstarter campaign suspensions are about 6.38% (= EXP (-0.066) - 1) less likely to be funded than the rest of the sample, i.e., campaigns launched on any other day during the observation period (see the coefficient for the dummy variable *Fraud Period*). Moreover, in Specifications (2) and (3), the pledged amounts (number of project backers) also decrease in an economically meaningful way. The results show that the predicted pledged amount in Specification (2) (predicted number of backers in Specification (3)) is approximately 9.6% (5.3%) lower for projects that are posted within fraud period, compared to the rest of the sample (see again the coefficient on *Fraud Period*).

For example, considering the average pledged amount of approximately USD $11,000,[1] campaigns posted within a fraud period lose on average about USD $1,000 on their pledged amounts. Note that the real effect is larger for raised amounts that are actually transferred, because we show that within-fraud period projects have a lower probability of success in general (i.e., reaching their goal amounts); and, in case of failure, the pledged amounts are not transferred to the project creators ("all-or-nothing" mechanism). The coefficient estimates of the control variables also show that *Duration, Daily Activity,* and Goal amount (*Log Goal*) negatively affect the success measures, while higher *Waiting Time* and being *Featured* by Kickstarter have a positive effect on campaign success measures.

In order to examine the sensitivity of our results to changes in the definition of the *Fraud Period* [*Post Fraud*] dummy (in the baseline, the number of days we consider is fourteen days around [*Pre/Post*] the suspension date), we extend the period day-wise to twenty-nine days. We then reduce it to seven days around the suspension date, and repeat the regressions from Table 2.9, plotting the coefficient for the variable of interest, *Fraud Period* [*Post Fraud*] in Figure 1, panel A [panel B]. Note that we expect to find the most negative coefficients when the platform-wide effects are most severe, i.e., when our sample of affected campaigns are in their first or last week of collecting funds.

[1] Note that the average pledged amount/number of backers reported in Table 2.2, Panel C, is the average of the log transformed variables.

Shortening or extending the observation period from the fourteen-day definition should result in higher coefficient estimates (i.e., lower absolute values of the *Fraud Period* and *Post-Fraud* negative coefficients). This is because an overly short period does not capture the effect in full, while an overly long period should dilute the platform-wide effect. This should result graphically in a V-shaped pattern.

From Figure 1, panel A [panel B], we observe that, in line with our reasoning, the effect is strongest for the thirteen days around the suspension date [thirteen days pre- and post-suspension announcement]. It fades slowly when we increase or decrease the number of days. The observed form reconciles with the V-shaped pattern. We interpret this as further support for the platform-wide negative consequences after the suspension of campaigns that slipped through Kickstarter's initial screening, received a certain level of attention by backers with substantial funding amounts, and were cancelled last minute.

—Please insert Figure 2.1 about here—

We find strong evidence for Hypothesis 4, that large public suspensions by Kickstarter (as identified by our filter criteria described above) have noticeably damaging effects on other funding activities, which can potentially hamper entrepreneurship and negatively affect the economy, employment, and innovation more generally. This also raises interesting policy implications, namely, that platforms' efforts to mitigate fraud should be focused more strongly on pre-screening mechanisms than on later suspension of projects.

Finally, in Table 2.10, we aim to relate the two parts of the empirical analyses, and examine the effect of identified signals of first-party enforcement (associated with lower probability of fraud) on success. In this table, we analyze the determinants of Success measured by Funded (logistic regression; coefficients are the logs of the odds ratios), Log Pledged (OLS regressions), and Log Backers (OLS regressions) in fraud vs. non-fraud period and examining the effect of main variables identified initially as determinants of fraud. In panel A, we include only the campaigns being affected by suspension announcements (Fraud Period =1) with goal amounts of at least $100 that were posted after January 1, 2010, and ended before December 31, 2018 (to a total of 41,229 affected campaigns). In panel B, we include only the campaigns not being affected by suspension announcements based on our definition (Fraud Period =0) to a total of 230,742 campaigns.

In summary, Table 2.10 provides further evidence for robustness of our results as we found that factors that are negatively (positively) associated with probability of observing fraudulent behavior, positively (negatively) predict campaign success, in both panels. Moreover, we show that coefficients are, in vast majority of cases (except the coefficient on *Facebook*), larger for the affected sample (Panel A) compared to the not affected sample (Panel B), possibly suggesting that signals of first-party enforcement play a more important role in determining backers' trust level, when there exists a signal of weak related-party enforcement.

—Please insert Table 2.10 about here—

2.7. Conclusion

This paper is the first to provide an in-depth examination of factors that are associated with a higher probability of observing fraudulent behavior in crowdfunding and analyze short-term consequences of breaches of trust in crowdfunding market. Furthermore, we provide evidence that legal enforcement by third parties such as the Federal Trade Commission or regional courts is very rare. Because the penalties are also very small, this puts the focus on the prescreening procedures and the liability of the crowdfunding platforms.

We contribute to the literature by providing a practical (albeit not legal) definition of fraud in the crowdfunding market, and identifying a comprehensive sample of campaigns associated with fraudulent behavior. We document campaign- and creator-related factors that tend to differ between fraudulent and a sample of non-fraudulent matched campaigns. We posit that these factors could be used by platforms to develop fraud-predicting models and fraud-preventing methods. We also provide the first empirical evidence of the effect of possible breaches of trust in the market, on crowdfunding success. In what follows, we discuss the implications of our findings for each specific party.

We first discuss the implications of our results for crowdfunding platforms. The evidence shows that not all scams have *ex ante* been detected. The lack of fraud detection might justify private or public regulation that requires platforms to offer standardized prescreening procedures and invest heavily in improving such detection models. However, such screenings could also become rapidly obsolete as fraudsters adapt and learn new methods to avoid detection. Therefore, and as an alternative way to increase trust in the market, platforms can design mechanism to hold project creators accountable after the funding is successful, by e.g., stopping the campaign once the funding goal is reached, and servicing any unmet demand in the aftermarket or retaining any funds raised in excess of the goal, as insurance for backers (see Belavina et al., 2020, for a theoretical discussion on these two options).

For policy makers, we believe regulators around the world are correct in their attempts to protect less sophisticated crowd members in this market. Until recently, most crowdfunding laws targeted specific branches of crowdfunding—primarily equity crowdfunding markets. Reward-based crowdfunding has not been regulated under specific laws except in a few jurisdictions, such as in Germany (Klöhn et al., 2016). Regulators may also require reward-based crowdfunding platforms to implement a minimum form of prescreening procedure that fulfills a more abstract catalog of quality requirements or require platforms not to allow a significant overcontribution to the campaigns (since the contribution amount is usually directly related to the fee that platforms charges, an intervention of regulatory bodies may be necessary in such cases). Once such dynamically adapting fraud detection models are implemented and there exists mechanisms to hold the campaign creators accountable, it may become safer to discuss the phenomenon of crowdfunding with old-fashioned securities lawyers *without* the need for a defibrillator!

For campaign creators, we emphasize the importance of signals of first-party enforcement, as well as project quality, in forming backer's trust level in their campaign, and the probability of successful funding. We show that the incidences of fraud and misconduct detection in the market, can be damaging to their campaign and they can mitigate this risk by reducing information asymmetry and provide difficult to mimic signals of project quality, and self confidence in the project. For crowdfunding backers, the identified factors can provide a basis for evaluating riskiness of the projects in terms of the probability of observing misconduct.

Our empirical analysis has some clear limitations. First, without doubt, we cannot rule out that some fraudulent campaigns go undetected. Undetected fraudulent campaigns (false negatives) will clearly result in an underestimation of the true probability of observing fraud, a challenge for all fraud prediction models, and remain as one of the limitations of this study. However, it appears to be very unlikely that large-scale fraudulent campaigns remain undetected on Kickstarter at least after some time. Small-scale fraud where few backers were engaged and were not reported, should be examined independently, given that their dynamics most likely differ from what we investigate in this article based on the defined criteria. Second, and more importantly, we cannot legally prove that any of the investigated campaigns were outright fraud. Whether campaign creators misappropriate funds, or develop low quality products because they put little effort in the development process, cannot be empirically tested in our context. Also, whether a judge would ultimately consider the "fraudulent" creators in our sample as being simply incapable of developing the project, or incompetent, cannot be clarified with certainty. Therefore, we use the words "fraud", "misconduct", and "fraudulent behavior" almost interchangeably throughout the study, and have aimed at being as strict as possible in defining the criteria for including a campaign in the fraud sample, while acknowledging the fact that no legal proof is involved.

Our study opens avenues for future research on developing and integrating new fraud detection models for reward-based crowdfunding, as well as other forms of crowdfunding (e.g., equity crowdfunding). In unreported tests, we examined whether concurrent projects in the same category where fraud occurred experienced more severe consequences. Our results reveal no evidence of statistically significant difference across different categories in response to fraud announcement in our sample. This might suggest that the borders between categories are somehow blurred in crowdfunding context (compared to e.g., publicly listed firms) and backers do not appear to differentiate between categories in response to visible suspensions, in a substantial way. This is, however, an interesting avenue for future research to shed light on backers' different reactions (in different categories) to fraud (or any other possible shock) in the crowdfunding market.

We posit that, once equity crowdfunding emerges more fully in the U.S., we will observe completely different twists in fraud. 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 to perhaps require different and more sophisticated detection mechanisms. Note that, under a reward-based model, fraud generally occurs because founders do not develop promised products and misuse the funds. 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 predictions will ultimately emerge, however, should be investigated empirically once the new market develops.

Acknowledgments

We thank Greg Shailer (field editor of *Journal of Business Ethics*) and two anonymous reviewers for many helpful comment and suggestions. We also thank Eliot Abrams, Ali Akyol, Yan Alperovych, Fabio Bertoni, Harjeet S. Bhabra, Martin Boyer, Steven Bradford, Shantanu Dutta, Philipp Geiler, Alexander Groh, Sofia Johan, Jonathan M. Karpoff, Yuri Khoroshilov, Maher Kooli, Iwan Meier, Fabio Moneta, Miwako Nitani, Juliane Proelss, Anita Quas, Rahul Ravi, Armin Schwienbacher, Silvio Vismara, Thomas Walker, Haoyong Zhou, and Tingyu Zhou, as well as seminar participants at the John Molson School of Business, Telfer School of Management, and HEC Montreal for many helpful comments. We are also grateful for the comments and suggestions gathered during the following conferences: Entrepreneurship, the Internet, and Fraud: Managerial and Policy Implications (Montreal, Canada); Munich Summer Institute 2017 (Munich, Germany); Corporate Governance Implications of New Methods of Entrepreneurial Firm Formation Workshop (Bergamo, Italy); 2nd Entrepreneurial Finance Conference (Ghent, Belgium); 32nd Annual Congress of the European Economic Association (Lisbon, Portugal); Financial Management Association Annual Meeting 2017 (Boston, USA); Crowdfunding Workshop at EMLYON Business School's Research Centre for Entrepreneurial Finance (Lyon, France); Annual Meeting of the Verein für Socialpolitik 2017 (Vienna, Austria); 36th International Conference of the French Finance Association (Quebec City, Canada); and the 5th Crowdinvesting Symposium (Berlin, Germany). This project has been supported by the Social Sciences & Humanities Research Council of Canada under the grant number 435-2015-1495.

References

Agrawal, A., Catalini, C., & Goldfarb, A. (2015). Crowdfunding: Geography, social networks, and the timing of investment decisions. *Journal of Economics and Management Strategy, 24*(2), 253–274.

Aggarwal, R., Faccio, M., Guedhami, O., & Kwok, C. (2016). Culture and finance: An introduction. *Journal of Corporate Finance, 100*(41), 466–474.

Ahlers, G. K., Cumming, D. J., Guenther, C., & Schweizer, D. (2015). Signaling in equity crowdfunding. *Entrepreneurship Theory and Practice, 39*(4), 955–980.

André, K., Bureau, S., Gautier, A., & Rubel, O. (2017). Beyond the opposition between altruism and self-interest: Reciprocal giving in reward-based crowdfunding. *Journal of Business Ethics, 146*(2), 313–332.

Assenova, V., Best, J., Cagney, M., Ellenoff, D., Karas, K., Moon, J., Neiss, S., Suber, R., & Sorenson, O. (2016). The present and future of crowdfunding. *California Management Review, 58*(2), 125–135.

Attig, N., Boubakri, N., El Ghoul, S., & Guedhami, O. (2016). Firm internationalization and corporate social responsibility. *Journal of Business Ethics, 134*(2), 171–197.

Belavina, E., Marinesi, S., & Tsoukalas, G. (2020). Rethinking crowdfunding platform design: Mechanisms to deter misconduct and improve efficiency. *Management Science, 66*(11), 4980–4997.

Belleflamme, P., Lambert, T., & Schwienbacher, A. (2013). Individual crowdfunding practices. *Venture Capital: An International Journal of Entrepreneurial Finance, 15*(4), 313–333.

Belleflamme, P., Lambert, T., & Schwienbacher, A. (2014). Crowdfunding: Tapping the right crowd. *Journal of Business Venturing, 29*(5), 585–609.

Berns, J. P., Figueroa-Armijos, M., da Motta Veiga, S. P., & Dunne, T. C. (2020). Dynamics of lending-based prosocial crowdfunding: Using a social responsibility lens. *Journal of Business Ethics, 161*(1), 169–185.

Bertoni, F., Colombo, M. G., & Grilli, L. (2011). Venture capital financing and the growth of high-tech start-ups: Disentangling treatment from selection effects. *Research Policy, 40*(7), 1028–1043.

Bradford, S. (2012). Crowdfunding and the federal securities laws. *Columbia Business Law Review, 1*, 1–150.

Brockman, P., El Ghoul, S., Guedhami, O., & Zheng, Y. (2020). Does social trust affect international contracting? Evidence from foreign bond covenants. *Journal of International Business Studies*, 1–34.

Chen, L. (2021). Investigating the impact of competition and incentive design on performance of crowdfunding projects: A case of independent movies. *Journal of Theoretical and Applied Electronic Commerce Research, 16*(4), 791–810.

Colombo, M. G., Franzoni, C., & Rossi-Lamastra, C. (2015). Internal social capital and the attraction of early contributions in crowdfunding projects. *Entrepreneurship Theory and Practice, 39*(1), 75–100.

Cumming, D. J., Leboeuf, G., & Schwienbacher, A. (2019a). Crowdfunding models: Keep-it-all vs. all-or-nothing. *Financial Management, 49*(2), 331–360.

Cumming, D. J., Meoli, M., & Vismara, S. (2019b). Does equity crowdfunding democratize entrepreneurial finance? *Small Business Economics*, *56*, 533–552.

Davidson III, W. N., & Worrel, D. L. (1988). The impact of announcements of corporate illegalities on shareholder returns. *Academy of Management Journal, 31*(1), 195–200.

Defazio, D., Franzoni, C., & Rossi-Lamastra, C. (2020). How pro-social framing affects the success of crowdfunding projects: The role of emphasis and information crowdedness. *Journal of Business Ethics*, *171*, 357–378.

Diamond, D. (1989). Reputation acquisition in debt markets. *Journal of Political Economy, 97*(4), 828–862.

Dupont, Q., & Karpoff, J. M. (2019). The trust triangle: Laws, reputation, and culture in Empirical finance research. *Journal of Business Ethics, 163,* 217–238.

Ehrlich, I. (1973). Participation in illegitimate activities: A theoretical and empirical investigation. *Journal of Political Economy, 81*(3), 521–565.

El Ghoul, S., Guedhami, O., Kwok, C.., & Shao, L. (2016). National culture and profit reinvestment: Evidence from SMEs. *Financial Management, 45*(1), 37–65.

El Ghoul, S., Guedhami, O., Nash, R., & Patel, A. (2019). New evidence on the role of the media in corporate social responsibility. *Journal of Business Ethics, 154*(4), 1051–1079.

Fischel, D. R. (1982). Use of modern finance theory in securities fraud cases involving actively traded securities. *Business Lawyer, 38*(1), 1–20.

Gino, F., Ayal, S., & Ariely, D. (2009). Contagion and differentiation in unethical behavior: The effect of one bad apple on the barrel. *Psychological Science, 20*(3), 393–398.

Grilli, L., & Murtinu, S. (2014). Government, venture capital and the growth of European high-tech entrepreneurial firms. *Research Policy, 43*(9), 1523–1543.

Hain, D., Johan, S., & Wang, D. (2016). Determinants of cross-border venture capital investments in emerging and developed economies: The effects of relational and institutional trust. *Journal of Business Ethics, 138*(4), 743–764.

Hainz, C. (2018). Fraudulent behavior by entrepreneurs and borrowers. In Cumming, D. J., & Hornuf, L. (Eds.), *The Economics of Crowdfunding* (pp. 79–99). London: Palgrave Macmillan.

Hornuf, L., Schmitt, M., & Stenzhorn, E. (2018). Equity crowdfunding in Germany and the UK: Follow-up funding and firm failure. *Corporate Governance: An International Review, 26*, 331–354.

Hornuf, L., & Schwienbacher, A. (2017). Should securities regulation promote equity crowdfunding? *Small Business Economics, 49*(3), 579–593.

Hornuf, L., & Schwienbacher, A. (2018). Market mechanisms and funding dynamics in equity crowdfunding. *Journal of Corporate Finance, 50*, 556–574.

Iyer, R., Khwaja, A. I., Luttmer, E. F., & Shue, K. (2015). Screening peers softly: Inferring the quality of small borrowers. *Management Science, 62*(6), 1554–1577.

Jiang, T. (2013). Cheating in mind games: The subtlety of rules matters. *Journal of Economic Behaviour and Organization, 93*, 328–336.

Karpoff, J. M., Lee, D. S., & Martin, G. S. (2008). The consequences to managers for cooking the books. *Journal of Financial Economics, 88*(88), 193–215.

King, G., & Zeng, L. (2001a). Logistic regression in rare events data. *Political Analysis, 9*(2), 137–163.

King, G., & Zeng, L. (2001b). Improving forecasts of state failure. *World Politics, 53*(4), 623–658.

Klöhn, L., Hornuf, L., & Schilling, T. (2016). The regulation of crowdfunding in the German small investor protection act: Content, consequences, critique, suggestions. *European Company Law, 13*, 56–66.

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models* (5th edition). New York: McGraw-Hill.

Lacker, J. M., & Weinberg, J. A. (1989). Optimal contracts under costly state falsification. *Journal of Political Economy, 97*(6), 1345–1363.

Lee, N., Sameen, H., & Cowling, M. (2015). Access to finance for innovative SMEs since the financial crisis. *Research Policy, 44*(2), 370–380.

Liang, T. P., Wu, S. P. J., & Huang, C. C. (2019). Why funders invest in crowdfunding projects: Role of trust from the dual-process perspective. *Information & Management, 56*(1), 70–84.

Lin, M., Prabhala, N. R., & Viswanathan, S. (2013). Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending, *Management Science, 59*(1), 17–35.

Mann, H., Garcia-Rada, X., Hornuf, L., Tafurt, J., & Ariely, D. (2016). Cut from the same cloth: Similarly dishonest individuals across countries. *Journal of Cross-Cultural Psychology, 47*(6), 858–874.

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self- concept maintenance. *Journal of Marketing Research, 45*(6), 633–644.

Mollick, E. (2014). The dynamics of crowdfunding: An exploratory study. *Journal of Business Venturing, 29*(1), 1–16.

Mollick, E., & Nanda, R. (2015). Wisdom or madness? Comparing crowds with expert evaluation in funding the arts. *Management Science, 62*(6), 1533–1553.

Parhankangas, A., & Renko, M. (2017). Linguistic style and crowdfunding success among social and commercial entrepreneurs. *Journal of Business Venturing, 32*(2), 215–236.

Perez, B., Machado, S. R., Andrews, J., & Kourtellis, N. (2020). I Call BS: Fraud detection in crowdfunding campaigns. *arXiv preprint*, arXiv:2006.16849.

Perino, M. A. (1998). Fraud and federalism: Preempting private state securities fraud causes of action. *Stanford Law Review, 50*, 273–338.

Rezaee, Z. (2005). Causes, consequences, and deterrence of financial statement fraud. *Critical Perspectives on Accounting, 16*(3), 277–298.

Shailer, G. (1999). Classificatory loan pricing as an incentive for signalling by closely held firms. *New England Journal of Entrepreneurship, 2*(1), 1–6.

Shalvi, S., Gino, F., Barkan, R., & Ayal, S. (2015). Self-serving justifications doing wrong and feeling moral. *Current Directions in Psychological Science, 24*(2), 125–130.

Siering, M., Koch, J. A., & Deokar, A. V. (2016). Detecting fraudulent behavior on crowdfunding platforms: The role of linguistic and content-based cues in static and dynamic contexts. *Journal of Management Information Systems, 33*(2), 421–455.

Signori, A., & Vismara, S. (2018). Does success bring success? The post-offering lives of equity-crowdfunded firms. *Journal of Corporate Finance, 50*, 575–591.

Simmonds, A. R., Sagat, K. A., & Ronen, J. (1992). Dealing with anomalies, confusion and contradiction in fraud on the market securities class actions. *Kentucky Law Journal, 81*, 123–186.

Sorenson, O., Assenova, V., Li, G. C., Boada, J., & Fleming, L. (2016). Expand innovation finance via crowdfunding: Crowdfunding attracts venture capital to new regions. *Science, 354*(6319), 1526–1528.

Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics, 87*(3), 355–374.

Thompson, S. B. (2011). Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics, 99*(1), 1–10.

Townsend, R. (1979). Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory, 21*(2), 265–293.

Vismara, S. (2016). Equity retention and social network theory in equity crowdfunding. *Small Business Economics, 46*(4), 579–590.

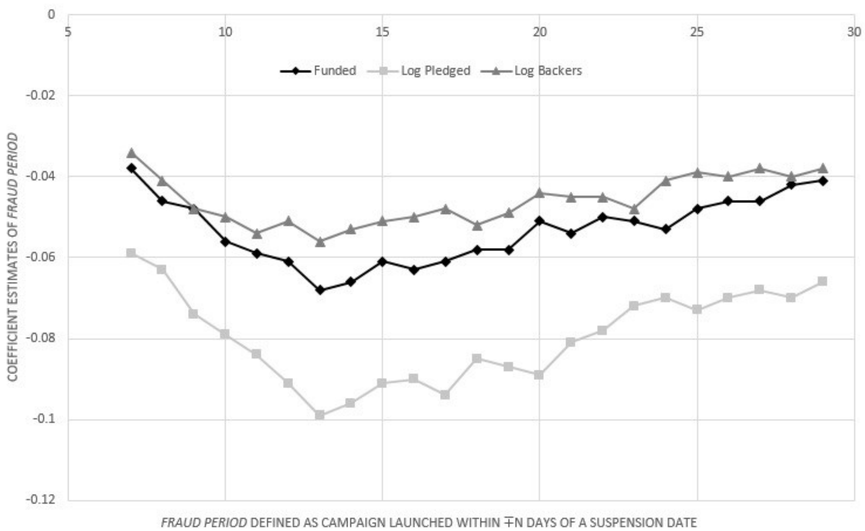
Yi, S., Kung, M., Steyvers, M., Lee, M. D., & Dry, M. J. (2012). The wisdom of the crowd in combinatorial problems. *Cognitive Science, 36*(3), 452–470.

Zott, C., & Huy, Q. N. (2007). How entrepreneurs use symbolic management to acquire resources. *Administrative Science Quarterly, 52*(1), 70–105.

Figure 2.1: Sensitivity Analysis

This figure in panel A [panel B], shows the estimated *Fraud Period* [*Post Fraud*] dummy variable regression coefficients, in specifications (1)-(3) of Table 2.9, panel A [Table 2.9, panel B] using alternative classifications schemes for campaigns being affected by suspension announcements (*Fraud Period* and *Post Fraud*) for the success measures as dependent variables (*Funded* (logit), *Log Pledged* (OLS), and *Log Backers* (OLS)). N (ranging from 7 to 29 days) determines the number of ∓ [+] days considered in the *Fraud Period* [*Post Fraud*] dummy variable definition (our default in all of the platform-wide consequences of fraud analyses is ∓ [+] 14 days). The sample includes all Kickstarter campaigns with goal amounts of at least $100 that were posted after January 1, 2010, and ended before December 31, 2018. The total sample in panel A includes 271,971 campaigns, and all control variables from Table 2.9, panel A (as well as the fixed effects variables) are available for all observations. The total sample in panel B depends on the "N" considered, and ranges from 20,674 (N=7) to 69,030 (N=29) since in Table 2.9, panel B regressions, we only keep projects that were either started within 14 days after fraud (*Post Fraud*=1), or ended on/within 14 days before the fraud (*Post Fraud*=0). All control variables from Table 2.9, panel B are available for all observations. All non-dummy variables are winsorized at the 1% level on both sides (see Table 2.1, panel B, for variable descriptions and calculation methods). Robust standard errors are one-way-clustered by campaign category. The reported coefficients for the "*Funded*" specification are the logs of the odds ratios. All calculated coefficients are all statistically significant at least at the 5% level.

Panel A



Panel B

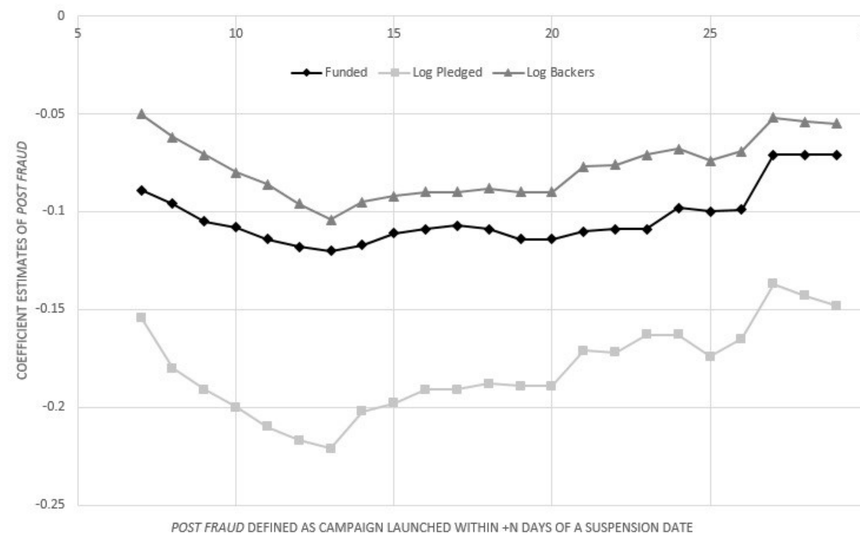


Table 2.1: Variable Definitions

This table gives a detailed description of the data-gathering process and calculation methods for all variables. Panel A includes only variables that are used in the determinants of fraud analyses; Panel B includes variables used in the platform-wide consequences of fraud analyses.

Panel A ("Determinants of Fraud" Analyses)

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.
	Creator(s)' Characteristics/Background
Creator-Backed Projects	Total number of projects backed by the creator since joining the platform.
Creator-Created Projects	Total number of projects created by the creator since joining the platform.
Waiting Time (months)	Number of months between the day that creator joined the platform (Kickstarter) and start date of the campaign.
Formal Name	Dummy variable that equals 1 if the project creator uses a formal profile name (i.e., [first name] [last name]), and 0 otherwise.
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.
	Social Media Affinity
# External Links	Total number of external links provided on campaign's page.
Facebook	Dummy variable that equals 1 if a personal Facebook / Facebook Page is linked to the project's web page on Kickstarter, 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.
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.
LinkedIn	Dummy variable that equals 1 if a link to a LinkedIn page of the creator(s) is provided, and 0 otherwise.
Log (FB Connections)	Natural logarithm of "the total friends of personal Facebook page linked to the project's web page on Kickstarter, plus the total likes of Facebook page associated with the campaign."
	Campaign Funding and Reward Structure
Duration	Number of days between the campaign's end date and start date.
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 category).
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.

Table 2.1: Variable Definitions—*continued*

Panel A ("Consequences of Fraud" Analyses)-*continued*

	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 understand.
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 understand.
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 understand.
GF	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, and the lower the number, the easier the text is to understand.
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.

Table 2.1: Variable Definitions—*continued*

Panel B ("Consequences of Fraud" Analyses)

	Dependent Variables (Success)
Funded	Dummy variable that equals 1 if the project reached its goal amount, and 0 otherwise.
Log Pledged	Natural logarithm of the (project's pledged amount in USD (regardless of the project's success) + 1).
Log Backers	Natural logarithm of the (project's total number of backers (regardless of the project's success) + 1).
	Independent Variables
Fraud Period	Dummy variable that equals 1 if the campaign's launch date is within ∓ 14 days of the suspension date of any of the identified suspended fraudulent campaigns and did not end before the announcement date of the suspended campaign, and 0 otherwise.
Post Fraud	Dummy variable that equals 1 if the campaign's launch date is within $+14$ days of the suspension date of any of the identified suspended fraudulent campaigns (i.e., *Post-Fraud*), 0 if the campaign's end date is within -14 days of the suspension date (i.e., *Pre-Fraud*), and omitted otherwise.
	Control Variables
Duration	Number of days between the campaign's end date and start date.
Waiting Time	Number of days between the campaign's start date and the date creator joined Kickstarter (i.e., created an account).
Featured	Dummy variable that equals 1 if the project is featured as "Projects We Love" by Kickstarter, and 0 otherwise.
Log Goal	Natural logarithm of the project's goal amount in USD.
Daily Activity	Average daily number of projects that were "live" during campaign's lifetime, divided by 1000.

Table 2.2: Derivation of Fraudulent Campaigns' Sample ("Determinants of Fraud" Analyses)

This table shows the derivation of the fraud sample (panel A), fraud categories and campaign status (panel B), and campaign categories, as well as the number of campaigns (No.) and the amounts raised in USD (Vol.) for each respective year for the fraudulent campaigns (panel C). Panel D shows the distribution of fraudulent campaigns across different countries. Panel A presents the number of identified fraud cases using the Kickscammed website and the news search. We dropped seven campaigns (data not available) from the initial fraud sample, because 1) there was missing information on at least one matching criteria, or 2) the campaign web page was no longer available on Kickstarter. In Panel C, "Failed" is defined as the goal amount not being met by the end date of the campaign, and "Successful" is defined as the goal amount being achieved (and neither suspended nor cancelled). The amounts raised in currencies other than USD are converted into USD using Federal Reserve System average foreign exchange rates in the year the campaign was launched.

Panel A

Identified Via	#
Kickscammed	181
News Search	19
Total (Initial Cases)	200
- Data Not Available	7
Total	**193**

Panel B

Fraud Category	Status	#
Detected Fraud	Pre-empted	19
	Attempted	25
Suspected Fraud	Rewards Changed	5
	Rewards Not Delivered	144
	Total	**193**

Table 2.2: Derivation of Fraudulent Campaigns' Sample—*continued*

Panel C

Category	**2010**	**Vol.**	**2011**	**Vol.**	**2012**	**Vol.**	**2013**	**Vol.**	**2014**	**Vol.**	**2015**	**Vol.**	**No.**	**Total:**
Art			1	32,017			1	14,651					2	46,668
Comics			2	21,875			3	66,068					5	87,943
Crafts							1	13,359	2	31,115			3	44,474
Design	1	87,407	4	631,294	17	1,913,405	23	3,953,543	7	723,299	2	25,710	54	7,334,658
Fashion					1	94,279	2	25,648	2	114,318	1	10,371	6	244,616
Film and Video			2	95,348	4	331,594	3	139,837	2	277,056			11	843,835
Food							4	208,084	1	13,355	1	20,780	6	242,219
Games			3	212,928	15	755,384	14	327,620	12	599,399	1	13,796	45	1,909,127
Music									2	18,452			2	18,452
Photography							1	8,047					1	8,047
Publishing	1	28,701							1	380,747			2	409,448
Technology					6	682,179	17	5,102,461	18	3,197,970	15	2,361,927	56	11,344,537
Total	**2**	**116,108**	**12**	**993,462**	**43**	**3,776,841**	**69**	**9,859,319**	**47**	**5,355,710**	**20**	**2,432,583**	**193**	**22,534,023**

Total Amount Raised	**22,534,023**
- Failed	69,294
- Detected	2,810,455
"Successful" Fraudulent Campaigns **(Total Amount)**	**19,654,274**

Table 2.2: Derivation of Fraudulent Campaigns' Sample—*continued*

Panel D

Country	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**Total**
Australia						1	1
Canada		1	1	3	1	1	7
China					1		1
Germany						1	1
Hong Kong						1	1
Israel				1	1		2
Spain					1		1
United Kingdom				4	2	2	8
United States	2	11	42	61	41	14	171
Total	**2**	**12**	**43**	**69**	**47**	**20**	**193**

Table 2.3: Derivation of Suspended Campaigns' Sample ("Consequences of Fraud" Analyses)

This table shows the derivation of the suspended campaigns sample. Panel A shows the number of suspended campaigns within each main category for each respective year. The amounts raised by suspended campaigns in USD (Vol.) for each respective year is presented in panel B (">0" indicates a volume less than 10 USD). Panel C shows the four filter criteria to derive the 14 suspended campaigns to identify the suspension dates. Panel D provides the list of the 14 main suspension dates, along with the suspended campaign name, main category, goal amount, pledged amount and number of backers. The amounts raised in currencies other than USD are converted into USD using static USD rate that is used by Kickstarter to show the amounts in local currencies.

Panel A

Num.	Main Category	2010	2011	2012	2013	2014	2015	2016	2017	2018	Total (#)
1	Art	4	7		2	7	32	8	23	13	96
2	Comics					8	6	2	1	6	23
3	Crafts			1		10	35	6	13	5	70
4	Dance					6	6		1	2	15
5	Design		1	13	7	22	51	40	47	33	214
6	Fashion		4	4	5	21	42	15	23	19	133
7	Film & video	3	7	7	7	11	39	17	22	8	121
8	Food		7	1	3	17	60	23	25	16	152
9	Games	1	1	5	5	24	85	30	39	30	220
10	Journalism	1		1		9	23	7	5	5	51
11	Music	2	10	3	1	34	52	23	22	9	156
12	Photography	2	3	1	1	2	26	7	5	2	49
13	Publishing	1	3	4	1	5	22	7	19	3	65
14	Technology	3	2	9	6	48	99	72	92	46	377
15	Theater	1	2			4	9	1		1	18
	Total	**18**	**47**	**49**	**38**	**228**	**587**	**258**	**337**	**198**	**1,760**

Table 2.3: Derivation of Suspended Campaigns' Sample—*continued*

Panel B

Num.	Main Category	2010	2011	2012	2013	2014	2015	2016	2017	2018	Total (Vol. in USD 1,000)
1	Art	>0	0.34		10.50	0.55	14.17	2.21	25.25	2.46	55.46
2	Comics					5.19	1.74	0.92	>0	4.93	12.77
3	Crafts			3.52		0.73	2.23	4.70	7.34	1.47	19.99
4	Dance					0.33	3.84		>0	0.61	4.78
5	Design		>0	149.26	73.77	162.62	664.84	956.65	1,456.08	389.22	3,852.44
6	Fashion		0.06	33.39	44.65	135.64	107.11	62.51	70.75	12.90	467.00
7	Film & video	0.05	0.60	48.38	41.63	65.25	28.38	30.51	0.88	16.26	231.92
8	Food		0.19	0.98	122.28	10.46	5.38	269.48	102.50	40.65	551.91
9	Games	>0	0.07	20.34	107.80	114.68	57.99	11.54	76.89	173.78	563.09
10	Journalism	0.05		>0		0.18	1.77	0.30	0.17	0.25	2.72
11	Music	>0	0.10	21.37	5.74	1.60	5.49	3.30	6.59	8.87	53.05
12	Photography	>0	>0	>0	>0	>0	1.60	6.23	2.73	1.94	12.49
13	Publishing	>0	0.02	5.14	0.92	0.51	12.69	1.71	55.94	20.14	97.05
14	Technology	0.10	>0	235.67	83.31	1,708.16	5,211.05	1,266.12	1,283.04	759.35	10,546.80
15	Theater	0.03	>0			0.46	0.02	0.01		>0	0.52
	Grand Total	0.22	1.36	518.04	490.58	2,206.34	6,118.29	2,616.17	3,088.16	1,432.85	**16,472**

Panel C

	Inclusion Criteria	#	Sub-Total
	Suspended Campaigns Sample	1,760	-
1	more than 20% of the campaign duration passed	*- 859*	901
2	less than 1 week remaining to the scheduled deadline	*- 689*	212
3	Number of backers more than 1,000	*- 198*	14
4	Pledged amount larger than USD 10,000	*- 0*	14
	Final Number of Suspended Campaigns		**14**

Table 2.3: Derivation of Suspended Campaigns' Sample—*continued*

Panel D

Num.	Suspension Date	Name	Main Category	Goal (USD)	Pledged (USD)	# Backers
1	2013-06-13	KOBE RED - 100% JAPANESE BEER FED KOBE BEEF JERKY	Food	2,374	120,309	3,252
2	2014-07-22	Areal	Games	50,000	64,928	1,090
3	2015-08-05	TrackerPad - Sticky GPS tracker pads	Technology	155,194	80,651	1,209
4	2015-08-11	Firestarter Survival Bracelet / Carabiner Paracord Keychain	Technology	10,000	477,462	9,139
5	2015-10-12	The Skarp Laser Razor: 21st Century Shaving	Technology	160,000	4,005,112	20,632
6	2016-01-27	TESLA – self-rechargeable, electronic lighter	Technology	5,000	118,693	3,605
7	2016-10-19	λ Chair - The Advanced Art of Seating	Design	25,000	614,382	1,531
8	2016-10-25	iLDOCK - charge and listen to iPhone 7 at the same time	Technology	5,000	212,459	9,895
9	2017-12-19	GARY 2.0 : Earphones & Cables Automatic Organizer	Technology	6,537	33,026	1,650
10	2018-02-02	YT TOUCH | Fast Aerospace Aluminium Defrosting Tray	Design	10,000	212,632	4,496
11	2018-02-24	Most functional Duffel bag ever	Design	5,764	108,781	1,316
12	2018-05-09	Zōk | Restore Calmness and Serenity to the Mind and Body	Technology	10,500	56,673	1,812
13	2018-06-22	amplify | The Ultimate Wireless Headphone Amplifier with DAC	Design	33,000	98,460	1,220
14	2018-07-18	Overturn Rising Sands	Games	34,133	114,380	1,093

Table 2.4: An Overview of Kickstarter Sample (2010-2018)

This table shows the derivation of the Kickstarter sample. Panel A shows the number of launched campaigns (either successful/funded or unsuccessful/failed) within Kickstarter's main categories for each respective year. "Unsuccessful/Failed" is defined as the goal amount not being reached by the end date of the campaign, and "Successful/Funded" is defined as the goal amount being achieved (and neither suspended nor cancelled). Panel B, shows the percentage of successful campaigns within each main category for each respective year. Panel C provides descriptive statistics (mean, standard deviation, min, and max) and Pearson correlation coefficients for all variables considered in the platform-wide consequences of fraud analyses (*indicates statistical significance at least at 5% level). All non-dummy variables are winsorized at the 1% level on both sides. See Table 2.1, panel B, for variable description and calculation method.

Panel A

Num.	Main Category	2010	2011	2012	2013	2014	2015	2016	2017	2018	Total (#)
1	Art	486	1,656	2,645	2,761	3,842	4,449	3,109	3,429	3,291	25,668
2	Comics	68	217	511	677	1,128	1,650	1,640	1,794	1,798	9,483
3	Crafts	23	62	143	295	1,407	2,090	1,517	1,307	915	7,759
4	Dance	123	385	487	525	658	569	406	341	212	3,706
5	Design	63	173	494	972	1,989	3,271	3,906	4,684	3,779	19,331
6	Fashion	2	10	265	717	2,544	3,930	3,245	3,459	3,043	17,215
7	Film & video	1,020	2,768	3,975	4,954	6,177	6,380	4,535	3,659	2,799	36,267
8	Food	35	53	141	340	3,582	4,425	2,759	2,400	1,729	15,464
9	Games	101	332	1,311	2,111	3,649	4,979	4,777	5,339	5,126	27,725
10	Journalism	95	101	146	129	710	1,188	678	497	332	3,876
11	Music	1,159	3,242	5,503	5,277	5,391	5,864	3,880	3,350	2,432	36,098
12	Photography	37	81	112	237	1,554	1,635	1,068	832	554	6,110
13	Publishing	307	1,071	3,042	4,061	5,582	5,869	4,467	4,270	3,188	31,857
14	Technology	140	240	458	1,218	4,393	6,849	5,345	4,712	3,011	26,366
15	Theater	21	67	82	273	1,102	1,309	916	742	534	5,046
	Total	**3,680**	**10,458**	**19,315**	**24,547**	**43,708**	**54,457**	**42,248**	**40,815**	**32,743**	**271,971**

Table 2.4: An Overview of Kickstarter Sample (2010-2018)—*continued*

Panel B

Num.	Main Category	2010	2011	2012	2013	2014	2015	2016	2017	2018	Total (#)
1	Art	51%	55%	51%	50%	37%	36%	40%	48%	57%	45%
2	Comics	54%	54%	50%	57%	57%	58%	65%	70%	77%	64%
3	Crafts	70%	71%	80%	68%	24%	23%	25%	28%	32%	29%
4	Dance	81%	76%	75%	74%	62%	51%	64%	63%	62%	66%
5	Design	56%	53%	48%	45%	38%	39%	48%	50%	55%	47%
6	Fashion	50%	50%	87%	69%	33%	26%	27%	34%	42%	34%
7	Film & video	46%	45%	40%	49%	43%	36%	41%	42%	47%	42%
8	Food	60%	64%	62%	58%	22%	21%	24%	27%	30%	25%
9	Games	38%	31%	27%	34%	30%	34%	41%	52%	60%	43%
10	Journalism	45%	43%	34%	43%	21%	18%	20%	24%	30%	24%
11	Music	45%	56%	59%	60%	52%	41%	47%	49%	56%	52%
12	Photography	43%	48%	46%	45%	25%	28%	40%	39%	47%	34%
13	Publishing	61%	58%	49%	44%	33%	30%	37%	39%	49%	39%
14	Technology	41%	49%	56%	48%	23%	22%	23%	25%	27%	25%
15	Theater	90%	82%	77%	66%	59%	58%	64%	61%	60%	61%
	Total	**49%**	**53%**	**50%**	**51%**	**36%**	**32%**	**38%**	**42%**	**50%**	**41%**

Panel C

Num.	Variable	# Obs.	Mean	Std. Dev.	Min	Max	1	2	3	4	5	6	7	8	9
1	Funded	271,971	0.41	0.49	0.00	1.00	1								
2	Log Pledged	271,971	5.99	3.23	0.00	11.96	0.67*	1							
3	Log Backers	271,971	2.81	1.88	0.00	7.52	0.71*	0.93*	1						
4	Fraud Period	271,971	0.15	0.36	0.00	1.00	-0.03	-0.03	-0.02	1					
5	Duration	271,971	33.16	11.55	8.00	60.00	-0.14*	-0.06*	-0.07*	-0.02	1				
6	Waiting Time	271,971	42.03	90.94	0.00	598.00	0.03	0.14*	0.13*	0.01	0.03	1			
7	Featured	271,971	0.10	0.30	0.00	1.00	0.27*	0.33*	0.39*	-0.02	-0.03	0.06	1		
8	Log Goal	271,971	8.55	1.59	5.02	12.61	-0.23*	0.12*	0.1*	0	0.21*	0.13*	0.12	1	
9	Daily Activity	271,971	3.86	1.40	0.69	6.71	-0.14*	-0.16*	-0.14*	0.15*	-0.06	0.03	-0.01	0.06	1

Table 2.5: Summary Statistics ("Determinants of Fraud" Analyses)

This table gives descriptive statistics (mean, standard deviation, min, and max) for the sample of our main analysis of fraud determinants (193 fraud and 193 non-fraud matched campaigns). All non-dummy variables are winsorized at the 2.5% level on both sides. See Table 2.1, Panel A, for variable description and calculation method.

Variable	# Obs.	Mean	Std. Dev.	Min	Max
		1. Creator(s)' Characteristics/Background			
(1) Creator-Backed Projects	386	12.63	21.55	0	109
(2) Creator-Created Projects	386	0.96	1.90	0	9
(3) Waiting Time (months)	379	10.15	11.29	0	42
(4) Formal Name	386	0.42	0.49	0	1
(5) Natural Person	386	0.49	0.50	0	1
		2. Social Media Affinity			
(6) # External Links	386	1.81	1.34	0	5
(7) Facebook	386	0.58	0.49	0	1
(8) Facebook_Page	386	0.24	0.43	0	1
(9) Facebook_Personal	386	0.47	0.50	0	1
(10) LinkedIn	386	0.03	0.17	0	1
(11) Log (FB Connections)	213	6.71	1.37	3.04	9.48
		3.1. Campaign Funding and Reward Structure			
(12) Duration	386	34.36	10.01	15	60
(13) Min. Pledge Amount	386	10.12	18.50	1	99
(14) No. of Pledge Categories	386	12.60	6.79	4	36
		3.2. Campaign Description Details			
(15) ARI	386	11.39	2.19	7.30	16.90
(16) CL	386	12.32	1.88	8.94	16.77
(17) FKG	386	9.21	1.74	6	13.4
(18) GF	386	8.63	1.22	6.40	11.60
(19) Video Pitch	386	0.93	0.26	0	1

Table 2.6: Mean Differences between Fraud and Matched Sample ("Determinants of Fraud" Analyses)

This table gives the comparison of means test for fraud (193) and non-fraud matched campaigns (193). All non-dummy variables are winsorized at the 2.5% level on both sides. See Table 2.1, panel A, for variable description and calculation method. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Fraud		Non-Fraud		
Variable	***# Obs.***	***Mean***	***# Obs.***	***Mean***	***Difference Test***
		1. Creator(s)' Characteristics/Background			
(1) Creator-Backed Projects	193	10.12	193	15.14	-5.02**
(2) Creator-Created Projects	193	0.60	193	1.32	-0.73***
(3) Waiting Time (months)	187	8.31	192	11.94	-3.63***
(4) Formal Name	193	0.41	193	0.43	-0.02
(5) Natural Person	193	0.49	193	0.49	0.00
		2. Social Media Affinity			
(6) # External Links	193	1.47	193	2.14	-0.67***
(7) Facebook	193	0.50	193	0.66	-0.16***
(8) Facebook_Page	193	0.18	193	0.31	-0.12***
(9) Facebook_Personal	193	0.40	193	0.54	-0.15***
(10) LinkedIn	193	0.04	193	0.03	0.01
(11) Log (FB Connections)	91	6.56	122	6.82	-0.26
		3.1. Campaign Funding and Reward Structure			
(12) Duration	193	35.52	193	33.21	2.31**
(13) Min. Pledge Amount	193	9.61	193	10.62	-1.01
(14) No. of Pledge Categories	193	13.37	193	11.82	1.55**
		3.2. Campaign Description Details			
(15) ARI	193	11.13	193	11.65	-0.52**
(16) CL	193	12.17	193	12.47	-0.30
(17) FKG	193	9.06	193	9.36	-0.31*
(18) GF	193	8.52	193	8.73	-0.21*
(19) Video Pitch	193	0.93	193	0.93	-0.01

Table 2.7: Multivariate Analysis of Determinants of Fraud

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 in specification (1) includes all fraud cases for which a one-to-one nearest neighbor propensity score matched non-fraud campaign can be found (without replacement). The total sample in specification (1) includes 386 campaigns (193 fraud + 193 non-fraud). The sample in specification (2) includes all fraud cases for which a one-to-one nearest neighbor propensity score matched non-fraud campaign can be found (with replacement). The total sample in specification (2) includes 321 campaigns (193 fraud + 128 non-fraud). The sample in specification (3) includes all fraud cases for which a one-to-two nearest neighbor propensity score matched non-fraud campaign can be found (without replacement). The total sample in specification (3) includes 579 campaigns (193 fraud + 386 non-fraud). The sample in specification (4) includes all fraud cases for which a one-to-two nearest neighbor propensity score matched non-fraud campaign can be found (with replacement). The total sample in specification (4) includes 424 campaigns (193 fraud + 231 non-fraud). Investigating the variance inflation factors (VIFs) reveals no evidence of multicollinearity, given the mean VIF range from 1.10 to 1.12 and that all individual values are well below the critical value of 5 (see Kutner et al., 2005). All non-dummy variables are winsorized at the 2.5% level on both sides (see Table 2.1, panel A, for variable descriptions and calculation methods). The reported coefficients are the logs of the odds ratios. Robust standard errors are one-way-clustered by campaign category. t-statistics are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
	1. Creator(s)' Characteristics/Background			
(1) # Creator-Backed Projects	-0.008	-0.002	-0.009**	-0.010
	(-1.40)	(-0.38)	(-2.11)	(-1.19)
(2) # Creator-Created Projects	-0.183**	-0.203*	-0.124**	-0.158**
	(-2.23)	(-1.86)	(-1.98)	(-2.30)
	2. Social Media Affinity			
(3) # External Links	-0.355***	-0.425***	-0.428***	-0.454***
	(-5.64)	(-7.53)	(-7.50)	(-10.59)
(4) Facebook	-0.606**	-0.672***	-0.279	-0.285
	(-2.31)	(-3.42)	(-1.44)	(-1.29)
	3.1. Campaign Funding and Reward Structure			
(5) Duration	0.031***	0.026**	0.026***	0.022**
	(3.37)	(2.52)	(2.61)	(2.12)
(6) Min. Pledge Amount	0.001	0.001	0.001	0.002
	(0.13)	(0.12)	(0.28)	(0.27)
(7) No. of Pledge Categories	0.054***	0.075***	0.042***	0.045***
	(3.09)	(8.28)	(2.59)	(3.66)
	3.2. Campaign Description Details			
(8) ARI	-0.116***	-0.079*	-0.131***	-0.123***
	(-3.09)	(-1.92)	(-3.76)	(-3.15)
(9) Video Pitch	0.156	0.019	-0.117	-0.055
	(0.60)	(0.07)	(-0.29)	(-0.10)
Constant	0.675*	0.817	0.610	1.182
	(1.66)	(1.15)	(0.94)	(1.51)
Replacement	No	Yes	No	Yes
# of Matching Campaigns	1:1	1:1	1:2	1:2
Mean VIF	1.11	1.12	1.10	1.10
Maximum VIF	1.24	1.26	1.24	1.24
Observations	386	321	579	424
Pseudo R^2	0.118	0.133	0.105	0.112

Table 2.8: Multivariate Analysis of Determinants of Fraud (Robustness Check)

In this table, we apply logistic regressions to analyze the determinants of fraud using alternative specifications and proxies, 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 nearest neighbor propensity score matched non-fraud campaign can be found (without replacement). The total sample includes 386 campaigns (193 fraud + 193 non-fraud) if data items are available. Control 1 (*creator(s)' characteristics/background*) includes *Creator-Backed Projects*, and *Creator-Created Projects*; Control 2 (*social media affinity*) includes *# External Links* and *Facebook*. Control 3 (*campaign characteristics*) including *Duration*, *Min. Pledge Amount, No. of Pledge Categories*, *ARI* and *Video Pitch* (see Table 2.1, panel B, for variable descriptions and calculation methods). In panel A, we focus on main and alternative measures of creator(s) characteristics/background. In panel B, we present the results based on main and alternative measures of social media affinity. In panel C, we focus on campaign funding and reward structure as well as campaign description details. Investigating the variance inflation factors (VIFs) reveals no evidence of multicollinearity, given the mean VIF and all individual values are well below the critical value of 5 (see Kutner et al., 2005). All non-dummy variables are winsorized at the 2.5% level on both sides (see Table 2.1, panel A, for variable descriptions and calculation methods). The reported coefficients are the logs of the odds ratios. Robust standard errors are one-way-clustered by campaign category. t-statistics are shown in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
(1) Creator-Backed Projects	-0.013**				
	(-2.48)				
(2) Creator-Created Projects		-0.215***			
		(-2.65)			
(3) Waiting Time (months)			-0.026**		
			(-2.45)		
(4) Formal Name				-0.078	
				(-0.40)	
(5) Natural Person					-0.012
					(-0.07)
Constant	0.308	0.727*	0.408	0.370	0.309
	(0.87)	(1.85)	(1.12)	(0.87)	(0.83)
Control 1	No	No	No	No	No
Control 2	Yes	Yes	Yes	Yes	Yes
Control 3	Yes	Yes	Yes	Yes	Yes
Mean VIF	1.07	1.06	1.06	1.08	1.08
Maximum VIF	1.12	1.11	1.12	1.14	1.15
Observations	386	386	379	386	386
Pseudo R^2	0.105	0.115	0.105	0.094	0.094

Table 2.8: Multivariate Analysis of Determinants of Fraud (Robustness Check)—*continued*

<u>Panel B</u>

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
(1) # External Links	-0.402***					
	(-6.06)					
(2) Facebook		-0.792***				
		(-3.00)				
(3) Facebook_Page			-0.864***			
			(-6.87)			
(4) Facebook_Personal				-0.650**		
				(-2.46)		
(5) LinkedIn					-0.017	
					(-0.05)	
(6) Log (FB Connections)						-0.108
						(-0.88)
Constant	0.689*	0.525	0.381	0.642	0.520	1.427
	(1.69)	(1.25)	(0.86)	(1.51)	(1.18)	(1.57)
Control 1	Yes	Yes	Yes	Yes	Yes	Yes
Control 2	No	No	No	No	No	No
Control 3	Yes	Yes	Yes	Yes	Yes	Yes
Mean VIF	1.10	1.10	1.09	1.09	1.09	1.11
Maximum VIF	1.24	1.23	1.23	1.23	1.23	1.28
Observations	386	386	386	386	386	213
Pseudo R^2	0.105	0.087	0.085	0.080	0.064	0.121

Table 2.8: Multivariate Analysis of Determinants of Fraud (Robustness Check)—*continued*

Panel C

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Duration	0.027***							
	(2.98)							
(2) Min. Pledge Amount		-0.004						
		(-0.64)						
(3) No. of Pledge Categories			0.053***					
			(2.90)					
(4) ARI				-0.103***				
				(-3.27)				
(5) CL					-0.088**			
					(-2.12)			
(6) FKG						-0.101**		
						(-2.40)		
(7) GF							-0.105**	
							(-2.26)	
(8) Video Pitch								0.149
								(0.48)
Constant	0.209	1.133***	0.546*	2.234***	2.156***	2.008***	1.980***	0.954***
	(0.67)	(5.27)	(1.91)	(6.40)	(4.48)	(5.95)	(3.99)	(3.22)
Control 1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control 2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control 3	No	No	No	No	No	No	No	No
Mean VIF	1.13	1.11	1.12	1.11	1.11	1.11	1.11	1.11
Maximum VIF	1.22	1.19	1.19	1.18	1.18	1.18	1.18	1.19
Observations	386	386	386	386	386	386	386	386
Pseudo R^2	0.089	0.080	0.097	0.087	0.083	0.084	0.081	0.079

Table 2.9: Multivariate Analysis of Platform-wide Consequences of Fraud

In this table, we analyze the determinants of *Success* measured by *Funded* (logistic regression; coefficients are the logs of the odds ratios), *Log Pledged* (OLS regressions), and *Log Backers* (OLS regressions). Campaigns being affected by suspension announcements are classified with the dummy variable, *Fraud Period* (panel A) or *Post Fraud* (panel B). In panels A and B, specifications (1)-(3) [specifications (4)-(6)] are based on a sample that includes all Kickstarter campaigns with goal amounts of at least $100 [$10,000] that were posted after January 1, 2010, and ended before December 31, 2018. Investigating the variance inflation factors (VIFs) reveals no evidence of multicollinearity, given the mean VIF is 1.01 to 1.04 in all models and all individual values are well below the critical value of 5 (see Kutner et al., 2005). All non-dummy variables are winsorized at the 1% level on both sides (see Table 2.1, panel B, for variable descriptions and calculation methods). Panel A includes main category, year, month of year (January to December), day of month (first day to last day of respective month), and day of week (Monday to Sunday) fixed effects. Moreover, in panel A, we control for a proxy of platform activity by calculating the average number of daily "live" campaigns during a project's lifetime, namely, *Daily Activity*. Panel B includes main category, and year fixed effects. The time fixed effects are all based on the campaign launch date. Robust standard errors are one-way-clustered by campaign category. *t*-statistics are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A

	(1)	(2)	(3)	(4)	(5)	(6)
	Funded	**Log Pledged**	**Log Backers**	**Funded**	**Log Pledged**	**Log Backers**
	Goal Amount > 99 USD			Goal Amount > 9,999 USD		
Fraud Period	-0.066***	-0.096***	-0.053***	0.014	-0.056*	-0.023
	(-4.60)	(-4.54)	(-4.82)	(0.63)	(-2.00)	(-1.50)
Duration	-0.020***	-0.019***	-0.012***	-0.018***	-0.018***	-0.011**
	(-6.64)	(-5.02)	(-4.10)	(-3.35)	(-3.46)	(-2.70)
Waiting Time	0.001***	0.003***	0.002***	0.001***	0.004***	0.002***
	(9.10)	(10.33)	(12.01)	(6.34)	(9.52)	(12.41)
Featured	2.397***	3.244***	2.280***	2.623***	4.012***	2.787***
	(19.14)	(16.09)	(19.65)	(18.31)	(17.63)	(24.58)
Log Goal	-0.385***	0.174***	0.076***	-0.729***	-0.423***	-0.258***
	(-23.11)	(6.38)	(4.21)	(-10.74)	(-5.33)	(-7.12)
Daily Activity	-0.188***	-0.334***	-0.169***	-0.161***	-0.369***	-0.187***
	(-8.65)	(-10.74)	(-8.92)	(-5.28)	(-10.47)	(-7.94)
Constant	3.441***	4.954***	2.159***	6.128***	10.127***	4.972***
	(13.61)	(16.75)	(13.95)	(8.52)	(10.82)	(10.89)
Category FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Month of Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Day of Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	271,971	271,971	271,971	98,702	98,702	98,702
Mean VIF	1.04	1.04	1.04	1.02	1.02	1.02
Maximum VIF	1.09	1.09	1.09	1.03	1.03	1.03
Adjusted R^2		0.202	0.260		0.266	0.336
Pseudo R^2	0.160			0.216		

Table 2.9: Multivariate Analysis of Platform-wide Consequences of Fraud—*continued*

Panel B

	(1)	(2)	(3)	(4)	(5)	(6)
	Funded	**Log Pledged**	**Log Backers**	**Funded**	**Log Pledged**	**Log Backers**
	Goal Amount > 99 USD			Goal Amount > 9,999 USD		
Post Fraud	-0.117***	-0.202***	-0.095***	-0.115***	-0.244***	-0.102**
	(-5.38)	(-6.72)	(-5.65)	(-2.61)	(-3.59)	(-2.56)
Duration	-0.019***	-0.020***	-0.012***	-0.019**	-0.021**	-0.012*
	(-4.54)	(-3.99)	(-3.13)	(-2.37)	(-2.83)	(-2.08)
Waiting Time	0.001***	0.003***	0.002***	0.001***	0.004***	0.002***
	(8.34)	(11.50)	(14.80)	(3.61)	(11.40)	(12.97)
Featured	2.637***	3.462***	2.417***	2.842***	4.250***	2.933***
	(17.84)	(14.23)	(18.94)	(16.00)	(15.47)	(22.39)
Log Goal	-0.360***	0.156***	0.070***	-0.740***	-0.513***	-0.314***
	(-27.64)	(4.96)	(3.26)	(-8.30)	(-6.24)	(-6.71)
Constant	3.456***	5.711***	2.669***	7.079***	12.548***	6.637***
	(21.74)	(20.38)	(15.06)	(8.43)	(16.97)	(17.90)
Category FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	37,255	37,255	37,255	13,978	13,978	13,978
Mean VIF	1.01	1.01	1.01	1.01	1.01	1.01
Maximum VIF	1.03	1.03	1.03	1.03	1.03	1.03
Adjusted R^2		0.208	0.271		0.270	0.340
Pseudo R^2	0.167			0.220		

Table 2.10: Multivariate Analysis of Platform-wide Consequences of Fraud (Robustness Check)

In this table, we analyze the determinants of *Success* measured by *Funded* (logistic regression; coefficients are the logs of the odds ratios), *Log Pledged* (OLS regressions), and *Log Backers* (OLS regressions) in fraud vs. non-fraud period. In panel A, we include only the campaigns being affected by suspension announcements (*Fraud Period* =1) with goal amounts of at least $100 that were posted after January 1, 2010, and ended before December 31, 2018 (i.e., the sample's launch date is within ∓14 days of the suspension date of one of the identified suspended fraudulent campaigns and did not end before the announcement date of the suspended campaign, to a total of 41,229 affected campaigns). In panel B, we include only the campaigns not being affected by suspension announcements based on our definition (*Fraud Period* =0) to a total of 230,742 campaigns. Investigating the variance inflation factors (VIFs) reveals no evidence of multicollinearity, given the mean VIF is 1.18 to 1.23 in all models and all individual values are well below the critical value of 5 (see Kutner et al., 2005). All non-dummy variables are winsorized at the 1% level on both sides. Controls include *Waiting Time, Featured, Log Goal,* and *Daily Activity* (see Table 2.1, for variable descriptions and calculation methods). All regressions include main category, and year fixed effects. Robust standard errors are one-way-clustered by campaign category. *t*-statistics are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A

	(1)	(2)	(3)
	Funded	**Log Pledged**	**Log Backers**
# Creator-Backed Projects	0.104**	0.026*	0.018*
	(2.35)	(2.01)	(2.38)
# Creator-Created Projects	0.636***	0.320***	0.196***
	(6.14)	(6.98)	(5.89)
# External Links	0.132***	0.085***	0.049***
	(6.25)	(3.42)	(2.97)
Facebook	0.664***	1.206***	1.985***
	(8.33)	(4.44)	(4.98)
Duration	-0.032***	-0.004***	-0.003***
	(-4.19)	(-2.98)	(-3.96)
Min. Pledge Amount	0.036	0.006*	0.004*
	(0.65)	(2.08)	(2.01)
No. of Pledge Categories	-0.270***	-0.045***	-0.030***
	(-3.46)	(-2.65)	(-3.24)
ARI	0.328**	0.241***	0.150***
	(2.22)	(2.98)	(3.65)
Video Pitch	0.225	0.082	0.030
	(0.65)	(0.48)	(0.85)
Constant	6.079***	11.528***	6.547***
	(6.41)	(12.24)	(17.70)
Controls	Yes	Yes	Yes
Category FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	41,229	41,229	41,229
Mean VIF	1.23	1.23	1.23
Maximum VIF	1.84	1.84	1.84
Adjusted R^2		0.222	0.268
Pseudo R^2	0.190		

Table 2.10: Multivariate Analysis of Platform-wide Consequences of Fraud (Robustness Check)—*continued*

Panel B

	(1)	**(2)**	**(3)**
	Funded	**Log Pledged**	**Log Backers**
# Creator-Backed Projects	0.102*	0.021	0.015
	(2.01)	(1.46)	(1.70)
# Creator-Created Projects	0.503***	0.303***	0.184***
	(7.47)	(9.54)	(7.32)
# External Links	0.111***	0.084***	0.046***
	(7.32)	(5.42)	(4.84)
Facebook	0.806***	1.065***	1.909***
	(8.75)	(5.19)	(5.65)
Duration	-0.027***	-0.002***	-0.001***
	(-6.21)	(-4.88)	(-4.96)
Min. Pledge Amount	0.035	0.010	0.006
	(0.87)	(1.24)	(1.51)
No. of Pledge Categories	-0.257***	-0.044***	-0.030***
	(-4.36)	(-3.65)	(-3.94)
ARI	0.304*	0.218**	0.137**
	(2.02)	(2.34)	(2.45)
Video Pitch	0.165	0.071	0.062
	(0.95)	(1.45)	(1.32)
Constant	3.056***	8.431***	4.449***
	(5.75)	(9.24)	(10.63)
Controls	Yes	Yes	Yes
Category FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	230,742	230,742	230,742
Mean VIF	1.18	1.18	1.18
Maximum VIF	1.73	1.73	1.73
Adjusted R^2		0.219	0.248
Pseudo R^2	0.188		

Appendix

Table 2.A1: Correlation Matrix

This table shows the Pearson correlation coefficients for all variables considered in the "Determinants of Fraud" Analysis. Variables are either used as main variables in Control 1, Control 2, and Control 3 (in italic), or as alternative proxies in robustness checks (see Table 2.1, Panel A, for variable descriptions and calculation methods). * indicates statistical significance at least at a 5% level.

		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
1	*Creator-Backed Projects*	1																		
2	*Creator-Created Projects*	0.38*	1																	
3	Waiting Time (months)	0.35*	0.35*	1																
4	Formal Name	0.12*	0.04	0.14*	1															
5	Natural Person	-0.11*	-0.03	-0.14*	-0.81*	1														
6	*# External Links*	0.13*	0.07	0.12*	0	0.01	1													
7	*Facebook*	0.08	0.05	0.07	0.15*	-0.17*	0.26*	1												
8	Facebook_Page	-0.01	0	-0.01	-0.09	0.09	0.42*	0.49*	1											
9	Facebook_Personal	0.09	0.06	0.11*	0.23*	-0.3*	0.11*	0.81*	0.1*	1										
10	LinkedIn	-0.03	-0.04	-0.01	0.18*	-0.14*	0.23*	0.03	0	0.04*	1									
11	Log (FB Connections)	0.11	-0.01	0.07	-0.07	0.01	0.36*	0.16*	0.44*	-0.14	0	1								
12	*Duration*	0.13*	-0.09	-0.08	0.01	-0.09	-0.01	0.08	0.06	0.07	0.06	0.07	1							
13	*Min. Pledge Amount*	-0.05	0.07	-0.02	-0.1*	0.1	-0.09	0	0.04	-0.02	0.01	0.01	-0.01	1						
14	*No. of Pledge Categories*	0.07	-0.04	0.05	0.03	0	0.11*	0.16*	0.12*	0.06	0.06	-0.03	0.03	-0.23*	1					
15	*ARI*	-0.01	-0.04	0.03	-0.15*	0.11*	0.12*	0.04	0.04	0	-0.06	0.04	0.1*	0.07*	-0.02	1				
16	CL	-0.03	-0.1	0	-0.18*	0.18*	0.1	0.06	0.04	0.02	0.02	0.04	0.1*	0.09*	-0.06	0.8*	1			
17	FKG	-0.05	-0.07	0.01	-0.15*	0.12*	0.1*	0.04	0.04	-0.01	-0.08	-0.01	0.08	0.08*	-0.05	0.95*	0.72*	1		
18	GF	0.01	0.06	0.08	-0.05	-0.01	0.09	0.03	0.03	-0.02	-0.14*	0.02	0.03	0.03	0.05	0.69*	0.15*	0.72*	1	
19	*Video Pitch*	0.07	-0.03	0.06	-0.07	0.1*	0.1	0.12*	0.13*	0.08	0.05	0.14*	0.09*	0.02	0.04	0.1*	0.15*	0.11*	-0.01	1

Chapter 3: Determinants of Success in Reward-Based Crowdfunding: The Case of Kickstarter

Moein Karami

John Molson School of Business, Concordia University, Montreal, Canada

Abstract

To date, numerous studies in empirical crowdfunding research have focused on determining factors affecting reward-based crowdfunding success. However, it is extremely difficult to compare results across papers as each use incompatible specifications, and different control variables. More importantly, many explanatory and control variables used in prior research are not defined, measured, and applied consistently across studies, and there are significant endogeneity concerns associated with their use in regressions aiming to explain crowdfunding success. Using a sample of 230,255 crowdfunding campaigns posted through 2013-2018 on Kickstarter, the largest global crowdfunding platform, and drawing upon previous empirical evidence, the statistically significant effect of five independent variables on campaign success is documented. These variables are selected based on three criteria: data availability, data reliability, and documented significant effect, therefore, could be utilized in future crowdfunding success research as control variables. The five factors aim to measure the intensity of competition, creator's crowdfunding experience, project quality & creator confidence, portal recognition, and project size. Furthermore, the current study sheds light on the effect of campaign creator's citizenship, as well as project location, on campaign dynamics and funding success.

3.1. Introduction

Crowdfunding refers to a novel form of fundraising whereby groups of people pool money, typically small individual contributions, to support a particular goal (Ahlers et al., 2015) or fund ventures without standard financial intermediaries (Mollick, 2014). It is categorized into four main forms: donation-based, reward-based, equity-based, and lending-based. In particular, reward-based crowdfunding (hereafter, crowdfunding), which is very similar to the idea of pre-selling a product, has attracted substantial academic attention as it does not involve offering securities or issuing debt. The total amount raised by crowdfunding activities has continued to grow dramatically, and the surge indicates how critical it has become for new entrepreneurial activities. Overall, the impact of crowdfunding on new product market reception in an open innovative environment is well-established, resulting in a greater funding efficiency for start-ups (Stanko and Henard, 2017).

Furthermore, and as Mollick (2018) points out, "crowdfunding campaigns lead to new organizations that ultimately generate billions in non-crowdfunding revenue and hire thousands of employees." For example, in February and March 2014, the Kickstarter campaign for The *Dash – Wireless Smart In Ear Headphones* raised about $3.4 million, which was followed by $22 million in subsequent financing from angel investors[1]. By the end of 2016, the company *Bragi*, which was behind the campaign, had sold more than 600,000 units, pushing revenues to $100 million[2]. Since the crowdfunding campaign ended, *Bragi* has successfully obtained over 30 patents, with another 150 in the application process, hired several new personnel, and created new businesses with partners[3]. Examples like this demonstrate that crowdfunding is a catalyst for entrepreneurship, a door opener for future financing, and an important new means for successful product development. The positive impact of crowdfunding has become evident for firm creation (Mollick, 2018) and subsequent venture capital investments (Sorenson et al., 2016). In sum, the positive impact of crowdfunding has become more and more evident for economic growth (Kitchens and Torrence, 2012), society (Lehner, 2013), and employment (Ramos and Gonzalez, 2018).

There are numerous empirical studies of the factors that might influence the success of a crowdfunding campaign, however, it is extremely difficult to compare results across these studies

[1] See: http://www.crowdfundinsider.com/2015/11/76751-bragi-receives-huge-22-million-investment-as-dash-prepares-to-ship
[2] See https://www.forbes.com/sites/amyfeldman/2016/04/14/ten-of-the-most-successful-companies-built-on-kickstarter/#7ccf66f969e8
[3] See https://www.forbes.com/sites/amyfeldman/2016/04/14/ten-of-the-most-successful-companies-built-on-kickstarter/#7ccf66f969e8

as each use different specifications of the explanatory variables and controls. Use of different control variables may result in finding a seemingly statistically significant effect of an explanatory variable of interest on crowdfunding success, while it is e.g., simply correlated with a relevant omitted variable. This problem has been previously discussed in Initial Public Offering (IPO) research (Butler et al., 2014). Using different methodologies, Butler et al. (2014) quantify the statistical robustness of variables used in prior IPO research to explain initial returns of IPOs, and aim to provide a benchmark regression specification similar to the ones in other important areas of finance, e.g., asset pricing (Fama and French, 1993). Similarly, the current study aims to provide a benchmark specification of control variables in crowdfunding success research using data obtained from Kickstarter website, the largest global crowdfunding platform. The data set includes 230,255 campaigns posted from January 1, 2013 to December 31, 2018, and is available for future research upon request (For a detailed description of dataset file, see Appendix: Table 3.A1).

More importantly, and considering limitations on accessing data, many explanatory and control variables used in prior research are not defined, measured, and applied consistently across papers, and there are significant endogeneity concerns associated with their use in regressions aiming to explain crowdfunding success. For example, following Mollick (2014) which is one of the first, and most cited, studies to use a comprehensive sample of Kickstarter crowdfunding campaigns (22,651 campaigns posted from April 2009 to July 2012) to determine factors affecting success, numerous researchers started using the number of Facebook friends of the campaign creator in order to examine the role of social media (see, e.g., Calic and Mosakowski, 2016; Fietkiewicz et al., 2018; Kromidha and Robson, 2016; Wessel et al., 2016) or to control for its effect. However, due to data collection limitations, this variable is almost always recorded as of the time of data collection[1] (rather than at the start date of campaign) and its use in regression models can result in endogeneity concerns related to reverse causality[2]. The current study also discusses such variables that are sometimes used in prior research as explanatory/control factors (e.g., number of updates

[1] Unless the researcher is collecting data at campaigns' launch date, first, and waits for the final outcome of campaigns, and then finalizes the data set (e.g., see Buttice et al., 2017) which is not the case at most of time.

[2] Mollick (2014) also notes this problem, however, he argues that since project creators tend to separate the project from their personal page, the number of friends is "less" likely (compared to number of likes of the Facebook "page" associated with the venture) to increase substantially as a result of campaign success. However, since these creators have linked their personal accounts to their Kickstarter page (and that is where the data comes from), I believe this is not a safe assumption to make, and the underlying reason for finding a significant positive effect of Facebook friends on campaign success, might stem from reverse causality, i.e., after a successful campaign, the number of Facebook connections of the creator increases significantly (e.g., many people who backed the campaign will also follow creator's personal Facebook). It should also be noted that some campaigns do not have a Facebook page associated with the venture, and only link their personal Facebook page to their Kickstarter campaign).

during campaign, or number of comments during the campaign), but considering data limitations, their use is problematic.

Overall, Mollick (2014) provides valuable first insights into dynamics of reward-based crowdfunding and provides a basis for control variables used in subsequent research. This paper aims to update these findings using a more comprehensive and more recent sample of Kickstarter campaigns, discuss limitations of some of the previously used variables, and introduce new factors that can reliably be utilized in future research as control variables. Using a sample of 230,255 crowdfunding campaigns posted through 2013-2018 on Kickstarter, and drawing upon previous empirical evidence, the statistically significant effect of five independent variables on campaign success is documented. These variables are selected based on three criteria: data availability, data reliability, and documented significant effect. The five factors aim to measure the intensity of competition (*Competition Int.*), creator's crowdfunding experience (*Membership Tenure*), project quality & creator confidence (*Duration*), portal recognition (*Featured*), and project size (*Log Goal*)[1]. The three latter variables are introduced in Mollick (2014) and have been extensively used in prior research as controls. Intensity of competition (*Competition Int.*) and creator's crowdfunding experience (*Membership Tenure*) are discussed in this study and suggested to be included in the set of control variables in future research[2]. Finally, to the best of my knowledge, this paper is the first to investigate the effect of project location, as well as creator's citizenship, on project success in a comprehensive sample.

The remainder of the paper proceeds as follows. Section 3.2 provides a review of the literature and presents the hypotheses, while section 3.3 describes data and methods. Section 3.4 summarizes empirical results, and section 3.5 concludes.

[1] A similar set of controls has been used in Cumming et al. (2020), however, the focus of that study is on fraud risks of crowdfunding and does not aim at determining robust determinants of success.
[2] The dataset including information on these variables along with many other useful information such as campaign name, campaign webpage, project location, creator citizenship etc. is available upon request for future research. See Appendix for more details.

3.2. Literature Review and Hypotheses

Shneor and Vik (2020) provide a systematic literature review on 88 academic papers published between 2010-2017 in 65 different journals on crowdfunding (in all of its forms) success factors. They identify 111 aggregated independent variables along with 6 main success indicators, and also show that reward-based crowdfunding has been the most popular context for crowdfunding success research. They show that vast majority of (reward-based) crowdfunding papers examined in their study (around 85%) use the indicator of reaching goal amount (variable *Funded* in this study), Amount Raised (variable *Log Pledged*, here), and Number of contributors (variable *Log Backers*, here) as dependent variables. We use all three in our subsequent analyses.

Prior to introducing the set of suggested control variables, I mention three set of independent variables that have been used in some prior crowdfunding success research as independent/control variables, discuss the practical problems associated with their inclusion in models, and suggest modifications that can mitigate the concerns associated with their use in the regressions. Furthermore, I will discuss some variables that are not included in the set of suggested controls here (considering the more difficult process of obtaining data, and a weaker documented effect on success), but could be added to the set of suggested controls depending on the necessity to control for their effect in some particular studies.

Three variables have been used in previous research as explanatory factors of crowdfunding success which aim at measuring the social capital of the creator, the creator's involvement in the course of campaign, and the backers' involvement during the campaign, however, there are practical considerations that need attention to ensure the reliability of results.

To proxy for social capital of the creator, personal Facebook of the creator (if linked to Kickstarter campaign webpage) has been used as the basis. Researchers usually use either a dummy variable that indicates whether the creator has a personal Facebook page linked to the campaign, or number of Facebook friends of the creator (e.g., Courtney et al., 2016; Kromidha and Robson, 2016; Mollick, 2014). The downside of using these variables, considering data limitations, is that it is collected as of the time of data collection and introduces biases related to reverse causality. However, these variables can be used if collected at the campaign launch (see Buttice et al., 2017) but for most researchers this is not feasible as they collect past data to conduct analyses. Moreover, number of updates posted by creator, and number of comments posted by backers, have been used in a number of prior studies as explanatory factors (e.g., see Cordova et al., 2015), however, these variables are also ex-post, and are collected as of the time of obtaining data, while creators/backers can post updates/comments before, during, and after the campaign and their inclusion in the regressions is associated with substantial endogeneity concerns. Mollick (2014) introduces a new variable, namely, Quick Updates, that only aims at capturing the number of updates in the first three days of the campaign which is adopted by a number of subsequent studies (see e.g., Courtney et al., 2017) which mitigates this concern, but it is more difficult to extract from Kickstarter website for large samples, as the researcher not only needs access to number of updates, but also requires information on timing of each update.

Number of spelling errors in the project description, number of the reward categories, and whether a Video pitch is available on the campaign webpage have also been used as control variables in regressions aiming at determining success factors, starting with Mollick (2014), and can be used in future research depending on the focus of study but the results on their effect and statistical significance is mixed.

As mentioned previously, it is difficult to compare results across crowdfunding success studies as most use different specifications of the controls and this paper aims to provide a basis for a benchmark regression specification that can be reliably used in future studies, considering data availability, data reliability, and documented significant effect on campaign success (across multiple studies, and in different samples). In what follows, five variables to proxy for the intensity of competition, creator's crowdfunding experience, project quality & creator confidence, portal recognition, and project size will be discussed and their effect on campaign success will be hypothesized based on theoretical background, and previous empirical evidence. The latter three variables are widely used in previous research and their effect is well-established.

3.2.1. Intensity of Competition

IPO research has already shown the importance of the intensity of competition on initial returns of offerings as marked by hot and cold IPOs (see, e.g., Helwege and Liang, 2004). In crowdfunding market, one would intuitively expect that as more concurrent projects are posted, the probability of reaching goal amount will decrease. However, this factor, along with any procyclical effect, is largely ignored in crowdfunding success research (Cumming et al., 2020, control for a proxy aiming at measuring the competition intensity and show its significant effect). Following their idea, and using 265,001 crowdfunding campaigns posted from 2012-10-01 to 2019-01-31 (including successful, failed, cancelled by creator, and suspended by Kickstarter campaigns) I construct a variable, namely, *Competition Int.*, which takes a distinctive value for each of the campaigns (a value between 2.01 to 7.67 after being winsorized at 1% level on both sides) in the 2013-2018 sample (which includes 235,255 observations (only successful and failed campaigns) posted that were posted after January 1, 2013, and ended on or before December 31, 2018).

The reason for collecting data in the 3 months before and after the sample period (and including cancelled and suspended campaigns) is to be able to count concurrent projects that were posted before 2013, or ended after 2018 to construct the *Competition Int.* measure for the 2013-2018 sample. Considering the maximum duration of 60 days for the campaigns, I ensure that no concurrent projects will be missed in calculating intensity of competition. The variable is defined as average daily number of projects that were "live" during campaign's lifetime, divided by 1000, and Hypothesis 1 is developed as follows:

> **Hypothesis 1:** *Competition intensity negatively affects the measures of crowdfunding success.*

3.2.2. Creator's Crowdfunding Experience

I use number of months between the date creator joined Kickstarter (i.e., created an account) and campaign's launch date (*Membership Tenure*) as a proxy for creator's crowdfunding experience (as an observer, backer, or previous project creator). Similar variables have been used in regressions in prior research (see, e.g., Calic and Mosakowski, 2016; Cumming et al., 2020), and it is a readily observable information on Kickstarter website, but overall many studies do not control for this effect, while they choose to use number of projects created/backed by creator as a proxy for creator's experience. While I agree that the latter variables can serve as a better proxy if obtained at the launch date of campaign, because of data collection limitations (i.e., these variables are shown on Kickstarter website based on today's date, and the number of projects created/backed by creator does not necessarily reflect those that they backed/created prior to the launch date of the observation campaign), and it is more difficult to collect data on the timing of each projects they backed/created. The *Membership Tenure* can serve as an easier to obtain control variable, with no endogeneity concerns, and at the same time indicative of how long a creator was a member of Kickstarter community before launching a campaign. Therefore, Hypothesis 2 is developed as follows:

Hypothesis 2: *Membership tenure of project creator positively affects the measures of crowdfunding success.*

3.2.3. Project Quality and Creator Confidence

There are multiple measures that can serve as a signal of project quality, however, one credible signal that previous empirical research has shown its significant effect on success is the campaign duration for reaching the goal amount set by the campaign creator which can be between 1-60 days (while Kickstarter suggests a 30-day funding period based on history of successful campaigns). Setting low durations can also signal the confidence of the creator in the project. This variable has been widely used in previous research (starting with Mollick (2014)), therefore, Hypothesis 3 is developed as follows:

Hypothesis 3: *Duration of funding set by project creator negatively affects the measures of crowdfunding success.*

3.2.4. Portal Recognition

There are thousands of projects live on Kickstarter at any time, and Kickstarter team is constantly following new launches for projects that really stand out. There are many factors that are taken into consideration before featuring a project on the homepage and giving it a "Project We Love" badge. This recognition is shown to have substantial positive effect on funding outcome and can serve as a credible signal of quality (Mollick, 2014) and has been controlled for in vast majority of crowdfunding success research. Therefore, Hypothesis 4 is developed as follows:

Hypothesis 4: *Being featured by Kickstarter as "Projects We Love" positively affects the measures of crowdfunding success.*

3.2.5. Project Size

Kickstarter follows an "all-or-nothing" model, i.e., the goal amount set by entrepreneur must be reached in order for the pledge money to be transferred to the creator. One intuitively expects that projects with higher goal amounts set, have a lower probability of reaching the goal amount, all else being equal. This variable is also controlled for in almost all crowdfunding success research starting with Mollick (2014), and is an important control variable to be included in the benchmark regression. Therefore, Hypothesis 5 is developed as follows:

Hypothesis 5: *Setting higher goal amount negatively affects the measures of crowdfunding success.*

3.2.6. Geography and Crowdfunding Success

To date, several studies have aimed at shedding light on the impact of geography on crowdfunding success (Agrawal et al., 2015; Dejean, 2019; Gallemore et al., 2019; Guenther et al., 2018; Lin and Viswanathan, 2016; Mollick, 2014). It is well-established in the literature that the success of entrepreneurial ventures seeking traditional forms of entrepreneurial funding is often highly constrained by geography (see e.g., Chen et al., 2009; Shane and Cable, 2002; Stuart and Sorenson, 2008), however, there are preliminary empirical evidence that crowdfunding substantially democratize access to funding worldwide (Cumming et al., 2019; Mollick and Robb, 2016).

According to Kickstarter website, the main project creator (individual legally associated with the campaign) should hold a valid passport of one of the following specific countries, that I classify into 7 categories: 1) US; 2) UK; 3) CA; 4) AU/NZ; 5) Asia (main) including *Hong Kong*, *Japan*, and *Singapore*; 6) EU (main) including *Germany*, *France, Spain, Italy, Belgium, Luxemburg, Netherlands, Austria, Switzerland, Scandinavia (Denmark, Norway, Sweden)*, and *Ireland*; and 7) Mexico. However, the project location can be anywhere in the world.

While remaining agnostic about the effects of creator citizenship and project location, this study aims at testing the effect of citizenship on the success, as well as the effect of project location being one of the accepted countries of citizenship on funding outcome. I also look at the interactions between the two sets, and e.g., look into the effects of having project locations outside the creator's country of citizenship. I also look at the success rates of projects that are located outside one of the accepted countries of citizenship. To the best of my knowledge, this study is the first to provide detailed results on the effect of creator's citizenship, as well as project location on success measures using a comprehensive sample of Kickstarter campaigns.

3.3. Data and Methods

I collected data directly from Kickstarter website using data crawling methods, and finalized a sample of 230,255 crowdfunding campaigns that were posted on or after January 1, 2013, and ended on or before December 31, 2018[1]. All variable definitions, and fixed effects used in subsequent empirical analyses can be found in Panel A to Panel C of Table 3.1. Panel D of Table 3.1 presents the derivation of the final sample of projects (either successful, or failed, in terms of reaching the goal amount) from the originally obtained dataset.

—Please insert Table 3.1 about here—

Table 3.2 presents an overview of the sample. Panel A shows the number of launched campaigns within Kickstarter's 15 main categories for each respective year, and results show that most projects are posted in "Film & Video", "Publishing", and "Music" categories. Panel B provides the success rates within each category by year, and shows that the highest success rate is among projects posted in "Comics", "Theater", and "Dance". Overall, results in both panels emphasize the importance of artistic and cultural ventures in crowdfunding market. In Panel C of Table 3.2, descriptive statistics for the sample along with correlation coefficients are presented.

—Please insert Table 3.2 about here—

Finally, Table 3.3 present an overview of Kickstarter sample based on Creator Country as well as project location. Panel A shows the number of launched campaigns within Kickstarter's main countries (based on creator citizenship) for each respective location (project location). Results show that most of the creators have US citizenship, followed by United Kingdom, Europe, and Canada. Same order holds when we look at project locations. Moreover, results show that, overall, most of the projects are located in the creator's country of citizenship. Panel B shows the percentage of successful campaigns within each country of citizenship for each respective location. Results show that highest success rate is among creators from Asian countries and United Kingdom. More interestingly, I found that not only projects have the highest success rate in Asia (main) locations, but also US citizens have the highest success rate if their projects are located in Asia. This provides preliminary evidence on how crowdfunding might democratize access to financing worldwide. Panel C [Panel D] provides descriptive statistics for all variables considered in the analyses for creator country [project location] in US vs. Non-US sample. I found that overall, US citizens have

[1] The dataset is available for future research upon request. See Appendix: Table 3.A1 for details.

a higher probability of success (41%) compared to non-US citizens (37%), and projects located in US also have higher probability of success (41%) compared to projects located in other countries (38%). However, univariate comparisons do not reveal any significant difference among independent variables (it should be noted that the dataset is winsorized at 1% level on both sides for all non-dummy variables).

—Please insert Table 3.3 about here—

The structure of our base logistic (and OLS) regression model in order to investigate determinants of success is as follows:

$$Success_i = \beta_0 + \beta_1 \cdot Competition\ Int._i + \beta_2 \cdot Membership\ Tenure_i + \beta_3 \cdot Duration_i + \beta_4 \cdot Featured_i + \beta_5 \cdot Log\ Goal_i + \varepsilon_i, \quad (1)$$

For each campaign i, *Success* represents the dummy variable *Funded* (Logistic), the variable *Log Pledged* (OLS), or the variable *Log Backers* (OLS). Category and year fixed effects will also be considered in all models, as well as "Month of Year" and "Week of Month" fixed effects in a robustness check (see Panel B of Table 3.1 for detailed definitions of fixed effects).

We also add creator citizenship dummies (with US citizens being the reference group), project location dummies (with US location being the reference group), and interactions of citizenship dummies with project location in order to investigate the effects of geography on funding outcome.

3.4. Results

3.4.1. Determinants of Funding Success

Table 3.4 presents the multivariate analysis of main determinants of campaign success (see Equation 1). Models (1) to (3) include the total sample, and models (4) to (6) only include projects with goal amounts of at least 10,000 (USD). All models control for category and year fixed effects. Results show that the *Competition Int.* significantly affects all measures of success. For example, model (1) shows that increasing the competition by 1 unit (i.e., 1000 more live concurrent projects), is associated with an average decreases in the odds of success by a factor of 15.1%. This effect is prevalent across all models and on all success measures, providing strong support for our Hypothesis 1.

The positive impact of *Membership Tenure* is also documented in all models, and illustrates that the longer the creator has been a member of Kickstarter community before the launch date of campaign, the higher the probability of success, all else being equal. This provides strong evidence in support of Hypothesis 2. The negative impact of *Duration* is also confirmed (in line with Hypothesis 3) and e.g., results presented in model (1) show that increasing duration of campaign by 1 day, is associated with an average decreases in the odds of success by a factor of 2.1%. Being *Featured* by Kickstarter has the largest economic impact on success measures and is highly statistically significant considering the t-statistics, providing strong support for Hypothesis 4. Finally, and in line with numerous previous findings, the negative effect of goal amount (*Log Goal*) on success measures is confirmed in line with our Hypothesis 5. In Panel B of Table 3.4, "Month of Year", and "Day of Month" fixed effects are added to capture any periodic effect that might affect the results, but I found that results remain significantly robust.

—Please insert Table 3.4 about here—

Overall, Table 3.4 results provide strong evidence of the significant effect of identified variables, and supports the idea that this benchmark regression can be reliably used (as a set of control variables) in future research considering the very low variance inflation factors (suggesting that multicollinearity is not an issue across these independent variables), and relatively large Pseudo and adjusted R^2 across all models (i.e., on average, almost 20% of variation in the success measures can be explained by the variation in this set of controls, along with the fixed effects).

Although there are many advantages to using large samples, one remaining concern is the p-value problem associated with it (Lin et al., 2013). We follow Lin et al. (2013) guidelines in using Coefficient/p-Value Charts (CPS) to mitigate the concern of large sample affecting our results and statistical inferences. Figure 3.1 to 3.5 illustrate the use of CPS charts for *Competition Int., Membership Tenure, Duration, Featured, and Log Goal* variables respectively, in an OLS regression with *Log Pledged* as dependent variable (year, and category fixed effects also considered in all models). They show that once the sample size is greater than 1000 in all figures, the p-value drops to near zero, and the coefficients approach their value as shown in column 2 of Table 3.4, Panel A. Given the size of our sample, there is little doubt that the results are statistically significant.

—Please insert Figure 3.1 to 3.5 about here—

3.4.2. Geography and Funding Outcome

Table 3.5 presents the multivariate analysis of the effect of creator country (i.e., citizenship) on measures of campaign success, using US creators as the reference group. Results across all models show that, all else being equal, the creators that are citizens of Asia (main) as a group, i.e., Hong Kong, Japan, and Singapore, have higher probability of success, raise more funds, and attract more backers, while all other groups underperform the US citizens (The results are in line with the previously discussed univariate analysis shown in Panel B of Table 3.3).

—Please insert Table 3.5 about here—

Moreover, and focusing on project locations, multivariate analyses presented in Table 3.6, shows that projects that are located in Asia (main) group, as well as projects located in other parts of world where their citizens cannot create campaigns on Kickstarter, outperform projects located in US that is used as reference group, all else being equal.

—Please insert Table 3.6 about here—

Finally, since the largest group of creators are US citizens, I dig deeper into multivariate analysis of success measures in a crowdfunding campaign, examining the interactions between creator US citizenship, and project location. Panel A of Table 3.7 shows that US citizens as project creators are less likely to succeed when the project is also located in US as shown by negative coefficient of interaction term "*C:US * L:US*" across all models. Finally, Panel B of Table 3.7 shows that the US citizens as campaign creators have a lower likelihood of success when the project is located in one of the main countries where citizens of the country are allowed to launch projects on Kickstarter themselves, all else being equal, and as shown by negative coefficient of interaction term "*C:US * L: Main*" across all models.

—Please insert Table 3.7 about here—

3.5. Conclusion

In this study, a basis for a benchmark regression specification that can be reliably used in future research, considering data availability, data reliability, and documented significant effect on campaign success is provided. Five main variables to proxy for the intensity of competition, creator's crowdfunding experience, project quality & creator confidence, portal recognition, and project size were discussed and their effect on campaign success was tested, using a sample of

230,255 crowdfunding campaigns posted through 2013-2018 on Kickstarter, the largest global crowdfunding platform. Furthermore, the current study sheds light on the effect of campaign creator's citizenship, as well as project location, on campaign dynamics and funding success. The dataset is available upon request for future research.

Acknowledgments

I thank my PhD supervisor, Dr. Denis Schweizer for his support and helpful feedback on research. I am also grateful to my PhD committee members, Dr. Maher Kooli and Dr. Thomas Walker, for many constructive comments and suggestions. All remaining errors are my own.

References

Agrawal, A., Catalini, C., & Goldfarb, A. (2015). Crowdfunding: Geography, social networks, and the timing of investment decisions. *Journal of Economics & Management Strategy, 24*(2), 253–274.

Ahlers, G. K., Cumming, D., Günther, C., & Schweizer, D. (2015). Signaling in equity crowdfunding. *Entrepreneurship Theory and Practice, 39*(4), 955–980.

Butler, A. W., Keefe, M. O. C., & Kieschnick, R. (2014). Robust determinants of IPO underpricing and their implications for IPO research. *Journal of Corporate Finance, 27*, 367–383.

Buttice, V., Colombo, M. G., & Wright, M. (2017). Serial crowdfunding, social capital, and project success. *Entrepreneurship Theory and Practice, 41*(2), 183–207.

Calic, G., & Mosakowski, E. (2016). Kicking off social entrepreneurship: How a sustainability orientation influences crowdfunding success. *Journal of Management Studies, 53*(5), 738–767.

Chen, H., Gompers, P., Kovner, A., & Lerner, J. (2010). Buy local? The geography of venture capital. *Journal of Urban Economics, 67*(1), 90–102.

Cordova, A., Dolci, J., & Gianfrate, G. (2015). The determinants of crowdfunding success: evidence from technology projects. *Procedia-Social and Behavioral Sciences, 181*, 115–124.

Courtney, C., Dutta, S., & Li, Y. (2017). Resolving information asymmetry: Signaling, endorsement, and crowdfunding success. *Entrepreneurship Theory and Practice, 41*(2), 265–290.

Cumming, D. J., Hornuf, L., Karami, M., & Schweizer, D. (2020). Disentangling crowdfunding from fraudfunding. *Max Planck Institute for Innovation & competition research paper*, *Available at SSRN*: https://ssrn.com/abstract=2828919.

Cumming, D., Meoli, M., & Vismara, S. (2019). Does equity crowdfunding democratize entrepreneurial finance? *Small Business Economics, 56*, 533–552.

Dejean, S. (2019). The role of distance and social networks in the geography of crowdfunding: evidence from France. *Regional Studies, 54*(3), 329–339.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*, 3–56.

Fietkiewicz, K. J., Hoffmann, C., & Lins, E. (2018). Find the perfect match: the interplay among Facebook, YouTube and LinkedIn on crowdfunding success. *International Journal of Entrepreneurship and Small Business, 33*(4), 472–493.

Gallemore, C., Nielsen, K. R., & Jespersen, K. (2019). The uneven geography of crowdfunding success: Spatial capital on Indiegogo. *Environment and Planning A: Economy and Space, 51*(6), 1389–1406.

Guenther, C., Johan, S., & Schweizer, D. (2018). Is the crowd sensitive to distance? How investment decisions differ by investor type. *Small Business Economics, 50*(2), 289–305.

Kitchens, R., & Torrence, P. D. (2012). The JOBS Act-crowdfunding and beyond. *Economic Development Journal, 11*(4), 42.

Kromidha, E., & Robson, P. (2016). Social identity and signalling success factors in online crowdfunding. *Entrepreneurship & Regional Development, 28*(9-10), 605–629.

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). Applied linear statistical models (5th edition). New York: McGraw-Hill.

Lehner, O. M. (2013). Crowdfunding social ventures: a model and research agenda. *Venture Capital, 15*(4), 289–311.

Lin, M., Lucas Jr, H. C., & Shmueli, G. (2013). Research commentary—too big to fail: large samples and the p-value problem. *Information Systems Research, 24*(4), 906–917.

Lin, M., & Viswanathan, S. (2016). Home bias in online investments: An empirical study of an online crowdfunding market. *Management Science, 62*(5), 1393–1414.

Mollick, E. (2014). The dynamics of crowdfunding: An exploratory study. *Journal of business venturing, 29*(1), 1–16.

Mollick, E. (2018). Crowdfunding as a font of entrepreneurship: Outcomes of reward-based crowdfunding. In Cumming, D. J., & Hornuf, L. (Eds.), *The Economics of Crowdfunding* (pp. 133–150): Palgrave Macmillan, Cham.

Mollick, E., & Robb, A. (2016). Democratizing innovation and capital access: The role of crowdfunding. *California Management Review, 58*(2), 72–87.

Ramos, J., and González, B. (2018). Crowdfunding and employment: An analysis of the employment effects of crowdfunding in Spain. In Brüntje, D., & Gajda, O. (Eds.), *Crowdfunding in Europe State of the Art in Theory and Practice* (pp. 97–114): Springer, Cham.

Shane, S., & Cable, D. (2002). Network ties, reputation, and the financing of new ventures. *Management science, 48*(3), 364–381.

Shneor, R., & Vik, A. A. (2020). Crowdfunding success: a systematic literature review 2010–2017. *Baltic Journal of Management, 15*(2), 149–182.

Sorenson, O., Assenova, V., Li, G. C., Boada, J., & Fleming, L. (2016). Expand innovation finance via crowdfunding. *Science, 354*(6319), 1526–1528.

Stanko, M. A., & Henard, D. H. (2017). Toward a better understanding of crowdfunding, openness and the consequences for innovation. *Research Policy, 46*(4), 784–798.

Stuart, T., & Sorenson, O. (2003). The geography of opportunity: spatial heterogeneity in founding rates and the performance of biotechnology firms. *Research policy, 32*(2), 229–253.

Wessel, M., Thies, F., & Benlian, A. (2016). The emergence and effects of fake social information: Evidence from crowdfunding. *Decision Support Systems, 90*, 75–85.

Figure 3.1: CPS Chart for Competition Intensity: Coefficient and p-Value vs. Sample Size

This figure summarizes the Coefficient/p-Value Chart for the variable *Competition Int.*, in the OLS regression with *Log Pledged* as dependent variable and "*Membership Tenure, Duration, Featured, and Log Goal*" as remaining regressors. Year, and Category fixed effects are considered in all models. The upper figure illustrates the coefficient vs. sample size, and the lower figure illustrates the p-value vs. sample size. The figures are zoomed into n < 5000 for illustration.

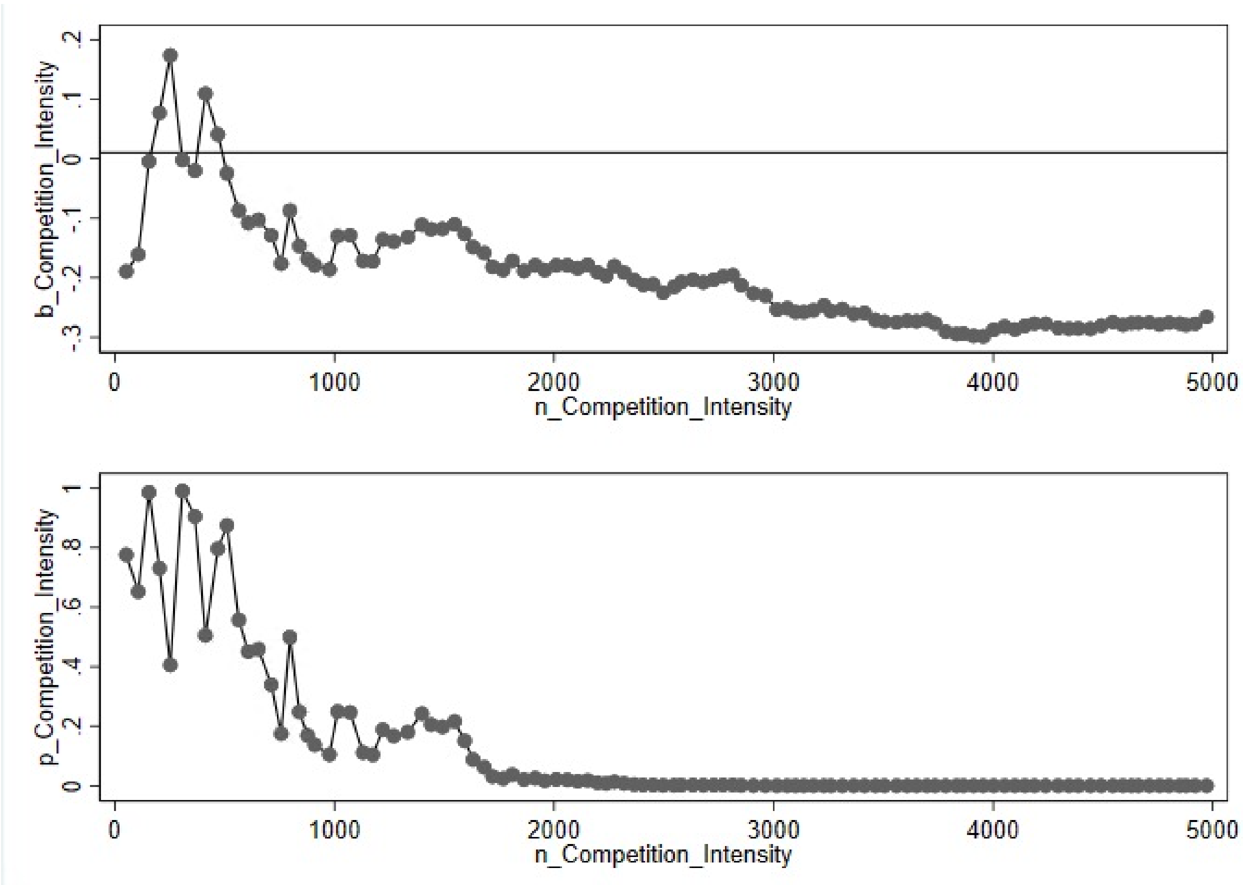


Figure 3.2: CPS Chart for Membership Tenure: Coefficient and p-Value vs. Sample Size

This figure summarizes the Coefficient/p-Value Chart for the variable *Membership Tenure*, in the OLS regression with *Log Pledged* as dependent variable and "*Competition Int., Duration, Featured, and Log Goal*" as remaining regressors. Year, and Category fixed effects are considered in all models. The upper figure illustrates the coefficient vs. sample size, and the lower figure illustrates the p-value vs. sample size. The figures are zoomed into n < 5000 for illustration.

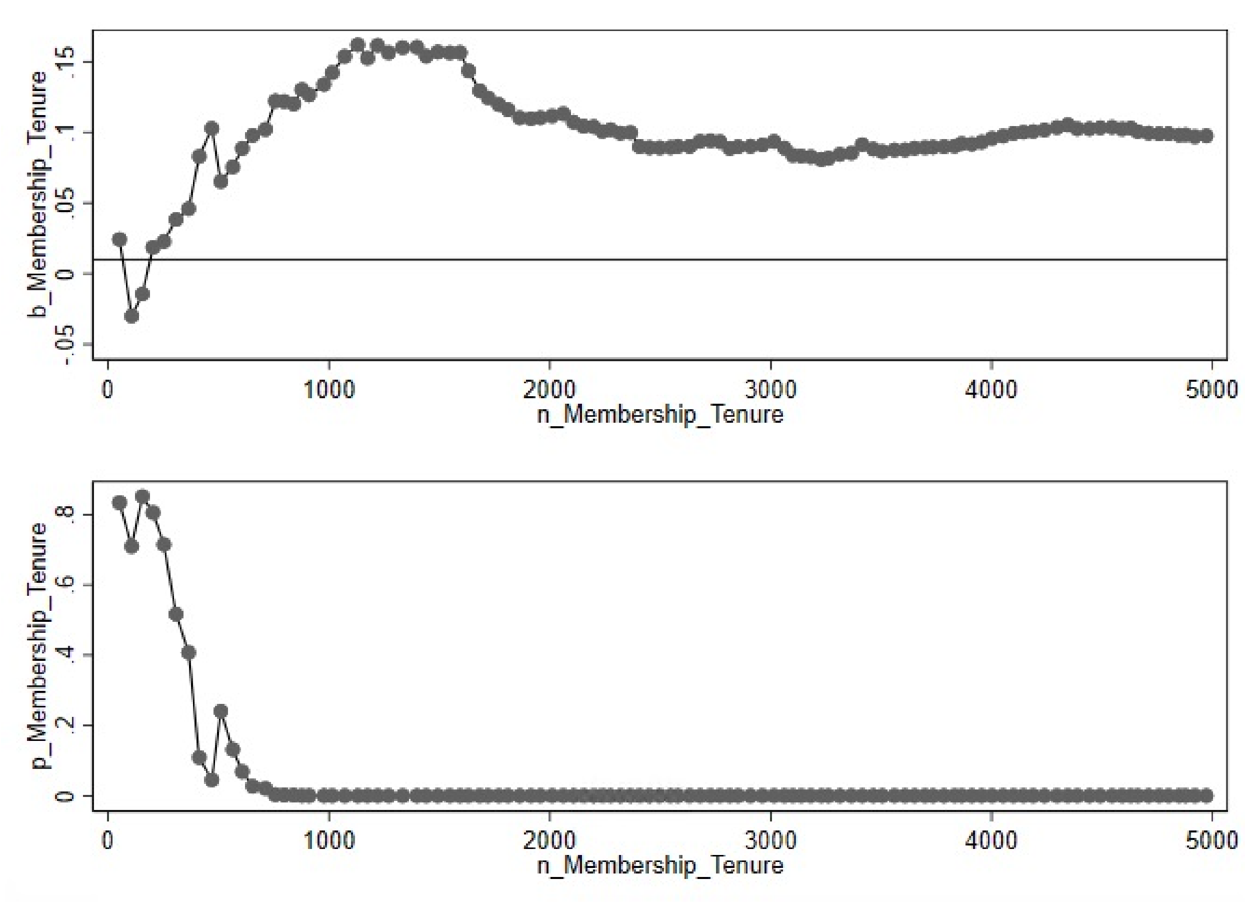


Figure 3.3: CPS Chart for Duration: Coefficient and p-Value vs. Sample Size

This figure summarizes the Coefficient/p-Value Chart for the variable *Duration*, in the OLS regression with *Log Pledged* as dependent variable and "*Competition Int., Membership Tenure, Featured, and Log Goal*" as remaining regressors. Year, and Category fixed effects are considered in all models. The upper figure illustrates the coefficient vs. sample size, and the lower figure illustrates the p-value vs. sample size. The figures are zoomed into n < 5000 for illustration.

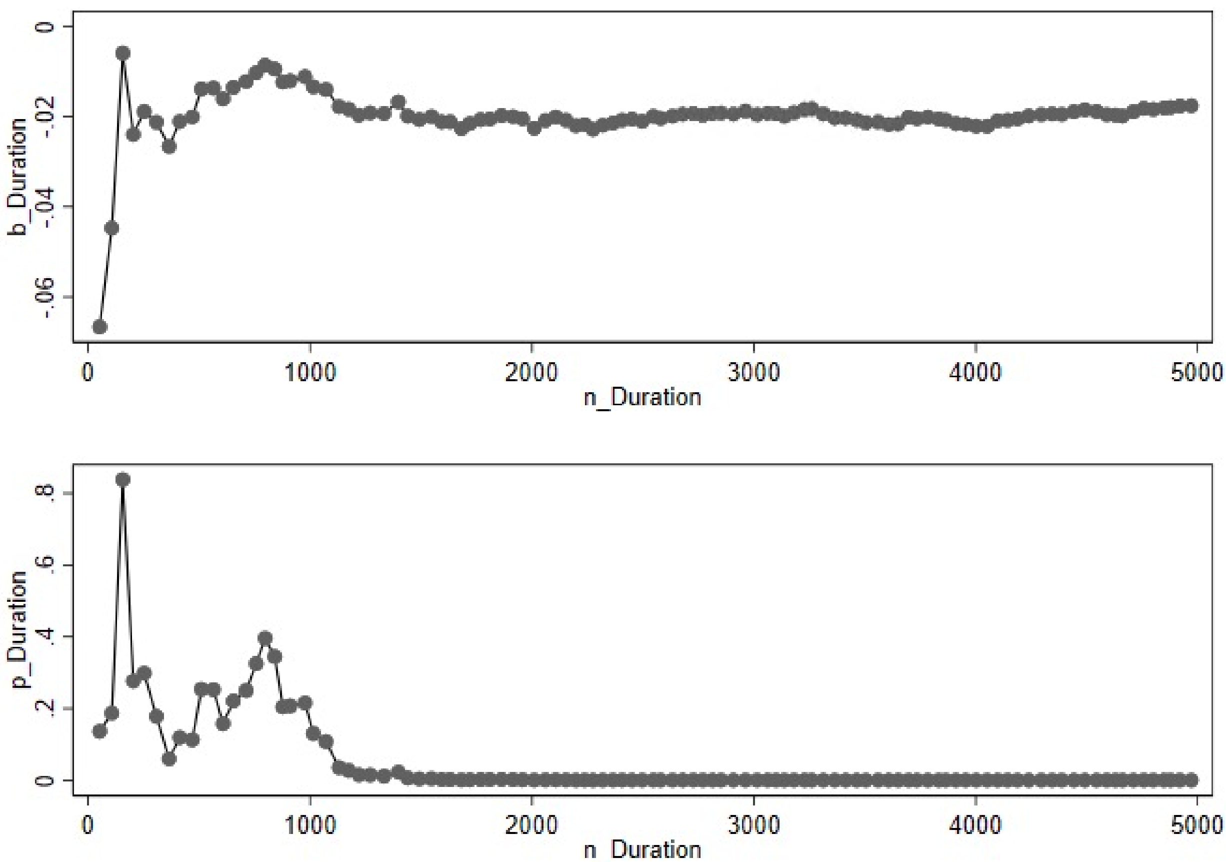


Figure 3.4: CPS Chart for Featured: Coefficient and p-Value vs. Sample Size

This figure summarizes the Coefficient/p-Value Chart for the variable *Featured,* in the OLS regression with *Log Pledged* as dependent variable and "*Competition Int., Membership Tenure, Duration*, *and Log Goal*" as remaining regressors. Year, and Category fixed effects are considered in all models. The upper figure illustrates the coefficient vs. sample size, and the lower figure illustrates the p-value vs. sample size. The figures are zoomed into n < 5000 for illustration.

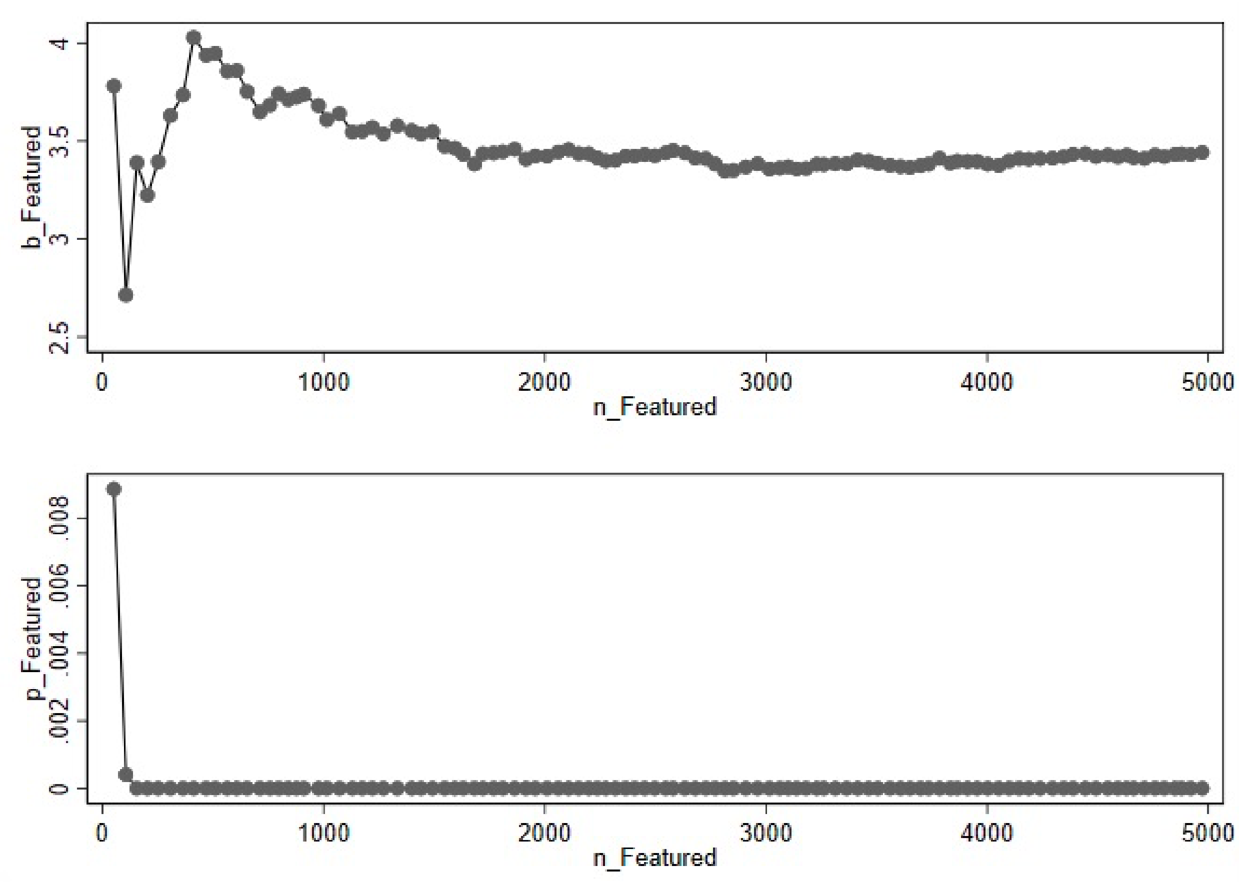


Figure 3.5: CPS Chart for Log Goal: Coefficient and p-Value vs. Sample Size

This figure summarizes the Coefficient/p-Value Chart for the variable *Log Goal,* in the OLS regression with *Log Pledged* as dependent variable and "*Competition Int., Membership Tenure, Duration*, *and Featured*" as remaining regressors. Year, and Category fixed effects are considered in all models. The upper figure illustrates the coefficient vs. sample size, and the lower figure illustrates the p-value vs. sample size. The figures are zoomed into n < 5000 for illustration.

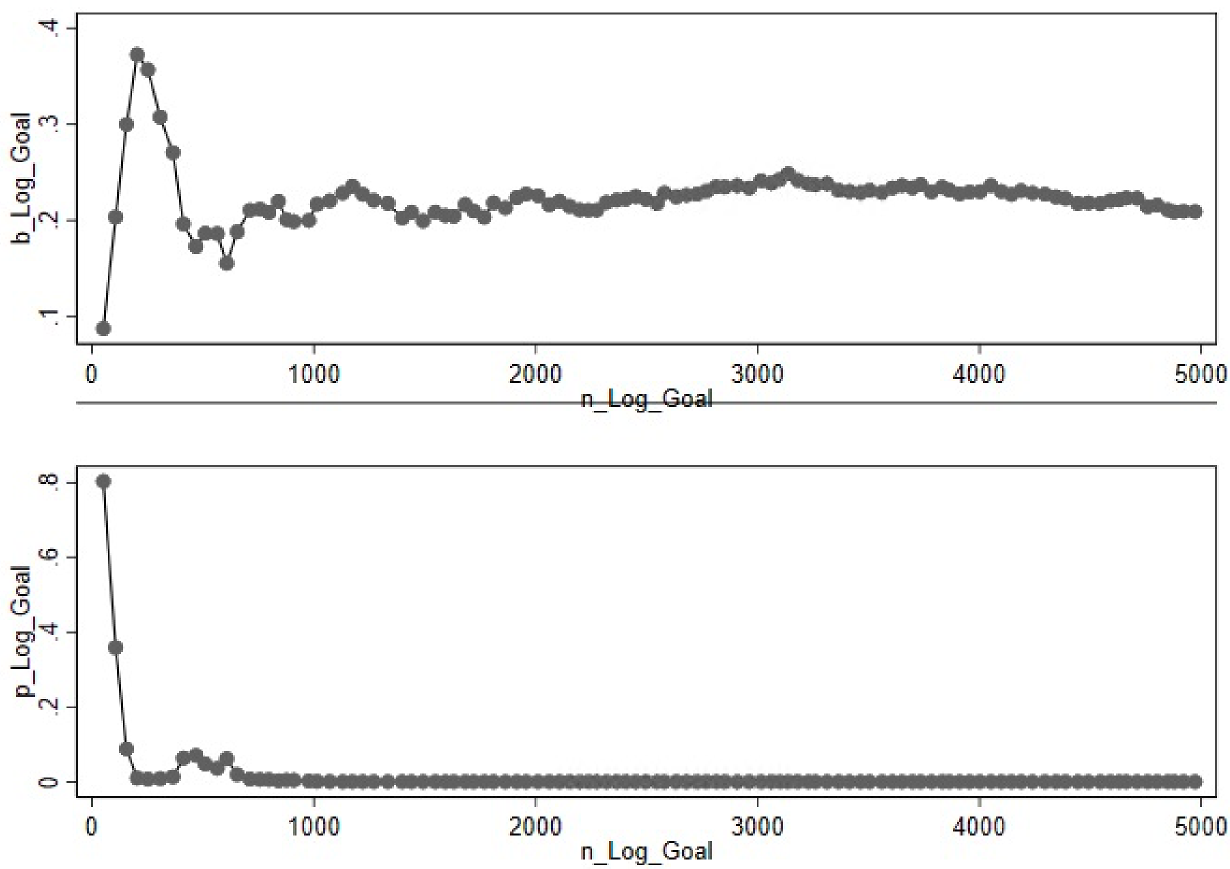


Table 3.1: Variable Definitions

This table gives a detailed description of the data-gathering process and calculation methods for all variables. Panel A includes dependent variables, and, explanatory variables that are used in all regression models. Panel B describe the fixed effect variables. Panel C includes definitions of variables related to creator country and project location. Panel D presents an overview of the Kickstarter data set, and the derivation of the final total sample of 230,255 (funded + failed) campaigns.

Panel A (Variables)

Variable Name	Description and Calculation
	Dependent Variables (Success)
Funded	Dummy variable that equals 1 if the project reached its goal amount, and 0 otherwise.
Log Pledged	Natural logarithm of the (project's pledged amount in USD (regardless of the project's success) + 1).
Log Backers	Natural logarithm of the (project's total number of backers + 1).
	Independent Variables
Competition Int.	Average daily number of projects that were "live" during campaign's lifetime, divided by 1000. This variable is designed to gauge the competition intensity.
Membership Tenure	Number of months between the date creator joined Kickstarter (i.e., created an account) and campaign's start date.
	Control Variables
Duration	Number of days between the campaign's start date and end date.
Featured	Dummy variable that equals 1 if the project is featured as "Projects We Love" by Kickstarter, and 0 otherwise.
Log Goal	Natural logarithm of the project's goal amount in USD.

Panel B (Fixed Effects)

Variable Name	Description and Calculation
	Base Fixed Effects
Category	15 dummy variables (C1, C2, etc.) that corresponds to the main Kickstarter category of the campaign.
Year	6 dummy variables (Y2013, Y2014, etc.) that corresponds to the year that campaign is posted (based on campaign launch date).
	Periodic Fixed Effects based on Launch Date
Month of Year	12 dummy variables (January, February, etc.) for each campaign that corresponds to the month it is posted (based on campaign launch date).
Week of Month	4 dummy variables (W1, W2, W3, and W4) that corresponds to the respective week of the month the campaign is posted.
Month-Year FE	72 dummy variables = 6 years × 12 months; D(2013-Jan), D(2013-Feb), …, D(2018-Dec) that corresponds to the respective year-month the campaign is posted.

Table 3.1: Variable Definitions)—continued

Variable Name	Description and Calculation

Panel C (Creator Country/ Project Location Variables)

	Creator Country
C: "[X]"	Dummy variable that equals 1 if the creator country = [X], and 0 otherwise. According to Kickstarter website, the main project creator (individual legally associated with the campaign) should hold a valid passport of one of the following specific countries, i.e., X can take one of the following 7 values: *1) US; 2) UK; 3) CA; 4) AU/NZ; 5) Asia (main)* including Hong Kong, Japan, and Singapore; *6) EU (main)* including Germany, France, Spain, Italy, Belgium, Luxemburg, Netherlands, Austria, Switzerland, Scandinavia (Denmark, Norway, Sweden), and Ireland; and *7) Mexico*.
	Project Location
L: "[X]"	Dummy variable that equals 1 if the project location = [X], and 0 otherwise. The project location can be anywhere in the world, so X can take one of the following 8 values: *1) US; 2) UK; 3) CA; 4) AU/NZ; 5) Asia (main)* including Hong Kong, Japan, and Singapore; *6) EU (main)* including Germany, France, Spain, Italy, Belgium, Luxemburg, Netherlands, Austria, Switzerland, Scandinavia (Denmark, Norway, Sweden), and Ireland; *7) Mexico*; and *8) Other.*
L: Main	Dummy variable that equals 1 if the project location is one of the acceptable countries of citizenship for project creators (i.e., *1) US; 2) UK; 3) CA; 4) AU/NZ; 5) Asia (main); 6) EU (main); 7) Mexico*), and 0 otherwise.

Panel D (Derivation of Kickstarter Sample)

	Exclusion Criteria	#	**Sub-Total**
	Dataset (*2012-10-01 to 2019-01-31*)	265,001	-
1	End date after Dec. 31, 2018	*4,881*	260,120
2	Launched date before Jan. 1, 2013	*4,204*	255,916
3	Cancelled by Creator	*24,217*	231,699
4	Suspended by Kickstarter	*1,444*	230,255
	Total Sample (Funded + Failed)		**230, 255**

Table 3.2: An Overview of Kickstarter Sample (Category – Year)

This table presents an overview of the Kickstarter sample. Panel A shows the number of launched campaigns (either successful/funded or unsuccessful/failed) within Kickstarter's main categories for each respective year. "Unsuccessful/Failed" is defined as the goal amount not being reached by the end date of the campaign, and "Successful/Funded" is defined as the goal amount being achieved (and neither suspended nor cancelled). Panel B, shows the percentage of successful campaigns within each main category for each respective year. Panel C provides descriptive statistics (mean, standard deviation, min, and max) and Pearson correlation coefficients for all variables considered in the analyses (*indicates statistical significance at least at 5% level). All non-dummy variables in panel C are winsorized at the 1% level on both sides. See Table 3.1, panel A, for variable description and calculation methods.

Panel A

Num.	Main Category	2013	2014	2015	2016	2017	2018	Total (#)
1	Art	2,588	3,499	3,831	2,552	2,781	2,610	17,861
2	Comics	202	951	1,542	1,517	1,662	1,655	7,529
3	Crafts	104	1,019	1,541	1,031	911	661	5,267
4	Dance	56	422	471	334	290	171	1,744
5	Design	947	1,948	2,921	3,641	4,599	3,735	17,791
6	Fashion	359	2,422	3,901	3,225	3,467	3,031	16,405
7	Film & video	5,001	6,281	6,496	4,600	3,725	2,841	28,944
8	Food	220	3,676	4,390	2,730	2,366	1,710	15,092
9	Games	2,127	3,711	5,083	4,852	5,418	5,210	26,401
10	Journalism	58	636	1,103	646	470	311	3,224
11	Music	5,290	5,433	5,932	3,941	3,396	2,456	26,448
12	Photography	260	1,592	1,667	1,077	839	556	5,991
13	Publishing	4,078	5,640	5,934	4,512	4,312	3,216	27,692
14	Technology	1,139	4,360	6,808	5,173	4,567	2,911	24,958
15	Theater	274	1,110	1,322	923	744	535	4,908
	Total	**22,703**	**42,700**	**52,942**	**40,754**	**39,547**	**31,609**	**230,255**

Table 3.2: An Overview of Kickstarter Sample (Category – Year)—*continued*

Panel B

Num.	Main Category	2013	2014	2015	2016	2017	2018	Total (#)
1	Art	0.52	0.39	0.36	0.41	0.48	0.56	0.44
2	Comics	0.60	0.56	0.57	0.64	0.69	0.76	0.65
3	Crafts	0.56	0.20	0.23	0.26	0.28	0.34	0.26
4	Dance	0.61	0.57	0.50	0.64	0.63	0.61	0.58
5	Design	0.44	0.38	0.32	0.44	0.49	0.55	0.45
6	Fashion	0.38	0.29	0.25	0.26	0.34	0.41	0.31
7	Film & video	0.49	0.43	0.36	0.42	0.43	0.47	0.43
8	Food	0.47	0.22	0.21	0.24	0.27	0.30	0.24
9	Games	0.34	0.31	0.34	0.42	0.53	0.60	0.44
10	Journalism	0.43	0.21	0.19	0.20	0.24	0.30	0.22
11	Music	0.60	0.52	0.41	0.47	0.49	0.57	0.51
12	Photography	0.46	0.25	0.28	0.40	0.39	0.48	0.34
13	Publishing	0.44	0.33	0.30	0.37	0.40	0.49	0.38
14	Technology	0.45	0.22	0.21	0.23	0.25	0.27	0.24
15	Theater	0.66	0.59	0.58	0.64	0.61	0.61	0.60
	Total	**0.49**	**0.36**	**0.32**	**0.38**	**0.43**	**0.50**	**0.40**

Panel C

Num.	Variable	# Obs.	Mean	Std. Dev.	Min	Max	1	2	3	4	5	6	7	8
1	Funded	230,255	0.40	0.49	0.00	1.00	1							
2	Log Pledged	230,255	5.89	3.30	0.00	12.05	0.66*	1						
3	Log Backers	230,255	2.77	1.91	0.00	7.57	0.71*	0.93*	1					
4	Competition Int.	230,255	4.60	1.36	2.01	7.67	-0.12*	-0.18*	-0.17*	1				
5	Membership Tenure	230,255	1.51	3.28	0.00	21.53	0.04*	0.14*	0.14*	-0.03	1			
6	Duration	230,255	32.77	11.33	7.00	60.00	-0.15*	-0.06*	-0.08*	-0.02	0.04	1		
7	Featured	230,255	0.10	0.30	0.00	1.00	0.27*	0.34*	0.40*	-0.01	-0.03	0.06	1	
8	Log Goal	230,255	8.57	1.67	4.16	12.72	-0.23*	0.13*	0.11*	-0.01	0.22*	0.12*	0.12*	1

Table 3.3: An Overview of Kickstarter Sample (Creator Country – Project Location)

This table present an overview of Kickstarter sample based on Creator Country as well as project location. Panel A shows the number of launched campaigns (either successful/funded or unsuccessful/failed) within Kickstarter's main countries (creator country) for each respective location (project location). "Unsuccessful/Failed" is defined as the goal amount not being reached by the end date of the campaign, and "Successful/Funded" is defined as the goal amount being achieved (and neither suspended nor cancelled). Panel B, shows the percentage of successful campaigns within each country for each respective location. Panel C [Panel D] provides descriptive statistics (mean, standard deviation, min, and max) for all variables considered in the analyses for creator country [project location] being US vs Non-US sample. All non-dummy variables in panels C, and D, are winsorized at the 1% level on both sides. See Table 3.1, for variable description and calculation method.

Panel A

		Project Location								
Num.	**Creator Country**	**US**	**UK**	**CA**	**AU/NZ**	**Asia (main)**	**EU (main)**	**Mexico**	**Other**	**Total (#)**
1	United States	139,005	349	386	170	532	1,231	220	3,626	145,519
2	United Kingdom	183	32,153	18	42	97	492	14	884	33,883
3	Canada	267	16	13,696	11	68	83	16	287	14,444
4	Australia & NZ	88	31	14	8,659	83	41	6	251	9,173
	Asia (main):									
5	Hong Kong	13	5	11	1	898	15	0	35	978
6	Japan	5	1		0	246	1	0	8	261
7	Singapore	18	2		0	751	2	0	34	807
	EU (main):									
8	Germany	37	26	3	10	11	4,091	2	225	4,405
9	France	31	19	4	6	16	3,105	4	120	3,305
10	Spain	25	19		0	4	2,411	11	92	2,562
11	Italy	40	23	1	3	5	2,971	4	87	3,134
12	Belgium, LU & NL	62	14	3	5	15	3,233	4	162	3,498
13	Austria	7	2	0	0	1	587	1	41	639
14	Switzerland	10	4	0	0	4	614	1	34	667
15	Scandinavia & IE	68	38	1	7	12	4,260	3	164	4,553
16	Mexico	16	2	3	0	2	12	2,378	14	2,427
	Total	**139,875**	**32,704**	**14,140**	**8,914**	**2,745**	**23,149**	**2,664**	**6,064**	**230,255**

Table 3.3: An Overview of Kickstarter Sample (Creator Country – Project Location)—*continued*

Panel B

		Project Location								
Num.	**Creator Country**	**US**	**UK**	**CA**	**AU/NZ**	**Asia (main)**	**EU (main)**	**Mexico**	**Other**	**Total (#)**
1	United States	0.41	0.55	0.63	0.53	0.59	0.56	0.45	0.47	0.41
2	United Kingdom	0.38	0.44	0.39	0.38	0.44	0.42	0.50	0.32	0.43
3	Canada	0.35	0.56	0.36	0.18	0.37	0.33	0.38	0.31	0.36
4	Australia & NZ	0.28	0.35	0.36	0.33	0.37	0.39	0.17	0.42	0.33
	Asia (main):									
5	Hong Kong	0.85	0.00	0.55	1.00	0.53	0.53	-	0.63	0.54
6	Japan	0.40	1.00	-	-	0.47	1.00	-	0.60	0.48
7	Singapore	0.61	0.50	-	-	0.41	0.00	-	0.33	0.41
	EU (main):									
8	Germany	0.22	0.19	0.33	-	0.18	0.30	0.50	0.32	0.30
9	France	0.42	0.37	0.25	0.17	0.50	0.38	0.75	0.32	0.38
10	Spain	0.20	0.37	-	-	0.25	0.30	0.36	0.30	0.30
11	Italy	0.13	0.09	0.00	0.33	0.20	0.21	0.75	0.31	0.21
12	Belgium, LU & NL	0.24	0.50	0.33	0.00	0.33	0.28	0.25	0.30	0.28
13	Austria	0.29	0.50	-	-	1.00	0.28	1.00	0.21	0.28
14	Switzerland	0.10	0.25	-	-	0.25	0.33	0.00	0.15	0.31
15	Scandinavia & IE	0.15	0.50	0.00	0.14	0.33	0.36	0.33	0.32	0.36
16	Mexico	0.31	0.00	0.00	-	0.50	0.33	0.30	0.29	0.30
	Total	**0.41**	**0.44**	**0.37**	**0.33**	**0.49**	**0.33**	**0.31**	**0.41**	**0.40**

Table 3.3: An Overview of Kickstarter Sample (Creator Country – Project Location)—*continued*

Panel C

		Creator Country = US					Creator Country = Non-US				
Num.	**Variable**	**# Obs.**	**Mean**	**Std. Dev.**	**Min**	**Max**	**# Obs.**	**Mean**	**Std. Dev.**	**Min**	**Max**
1	Funded	145,519	0.41	0.49	0.00	1.00	84,736	0.37	0.48	0.00	1.00
2	Log Pledged	145,519	6.01	3.31	0.00	12.05	84,736	5.68	3.29	0.00	12.05
3	Log Backers	145,519	2.84	1.93	0.00	7.57	84,736	2.65	1.87	0.00	7.57
4	Competition Int.	145,519	4.60	1.42	2.01	7.67	84,736	4.58	1.25	2.01	7.67
5	Membership Tenure	145,519	1.66	3.57	0.00	21.53	84,736	1.24	2.70	0.00	21.53
6	Duration	145,519	32.66	11.23	7.00	60.00	84,736	32.96	11.50	7.00	60.00
7	Featured	145,519	0.10	0.30	0.00	1.00	84,736	0.09	0.29	0.00	1.00
8	Log Goal	145,519	8.61	1.65	4.16	12.72	84,736	8.50	1.72	4.16	12.72

Panel D

		Project Location = US					Project Location = Non-US				
Num.	**Variable**	**# Obs.**	**Mean**	**Std. Dev.**	**Min**	**Max**	**# Obs.**	**Mean**	**Std. Dev.**	**Min**	**Max**
1	Funded	139,875	0.41	0.49	0.00	1.00	90,380	0.38	0.49	0.00	1.00
2	Log Pledged	139,875	5.95	3.30	0.00	12.05	90,380	5.80	3.31	0.00	12.05
3	Log Backers	139,875	2.80	1.92	0.00	7.57	90,380	2.72	1.90	0.00	7.57
4	Competition Int.	139,875	4.61	1.41	2.01	7.67	90,380	4.57	1.27	2.01	7.67
5	Membership Tenure	139,875	1.66	3.58	0.00	21.53	90,380	1.28	2.76	0.00	21.53
6	Duration	139,875	32.66	11.25	7.00	60.00	90,380	32.95	11.46	7.00	60.00
7	Featured	139,875	0.10	0.30	0.00	1.00	90,380	0.10	0.30	0.00	1.00
8	Log Goal	139,875	8.59	1.65	4.16	12.72	90,380	8.53	1.71	4.16	12.72

Table 3.4: Multivariate Analysis of Main Determinants of Campaign Success

In this table, we analyze the determinants of *Success* measured by *Funded* (logistic regression; coefficients are the logs of the odds ratios), *Log Pledged* (OLS regressions), and *Log Backers* (OLS regressions). In all panels, specifications (1)-(3) [specifications (4)-(6)] are based on a sample that includes all Kickstarter campaigns [campaigns with goal amounts of at least $10,000] that were posted on or after January 1, 2013, and ended before December 31, 2018. Investigating the variance inflation factors (VIFs) reveals no evidence of multicollinearity, given the mean VIF is 1.01 to 1.04 in all models and all individual values are well below the critical value of 5 (see Kutner et al., 2005). All non-dummy variables are winsorized at the 1% level on both sides (see Table 3.1, panel A, for variable descriptions and calculation methods). Panel A includes main category, and year fixed effects. Panel B includes main category, year, month of year (January to December), and week of month (first week to fourth week of respective month) fixed effects. The time fixed effects are all based on the campaign launch date. Robust standard errors are one-way-clustered by campaign category. *t*-statistics are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	Funded	**Log Pledged**	**Log Backers**	**Funded**	**Log Pledged**	**Log Backers**
	Total Sample			Goal Amount > 9,999 USD		
Competition Int.	-0.151***	-0.283***	-0.137***	-0.097***	-0.271***	-0.129***
	(-25.95)	(-38.60)	(-33.68)	(-8.40)	(-19.70)	(-16.72)
Membership Tenure	0.037***	0.097***	0.050***	0.029***	0.106***	0.052***
	(25.57)	(51.10)	(47.77)	(13.10)	(35.48)	(31.14)
Duration	-0.021***	-0.022***	-0.013***	-0.020***	-0.020***	-0.012***
	(-45.95)	(-39.06)	(-41.41)	(-22.54)	(-19.86)	(-21.36)
Featured	2.428***	3.415***	2.376***	2.646***	4.170***	2.872***
	(130.15)	(163.52)	(204.54)	(105.94)	(130.74)	(160.65)
Log Goal	-0.364***	0.190***	0.081***	-0.703***	-0.438***	-0.268***
	(-110.19)	(47.45)	(36.41)	(-56.56)	(-34.44)	(-37.50)
Constant	3.639***	5.758***	2.720***	6.468***	11.710***	6.031***
	(95.85)	(121.82)	(103.45)	(47.07)	(79.23)	(72.80)
Category FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	230,255	230,255	230,255	86,399	86,399	86,399
Mean VIF	1.04	1.04	1.04	1.01	1.01	1.01
Maximum VIF	1.09	1.09	1.09	1.03	1.03	1.03
Adjusted R^2		0.208	0.267		0.265	0.333
Pseudo R^2	0.161			0.210		

Table 3.4: Multivariate Analysis of Main Determinants of Campaign Success—*continued*

Panel B

	(1)	(2)	(3)	(4)	(5)	(6)
	Funded	**Log Pledged**	**Log Backers**	**Funded**	**Log Pledged**	**Log Backers**
	Total Sample			Goal Amount > 9,999 USD		
Competition Int.	-0.185***	-0.335***	-0.169***	-0.144***	-0.339***	-0.172***
	(-29.59)	(-42.54)	(-38.48)	(-11.57)	(-22.79)	(-20.61)
Membership Tenure	0.036***	0.095***	0.049***	0.029***	0.104***	0.051***
	(25.06)	(50.43)	(47.08)	(12.74)	(35.01)	(30.65)
Duration	-0.021***	-0.022***	-0.013***	-0.020***	-0.020***	-0.012***
	(-45.73)	(-38.68)	(-41.13)	(-22.40)	(-19.59)	(-21.14)
Featured	2.428***	3.405***	2.371***	2.642***	4.147***	2.859***
	(130.01)	(163.32)	(204.39)	(105.59)	(130.23)	(160.18)
Log Goal	-0.368***	0.185***	0.079***	-0.705***	-0.436***	-0.266***
	(-110.92)	(46.31)	(35.34)	(-56.60)	(-34.32)	(-37.41)
Constant	3.601***	5.649***	2.691***	6.393***	11.457***	5.933***
	(86.45)	(108.44)	(92.85)	(44.92)	(74.38)	(68.71)
Month of Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Week of Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Category FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	230,255	230,255	230,255	86,399	86,399	86,399
Adjusted R^2		0.211	0.270		0.269	0.337
Pseudo R^2	0.162			0.212		

Table 3.5: Multivariate Analysis of the Effect of Creator Country on Campaign Success

In this table, we analyze the effect of creator country on *Success* measures of crowdfunding campaigns proxied by *Funded* (logistic regression; coefficients are the logs of the odds ratios), *Log Pledged* (OLS regressions), and *Log Backers* (OLS regressions). specifications (1)-(3) [specifications (4)-(6)] are based on a sample that includes all Kickstarter campaigns [campaigns with goal amounts of at least $10,000] that were posted on or after January 1, 2013, and ended before December 31, 2018. All non-dummy variables are winsorized at the 1% level on both sides (see Table 3.1, for variable descriptions and calculation methods). The reference group is *C: US* and all models include main category, and year fixed effects. Robust standard errors are one-way-clustered by campaign category. *t*-statistics are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Funded	**Log Pledged**	**Log Backers**	**Funded**	**Log Pledged**	**Log Backers**
	Total Sample			Goal Amount > 9,999 USD		
C: UK	-0.068***	-0.114***	-0.080***	-0.337***	-0.258***	-0.228***
	(-4.91)	(-6.41)	(-8.08)	(-11.72)	(-7.39)	(-11.65)
C: CA	-0.182***	-0.305***	-0.145***	-0.404***	-0.454***	-0.250***
	(-8.98)	(-11.90)	(-10.14)	(-9.49)	(-9.20)	(-9.06)
C: AU/NZ	-0.264***	-0.321***	-0.154***	-0.466***	-0.429***	-0.250***
	(-10.41)	(-10.12)	(-8.72)	(-8.95)	(-7.20)	(-7.49)
C: Asia (main)	0.099**	0.290***	0.102***	0.109	0.475***	0.083
	(2.01)	(4.38)	(2.78)	(1.30)	(4.01)	(1.24)
C: EU (main)	-0.423***	-0.541***	-0.343***	-0.543***	-0.473***	-0.369***
	(-24.15)	(-25.19)	(-28.75)	(-17.77)	(-13.25)	(-18.46)
C: Mexico	-1.209***	-1.959***	-1.014***	-1.722***	-2.730***	-1.413***
	(-23.49)	(-32.22)	(-29.96)	(-9.72)	(-16.80)	(-15.53)
Competition Int.	-0.155***	-0.290***	-0.142***	-0.104***	-0.279***	-0.135***
	(-26.65)	(-39.64)	(-34.85)	(-8.97)	(-20.28)	(-17.56)
Membership Tenure	0.034***	0.092***	0.048***	0.026***	0.102***	0.049***
	(23.57)	(48.84)	(45.38)	(11.54)	(34.15)	(29.44)
Duration	-0.020***	-0.020***	-0.012***	-0.020***	-0.019***	-0.012***
	(-43.98)	(-36.66)	(-39.09)	(-22.15)	(-19.14)	(-20.77)
Featured	2.447***	3.428***	2.383***	2.661***	4.166***	2.870***
	(130.75)	(164.65)	(205.76)	(106.07)	(130.97)	(161.07)
Log Goal	-0.369***	0.185***	0.079***	-0.689***	-0.426***	-0.258***
	(-110.86)	(46.22)	(35.32)	(-55.41)	(-33.54)	(-36.25)
Constant	3.689***	5.814***	2.753***	6.422***	11.664***	6.003***
	(95.95)	(122.31)	(104.10)	(46.73)	(79.11)	(72.69)
Category FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	230,255	230,255	230,255	86,399	86,399	86,399
Adjusted R^2		0.214	0.273		0.270	0.339
Pseudo R^2	0.165			0.216		

Table 3.6: Multivariate Analysis of the Effect of Project Location on Campaign Success

In this table, we analyze the effect of project location on *Success* measures of crowdfunding campaigns proxied by *Funded* (logistic regression; coefficients are the logs of the odds ratios), *Log Pledged* (OLS regressions), and *Log Backers* (OLS regressions). Specifications (1)-(3) [specifications (4)-(6)] are based on a sample that includes all Kickstarter campaigns [campaigns with goal amounts of at least $10,000] that were posted on or after January 1, 2013, and ended before December 31, 2018. All non-dummy variables are winsorized at the 1% level on both sides (see Table 3.1, for variable descriptions and calculation methods). The reference group is *L: US* and all models include main category, and year fixed effects. Robust standard errors are one-way-clustered by campaign category. *t*-statistics are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Funded	**Log Pledged**	**Log Backers**	**Funded**	**Log Pledged**	**Log Backers**
	Total Sample			Goal Amount > 9,999 USD		
L: UK	-0.032**	-0.069***	-0.049***	-0.280***	-0.199***	-0.182***
	(-2.27)	(-3.80)	(-4.84)	(-9.53)	(-5.54)	(-9.05)
L: CA	-0.129***	-0.220***	-0.091***	-0.278***	-0.283***	-0.140***
	(-6.30)	(-8.49)	(-6.32)	(-6.64)	(-5.69)	(-5.02)
L: AU/NZ	-0.239***	-0.256***	-0.117***	-0.390***	-0.293***	-0.163***
	(-9.29)	(-7.96)	(-6.53)	(-7.47)	(-4.85)	(-4.83)
L: Asia (main)	0.265***	0.620***	0.331***	0.416***	0.971***	0.459***
	(6.18)	(10.88)	(10.45)	(6.02)	(9.90)	(8.35)
L: EU (main)	-0.320***	-0.378***	-0.247***	-0.385***	-0.264***	-0.230***
	(-18.59)	(-17.77)	(-20.90)	(-13.03)	(-7.48)	(-11.61)
L: Mexico	-1.043***	-1.710***	-0.888***	-1.324***	-2.154***	-1.119***
	(-21.49)	(-29.46)	(-27.50)	(-9.29)	(-14.94)	(-13.85)
L: Other	0.114***	0.274***	0.143***	0.086*	0.276***	0.177***
	(3.78)	(7.08)	(6.63)	(1.78)	(4.39)	(5.02)
Competition Int.	-0.154***	-0.287***	-0.140***	-0.102***	-0.275***	-0.133***
	(-26.45)	(-39.26)	(-34.44)	(-8.87)	(-19.99)	(-17.22)
Membership Tenure	0.035***	0.094***	0.049***	0.027***	0.103***	0.050***
	(24.14)	(49.58)	(46.17)	(12.09)	(34.77)	(30.20)
Duration	-0.020***	-0.021***	-0.012***	-0.020***	-0.020***	-0.012***
	(-44.41)	(-37.24)	(-39.67)	(-22.35)	(-19.45)	(-21.10)
Featured	2.447***	3.430***	2.385***	2.662***	4.171***	2.874***
	(130.75)	(164.62)	(205.73)	(106.11)	(130.99)	(161.07)
Log Goal	-0.368***	0.185***	0.079***	-0.692***	-0.432***	-0.262***
	(-110.64)	(46.21)	(35.36)	(-55.65)	(-33.94)	(-36.70)
Constant	3.680***	5.806***	2.747***	6.449***	11.692***	6.020***
	(95.75)	(122.02)	(103.76)	(46.92)	(79.26)	(72.83)
Category FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	230,255	230,255	230,255	86,399	86,399	86,399
Adjusted R^2		0.213	0.272		0.269	0.337
Pseudo R^2	0.164			0.214		

Table 3.7: Multivariate Analysis of Main Determinants of Campaign Success Considering Creator Country and Project Location Interaction

In this table, we analyze the effect of creator country/project location being US and their interaction term on *Success* measures of crowdfunding campaigns proxied by *Funded* (logistic regression; coefficients are the logs of the odds ratios), *Log Pledged* (OLS regressions), and *Log Backers* (OLS regressions). Specifications (1)-(3) [specifications (4)-(6)] are based on a sample that includes all Kickstarter campaigns [campaigns with goal amounts of at least $10,000] that were posted on or after January 1, 2013, and ended before December 31, 2018. All non-dummy variables are winsorized at the 1% level on both sides (see Table 3.1, for variable descriptions and calculation methods). All models include main category, and year fixed effects. Robust standard errors are one-way-clustered by campaign category. *t*-statistics are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A

	(1)	(2)	(3)	(4)	(5)	(6)
	Funded	**Log Pledged**	**Log Backers**	**Funded**	**Log Pledged**	**Log Backers**
	Total Sample			Goal Amount > 9,999 USD		
C: US	0.679***	1.098***	0.661***	0.957***	1.278***	0.895***
	(23.29)	(28.89)	(31.26)	(21.37)	(21.21)	(26.56)
L: US	-0.093	-0.481***	-0.242***	0.002	-0.769***	-0.336***
	(-1.13)	(-4.81)	(-4.36)	(0.01)	(-4.56)	(-3.56)
C: US * L: US	-0.388***	-0.339***	-0.255***	-0.567***	-0.178	-0.312***
	(-4.46)	(-3.17)	(-4.29)	(-3.60)	(-0.99)	(-3.12)
Competition Int.	-0.153***	-0.286***	-0.139***	-0.100***	-0.274***	-0.131***
	(-26.30)	(-39.07)	(-34.17)	(-8.68)	(-19.92)	(-17.05)
Membership Tenure	0.035***	0.094***	0.049***	0.026***	0.102***	0.050***
	(24.26)	(49.59)	(46.19)	(11.78)	(34.47)	(29.85)
Duration	-0.021***	-0.022***	-0.013***	-0.020***	-0.020***	-0.012***
	(-45.47)	(-38.56)	(-40.91)	(-22.63)	(-19.92)	(-21.47)
Featured	2.425***	3.396***	2.365***	2.643***	4.132***	2.846***
	(129.82)	(162.90)	(203.98)	(105.48)	(129.73)	(159.71)
Log Goal	-0.370***	0.183***	0.077***	-0.692***	-0.427***	-0.260***
	(-111.51)	(45.68)	(34.49)	(-55.61)	(-33.60)	(-36.49)
Constant	3.517***	5.560***	2.604***	6.037***	11.301***	5.732***
	(91.46)	(116.25)	(97.89)	(43.45)	(75.69)	(68.62)
Category FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	230,255	230,255	230,255	86,399	86,399	86,399
Adjusted R^2		0.212	0.271		0.270	0.340
Pseudo R^2	0.163			0.216		

Table 3.7: Multivariate Analysis of Main Determinants of Campaign Success Considering Creator Country and Project Location Interaction—*continued*

Panel B

	(1)	(2)	(3)	(4)	(5)	(6)
	Funded	**Log Pledged**	**Log Backers**	**Funded**	**Log Pledged**	**Log Backers**
	Total Sample			Goal Amount > 9,999 USD		
C: US	0.688***	0.985***	0.613***	0.937***	1.138***	0.774***
	(11.24)	(12.81)	(14.34)	(8.82)	(8.93)	(10.84)
L: Main	0.113**	0.037	0.062*	0.170*	0.126	0.074
	(2.30)	(0.61)	(1.84)	(1.86)	(1.21)	(1.27)
C: US * L: Main	-0.479***	-0.682***	-0.437***	-0.526***	-0.767***	-0.500***
	(-7.74)	(-8.75)	(-10.07)	(-4.87)	(-5.91)	(-6.89)
Competition Int.	-0.154***	-0.287***	-0.140***	-0.102***	-0.277***	-0.133***
	(-26.44)	(-39.25)	(-34.37)	(-8.88)	(-20.16)	(-17.33)
Membership Tenure	0.035***	0.094***	0.049***	0.026***	0.102***	0.050***
	(24.23)	(49.55)	(46.13)	(11.72)	(34.38)	(29.74)
Duration	-0.021***	-0.022***	-0.013***	-0.020***	-0.020***	-0.012***
	(-45.48)	(-38.58)	(-40.92)	(-22.62)	(-19.93)	(-21.47)
Featured	2.431***	3.410***	2.373***	2.653***	4.158***	2.863***
	(130.18)	(163.53)	(204.64)	(105.93)	(130.54)	(160.61)
Log Goal	-0.369***	0.185***	0.078***	-0.688***	-0.425***	-0.258***
	(-111.14)	(46.08)	(34.98)	(-55.42)	(-33.43)	(-36.23)
Constant	3.405***	5.528***	2.546***	5.846***	11.178***	5.660***
	(55.10)	(72.54)	(60.06)	(35.57)	(62.07)	(56.15)
Category FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	230,255	230,255	230,255	86,399	86,399	86,399
Adjusted R^2		0.211	0.270		0.268	0.338
Pseudo R^2	0.163			0.215		

Appendix

Table 3.A1: Dataset Description

This table gives a detailed description of the variables included in the "Kickstarter Dataset (2013-2018).dta" available for future research upon request. Panel A includes description of base variables, and Panel B includes calculation methods of variables considered in the study. Dataset includes information on 230,255 (funded + failed) campaigns that were posted on or after January 1, 2013, and ended before December 31, 2018 (i.e., projects cancelled by creator, and projects suspended by Kickstarter are excluded from the sample). The data is sorted based on campaign's launch date, and end date. All non-dummy *study variables* presented in Panel B are winsorized at the 1% level on both sides and then used in the analyses.

Variable Name	**Description and Calculation**
Panel A (Base Variables)	
	General Campaign Information
ID	Unique Identifier of each campaign in the dataset ranging from 1 to 230,255.
Campaign_Title	Title of the Kickstarter Campaign.
Main_Category	Campaign's category within Kickstarter's main 15 categories.
Sub_Category	Campaign's subcategory within the main category.
	Campaign Details
Creator_Country	Country of the citizenship for the individual legally entitled to campaign.
Project_Location	The location of the project (city, [state], country).
Launch_Date	Campaign's start date.
End_Date	Campaign's end-date.
Creator_Profile_Name	Profile name of the campaign creator.
Creator_Account_Date	The date creator joined platform.
Goal_Original_Currency	Goal amount in the original currency.
Pledged_Original_Currency	Pledged amount in the original currency.
Currency	Original currency.
Static_USD_Rate	The USD exchange rate used by Kickstarter website to show values in the local currency for specific campaign. This rate is used in order to calculate *usd-goal*, and *usd-pledged*.
usd_goal	Campaign's goal amount in USD.
usd_pledged	Campaign's pledged amount in USD.
Number_of_Backers	Campaign's number of total backers.

Table 3.A1: Dataset Description—*continued*

Panel B (Study Variables)

	Dependent Variables (Success)
Funded	Dummy variable that equals 1 if the project reached its goal amount, and 0 otherwise.
Log_Pledged	Natural logarithm of the (project's pledged amount in USD (regardless of the project's success) + 1).
Log_Backers	Natural logarithm of the (project's total number of backers + 1).
	Independent Variables
Competition Int.	Average daily number of projects that were "live" during campaign's lifetime, divided by 1000. This variable is designed to gauge competition intensity.
Membership_Tenure	Number of months between the date creator joined Kickstarter (i.e., created an account) and campaign's start date.
	Control Variables
Duration	Number of days between the campaign's start date and end date.
Featured	Dummy variable that equals 1 if the project is featured as "Projects We Love" by Kickstarter, and 0 otherwise.
Log_Goal	Natural logarithm of the project's goal amount in USD.

Chapter 4: Facial Masculinity, Testosterone, and Financial Decision-Making: Investors vs. Entrepreneurs

Moein Karami, Denis Schweizer

John Molson School of Business, Concordia University, Montreal, Canada

Abstract

The effect of biological factors on behavioral traits is forming the basis for a new strand of research that incorporates human biology into the study of managerial and investment behavior. Recently, there is also a significant increase in the number of papers on the biological factors affecting entrepreneurship. Previous research show that face structure, in general, and facial Width-to-Height Ratio (fWHR), in particular, are associated with various personality traits in men, such as aggression and status-striving. Moreover, many studies show that there is a positive correlation between fWHR and testosterone; a hormone which its role in describing behavioral patterns such as egocentrism, competitiveness and risk-taking is well-established. First, and focusing on several proxies for hedge funds' risk and performance (1994-2016), a significant positive relationship between fWHR of the hedge fund managers in the sample and their fund's risk is documented. The association between facial masculinity of male entrepreneurs and their fund-raising outcome is also investigated using a sample of ABC channel's "Shark Tank" show (2009-2014) where selected entrepreneurs and angel investors negotiate for a short period of time to make a deal and come to terms. The current study sheds light on the factors that are not incorporated in economic models, but may significantly affect financial risk-taking and performance, as well as entrepreneurial outcomes.

4.1. Introduction

Economic models assume rationality and utility maximization in explaining financial decisions of individuals. However, this assumption has been challenged by historical data in many different contexts. In response to the observed financial decision-making process, which is sometimes not strictly rational, behavioral models come into the picture. Many behavioral economic models aim to find psychological, physiological and hormonal reasons to explain individual's financial performance and decision-making process.

Previous studies have shown the capability of facial characteristics to influence personality traits in men (Lewis et al., 2012). Relatedly, many studies show a positive link between facial metrics and the level of testosterone in men; a hormone that is considered to be associated with a wide range of behavioral traits (see e.g., Lefevre et al., 2013). Testosterone level is negatively associated with self-regulation and impulse control. Furthermore, high levels of testosterone results in reducing fear, increasing the willingness to engage in competition, aggression and risk-taking (Zuckerman & Kuhlman, 2000). This study, in the broadest sense, is related to a growing literature on how biochemical factors in general, and face structure, in particular, influence behavioral traits in men (e.g., Carre and McCormick, 2008; Carre et al., 2009; Lefevre et al., 2013; Lewis et al., 2012).

A recent strand of literature focuses on the effect of physiological factors on financial decision-making, risk-taking, economic behavior and corporate success (Barber and Odean, 2001; Coates and Herbert, 2008; Coates et al., 2009; Jia et al., 2014; Sapienza et al., 2009; and Wong et al., 2011). There are also multiple studies that aim at shedding light on the biological underpinnings of entrepreneurship (e.g., Shane and Nicolaou, 2015). Research in this area has investigated the role of genetics (Nicolaou et al., 2008), physiology (White et al., 2006) and biology (Nicolaou and Shane, 2014) in entrepreneurial activities, and have examined the effect of hormones such as testosterone (Nicolaou et al., 2018; Unger et al., 2015; White et al., 2006) on entrepreneurship. In sum, the role of biology in management is forming the basis for a new school of thought that incorporates human biology into the study of managerial behavior (Nofal et al., 2018).

In the first part of our study, we examine the effect of facial Width-to-Height Ratio (fWHR) on the risk-taking behavior of male hedge fund managers. Although hedge fund industry is a well-established business sector, but the question that what drives the performance and investment behavior of hedge funds, is still not fully answered. Moreover, hedge funds are different from other institutional investors, since they have more incentive to generate higher returns because they receive significant proportion of excess returns as incentive fees, while the compensation structure of mutual funds or pension funds usually does not permit managers to claim such significant fees (Partnoy, 2015). Furthermore, hedge funds, as delegated portfolio managers, are competing for investor flow, i.e., their survival depends on the approval of the investors who finance them. Such "flow-performance" relationship provides stronger incentives to make profit (Burkart and Dasgupta, 2014). Less regulatory constraints, more concentrated portfolios, fewer conflicts of interest and the ability to impose restrictions on investor redemptions (which gives them time to pursue their investment strategies) are some other features of hedge funds that make them an interesting sample for this study.

Focusing on both risk (e.g., standard deviation of monthly returns) and performance (e.g., annualized average HF return) measures, we aim to detect the relationship between fWHR and HF managers' risk-taking behavior, controlling for fund's relevant characteristics (e.g., management fee, incentive fee, fund age, etc.). This link is extremely interesting for academia, because HFs are rumored to be holding system-relevant risks, potentially endangering the financial system in case of crises. We extend the literature on previously documented links between facial metrics and behavioral traits in a behavioral finance context. More importantly, this study is one of the first to provide evidence on how testosterone level (approximated by the fWHR measure) of "skilled investors" is related to risk-taking behavior and subsequent performance (see Lu and Teo, 2021; for a similar study).

Using CISDM Morningstar data base from 1994 to 2016, and deploying a methodical search for HF managers' names and photos, we show that, overall, HF managers with higher facial masculinity, manage more risky funds. Moreover, we show that funds under their management tend to be riskier than their peers' with a similar investment strategy. We also show that although the higher fWHR managers take more risk on average, it does not necessarily result in a higher return profile. We conclude that HF mangers with higher fWHR have lower risk-adjusted performance (proxied by e.g., Sharpe ratio) in our sample.

In the second part of our study, and in an aim to investigate the effect of facial masculinity on entrepreneurial outcomes, we analyze data from the most public, high-stakes pitch competition in the USA: ABC channel's Shark Tank show. We construct a sample comprising all firms that have at least one male entrepreneur in the management team, and have aired through 2009-2014 during the first five seasons of the show. Shark Tank is a pitch competition couched in a TV show that first aired on August 9, 2009 and the show's premise is one where entrepreneur(s) contestants pitch to a panel of five judges (known as "sharks") who potentially make competing offers (Smith and Viceisza, 2018). This show is an interesting environment to test our hypothesis, since it provides a high-stake competition observable to general public and potential future investors, and the funding outcome largely depends on bargaining skills of the entrepreneur. Our results show that male entrepreneurs with higher levels of fWHR are more successful in coming to terms with the sharks and make a deal[1].

The remainder of this study is organized as follows. In section 4.2, a review of the literature is presented, and hypotheses are developed. Section 4.3 provides an overview of the data set, sample and methodology. Results are reported in section 4.4, and section 4.5 concludes.

4.2. Literature Review and Hypotheses

Sell et al. (2009) argue that testosterone is correlated with the perception of masculinity shown on a photograph. Facial width-to-height ratio (fWHR) is a novel measure of facial masculinity which is defined as the ratio between the bizygomatic width and the height of the upper face (see e.g., Carre and McCormick, 2008). Lefevre et al. (2013) examine the relationship between testosterone levels and fWHR. They conclude that facial width (scaled by two measures of facial height) is positively associated with testosterone levels.

Recently, fWHR has gained a lot of academic attention since it is a more feasible measure compared to other proxies of testosterone level which mostly need laboratory conditions. Moreover, regardless of its association with testosterone, facial masculinity is believed to have effect on numerous personality traits including status-striving (Lewis et al., 2012), aggression (Carre and McCormick, 2008) and trustworthiness (Jia et al., 2014). For example, Jia et al. (2014) study the effect of facial masculinity on financial misreporting and find out that a CEO's facial

[1] Due to our data limitations, we have only focused on male investors/entrepreneurs in our analyses given that our aim is to investigate behavioral traits affecting financial decision-making and outcomes that are linked to testosterone hormone, which could be proxied for using fWHR for males, according to previous literature. Future research can shed more light on the issue by including females in the sample and examining the effect of different hormonal/physiological factors that could be proxied for, in a feasible way, for both male and female investors/entrepreneurs.

masculinity predicts his firm's likelihood of being subject to an SEC enforcement action. Furthermore, Wong et al. (2011) use a sample of 55 CEOs of publicly traded Fortune 500 companies, and report that higher facial masculinity of CEO is positively related to firm's financial performance, as well as CEO's leadership effectiveness. In what follows, we provide a brief literature review on the relationship between testosterone and investment behavior, as well as its link to entrepreneurial activities and develop our hypotheses.

4.2.1. Testosterone and Investment Behavior

Barber and Odean (2001) analyze the common stock investments of men and women and document that men trade more excessively which is a sign that men are more overconfident than women in financial matters; an overconfidence that leads to more risk-taking behavior. A potential cause for gender differences in investment behavior is hormonal differences. By studying a small sample of male traders on a London trading floor, Coates and Herbert (2008) conclude that higher testosterone may contribute to economic return. Apicella et al. (2008) examine the specific relationship between testosterone and financial risk preferences in a sample of 98 men. They conclude that risk-taking in an investment game designed for the purpose of the study, is positively correlated with testosterone levels.

Coates et al. (2009) point out that successful trading in financial markets requires more than correct valuation of securities and a successful trader needs to have enough confidence to place their bets high enough in order to overperform their competitors. They show that higher levels of prenatal androgens (proxied by 2D:4D; second-to-fourth digit length ratio defined as the ratio between the length of index and ring fingers) is associated with increased risk-taking and long-term profitability in a sample of high-frequency traders, suggesting that "financial markets may select for biological traits rather than rational expectations". Other studies show the correlation of the same proxy with testosterone levels (see Manning et al., 2011).

Stenstrom et al. (2011) use a similar measure (the length of the second finger relative to the sum of the lengths of all four fingers) as a proxy for prenatal testosterone and find that higher levels of testosterone is predictive of greater financial, social, and recreational risk-taking. Even simpler metrics such as a person's height is shown to be of value in explaining economic behavior of financial agents (Addoum et al., 2015). The question that why physical factors are associated with personality traits and behaviors is still not well-understood. This strand of research, particularly in the neurosciences field, attempts to unravel the underlying mechanisms.

Based on the empirical background, we are particularly interested in examining the association of fWHR and risk-taking behavior of hedge fund managers. Prior work, discussed above, mostly relies on small size samples in controlled conditions (e.g., investment games). We aim at examining this association in a large sample of highly skilled investors over a long period of time. Also, unlike many previous studies, our sample consists of experts in the financial market. Their main compensation depends heavily on fund performance and in case of failure, HF managers bear a substantial reputation damage and career risk that makes our sample an interesting environment for studying this relationship. In sum, literature suggests that higher facial masculinity is positively related to risk-taking, and performance. Therefore, we derive Hypothesis 1.A and Hypothesis 1.B as follows.

Hypothesis 1.A: *Facial masculinity of HF manger is positively associated with the risk of the fund under management.*

Hypothesis 1.B: *Facial masculinity of HF manger is positively associated with the performance of the fund under management.*

The basic implication underlying our first hypothesis is that fund managers with higher facial masculinity choose more risky investment styles in the first place. Otherwise, it is well established in the literature that different investment strategies across hedge funds are associated with different risk-return profiles (see Agarwal et al., 2009; Brown and Goetzmaan, 2003; and Fung and Hsieh, 1997), hence, their ex-post comparison might not be fully appropriate for the purpose of our study. To address this concern, we develop the following two hypotheses and compare each manager to other managers with a similar investment strategy.

Hypothesis 2.A: *Higher facial masculinity of HF manger is associated with a higher risk-taking behavior compared to peers with a similar investment strategy.*

Hypothesis 2.B: *Higher facial masculinity of HF manger is associated with a superior performance compared to peers with a similar investment strategy.*

We remain agnostic about the risk-adjusted performance, since the results in literature is mixed and a valid speculation that whether the level of risk-taking of individuals with high facial masculinity is optimal or stems from e.g., aggression or impulsive behavior is not possible in an academically rigorous manner. However, we also conduct tests examining the effect on risk-adjusted performance and report the results for our sample.

4.2.2. Testosterone and Entrepreneurship

De Holan (2014) argues that entrepreneurship researchers have been reluctant to embrace neuroscience for reasons such as the difficulty in conducting multidisciplinary research, the cost involved, and the complexity of the endeavor, however, and while agreeing with the challenges, Nicolaou and Shane (2014) point out that scholars in the entrepreneurship field would need to accept more paradigmatic pluralism for the neuroscience and biology perspective, in order to deepen our knowledge on the entrepreneurship phenomenon.

Assuming that the entrepreneur is a key component of entrepreneurship, it is logical to extrapolate that individual differences have a significant impact on entrepreneurial outcomes (Carland et al., 1988). Nicolaou et al. (2008) argue that the failure to develop a comprehensive understanding of why people engage in entrepreneurial activity has occurred, in part, because researchers have failed to examine genetics as an important explanatory factor. Prior work suggests that entrepreneurship may have a biological basis (see, e.g., Zhang et al., 2009). As suggested previously, one of the most widely studied biological influences on attitudes and behaviors is found in the hormone testosterone.

Furthermore, evolutionary psychology as the theoretical basis for investigating the relationship between a heritable biological characteristic (testosterone level) and an important business behavior such as new venture creation has been deployed in prior research (see White et al., 2006). More specifically, a strand of the biological perspective on entrepreneurship has investigated the effect of hormones on entrepreneurial activities and outcomes. White et al. (2006) found that individual differences in testosterone levels were associated with entrepreneurial activity. Similarly, it has been documented that entrepreneurs with a lower digit ratio (as a proxy for greater exposure to prenatal testosterone) had more employees, higher revenues, and managed firms that grew faster (Guiso and Rustichini, 2011). Nicolaou et al. (2018) also find that serum testosterone levels are positively associated with self-employment for males.

Testosterone levels have been shown to increase during competitive situations (Carre et al. 2013). Additionally, the "biosocial model" of testosterone influence, suggests that "winners" of such competitive situations will experience higher levels of testosterone than will "losers" (Nicolaou et al., 2018). Building upon prior research, and using a high-stake competition observable to general public and potential future investors (ABC channel's Shark Tank show) as the research environment in a sample of male entrepreneurs, we derive Hypothesis 3 as follows:

Hypothesis 3: *Higher facial masculinity of the entrepreneur (as a proxy for the level of testosterone) is associated with a higher probability of deal success.*

4.3. Data and Methods

4.3.1. HF Sample Construction

We obtained data from Morningstar CISDM HF database (hereafter CISDM). According to Agarwal et al. (2009), CISDM covers about 40% of the hedge fund universe which is the highest coverage amongst four main hedge fund data sources (namely, CISDM, TASS, HFR, and MSI). CISDM provides information on HF monthly net-of-fee returns (hereafter monthly returns), as well as important fund characteristics such as investment strategy, management fee, incentive fee, fund domicile, and inception date. Another unique feature of this data source is that it provides information on HF manager names which is crucial to our study.

The database contains information on 20,455 active (4,941) and defunct (15,514) funds as of June, 2016. The database keeps track of defunct funds from January 1994, onwards. Therefore, our investigation period starts from January, 1994 and ends in June, 2016. Since we are interested in comparing the risk of different hedge fund managers, and distinct investment strategy factors directly affect HF returns (see Fung and Hsieh, 1997), funds with no information on investment strategy are excluded from the sample. Moreover, we exclude fund of funds since their managers are not directly involved in the investment decisions of hedge funds that they invest in.

One potential bias that may significantly affect the results is backfilling or instant-history bias (see Agarwal et al., 2009) that happens because hedge funds can choose when to start reporting to the database, which presumably will be after a period of good performance. The returns occurred on this period also may be reported by the data vendor. We follow the common procedure in literature and remove the first 12 monthly returns available during the sample period (e.g., see Lim et al., 2016). Finally, we exclude any HF with less than 18 months of observations. This leaves us with 7,549 hedge funds. Panel A of Table 4.1, provides more details on the derivation of the dataset.

CISDM provides information on HF investment strategies and classifies them into 15 main categories. However, previous studies have shown that there are a few distinct strategy categories that affect hedge fund returns (see Brown and Goetzmaan, 2003; and Fung and Hsieh, 1997). Motivated by these studies, Agarwal et al. (2009) classify different HF strategies provided by data vendors into four broad strategies. We follow their suggestion and categorize CISDM investment strategies into directional traders (DT), relative value (RV), security selection (SS), and multiprocess (MP). Furthermore, we consider "long-only" investment category (LO), as a separate broad strategy. Calculation methods and description of all of the variables used in this study is reported in Appendix (see Table 4.A1, Panel A). Details on the mapping between CISDM classification of investment categories and the five broad strategies is also reported in the Appendix (see Table 4.A1, Panel B).

To construct our sample of hedge fund managers, we started with the 7,549 observations. Manager names are available for 2,202 funds (1,235 unique "Management Team", since some of them manage multiple HFs). We do not use "Manager History" data, in order to make sure all of the reported returns in the data set, fall into the tenure of the HF manager in our sample. If the fund is managed by a management team, we use the first non-female name on the list. However, we control for the effect of management team in our analyses, in case the fund is not managed by only one person.

Two independent research assistants have searched online to obtain full face, frontal view photos of the 1,235 managers (managing 2,202 funds) using HF webpages, LinkedIn profiles, and google Image. For a picture to get approved, the main condition is full face frontal view and fairly neutral facial expression. The quality control of pictures is done through a standard procedure (see Carre and McCormick, 2008; and Wong et al., 2011 for similar procedures). We were able to obtain and confirm photos of managers of 358 distinct funds. A scoring system is developed to eliminate the pictures that do not meet the requirements suggested in literature: A third research assistant rated the photos on "neutral facial expression" (on a 1 to 5 scale). We removed the photos that were rated below 3 from our sample. Our final sample consists of 320 HFs. Details of sample selection is described in Table 4.1, Panel B.

—Please insert Table 4.1 about here—

The proxy for facial masculinity (fWHR) which is defined as the ratio between the bizygomatic width and the height of the upper face (Carre and McCormick, 2008; Lefevre et al., 2013) is calculated using "Adobe Photoshop" software by two independent research assistants. In case the measurements are different, the average value of the two measurements is used (see Jia et al., 2014). Figure 1 shows the calculation method.

—Please insert Figure 4.1 about here—

After identifying the HFs with available fWHR data, we use our dataset of 7,549 funds to calculate the HF risk and performance measures for subsequent analyses. Following Ackermann et al. (1999), we use the standard deviation of monthly returns (logged) as our first proxy for the funds' overall risk (*Log (SD)*). As they point out, the use of monthly return data to calculate standard deviation of returns, enhances the accuracy of the risk estimation since it is calculated based on more observations (compared to quarterly, or annual returns).

However, since we are more interested in the risk of the funds, compared to other funds within the same broad strategy, we define four other proxies of risk, adjusted for differences in broad investment strategies. We calculate standard deviation of monthly returns (SD), mean absolute deviation of returns (MAD), maximum monthly return (Max), and minimum monthly return (Min) of each HF, and adjust each variable for its average value across all of the funds within the same broad strategy (for a similar methodology, see Kouwenberg and Ziemba, 2007).

Equation 1, defines $\alpha_{i,j}\_A$ as the adjusted risk proxy $j \in \{SD, MAD, Max, Min\}$, for fund i, belonging to broad strategy $S \in \{DT, RV, SS, MP, LO\}$, measured relative to the average of risk proxy "j" within fund i's broad strategy group ($\bar{\alpha}_{j,S}$).

$$\alpha_{i,j}\_A = \alpha_{i,j} - \bar{\alpha}_{j,s} \tag{1}$$

Similarly, we use *Mean Return* (annualized) and *Sharpe Ratio* as our main proxies for HF performance. However, to take into account the intrinsic differences between broad investment strategies, we introduce two adjusted performance measures based on *Mean Return* (Mean) and *Sharpe Ratio* (Sharpe) as described in Equation 2.

$$\beta_{i,j}\_A = \beta_{i,j} - \bar{\beta}_{j,s} \tag{2}$$

Where $\beta_{i,j}\_A$ is the adjusted performance proxy $j \in \{Mean, Sharpe\}$, for fund i, belonging to broad strategy $S \in \{DT, RV, SS, MP, LO\}$, measured relative to the average of performance proxy "j" within fund i's broad strategy group ($\bar{\beta}_{j,S}$).

Moreover, following Ackermann et al. (1999), we use *Incentive Fee*, *Management Fee*, *Fund Age*, and *Offshore* dummy which takes the value of 1 if the fund is non-US-domiciled, as control variables. We also control for *Management Team* dummy which takes the value of 1, if the fund is managed by more than one manager. Detailed description and calculation methods of the variables are presented in Appendix (see Table 4.A1).

Furthermore, we winsorized all non-dummy "Dependent" and "Control" variables at 1% level on both sides, to mitigate the effect of extreme outliers on the results. Table 4.2, Panel A, reports the summary statistics for the data set. Finally, we only kept those funds with available information on the manager's fWHR. Our final sample consists of 320 HFs. Table 4.2, Panel B reports the summary statistics for the current study's sample, and presents the correlation matrix of our independent variables, in Panel C. A comparison of the values reported in Panel A, and Panel B of Table 4.2, shows that the sample is highly representative of the data set. The mean values of variables are very close in the two panels. More importantly, comparing the minimum and maximum values of the variables, shows that our sample includes observations in both tails of the distribution for most of the variables of interest.

—Please insert Table 4.2 about here—

In the empirical analyses, we are interested in examining the effect of facial masculinity of HF manager, as proxied by fWHR, on the risk and performance of the fund. We apply OLS regression models to analyze the determinants of HF risk and Performance, in general, and examine the effect of facial masculinity of fund manager, in particular. Our base models are described in Equations 3 and 4, respectively.

$$Risk\ Measure = \alpha + \beta . fWHR + \sum_{i=1995}^{2016} \gamma_i . Y_i + \varepsilon \quad (3)$$

$$Performance\ Measure = \alpha + \beta . fWHR + \sum_{i=1995}^{2016} \gamma_i . Y_i + \varepsilon \quad (4)$$

Where $Risk\ Measure \in \{Log(SD), SD\_A, MAD\_A, Max\_A, Min\_A\}$; γ_i takes the value of 1, if the fund has reported monthly returns in Year "i", and zero, otherwise. $Performance\ Measure \in$

$\{Mean\ Return, Sharpe\ Ratio, Mean\_A, Sharpe\_A\}$. We do not include the control variables in the base models in order to avoid reducing number of observations due to data availability. However, we repeat our analyses, including the control variables, as described in Equations 5 and 6, Where $Control_{1-5}$ = *Incentive Fee*, *Management Fee*, *Fund Age*, *Offshore*, and *Management Team*, respectively.

$$Risk\ Measure = \alpha + \beta . fWHR + \sum_{n=1}^{5} \rho_n . Control_n + \sum_{i=1995}^{2016} \gamma_i . Y_i + \varepsilon \quad (5)$$

$$Performance\ Measure = \alpha + \beta . fWHR + \sum_{n=1}^{5} \rho_n . Control_n + \sum_{i=1995}^{2016} \gamma_i . Y_i + \varepsilon \quad (6)$$

We also check the robustness of the results using an alternative fWHR measure as a dummy variable (*D_fWHR*) which takes the value of 1, if the fWHR of HF manager is more than or equal to median fWHR of fund managers in the sample (1.88), and 0 otherwise. Furthermore, we re-examine models presented in equations 3 to 6, using *D_fWHR* as our proxy for facial masculinity instead of *fWHR*.

4.3.2. Shark Tank Sample Construction

Shark Tank's internal data are not available to the public; therefore, we construct a novel hand-collected dataset comprising all entrepreneur/firms that have aired on the show from Season 1 to Season 5 from 2009 through 2014 (N = 379). After excluding 96 firms that "only" had female entrepreneurs in the team, our final sample includes 283 firms that had at least one male member in the team. We collect one dependent variable ("Deal" which is a dummy factor that equals 1 if the deal is successful, and 0 if no deals are made), and collect/create five independent variables (namely, *fWHR*, *Team*, *Mixed Team*, *Ask Equity*, and *Ask Valuation*) by combining publicly available data from three sources (1) Show episodes[1], (2) Shark Tank Wikipedia page[2], (3) Halle Tecco's database[3]. For detailed description and calculation methods of variables, see Table 4.A1 (Panel A) in the Appendix.

We obtained frontal neutral face photos of entrepreneurs from show episodes (screenshots of same size, and quality). To measure the fWHR, we used Python face recognition package and the code developed by Ties de Kok[4]. It should be noted that for the HF sample, we needed to measure

---

[1] See https://abc.com/shows/shark-tank
[2] See https://en.wikipedia.org/wiki/List_of_Shark_Tank_episodes
[3] See https://docs.google.com/spreadsheets/d/1Lr0gi_QJB_JU0lBMjJ7WiBRxA0loml1FlM-KlmKsaEY/edit#gid=0
[4] See https://www.tiesdekok.com/calculatefwhr/

fWHR manually using Adobe Photoshop because it would have resulted in more accurate measurements considering the alignment of photos and different picture qualities, however, in the Shark Tank sample, we can ensure the necessary conditions for using the code are satisfied[1]. In case multiple male entrepreneurs are present in the management team, we use the average fWHR across male entrepreneurs. We also control for the following variables: *Control 1*: If the management is a team rather than one person (*Team*), *Control 2*: If at least one female is present in the entrepreneurial team (*Mixed Team*), *Control 3*: The percentage of equity that the entrepreneur(s) are offering to the sharks regardless of final deal (*Ask Equity*), and *Control 4*: the total valuation of the firm set by entrepreneur regardless of final deal (*Ask Valuation*). Table 4.3 reports the summary statistics for the Shark Tank sample (Panel A), and presents the correlation matrix, in Panel B.

—Please insert Table 4.3 about here—

In the empirical analyses, we are interested in examining the effect of facial masculinity of the entrepreneur, as proxied by fWHR, on the deal outcome. We apply Logistic regression models to analyze the determinants of deal success, in general, and examine the effect of facial masculinity of the entrepreneur, in particular. Our full model is described in Equations 7, below.

$$Deal_i = \alpha_i + \beta . fWHR_i + \sum_{n=1}^{4} \rho_n . Control_{n(i)} + \varepsilon_i \tag{7}$$

[1] The conditions are as follows: pictures should be (1) of high quality (i.e. high face area resolution) (2) subject's nose is pointing directly towards the lens, and (3) subject's face is not tilted and eyes are aligned.

4.4. Results

4.4.1. HF Sample

Our results show that fWHR of HF manager significantly affects the risk of the fund. Table 4.4, Panel A, reports the results of regressions of our five dependent variables (risk measures) on fWHR, considering the year effects, only (see Equation 3). In Panel B of Table 4.4, we add the control variables to the regressions (see equation 5) and show that our results remain unchanged with the inclusion of control variables. However, we lose 48 observations (Number of observations is reduced to 272 from 320) due to the missing data on sample's control variables.

Model (1) of Table 4.4 shows that fWHR has a significant positive relationship with the risk of fund as measured by standard deviation of monthly returns (logged), regardless of the broad strategy. Thus, we cannot reject Hypothesis 1.A that overall, HF managers with higher facial masculinity, take more risk regardless of their investment strategy. A possible explanation is that managers with high fWHR choose more risky investment strategies in the first place, and that might be the reason of the high standard deviation of monthly returns.

Furthermore, models (2) and (3) show that fWHR also has a positive effect on the adjusted risk, as measured by adjusted standard deviation (*SD_A*) and adjusted mean absolute deviation (*MAD_A*). This means HF managers with higher fWHR show more risk-taking behavior compared to their peers within same broad strategy, providing support for Hypothesis 2.A. Model (4) shows that the historical "maximum monthly return" of HF managers (adjusted for the average "maximum monthly return" of all of the funds within same broad strategy) with higher facial masculinity is significantly higher than those of managers with lower fWHR. However, their historical minimum is also significantly lower (as shown in Model (5)), which confirms their higher risk-taking behavior, i.e., higher maximums and lower minimums compared to their peer funds.

—Please insert Table 4.4 about here—

In sum, Table 4.4 provides support for both Hypothesis 1.A and Hypothesis 2.A. Results show that managers with higher facial masculinity tend to take more risk, and their risk-taking behavior deviates (i.e., it is higher) from their peers with a similar stated investment strategy. An important question that follows is if this risk-taking behavior is optimal and is reflected in their return profile or it largely stems from managers' over-confidence and does not lead to higher returns over time.

Table 4.5's results are in favor of the latter proposition, i.e., while managers with higher fWHR show more risk-taking behavior, their performance is not significantly better, and considering the risk that they are taking, their risk-adjusted performance is significantly lower.

Table 4.5, Panel A, reports the results of regressions of our four dependent variables (performance measures) on fWHR, considering the year effects, only (see Equation 4). In Panel B of Table 4.5, we add the control variables to the regressions (see equation 6) and show that our results remain unchanged with the inclusion of control variables. However, as mentioned previously, our sample is reduced to 272 observations due to the missing data on sample's control variables.

Model (1) of Table 4.5 show that fWHR does not have a statistically significant relationship with average HF return (annualized) as measured by *Mean Return*. The effect is non-existent even when *Mean Return* is adjusted for the average value within the same broad strategy (Model (3)). This result is interesting in the sense that while HF managers with high facial masculinity, manage riskier funds, but, on average, they do not perform better. They also do not perform better compared to their peers within the same broad strategy. Hence, Hypothesis 1.B and Hypothesis 2.B are not supported. However, Model (2) and Model (4) show that, if risk is taken into account, HF mangers with higher fWHR have significantly lower risk-adjusted returns as measured by *Sharpe Ratio* (Model (2)), and lower risk-adjusted returns compared to their peer following the same broad strategy (Model (4)). This is in line with the idea that the risk-taking behavior of HF managers with high facial masculinity is not necessarily optimal and may root in e.g., personal or psychological factors instead of strict economic rationality.

—Please insert Table 4.5 about here—

In Table 4.6, we check the robustness of our results regarding the risk measures, comparing two groups of low fWHR (less than median in our sample) and high fWHR HF managers. We define a dummy variable, *D-fWHR*, that equals 1 if the fWHR of manager is more than or equal to the sample median. A comparison between Table 4.4 and Table 4.6 shows that our results remain robust to the alternative proxy for facial masculinity. Models (1) shows that the high fWHR group take significantly more risk and Models (2) and (3) provide evidence that high fWHR group take more risk compared to their peers with the same broad strategy. Moreover, Models (4) and (5) support our initial results that high facial masculinity will result in higher historical maximum returns and lower minimum returns when compared to peers within same broad strategy.

—Please insert Table 4.6 about here—

Table 4.7, checks the robustness of our results regarding the performance measures, comparing two groups of low fWHR and high fWHR HF managers. By comparing Table 4.5 and Table 4.7, we show that our results remain robust to the alternative proxy of facial masculinity. While no significant effect on *Mean Return* and *Mean_A* is reported, however, results show that high fWHR group have lower Sharpe ratios, overall; and also when compared to peers within the same broad strategy.

Overall, we provide evidence supporting our Hypothesis 1.A, and Hypothesis 2.A, however, our empirical evidence does not support Hypothesis 1.B, and Hypothesis 2.B, showing that although HF managers with higher facial masculinity, take more risk, it is not necessarily an optimal level of risk taking and does not result in a superior performance. This latter result is in line with Lu and Teo (2021) that conclude that hedge funds operated by high-fWHR managers underperform those operated by low-fWHR managers, bear greater downside risk, and are more susceptible to failure.

—Please insert Table 4.7 about here—

4.4.2. Shark Tank Sample

Using our Shark Tank sample, we provide preliminary empirical evidence showing that fWHR of the male entrepreneur-contestant significantly affects the probability of deal success. Table 4.8, reports the results of logistic regressions of *Deal* dummy variable on fWHR. In Model (1), we only include fWHR as our dependent variable and document the significant positive effect on deal success. We eventually add the control variables to the regressions (see equation 7) through Models (2) to (5) and show that our results remain unchanged with the inclusion of control variables, economically and statistically.

—Please insert Table 4.8 about here—

Considering the full model (Model (5)), our sample results show that a 0.10 increase in fWHR (in absolute terms) increases the probability of coming to terms with the investors (i.e., sharks) by 32.84%, all else being equal. Our results also show that having an entrepreneurial team (rather than one person) positively affects the probability of successful deal, while having a mixed team of males and females negatively affects the probability of making a deal. This latter result, however, is only statistically significant at 10% level. Moreover, and as one intuitively predicts, when the entrepreneur is willing to give away a larger proportion of the venture to the investors, the probability of successful deal decreases. Finally, results show that setting a higher initial valuation negatively predicts deal success. Our results provide strong evidence in support of Hypothesis 3, and are in line with previously documented relationship between testosterone level, facial metrics, and the entrepreneurial outcomes.

4.5. Conclusion

We provide evidence of a positive association between facial masculinity of male HF managers (as proxied by fWHR) and the risk of the fund under their management. We also determined the relationship between facial masculinity and risk-adjusted performance, showing that although HF managers with high facial masculinity take more risk, but it does not translate into a higher return profile. We showed that HF managers with high fWHR tend to have lower risk-adjusted performance. Furthermore, we also document the positive link between facial masculinity of male entrepreneurs in a sample of ABC channel's Shark Tank show pitches, and the probability of coming to terms with investors and making a successful deal.

We also provide a brief inter-disciplinary literature review on the subject in order to introduce some avenues for future research in finance and entrepreneurship, on similar topics. Our results shed more light on recent works linking facial metrics and physiological factors to economic behavior, and extend our understanding on biological reasons affecting financial risk-taking, performance, and entrepreneurship.

Acknowledgments

We are grateful for the comments and suggestions gathered during the Financial Management Association Annual Meeting 2018 (San Diego, USA). This research was funded by Concordia's Office of the Vice President, Research & Graduate Studies (OVPRGS under grant number VS0271).

References

Ackermann, C., McEnally, R., & Ravenscraft, D. (1999). The performance of hedge funds: Risk, return, and incentives. *Journal of Finance, 54*(3), 833–874.

Addoum, J. M., Korniotis, G. M., & Kumar, A. (2015). Stature, obesity, and portfolio choice. *Available at SSRN:* http://papers.ssrn.com/abstract=1509173.

Agarwal, V., Daniel, N. D., & Naik, N. Y. (2009). Role of managerial incentives and discretion in hedge fund performance. *Journal of Finance, 64*(5), 2221–2256.

Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics, 116*(1), 261–292.

Brown, S. J., & Goetzmann, W. N. (1995). Performance persistence. *Journal of Finance, 50*(2), 679–698.

Burkart M, & Dasgupta A. (2014). Activist funds, leverage, and procyclicality. Working Paper, Stockholm Sch. of Economics/ London Sch. of Economics.

Carland, J. W., Hoy, F., & Carland, J. A. C. (1988). "Who is an entrepreneur?" is a question worth asking. *American journal of small business, 12*(4), 33–39.

Carré, J. M., & McCormick, C. M. (2008). In your face: Facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society B: Biological Sciences, 275*(1651), 2651–2656.

Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial structure is a reliable cue of aggressive behavior. *Psychological Science, 20*(10), 1194–1198.

Carré, J. M., Campbell, J. A., Lozoya, E., Goetz, S. M., & Welker, K. M. (2013). Changes in testosterone mediate the effect of winning on subsequent aggressive behaviour. *Psychoneuroendocrinology, 38*(10), 2034–2041.

Coates, J. M., Gurnell, M., & Rustichini, A. (2009). Second-to-fourth digit ratio predicts success among high-frequency financial traders. *Proceedings of the National Academy of Sciences, 106*(2), 623–628.

Coates, J. M., & Herbert, J. (2008). Endogenous steroids and financial risk taking on a London trading floor. *Proceedings of the National Academy of Sciences, 105*(16), 6167–6172.

Fung, W., & Hsieh, D. A. (1997). Empirical characteristics of dynamic trading strategies: The case of hedge funds. *Review of Financial Studies, 10*(2), 275–302.

Guiso, L., & Rustichini, A. (2011). Understanding the size and profitability of firms: The role of a biological factor. *Available at SSRN*: https://ssrn.com/abstract=1749846.

Jia, Y., Lent, L. V., & Zeng, Y. (2014). Masculinity, testosterone, and financial misreporting. *Journal of Accounting Research, 52*(5), 1195–1246.

Kouwenberg, R., & Ziemba, W. T. (2007). Incentives and risk taking in hedge funds. *Journal of Banking & Finance, 31*(11), 3291–3310.

Lefevre, C. E., Lewis, G. J., Perrett, D. I., & Penke, L. (2013). Telling facial metrics: Facial width is associated with testosterone levels in men. *Evolution and Human Behavior, 34*(4), 273–279.

Lewis, G. J., Lefevre, C. E., & Bates, T. C. (2012). Facial width-to-height ratio predicts achievement drive in US presidents. *Personality and Individual Differences, 52*(7), 855–857.

Lim, J., Sensoy, B. A., & Weisbach, M. S. (2016). Indirect incentives of hedge fund managers. *Journal of Finance, 71*(2), 871–918.

Lu, Y., & Teo, M. (2021). Do alpha males deliver alpha? Facial width-to-height ratio and hedge funds. *Journal of Financial and Quantitative Analysis*, forthcoming.

Manning, J. T., Scutt, D., Wilson, J., & Lewis-Jones, D. I. (1998). The ratio of 2nd to 4th digit length: a predictor of sperm numbers and concentrations of testosterone, luteinizing hormone and oestrogen. *Human Reproduction (Oxford, England), 13*(11), 3000–3004.

De Holan, P. M. (2014). It's all in your head: Why we need neuroentrepreneurship. *Journal of Management Inquiry, 23*(1), 93–97.

Nicolaou, N., & Shane, S. (2014). Biology, neuroscience, and entrepreneurship. *Journal of Management Inquiry, 23*(1), 98–100.

Nicolaou, N., Shane, S., Cherkas, L., Hunkin, J., & Spector, T. D. (2008). Is the tendency to engage in entrepreneurship genetic? *Management Science, 54*(1), 167–179.

Nicolaou, N., Patel, P. C., & Wolfe, M. T. (2018). Testosterone and tendency to engage in self-employment. *Management Science, 64*(4), 1825–1841.

Nofal, A. M., Nicolaou, N., Symeonidou, N., & Shane, S. (2018). Biology and management: A review, critique, and research agenda. *Journal of Management, 44*(1), 7–31.

Partnoy, F. (2015). US hedge fund activism. In *Research Handbook on Shareholder Power*. Edward Elgar Publishing.

Sapienza, P., Zingales, L., & Maestripieri, D. (2009). Gender differences in financial risk aversion and career choices are affected by testosterone. *Proceedings of the National Academy of Sciences, 106*(36), 15268–15273.

Sell, A., Cosmides, L., Tooby, J., Sznycer, D., Von Rueden, C., & Gurven, M. (2009). Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proceedings of the Royal Society B: Biological Sciences, 276*(1656), 575–584.

Shane, S., & Nicolaou, N. (2015). The biological basis of entrepreneurship. In S. M. Colarelli & R. D. Arvey (Eds.). *The Biological Foundations of Organizational Behavior*: University of Chicago Press.

Smith, B., & Viceisza, A. (2018). Bite me! ABC's Shark Tank as a path to entrepreneurship. *Small Business Economics, 50*(3), 463–479.

Stenstrom, E., Saad, G., Nepomuceno, M. V., & Mendenhall, Z. (2011). Testosterone and domain-specific risk: Digit ratios (2D: 4D and rel2) as predictors of recreational, financial, and social risk-taking behaviors. *Personality and Individual Differences, 51*(4), 412–416.

Unger, J. M., Rauch, A., Weis, S. E., & Frese, M. (2015). Biology (prenatal testosterone), psychology (achievement need) and entrepreneurial impact. *Journal of Business Venturing Insights, 4*, 1–5.

White, R. E., Thornhill, S., & Hampson, E. (2006). Entrepreneurs and evolutionary biology: The relationship between testosterone and new venture creation. *Organizational Behavior and Human Decision Processes, 100*(1), 21–34.

Wong, E. M., Ormiston, M. E., & Haselhuhn, M. P. (2011). A face only an investor could love: CEOs' facial structure predicts their firms' financial performance. *Psychological Science, 22*(12), 1478–1483.

Zhang, Z., Zyphur, M. J., Narayanan, J., Arvey, R. D., Chaturvedi, S., Avolio, B. J., Lichtenstein, P., & Larsson, G. (2009). The genetic basis of entrepreneurship: Effects of gender and personality. *Organizational behavior and human decision processes, 110*(2), 93–107.

Zuckerman, M., & Kuhlman, D. M. (2000). Personality and risk-taking: common bisocial factors. *Journal of personality, 68*(6), 999–1029.

Figure 4.1: facial Width to Height Ratio (fWHR) Calculation

This figure shows the fWHR calculation method. The Points determine the facial landmarks used to derive facial masculinity proxy: fWHR = a/b.

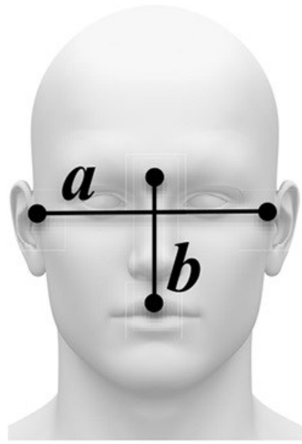


Table 4.1: HF Sample Selection

This table shows the derivation of the HF dataset in Panel A (All of the dependent variables are constructed based on this data set). Panel B shows details of our HF sample construction. Morningstar categories are classified into five broad strategies: *Directional Traders **(DT)***, *Relative Value **(RV)***, *Security Selection **(SS)**, Multiprocess **(MP)*** and *Long-Only (**LO**)*. Fund of Funds (***FOF*s**) are removed from the sample.

Panel A

	##	#	Active	Defunct
Morningstar CISDM Database (as of June, 2016)	**20,455**	-	4,941	15,514
- Duplicate observations (Identical ID in database)	-	3	-	-
- No Morningstar Category Specified	-	4,132	-	-
Subtotal	**16,320**	-	4,366	11,954

			Broad Strategy					
	##	#	**DT**	**RV**	**SS**	**MP**	**LO**	**FOF**
Dataset-Initial	**16,320**	-	2,305	1,812	5,364	1,470	398	4,971
- Fund of Funds	-	4,971	-	-	-	-	-	-
- No data on monthly returns	-	1,428	-	-	-	-	-	-
- Less than 18 entries of monthly returns[*]	-	2,372	-	-	-	-	-	-
Total	**7,549**	-	**1,636**	**1,282**	**3,293**	**1,048**	**290**	**0**
			(10.02%)	(7.85%)	(20.17%)	(6.42%)	(1.77%)	(0%)
Dataset-Total	**7,549**	**Active**	431	337	1,282	294	162	0
		Defunct	1,205	945	2,011	754	128	0

(continued)

[*] The first 12 "monthly return" entries are deleted (to take account of potential backfill bias). Thenceforth, any fund with less than 18 entries on "monthly return" is removed from the sample.

Table 4.1: HF Sample Selection—*continued*

Panel B

	##	#	Active	Defunct
Final Dataset	**7,549**	**-**	2,506	5,043
- No Information on "Manager Name"	**-**	5,205	**-**	**-**
Subtotal	**2,344**			
- "Manager Name" recorded as "Management Team"	**-**	50	**-**	**-**
- "Manager Name" recorded as "Not Disclosed"	**-**	92	**-**	**-**
Subtotal	**2,202**	-	4,366	11,954
- No Photo available	**-**	1,844	**-**	**-**
Subtotal	**358**	-	308	50
- Low-quality Photos	**-**	38	**-**	**-**
Sample-Total	**320**	-	276	44

			Broad Strategy					
	##	#	**DT**	**RV**	**SS**	**MP**	**LO**	**FOF**
Total	**320**	**-**	**86**	**42**	**125**	**58**	**9**	**0**
			(26.87%)	(13.12%)	(39.06%)	(18.12%)	(2.81%)	(0%)
Sample-Total	**320**	**Active**	64	35	121	47	9	0
		Defunct	22	7	4	11	0	0

Table 4.2: HF Sample Summary Statistics

This table gives descriptive statistics (mean, standard deviation, min, and max) for the HF data set. All variables are winsorized at the 1% level on both sides (Panel A). The HF sample is derived from the data set described in Panel A. Panel B shows the summary statistics for the HF sample (see Table 4.1 for details of sample construction). All variables are considered in subsequent HF sample analyses (see Table 4.A1 for variable descriptions and calculation methods). Panel C shows Pearson correlation coefficients for all of the independent variables considered in subsequent HF analyses (* indicates statistical significance at a 5% level or below).

Panel A

Variable	# Obs.	Mean	Std. Dev.	Min	Max
Dependent Variables: Risk					
Log(SD [of monthly returns])	7,549	-3.427	0.734	-5.783	-1.935
SD_A	7,549	-0.001	0.027	-0.040	0.097
MAD_A	7,549	0.000	0.019	-0.029	0.068
Max_A	7,549	-0.004	0.098	-0.124	0.442
Min_A	7,549	0.002	0.096	-0.389	0.124
Dependent Variables: Performance					
Mean Return [Annualized]	7,549	0.069	0.099	-0.227	0.432
Sharpe Ratio	7,549	0.714	1.075	-1.282	6.784
Mean_A	7,549	-0.002	0.098	-0.296	0.355
Sharpe_A	7,549	-0.149	1.086	-2.529	5.581
Control Variables					
Incentive Fee [%]	6,614	18.564	5.135	0	30
Management Fee [%]	6,777	1.509	0.528	0	3
Fund Age [months]	2,755	114.553	66.489	33	340
Offshore	7,549	0.319	0.466	0	1
Management Team	2,344	0.328	0.470	0	1

(*continued*)

Table 4.2: HF Sample Summary Statistics—*continued*

Panel B

Variable	# Obs.	Mean	Std. Dev.	Min	Max
Dependent Variables: Risk					
Log(SD [of monthly returns])	320	-3.419	0.607	-4.939	-1.935
SD_A	320	-0.003	0.023	-0.040	0.097
MAD_A	320	-0.002	0.016	-0.029	0.068
Max_A	320	0.008	0.097	-0.124	0.442
Min_A	320	-0.002	0.090	-0.389	0.110
Dependent Variables: Performance					
Mean Return [Annualized]	320	0.080	0.063	-0.227	0.432
Sharpe Ratio	320	0.781	0.608	-1.060	3.399
Mean_A	320	0.009	0.064	-0.294	0.355
Sharpe_A	320	-0.044	0.630	-2.106	2.929
Independent Variables					
fWHR	320	1.874	0.131	1.6	2.2
D-fWHR	320	0.484	0.501	0	1
Control Variables					
Incentive Fee [%]	304	19.289	4.017	0	30
Management Fee [%]	315	1.556	0.556	0	3
Fund Age [months]	291	153.997	67.165	35	340
Offshore	320	0.369	0.483	0	1
Management Team	320	0.484	0.501	0	1

Panel C

		(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	fWHR	1						
(2)	D-fWHR	0.831*	1					
(3)	Incentive Fee	-0.132*	-0.088	1				
(4)	Management Fee	0.059	0.084	0.056	1			
(5)	Fund Age	0.000	-0.055	-0.022	-0.047	1		
(6)	Offshore	-0.043	-0.119*	-0.07	-0.042	0.050	1	
(7)	Management Team	-0.013	0.074	0.022	0.021	-0.087	0.179*	1

Table 4.3: Shark Tank Sample Summary Statistics

This table gives descriptive statistics (mean, standard deviation, min, and max) for the Shark Tank sample. All variables are considered in subsequent Shark Tank sample analyses (see Table 4.A1 for variable descriptions and calculation methods). The sample includes 283 pitched that had at least one male member in the entrepreneurial team. Panel B shows Pearson correlation coefficients for all of the variables considered in subsequent Shark Tank analyses (* indicates statistical significance at a 5% level or below).

Panel A

Variable	# Obs.	Mean	Std. Dev.	Min	Max
Dependent Variable: Deal					
Deal	283	0.473	0.500	0	1
Independent Variables					
fWHR	283	1.686	0.104	1.4	2.08
Team	283	0.413	0.493	0	1
Mixed Team	283	0.187	0.390	0	1
Ask Equity [%]	283	17.930	10.544	3	100
Ask Valuation [$ Million]	283	2.074	3.295	0.04	20

Panel B

		(1)	(2)	(3)	(4)	(5)	(6)
(1)	Status	1					
(2)	fWHR	0.147*	1				
(3)	Team	0.181*	0.051	1			
(4)	Mixed Team	0.034	0.062	0.572*	1		
(5)	Ask Equity	-0.109	0.051	-0.097	-0.015	1	
(6)	Ask Valuation	-0.141*	-0.014	-0.019	-0.103	-0.305*	1

Table 4.4: Multivariate Analysis of HF Risk Measures

In this table, we apply OLS regressions to analyze the effect of fWHR on HFs' risk. We classified Morningstar CISDM HF categories into five broad strategies: *Directional Traders*, *Relative Value*, *Security Selection, Multiprocess,* and *Long-only* (following Agarwal et al., 2009). The sample includes 320 HFs with available data on fund manager's fWHR. All of the models control for *Year Effects*, i.e., each year dummy takes the value of 1, if the HF has recorded data on monthly returns, in a given year. Panel A reports the effect of fWHR on risk measures, controlling only for *Year Effects*. In Panel B, common control variables related to HF risk and return are added (see Ackermann et al., 1999). t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Panel A				
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
	Log(SD)	*SD_A*	*MAD_A*	*Max_A*	*Min_A*
fWHR	1.290***	0.036***	0.024***	0.108***	-0.117***
	(5.13)	(3.79)	(3.42)	(2.65)	(-3.29)
Constant	-6.219***	-0.084***	-0.057***	-0.291***	0.287***
	(-12.81)	(-4.56)	(-4.21)	(-3.71)	(4.17)
Year Effects	Yes	Yes	Yes	Yes	Yes
Observations	320	320	320	320	320
Adjusted R^2	0.143	0.122	0.071	0.117	0.224

	Panel B				
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
	Log(SD)	*SD_A*	*MAD_A*	*Max_A*	*Min_A*
fWHR	1.189***	0.030***	0.019**	0.084*	-0.094**
	(4.39)	(2.93)	(2.54)	(1.88)	(-2.41)
Incentive Fee	0.011	0.001*	0.000	0.003**	-0.001
	(1.33)	(1.78)	(1.54)	(2.01)	(-0.79)
Management Fee	0.260***	0.007***	0.006***	0.024**	-0.013
	(3.93)	(2.87)	(3.32)	(2.19)	(-1.41)
Fund Age	-0.002	-0.000	-0.000	-0.001*	0.001
	(-0.43)	(-0.62)	(-0.22)	(-1.81)	(1.12)
Offshore	-0.189**	-0.007**	-0.006***	-0.035***	0.015
	(-2.39)	(-2.51)	(-2.82)	(-2.66)	(1.33)
Management Team	-0.066	-0.003	-0.001	-0.012	0.007
	(-0.89)	(-0.97)	(-0.29)	(-1.01)	(0.66)
Constant	-6.452***	-0.084***	-0.058***	-0.265***	0.256***
	(-10.59)	(-3.65)	(-3.43)	(-2.65)	(2.92)
Year Effects	Yes	Yes	Yes	Yes	Yes
Observations	272	272	272	272	272
Adjusted R^2	0.218	0.186	0.131	0.186	0.257

Table 4.5: Multivariate Analysis of HF Performance Measures

In this table, we apply OLS regressions to analyze the effect of fWHR on HFs' performance. We classified Morningstar CISDM HF categories into five broad strategies: *Directional Traders*, *Relative Value*, *Security Selection, Multiprocess,* and *Long-only* (following Agarwal et al., 2009). The sample includes 320 HFs with available data on fund manager's fWHR. All of the models control for *Year Effects*, i.e., each year dummy takes the value of 1, if the HF has recorded data on monthly returns, in a given year. Panel A reports the effect of fWHR on performance measures, controlling only for *Year Effects*. In Panel B, common control variables related to HF risk and return are added (see Ackermann et al., 1999). t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Panel A			
	(1)	*(2)*	*(3)*	*(4)*
	Mean Return	*Sharpe Ratio*	*Mean_A*	*Sharpe_A*
fWHR	-0.027	-1.236***	-0.027	-1.027***
	(-1.00)	(-4.90)	(-0.98)	(-3.83)
Constant	0.100*	3.058***	0.040	1.789***
	(1.96)	(6.28)	(0.77)	(3.46)
Year Effects	Yes	Yes	Yes	Yes
Observations	320	320	320	320
Adjusted R^2	0.112	0.141	0.099	0.097

	Panel B			
	(1)	*(2)*	*(3)*	*(4)*
	Mean Return	*Sharpe Ratio*	*Mean_A*	*Sharpe_A*
fWHR	-0.035	-1.264***	-0.034	-1.061***
	(-1.25)	(-4.51)	(-1.18)	(-3.62)
Incentive Fee	-0.000	-0.002	0.000	0.010
	(-0.21)	(-0.29)	(0.09)	(1.12)
Management Fee	0.007	-0.113	0.007	-0.015
	(0.97)	(-1.65)	(0.95)	(-0.20)
Fund Age	-0.000	-0.000	-0.000	0.001
	(-0.42)	(-0.07)	(-0.44)	(0.20)
Offshore	-0.020**	0.038	-0.020**	0.074
	(-2.42)	(0.46)	(-2.40)	(0.87)
Management Team	-0.003	-0.031	-0.000	-0.112
	(-0.41)	(-0.41)	(-0.06)	(-1.40)
Constant	0.150**	3.707***	0.084	2.009***
	(2.36)	(5.88)	(1.29)	(3.05)
Year Effects	Yes	Yes	Yes	Yes
Observations	272	272	272	272
Adjusted R^2	0.181	0.185	0.166	0.139

Table 4.6: Robustness Check Analysis of HF Risk Results to High/Low fWHR

In this table, we apply OLS regressions to analyze the effect of D-fWHR, dummy variable that equals 1 if the fWHR of the fund manager is more than median fWHR in our sample on HFs' risk. We classified Morningstar CISDM HF categories into five broad strategies: *Directional Traders*, *Relative Value*, *Security Selection, Multiprocess,* and *Long-only* (following Agarwal et al., 2009). The sample includes 320 HFs with available data on fund manager's fWHR. All of the models control for *Year Effects*, i.e., each year dummy takes the value of 1, if the HF has recorded data on monthly returns, in a given year. Panel A reports the effect of D-fWHR on risk measures, controlling only for *Year Effects*. In Panel B, common control variables related to HF risk and return are added (see Ackermann et al., 1999). t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

			Panel A		
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
	Log(SD)	*SD_A*	*MAD_A*	*Max_A*	*Min_A*
D-fWHR	0.362***	0.011***	0.007***	0.047***	-0.039***
	(5.49)	(4.61)	(4.10)	(4.43)	(-4.21)
Constant	-4.070***	-0.024***	-0.017***	-0.120***	0.095***
	(-24.42)	(-3.87)	(-3.77)	(-4.52)	(4.02)
Year Effects	Yes	Yes	Yes	Yes	Yes
Observations	320	320	320	320	320
Adjusted R^2	0.153	0.141	0.086	0.153	0.241

			Panel B		
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
	Log(SD)	*SD_A*	*MAD_A*	*Max_A*	*Min_A*
D-fWHR	0.326***	0.010***	0.006***	0.043***	-0.032***
	(4.52)	(3.69)	(3.16)	(3.69)	(-3.05)
Incentive Fee	0.010	0.001*	0.000	0.003**	-0.001
	(1.24)	(1.79)	(1.55)	(2.17)	(-0.80)
Management Fee	0.264***	0.007***	0.006***	0.022**	-0.013
	(4.01)	(2.87)	(3.33)	(2.06)	(-1.39)
Fund Age	-0.000	-0.000	-0.000	-0.001*	0.000
	(-0.06)	(-0.39)	(-0.02)	(-1.74)	(0.94)
Offshore	-0.157**	-0.006**	-0.006**	-0.029**	0.012
	(-1.98)	(-2.16)	(-2.52)	(-2.29)	(1.04)
Management Team	-0.090	-0.003	-0.001	-0.016	0.010
	(-1.21)	(-1.26)	(-0.53)	(-1.33)	(0.90)
Constant	-4.413***	-0.033**	-0.025**	-0.120*	0.095
	(-11.03)	(-2.18)	(-2.28)	(-1.86)	(1.65)
Year Effects	Yes	Yes	Yes	Yes	Yes
Observations	272	272	272	272	272
Adjusted R^2	0.221	0.202	0.143	0.218	0.268

Table 4.7: Robustness Check Analysis of HF Performance Results to High/Low fWHR

In this table, we apply OLS regressions to analyze the effect of D-fWHR, dummy variable that equals 1 if the fWHR of the fund manager is more than median fWHR in our sample on HFs' performance. We classified Morningstar CISDM HF categories into five broad strategies: Directional Traders, Relative Value, Security Selection, Multiprocess, and Long-only (following Agarwal et al., 2009). The sample includes 320 HFs with available data on fund manager's fWHR. All of the models control for Year Effects, i.e., each year dummy takes the value of 1, if the HF has recorded data on monthly returns, in a given year. Panel A reports the effect of D-fWHR on performance measures, controlling only for Year Effects. In Panel B, common control variables related to HF risk and return are added (see Ackermann et al., 1999). t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Panel A			
	(1)	*(2)*	*(3)*	*(4)*
	Mean Return	*Sharpe Ratio*	*Mean_A*	*Sharpe_A*
D-fWHR	0.003	-0.273***	0.004	-0.210***
	(0.44)	(-4.06)	(0.56)	(-2.94)
Constant	0.051***	0.959***	-0.010	0.038
	(2.85)	(5.64)	(-0.55)	(0.21)
Year Effects	Yes	Yes	Yes	Yes
Observations	320	320	320	320
Adjusted R^2	0.109	0.120	0.097	0.079

	Panel B			
	(1)	*(2)*	*(3)*	*(4)*
	Mean Return	*Sharpe Ratio*	*Mean_A*	*Sharpe_A*
D-fWHR	0.003	-0.246***	0.005	-0.190**
	(0.41)	(-3.23)	(0.62)	(-2.39)
Incentive Fee	-0.000	-0.001	0.000	0.012
	(-0.03)	(-0.06)	(0.28)	(1.30)
Management Fee	0.005	-0.128*	0.005	-0.029
	(0.75)	(-1.84)	(0.71)	(-0.41)
Fund Age	-0.000	-0.002	-0.000	-0.001
	(-0.57)	(-0.48)	(-0.59)	(-0.14)
Offshore	-0.019**	0.019	-0.019**	0.061
	(-2.28)	(0.22)	(-2.24)	(0.69)
Management Team	-0.004	-0.016	-0.001	-0.100
	(-0.48)	(-0.20)	(-0.15)	(-1.23)
Constant	0.090**	1.546***	0.026	0.196
	(2.16)	(3.65)	(0.62)	(0.45)
Year Effects	Yes	Yes	Yes	Yes
Observations	272	272	272	272
Adjusted R^2	0.176	0.153	0.163	0.113

Table 4.8: Multivariate Analysis of the Effect of fWHR on Entrepreneurial Outcomes

In this table, we apply Logistic regressions to analyze the effect of fWHR on entrepreneurs' performance (deal success). The dependent variable in all specifications is *Deal*, a dummy variable that equals 1 if the deal is successful and 0 otherwise. The sample includes 283 Shark Tank pitches that had at least one male entrepreneur in the team. t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	*(2)*	*(3)*	*(4)*	*(5)*
fWHR	2.889**	2.782**	2.858**	2.986**	3.284***
	(2.45)	(2.34)	(2.40)	(2.50)	(2.68)
Team		0.727***	0.995***	0.948***	1.030***
		(2.93)	(3.24)	(3.05)	(3.16)
Mixed Team			-0.583	-0.559	-0.746*
			(-1.52)	(-1.45)	(-1.87)
Ask Equity				-0.020	-0.037**
				(-1.61)	(-2.56)
Ask Valuation					-0.178***
					(-2.96)
Constant	-4.980**	-5.103**	-5.230***	-5.080**	-4.940**
	(-2.50)	(-2.54)	(-2.59)	(-2.51)	(-2.40)
Observations	283	283	283	283	283
Pseudo R^2	0.016	0.038	0.044	0.051	0.083

Appendix

Table 4.A1: Variable Definitions and HF Broad Strategy Classification

This table gives a detailed description of the data-gathering process and calculation methods for all variables in Panel A. Panel B shows the details of broad strategy classification (following Agarwal et al., 2009).

Panel A

Variable Name	Description and Calculation
	Dependent Variables (HF Sample): Risk
Log(SD)	Natural logarithm of HF's standard deviation of monthly returns.
SD_A	Standard deviation of HF monthly returns minus the average standard deviation of all funds with the same broad strategy.
MAD_A	Mean absolute deviation of HF monthly returns minus the average "mean absolute deviation" of all funds with the same broad strategy.
Max_A	Maximum monthly return of HF minus the average "maximum return" of all funds with the same broad strategy.
Min_A	Minimum monthly return of HF minus the average "minimum return" of all funds with the same broad strategy.
	Dependent Variables (HF Sample): Performance
Mean Return	Average Return of HF (annualized) during observation period.
Sharpe Ratio	Sharpe Ratio of HF during observation period, defined as "annualized average return" over "annualized standard deviation of monthly returns".
Mean_A	Mean Return of HF (annualized) minus the average "*Mean Return*" of all funds with the same broad strategy.
Sharpe_A	Sharpe Ratio of HF minus the average "*Sharpe Ratio*" of all funds with the same broad strategy.
	Dependent Variables (Shark Tank Sample)
Deal	Dummy variable that equals 1 if the deal is successful, and 0 otherwise.
	Independent Variables
fWHR	Facial Width-to Height Ratio to proxy for facial masculinity of HF manager/entrepreneur. The measure is defined as the ratio between the bizygomatic width and the height of the upper face (see Figure 4.1 for more details). In case of multiple male managers/entrepreneurs in the team, average fWHR is considered.
D-fWHR	Dummy variable that equals 1 if the fWHR of HF manager is more than or equal to the median fWHR in the sample, and 0 otherwise.
	Control Variables (HF Sample)
Incentive Fee	Incentive fee of the HF (%). Percentage of annual profits (over some benchmark) captured by HF management.
Management Fee	Management fee of the HF (%). Periodic charge to compensate for management services as a percentage of fund's net assets under management.
Fund Age	Number of the months (rounded up to the nearest integer) that the HF has been active from inception as of June, 2016.

Table 4.A1: Variable Definitions and Broad Strategy Classification—*continued*

Offshore	Dummy variable that equals 1 for non-US domiciled funds, and 0 for US-domiciled funds.
Management Team	Dummy variable that equals 1 if more than one manager is managing the fund, and 0 otherwise.
	Control Variables (Shark Tank Sample)
Team	Dummy variable that equals 1 if there are multiple entrepreneurs in the team, and 0 otherwise.
Mixed Team	Dummy variable that equals 1 if there is at least one female in the entrepreneurial team, and 0 otherwise.
Ask Equity	Percent equity that the entrepreneur initially offers to the sharks.
Ask Valuation	Total valuation of the business set initially by the entrepreneur in million dollars.

Panel B

Broad Strategy	**Morningstar Category**
Directional Trader (DT)	Emerging Markets; Global Macro; Currency; Systematic Futures
Relative Value (RV)	Convertible/Debt/Diversified/Merger Arbitrage; Equity Market Neutral; Long/Short Debt
Security Selection (SS)	Long/Short Equity; Bear Market Equity; Volatility
Multiprocess (MP)	Distressed Securities; Event Driven; Multi-Strategy
Long-Only (LO)	Long-Only Equity/Debt/Other