Three Essays on Investors' Behavior in FinTech Markets

by

Gabriel Rodríguez Garnica

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Business and Finance

Universidad Carlos III de Madrid

Advisors:
María Gutiérrez Urtiaga
Josep A. Tribó Giné

Tutor:
María Gutiérrez Urtiaga

July 2023

A mi familia

Agradecimientos

Todos estos años dedicados al desarrollo de esta tesis van a ser inolvidables. El tiempo pasa rápido y, cuando menos te lo esperas, estás a punto de redactar las últimas palabras de la tesis para así finalizar tus estudios de doctorado. Con la defensa de la misma, culmina una de las etapas más enriquecedoras que he experimentado, no solo a nivel académico sino también a nivel personal.

Estas últimas palabras me gustaría dedicarlas a los que han hecho que llegar hasta donde estoy hoy haya sido no solo un hecho, sino también un camino por el que poder avanzar, parar y disfrutar.

En primer lugar, me gustaría expresar mi más profundo agradecimiento a mis directores y coautores María Gutiérrez y Josep A. Tribó por todo su tiempo, conocimiento, y atención. Esto no hubiese sido posible sin su inconmensurable apoyo desde el primer día hasta el último. No tengo palabras suficientes para expresar mi gratitud por todo lo que me han aportado durante este proceso.

También quiero agradecer a Fernando Zapatero por su simpatía, disponibilidad y todos sus consejos, así como por su invitación a realizar una estancia de investigación en Questrom School of Business, Boston University.

Asimismo, deseo expresar mi gratitud a Anna Toldrá, Jesper Rüdiger, y Mónica López-Puertas por todo el feedback aportado como miembros del comité interno de evaluación anual. También me gustaría agradecer a David Moreno y Eduardo Melero por sus comentarios, valiosa ayuda y disponibilidad, así como a Pablo Ruíz, Pedro Serrano, José Penalva, Marco Giometti, David Martínez, José Marín, y Juan Manuel García. Además, agradezco a mis coautores Hugo Benedetti, Itziar Catello y Jordi Surroca por su confianza y estrecha colaboración durante estos últimos años.

Sin lugar a dudas, una de las cosas más valiosas que me ha dado el doctorado es haber conocido a gente extraordinaria, no solo los mencionados anteriormente, sino también a todos los compañeros con los que he coincidido en nuestra oficina 7.1.02, y con los que espero seguir compartiendo grandes momentos. En particular, quiero agradecer a Blanca, Jesús García, y a los integrantes de "Mojito's Team"; Jesús Diego y

Alex, nuestro primer grupo de trabajos creado al comienzo del máster. Son ellos los que desde el inicio han puesto la guinda a esta aventura. Estoy eternamente agradecido a mis colegas Alex y Jesús por todos esos días de conversaciones intensas, compartiendo ideas y preocupaciones, pero también deporte, comidas y viajes. Solo espero seguir colaborando, coincidir en congresos, viajar y encontrarme con todos ellos allá donde estemos.

También agradezco el apoyo de mis amigos de Logroño, quienes me han hecho mejor persona y han estado apoyándome tanto cuando estaba en la "tierra con nombre de vino" como en la distancia.

Moitísimas gracias á miña compañeira de viaxe, Andrea, quien me ha acompañado durante toda esta etapa en el camino, sin soltarme la mano cuando estaba a punto de caer. Ella ha sido capaz de entenderme cuando ni yo mismo me entendía, ha escuchado todas mis preocupaciones, me ha hecho parar cuando necesitaba un respiro, y me ha brindado la fuerza, energía y motivación necesarias para seguir adelante en todo momento.

Y finalmente, quiero expresar mi más profundo agradecimiento a mi familia, a quienes dedico esta tesis. Nunca podré devolver todo lo que han dado por mi. Son mis abuelos, quienes, con su ejemplo, me han inculcado el espíritu del esfuerzo, trabajo y superación desde muy pequeño, así como los mejores valores. Mis padres, quienes me han enseñado a estar siempre despierto, a desarrollar todo mi potencial, a establecer metas exigentes pero asequibles, y sobre todo a equilibrar mi vida profesional y personal, con el claro objetivo de disfrutar con cada reto. Mi tato, Sergio, el mejor regalo que mis padres han podido darme, con quien he crecido y compartido tantos momentos, tenerle siempre a mi lado es lo mejor y a quien le debo tanto.

Gracias a todos por formar parte de esta tesis.

Acknowledgments

I acknowledge that this thesis has received financial support from the Spanish Ministry of Science, Innovation and Universities (Grant Ref. FPU18/01051), MEC PGC2018-097187-B-100 and WRDS-UC3M: Infrastructure for large scale data processing, FEDER UNC315-EE-3636, Comunidad de Madrid (EARLYFIN-CM, #S2015/HUM-3353), the Comunidad de Madrid-Spain under the Multiannual Agreement with UC3M in the line of Excellence of University Professors (EPUC3M12), as well as the V PRICIT (Regional Programme of Research and Technological Innovation).

I also gratefully acknowledge the hospitality of Questrom School of Business at Boston University during my research visit. In particular, I am immensely thankful to Fernando Zapatero and the Finance Department for their unwavering support, invaluable feedback, and continued assistance throughout my stay.

Published and Submitted Content

- **Chapter 1:**

Contents partially submitted to the repository *Universidad Carlos III de Madrid, Trabajo del Estudiante*, as part of the master thesis. This document is a former version and proposal for the PhD chapter 1 (master thesis) submitted by the candidate in September 2019.

- The material from this source included in this thesis is not singled out with typographic means and references.

Other Research Merits

- **Chapter published:**

Benedetti, H. and Rodríguez-Garnica, G. (2023), "Tokenized Assets and Securities", Baker, H.K., Benedetti, H., Nikbakht, E. and Smith, S.S. (Ed.) The Emerald Handbook on Cryptoassets: Investment Opportunities and Challenges, Emerald Publishing Limited, Bingley, pp. 107-121. ISBN: 978-1-80455-321-3.
Available at: https://doi.org/10.1108/978-1-80455-320-620221008

Contents

Abstract

This Doctoral Thesis is composed of three chapters on investors' behavior in FinTech and digital financial markets. Specifically, chapter 1 looks at the investment behavior of backers in reward-based crowdfunding. Chapter 2 analyzes the behavior of backers in reward-based crowdfunding markets, which stigmatize by social sanctioning new and repeating entrepreneurs for their personal interest, so that they increase their status and influence capacity or decisiveness in the community. Finally, chapter 3 studies the strategic use of on-chain public transactions in the Ethereum blockchain in order to produce signals and make use of this information to make token use and investment decisions. This helps us better understand how the behavior of market participants, particularly users and investors, affects token prices.

Introduction

New digital financial markets and FinTech markets are evolving very quickly. Not only traditional problems and questions emerge in these markets, but also new problems intrinsic or not to the technology that these markets use. This requires innovative thinking and development, as well as new research and development of the existing literature.

Digital financial markets make it easier to reduce moral hazard problems, being able to provide funding to more unknown but good entrepreneurs that deserve it. Hence, concepts such as the democratization of finance arise. These markets are typically characterized by having a larger set of individual, non-professional, non-accredited investors. Hence, the problem that we can find in these markets is a significantly higher level of information asymmetries caused by the huge heterogeneity of market participants. Therefore, studying investors' behavior in these new emerging, developing markets is really interesting.

In this doctoral thesis, I study FinTech and Digital Financial markets with a particular emphasis on new blockchain and platform-enabled financing methods, using empirical strategies on big databases. Topics covered relate to asset pricing, corporate finance, and management. Precisely, behavioral finance is an important matter to properly adopt policies and regulations to new economic environments originating from the interaction between Finance and Technology. Therefore, understanding entrepreneurs' and investors' behavior and its consequences in the digital finance ecosystem is central to this doctoral thesis.

In the first chapter "**Signaling and herding to raise funds in reward-based crowdfunding**", joint with María Gutiérrez Urtiaga and Josep A. Tribó, we empirically investigate the complex connection between signaling and herding behavior in reward-based crowdfunding markets platforms like Kickstarter, allowing it to provide recommendations on how best to raise funds using these platforms. Using a panel of data from 3,500 projects crawled from Kickstarter, we find that during the initial phase of the campaign, the funding decisions of a reduced number of early backers are based on quality signals offered by the creator. However, during the second phase, signaling is substituted by the herding behavior of a large number of late backers, imitating early backers. Therefore, although they are observed sequentially, herding and signaling complement each other as the crowdfunding campaign evolves, increasing the size and attractiveness of the reward-based crowdfunding market. In terms of theoretical and academic implications, the results suggest that backers self-select into early or late backers depending on their ability to process the quality signals of the creator, so that herding after signaling ameliorates asymmetric information problems. In terms of practitioner and policy implications, our findings indicate that (i) creators will obtain better results by targeting their campaigns at more sophisticated potential customers, and (ii) in this market, backers' self-selection and herding work together to protect uninformed backers from fraud and deception, supporting a policy of minimum intervention in the regulation of reward-based crowdfunding markets.

Both the first and the second chapters investigate the dynamics of investor behavior in crowdfunding markets and the important consequences that this behavior implies for borrowers, lenders, and regulators. In particular, in the second chapter, "**Tough love: a story of stigmatization in crowdfunding communities**", joint with Josep A. Tribó, Jordi Surroca, and Itziar Castello, entitled the focus lies on the study of the investors' and evaluators´ behavior through the comments they post, as well as the consequences they induce. Specifically, we try to answer the reasons for entrepreneurs entering a new community might be stigmatized. The novelty of this study resides in that it is the first one that empirically studies the social sanctioning and stigma in crowdfunding communities by focusing on the evaluator's perspective and developing a stigmatization strategy. Therefore we focus on backers' private gains in the stigmatization process. This is different from what previous studies have analyzed. Existing literature has been oriented toward the community gains of the valuators. In this project, we also use a Kickstarter dataset, that in this case is a larger set that includes all projects launched on the platform between 2013 and 2017 included. Results show that stigmatization and social sanctioning have a positive impact on valuators' decisiveness, hence on their status in the community. Nevertheless, there exist limits on persistent social sanctioning and stigmatization. Backers' status starts to decrease after they stigmatize too much in the community, creating what we call a "boomerang effect".This effect not only reduces backers' own status in the community but also the attractiveness to newcomers of the community.

The third chapter, "**Token use and CeDex transactions' informativeness in token markets. Are transactions being used strategically?**" (solo paper) relates to behavior and information gathering in blockchain-enabled markets. In this chapter, I analyze whether higher token use increases token market value, and how token public on-chain transactions impact token value through the release of information about market sentiment. Whereas extant research has focused on token adoption and investors' attention from the network point of view, the novelty of this study resides in that it is the first that empirically measures real token use intensity and its interaction with token prices. Additionally, my paper advances the literature on crypto-asset valuation by being the first that analyzes market sentiment through public on-chain transactions. Therefore, I develop two new proxy variables to measure token use and investors' sentiment using public on-chain transaction data from Ethereum blockchain-based tokens. I use these proxies to analyze the relationships between token use transactions and token exchange transactions and their impact on token prices. Results show that token use in its ecosystem does not depend on exchange token prices or transaction volume. Neither does the use of the token affect its market value. On the contrary, market investors do observe and use strategically the information on other investors' exchange-related transactions and previous prices to make trading decisions, hence confirming the existence of feedback and momentum trading strategies in token markets. Furthermore, I disentangle the channels that impact prices through investors' exchange-related transactions. Therefore, in this study, I provide new proxies that can be used in the coming literature for analyzing token market dynamics. All in all, my results contribute to a better understanding of investors' behavior and market sentiment in crypto markets.

Chapter 1

Signaling and herding behavior to raise funds in reward-based crowdfunding

1.1. Introduction

Creative ideas are often generated by individuals with reduced financial resources and no access to banks, angel investors, or venture capital funds (Schwienbacher and Larralde, 2012; Agrawal et al., 2015). In this setting, reward-based crowdfunding appears as an innovative way of overcoming difficulties in early-stage funding (Cosh et al., 2009), allowing entrepreneurs to bypass traditional financial investors, and raise funds from large, online communities that meet on crowdfunding platforms such as Kickstarter (Schwienbacher and Larralde 2012; Belleflamme et al., 2014; Agrawal et al., 2015; Kuppuswamy and Bayus, 2018b).

On these platforms, entrepreneurs are known as “creators” and the funds are provided by “backers”. Backers are usually small, unsophisticated investors and there is doubt concerning the soundness of their funding decisions relative to those made by business angels, venture capitalists, and banks in the traditional financial industry (Agrawal et al., 2013). Nevertheless, there is empirical evidence showing that both creators and backers find ways to reduce the asymmetric information that characterizes this market for early funding.

On the one hand, creators use costly signals (signaling effect) to convince backers of the quality of their projects, and backers react to these signals. Mollick (2014) initiated the early literature on reward-based crowdfunding focusing on the static analyses of the determinants of the success of the campaign and looking for ex-ante quality signals that creators can use to influence a project’s success. On the other hand, backers also pay attention to the funding decisions made by other backers and exhibit herding behavior, with late backers imitating early backers (herding). Using panel data to observe the funding dynamics over time, authors such as Colombo et al. (2015) and

Gangi and Daniele (2017) show that late backers exhibit herding behavior and imitate previous backers in reward-based crowdfunding.

In this paper, we study the way in which signaling and herding behavior interplay, working as complements or substitutes as the funding campaign evolves in the reward-based crowdfunding market. This is important because herding can produce negative outcomes if late backers blindly follow the whims of early backers. Nonetheless, to date, there has been no comprehensive discussion regarding this interplay and its effects on the financing outcome of the campaign.

Our basic claim is that, in the reward-based crowdfunding market, backers self-select into early or late backers depending on their ability to process the quality signals of the creator. Backers who do not have the ability to process quality signals prefer to come late and rely on herding, while backers that are confident of their information arrive early and make their decisions based on the signals of the creator.

Using a unique, granular dataset with daily information on funding dynamics, we test this idea as follows. First, we prove that the progress of the funding campaign over time depends both on quality signals and on the information coming from early backers' behavior, showing that both signaling and herding impact the outcome. Secondly, we prove that early backers do not act based on whims, fashions or fads because they rely heavily on quality signals to make their funding decisions. Lastly, we show that as the funding campaign progresses, late backers pay less attention to quality signals and more to the behavior of early backers, so that their herding becomes more pronounced. Remarkably, we show that our results are robust once we tackle endogeneity concerns related to omitted variables problems that may lead to a spurious connection between early backers' interests and the behavior of late backers (*herding behavior)*.

Therefore, our findings indicate that signaling and herding both complement and substitute each other in interesting ways as the reward-based crowdfunding campaign evolves. First, signaling effects dominate the early phase of the campaign, influencing the behavior of a small minority of investors and potential consumers (*i.e. early* backers). However, as the funding campaign evolves, later backers pay less attention to these quality signals (signaling effect) and exhibit a more pronounced herding behavior. In this sense, there is a substitution between signaling and herding. Second, the herding behavior of late backers benefits from the quality of the initial signaling analyzed by early backers. Projects that offer high-quality signals quickly attract early backers' interest in the project. This initial interest is directly observed by late backers. In this sense, there is an intertemporal complementarity between signaling and herding.

Together these results show that over campaign time, signaling and herding substitute each other because they are observed sequentially: first signaling then herding. However, considering the total duration of the campaign, signaling and herding are complements because the herding behavior of late backers, who provide the bulk of the funding, follows the decisions of early backers who relied on quality signals.

Hence, the overall result of the campaign will depend critically on the ability of early backers to both interpret the signals of the creator and trigger positive reinforcing herding behavior. In particular, our results show that a significant proportion of the funding in a reward-based crowdfunding campaign comes from backers that do not pay much attention to quality signals, but rely instead on the observable behavior of the earlier backers, that make their decisions based on the quality signals provided by the creator. Therefore, potential poorly informed backers can benefit from herding and following the decisions of early backers, who are shown to have the ability to analyze the quality signals offered by the creator.

These findings have important policy implications for creators and regulators and for our understanding of the reward-based crowdfunding market. Regarding entrepreneurs/creators, our results indicate that they will obtain better results from reward-based crowdfunding campaigns if they target their products and campaigns at more sophisticated potential customers. If these sophisticated customers become early backers, they can better interpret the signals of the entrepreneur and trigger positive reinforcing herding behavior so that other investors will follow confidently. Regulators can also find these results useful. Some legal scholars have expressed concerns that irrational herding behavior can result in fraud in reward-based crowdfunding because consumer protection is scarce in these markets (Griffin, 2013; Bradford, 2012; Hazen, 2012). However, our results indicate that, although late backers do herd, this herding can result in better overall funding decisions. This implies that, first, large-scale fraud and deception are unlikely to occur in reward-based crowdfunding markets. And second, the interplay between signaling and herding increases both the funding opportunities and the quality of decisions so that the reward-based crowdfunding market becomes more attractive. Rational uninformed backers would not participate in reward-based crowdfunding if they could not choose to arrive late and herd on the behavior of better-informed backers who arrived earlier and paid attention to the quality signals of the creator.

The structure of the chapter is as follows. Section 2 presents the theoretical background and develops the hypotheses. Section 3 describes the data and the methodology used. Section 4 presents the results. Section 5 conducts some robustness tests. Section 6 discusses the results obtained. The chapter finishes with some concluding remarks and some guidelines for future research.

1.2. Theoretical Background, Literature Review, and Hypothesis Development

1.2.1. Asymmetric information in reward-based crowdfunding

The existence of high information asymmetries between entrepreneurs and early investors has always been a huge barrier to innovation, and it significantly reduces the number of new products reaching the market (Cosh et al., 2009). Crowdfunding can help entrepreneurs to overcome this difficulty.

According to Mollick (2014), crowdfunding "refers to the efforts by entrepreneurial individuals and groups –cultural, social, and for-profit – to fund their ventures by drawing on relatively small contributions from a relatively large number of individuals using the internet, without standard financial intermediaries". These contributions accumulate over a fixed period, which is generally a few weeks (Kuppuswamy and Bayus, 2018a). While these funds might come from accredited (high net worth) individuals and specialized investors, such as business angels, banks, or venture capital investors, they typically come from non-accredited individuals that provide small monetary contributions. Crowdfunding widens the possibilities for an innovative idea to get off the ground, and it allows for an interactive development process whereby the entrepreneur receives comments, requests for information, and feedback from the investors through the online platform (Belleflamme et al., 2014; Agrawal et al., 2015; Kuppuswamy and Bayus, 2018b; Walthoff-Borm, et al., 2018; Eiteneyer et al., 2019).

Crowdfunding can be classified depending on the type of contract used into equity, debt, charity, and reward-based crowdfunding. Typically, equity and debt crowdfunding are used to finance for-profit businesses using standard equity and debt contracts, respectively. Charity crowdfunding is a cheap way of raising money through donations and focusing on non-for-profit projects. Finally, reward-based crowdfunding raises money ("pledges") in exchange for rewards that vary from mere acknowledgment to the promise of delivery of a product or service.

Reward-based crowdfunding is particular in that the funding comes from backers, who are investors with a stake in the success of the project instead of on the returns generated. In case of project success, such backers will become potential customers[1]. This implies that reward-based crowdfunding unites a funding campaign with a marketing campaign, offering a low-cost way to promote products in the market and gather information from the reactions of customers. The combination of cheap financing, information gathering, and low-cost promotion makes this type of financing particularly efficient. This is confirmed by the market size, which is expected to grow at an 11% compound rate and duplicate by 2028 (according to GlobeNewswire 2022). Also, there is a large number of businesses successfully financed through reward-based crowdfunding platforms that, at a later stage, have become high-growth ventures (Schwienbacher and Larralde, 2012; Kuppuswamy and Mollick, 2014; Greenberg and Mollick, 2017).

Perhaps the most interesting characteristic of reward-based crowdfunding is the existence of very significant information asymmetries. First, while equity and debt crowdfunding benefit from the presence of specialized and high net-worth investors, almost all backers in reward-based crowdfunding are not specialized, they are small and more dispersed. Second, typically, the entrepreneurs that initiate reward-based crowdfunding campaigns do not have a firm that can issue equity or debt and have little experience in moving products from their initial concept to the market. Also, they are

[1] According to the Cambridge Centre for Alternative Finance, in year 2020 reward-based crowdfunding accounted for $876.8 million globally, with only 3% of this funding coming from institutional investors.

relatively unknown to potential backers (Ganatra, 2016). Third, projects are usually launched on reward-based crowdfunding platforms at the development stage, when they are only prototypes, and this exacerbates the problem of evaluating their quality (Mollick, 2014; Sewaid et al. 2021).

Notwithstanding all these informational problems, reward-based crowdfunding has the differential characteristic of raising funds from preselling the product. This allows backers to focus their evaluation on the utility that the product can give them and the ability of the creator to deliver it successfully (Gutiérrez and Sáez, 2018). This is clearly different from evaluating the future cash flows that the product may generate. This also explains why regulators do not limit the participation of non-accredited investors in reward-based crowdfunding, while there are restrictions on the amounts that investors can invest in equity or debt crowdfunding[2].

1.2.2. Literature review

Most of the studies on reward-based crowdfunding have focused on the study of the information asymmetries that are important in this market. Early studies on reward-based crowdfunding focused on signaling, analyzing the type of information that creators can credibly convey to the backers leading to project success. Later studies have focused on herding behavior, studying how backers make decisions following peer behavior.

[2] Crowdfunding is regulated in the US by the JOBS act. This act came into effect on May 16, 2016 and separates accredited from non-accredited investors and sets limits of investment for each type of investor in equity and debt crowdfunding. These measures are aimed to protect the survival of security-based crowdfunding platforms, that may be put at risk when investors that participate in them don’t have a minimum financial literacy (Meoli et al., 2022). National governments of the European Union have also introduced regulations for equity and debt crowdfunding in their markets like the German Retail Investor's Protection Act (Kleinanlegerschutzgesetz-KASG), which came into force on July 10, 2015.

The initial studies showed that costly quality signals about the project or the entrepreneur, such as the quality of the description provided and previous experience, increase the probability of success (Mollick, 2014; Zheng et al., 2014). Moreover, funding success is highly correlated with ultimate product delivery and project success in other markets[3]. This line of research has provided important insights into the information that backers pay attention to (Allison et al., 2015; Courtney et al., 2017; Davis et al., 2017; Block et al., 2018, Gafni et al., 2019; Hellman et al., 2019; Deichmann et al. 2021).

The more recent empirical studies have focused on the dynamics of reward-based crowdfunding as another important mechanism that allows backers to obtain information, not from the creator, but from the behavior of previous backers. The general idea is that of herding, dating back to Schelling (1978), implying that the existence of many early participants triggers even more participation. This idea was formalized in the observational learning literature (Banerjee, 1992; Bikhchandani et al., 1992; Devenow and Welch, 1996), showing that sequential decision-makers mitigate uncertainty by observing and imitating what others have previously done. As a consequence, there is a sequence of informational cascades, which result in observational learning among investors over time. Nevertheless, herding may lead to good or bad results depending on what imitators are learning. If the early decision-makers are not well informed the whole process may lead to an uninformed, inefficient outcome (Croson and Shang, 2008; Simonsohn and Ariely, 2008).

[3] According to Mollick (2014) and Mollick and Kuppuswamy (2014) in reward-crowdfunding: (i) 37% of the projects that get funding go over budget and many are delayed; however (ii) only 5% to 14% of the funded projects fail to deliver the reward, with 50% of rewards delivered late; and additionally (iii) 90% of the projects that get funding turn out into ongoing organizations that are later able to raise additional money from venture capitalists, angel investors, or banks after the campaign concluded. Conversely, less than 25% of the projects that fail to raise funds from backers are ever completed.

By now, the existence of herding in crowdfunding is well established. Testing for informational cascades and herding seems natural in equity crowdfunding because this market is characterized by the coexistence of both sophisticated and non-sophisticated investors. Vismara (2018) presents evidence of informational cascades in equity crowdfunding that runs from investors that disclose public profile information to other investors (without a public profile), who imitate them. Moreover, Wang et al. (2019) also find that in equity crowdfunding, information flows from business angels to the crowd. However, in reward-based crowdfunding, all backers are small individuals, but a difference can be drawn between early backers and late backers. Colombo et al. (2015) are the first to test for herding in reward-based crowdfunding. They show that the creator's patient development of social capital (by contributing to a number of campaigns before launching their own) is highly valued by early backers. Gangi and Daniele (2017) find that the number of backers the campaign attracts at the beginning of the tail of the campaign is positively associated with its final success. Kuppuswamy and Bayus (2018a) focus on the number of new backers each day in reward-based crowdfunding campaigns and show that early backers tend to attract subsequent backers, but this effect is only strong at the end of the campaign. Finally, Chan et al. (2020) show that the total pledged amount (to date) exhibits a U-shaped relationship with the daily pledged amount. Also, they investigate the moderating effects of visual-media quality signals and situational urgency on this herding behavior.

Therefore, according to the results from all this previous literature, we expect to observe that both signaling and herding can contribute to the success of the crowdfunding campaigns in our sample. However, we still do not know whether herding is leading to better or worst outcomes. On the positive side, Colombo et al. (2015) and Chan et al. (2020) show that early backers pay attention to social capital and the quality of visual-media. However, on the negative side, we have to consider that, precisely because backers are consumers, they are also likely to demand fashions and fads that only bring utility because of their novelty rather than from their objective quality (Stigler and Becker, 1977). Moreover, Jiang et al. (2021) show that the decision to back a crowdfunding project depends critically not only on the expected utilitarian value but also on the socioemotional value and participatory value of the backers. These extra sources of value for the backers may easily be driven by irrational impulses. Additionally, Allison et al. (2017) find those inexperienced first-time funders are likely to be influenced by subjective cues about group identity and pay little attention to objective information. Therefore, we have to consider the possibility that the behavior of early backers is based on fads and, if late backers passively mimic the behavior of early backers, the result may be an uninformed market.

Answering this question requires determining the quality of the information of early backers and the interaction between signaling and herding behavior. Are they complements? Are they substitutes? How do they evolve over time?

1.2.3. Hypothesis development

The critical question we face is whether early backers make their decisions based on objective quality or only on subjective and irrational impulses.

Our hypothesis is that the decisions of early backers will be based on their analysis of objective quality signals. Our argument in favor of this hypothesis is that the arrival of backers to the market is endogenous and depends on their information set. This is because in reward-based crowdfunding, backers have the option to wait and see, and this option is more valuable when a backer has less confidence in her/his information and believes other backers to be better informed. Specifically, some backers will have more private information taken from their experience at evaluating previous projects or from better knowledge of the type of product or service offered in the campaign and will therefore feel more capable of analyzing the public information disclosed by the entrepreneur. Moreover, backers with a higher information set have incentives to provide funds earlier since they can benefit from a wider menu of funding options and additional perks (e.g. two products at a discounted amount of funding for the first 50 backers). In addition, the fear of losing out on the opportunity to provide funds to a project incentivizes these backers. Thus, we expect backers who believe they have better (worse) information to be early (late) backers. Hence our first hypothesis is:

> ***Hypothesis 1****: In the early stage of reward-based crowdfunding campaigns, the funding decisions of early backers are determined by the quality signals provided by the creator.*

Regarding the choices of less informed backers, we expect that they choose to wait and see how many early backers are interested in and back the project. They either choose to become followers or refrain from backing the project depending on the strength of the early backers' interest in the project. We know that following the behavior of the preceding individuals and disregarding his/her own information can be optimal for the poorly informed individual (Bikhchandani et al., 1992). Besides, in our particular context, this effect will be reinforced due to self-selection: late backers can infer the pattern of behavior of early backers and realize these backers are better informed. This leads us to formulate our second hypothesis as follows:

> ***Hypothesis 2****: In the later stages of reward-based crowdfunding campaigns, herding behavior will substitute the creator's signaling as the main determinant of the funding decisions of late backers.*

The interplay between the hypotheses is summarized in Figure 1.1, which shows how herding substitutes signaling as the campaign evolves but also how the outcome of the herding phase depends critically on initial signaling in such a way that the quality of signaling complements herding decisions once we consider an intertemporal view.

During the campaign, there is an initial signaling phase (starting at t=0) when early backers make their funding decisions by analyzing the signals provided by the creator, and then there is a second phase of herding that starts when the decisions of some early backers are already observable. Late backers decide to be late because of their lower ability to analyze quality signals. Therefore, they pay little attention to these signals and herd on the behavior of early backers. Over time, the information on earlier backers accumulates and gains weight so that at the end of the campaign (close to t=T), herding completely dominates backers' funding decisions.

1.3. Data and Methods

1.3.1. Data

We used a unique and granular dataset crawled from Kickstarter. In particular, we gathered data on all the projects launched on Kickstarter from November 15, 2017, to December 20, 2017 (3,923 Kickstarter projects) to build a panel of data with daily information on the funding dynamics until February 13, 2018, when the last project campaign finalizes. In the cleaning procedure, we removed a total of 423 projects, either because there was no precise ex-ante information about their duration or because the end date of the campaign was changed once the campaign had started. A small portion of these projects was directly canceled by Kickstarter as an antifraud measure. Our final sample includes a total of 3,500 projects.

The initial exploration of the data shows no significant differences in the key variables of the projects when we compare our sample period with other periods. For example, in the period analyzed, the success rate was 40%, while for the overall period of projects financed through Kickstarter (201,400 projects from April 28, 2009, until May 2021), the success rate is 38.7%. It is important to highlight that we work with a sample composed of the population of all projects launched during the mentioned period, hence, reducing sample selection issues.

1.3.2. Variables

Dependent variables

To measure backers' funding behavior, we use different variables that capture funding outcomes. The signaling literature generally uses a dummy variable, *Success*, indicating whether the campaign's funding goal has been reached at the closing date. However, this variable does not separate the behavior of early and late backers. The behavior of early backers is captured by using as a dependent variable *Time to reach 10% goal*, defined as the ratio between the number of days that the project takes to reach 10% of the funding goal over the total duration of the campaign. A low value of this variable indicates strong early funding interest in the project[4].

To capture the behavior of late backers we use as dependent variables *Time to move from one % to the next* that capture the time it takes to move from one percent of the funding goal to another (e.g., from 10% to 20%). In these latter specifications, we test for herding in the behavior of late backers and use *Time to reach 10% goal* as an additional explanatory variable because late backers are able to observe this information

[4] Projects that never reach 10% of their funding goal take the value of 100% for the *Time to reach 10% goal.* In the Robustness Checks section we use alternative thresholds other than 10% to construct the variable that separates early from late backers' interest.

on reward-based crowdfunding platforms before making their decisions. Herding behavior would be consistent with a negative impact of this variable on the funding provided by late backers. Thus, reductions in the time to reach 10% goal stimulate the interests of late backers in providing funds for the project.

Independent variables

We follow the existing literature on signaling and construct several variables that capture ex-ante information on the characteristics of either the project or the entrepreneur.[5] This information is ex-ante because it is released and available for all potential investors when the project is launched. Specifically, we use three sets of quality variables.

First, according to Mollick (2014) and Chen et al. (2009), investors extrapolate the quality of the project by measuring *"preparedness"*, through the following variables:

Project funding goal: Amount of funding (in logs) that the entrepreneur intends to obtain in order to develop the project. Projects with high funding goals are more likely to fail to reach their goal, and their chances of success are lower. Nevertheless, the funding goal may also depend on the technical needs of the project or on the funding that the entrepreneur has obtained before, which in turn, may depend on the quality of the project and/or the entrepreneur. In our sample, the funding goals ranged from $50 to $5,000,000.

Duration: The number of days during which the project's funding campaign will be active. A longer duration makes it more likely to reach the funding goal but may

[5] To simplify the model we get rid of the possible interactions between the variables of project quality with those of entrepreneur's quality although some authors have shown the existence of complementarities among these variables to explain project success (Huang, Pickernell, Battisti, and Nguyen, 2022).

signal low confidence in the project. The maximum duration that Kickstarter allows is currently 60 days.

Number of webs: The number of links to the project's and to the entrepreneur's webpages. Providing this information is important for developing a social community and for better informing backers of the main objectives and characteristics of the project.

Quick update: Kickstarter strongly suggests posting information on any new development or idea, or missing information about the project based on the feedback obtained within the first day of the campaign from backers or other sources. Therefore, we define the quick update measure with a value of one if updates are provided within the first day.

Second, we introduce a set of ex-ante quality variables related to the entrepreneur and her/his previous experience.

Created projects: The number of projects a creator has previously created and launched on Kickstarter. The more projects the entrepreneur has launched previously, the more experience s/he has in obtaining funds.

Backed projects: The number of projects launched by other entrepreneurs that the entrepreneur has backed prior to launching her/his own project on Kickstarter.

Creator indirect experience: The time, in years, since the entrepreneur has had an active profile on Kickstarter. Some entrepreneurs may spend months studying the platform before launching their project to learn how the platform and backers behave.

Our last set of quality variables includes new variables intended to capture the subjective quality of the project's description.[6]

Sentiment of risk description: Lower-risk and/or better-prepared projects should have a more positive sentiment measure for their risk descriptions, and this should be an important variable for backers' funding decisions. In particular, *Sentiment of risk description* = (# positive net words + 1) / (# negative net words + 1). The variable takes the value between 0 and 1 if the overall tone is negative, and greater than 1 otherwise. We use the Harvard IV dictionary for determining negative and positive words.

Project description length: A more detailed description provides more information to the potential backer. However, a more complex project will require a longer description, therefore, the net impact of the length of the description requires controlling carefully for other project characteristics. This variable is constructed as the natural logarithm of one plus the number of *"net"* words (after cleaning prepositions and conjunctions) used in the funding campaign description.

Finally, as control variables, we also use Kickstarter's categories that classify products in 15 industries according to their characteristics, so we can control for possible existing trends in an industry.

The summary statistics of our data can be found in Table 1.1. The table shows that the mean values of the variables that measure the level of preparedness directly connected to a project's quality (*Number of webs, Quick update*), as well as those of entrepreneur's experience (*Created projects*, *Backed projects*, *Creator indirect experience*) are higher for projects that end up being successful compared to those that are unsuccessful.

[6] Kaminsky and Hopp (2020) show among others the predicting capacity of the text message in crowdfunding campaigns.

These results are also confirmed in the correlation matrix, where the correlations of the previous variables with *Success* are positive (see Table 1.2). Also notable is the significant negative correlation between *Time to reach 10% goal* and *Success.* This result highlights the relevance of early backers' interest in ensuring financial success, which is consistent with the proposal that early backers rely on signaling by the entrepreneur to make their decisions.

1.3.3. Methodology

We study the information that early and late backers use to make their funding decisions. We do so by using the following regression as our general specification:

$$\textbf{\textit{Funding Outcome}}_i = \alpha + \beta(\textit{Project Quality Signals})_i + \gamma(\textit{Entrepreneur Quality Signals}) + \delta(\text{Early Backers Interest in the Project})_i + \theta(\textit{Category controls})_i + \varepsilon_i \qquad (1)$$

When implementing this general specification, we run a logistic regression model when we use the dichotomous variable *Success* as our dependent variable, and a linear regression model when we use *Time to reach 10% goal* and other linear funding outcomes as dependent variables.

We will initially focus on the *Success* variable. Significant results for the explanatory variables capturing the quality of the project and/or entrepreneur ($\beta > 0$, $\gamma > 0$) indicate signaling effects, while significant results for early backers' interest in the project ($\delta > 0$) indicate that herding behavior exists.

To test Hypothesis 1 regarding the reliance of early backers on signaling, we will use *Time to reach 10% goal* as a dependent variable (inversely correlated with success).

By construction, the early backers' interest in the project measure is not included in the explanatory variables in this case. Negative and significant coefficients for the quality signals ($\beta < 0$, $\gamma < 0$) will indicate that early backers act as informed investors that rely on quality signals and are not simply following whims, fashions, or fads.

To test Hypothesis 2 we will use alternative dependent variables, *Time to move from one % to the next*, capturing the time it takes to move from one funding level to another. According to Hypothesis 2, quality signals should become less significant and early backers' interest more significant (δ>0) as we move to the later stages of the campaign (reductions in the time to reach 10% goal will lead to reductions in the time to get further percentages of funding goals).

1.4. Results

1.4.1. Descriptive results

Considering all the projects launched on Kickstarter during the period studied, we observe that the projects that fail tend to miss their funding threshold by large amounts. Conversely, projects that succeed tend to pass their funding threshold by very small amounts. This is reflected in the frequencies table of the projects' achieved funding level as a discontinuity around the 100% threshold (see Figure 1.2).

This result can be partially explained by the "bystander effect" among funders (Kuppuswamy and Bayus, 2018a). Once projects achieve their funding goal, backers tend to ignore these projects. Also, the "Matthew Effect" (Merton, 1957), which states that "the rich get richer, and the poor get poorer" explain the bi-modal distribution shown in Figure 1.2. This distribution would not be observed in a setting in which the funds provided by previous backers were not public information.

These initial findings are consistent with our basic hypothesis that projects with high-quality signals attract the interest of early backers, and the interest of early backers attracts the herding of late backers. These observables increase the interest of new backers, so these projects continue attracting funds until the funding campaign is closed.

Another interesting phenomenon can be seen in Figure 1.3, which shows the evolution of funding relative to the funding goal (from 0% to 25%). Successful projects and failures can be distinguished early on by looking at the daily funding they attract. The projects that will end up being successful, on average, and from the very beginning of the campaign, obtain higher daily pledges compared to projects that will not succeed. Hence, the speed of funding at the very beginning of the project campaign can be used as a proxy for backers' interest that correlates with the project's chances of success.

1.4.2. Signaling and herding as determinants of overall project success

The previous literature has found evidence of signaling and herding behavior in reward-based crowdfunding. We start by replicating these findings, taking *Success* as the dependent variable in our specifications, estimated using a Logistic regression. The results are shown in Table 1.3.

Starting with signaling (columns 1 to 3), the complete specification (column 3) provides strong evidence of signaling, consistent with Mollick (2014) and Colombo et al. (2015). The influence of quality signals on a project's success indicates that the decisions made by the crowd are similar to the decisions of rational and highly sophisticated investors, such as angels and venture capital funds (Cardon et al., 2009; Chen et al., 2009).

We observe, first, that increasing the funding target goal is negatively associated with success. Second, increasing the duration of the fundraising campaign is also negatively associated with success. This is not surprising given that a good project should not need many campaign days to reach its goal. These results are comparable with those found by Mollick (2014). Third, the number of links to webpages provided by the project, demonstrating preparedness, is positively connected with the probability of success. Fourth, the variable *Quick update* as a measure of preparedness and project quality is positively and significantly associated with the probability of success. Fifth, the entrepreneur's previous experience in launching campaigns for previous projects (*Created projects*) significantly increases the probability of success. This result is explained by investors' positive perception of the entrepreneur as competent and trustworthy, which is highly relevant in crowdfunding (Li and Martin, 2019). Sixth, the variable *Backed projects* have a positive impact on the chances of the focal project's success. In fact, entrepreneurs that are looking for funds on a crowdfunding platform and are also backing other projects might be signaling that they are active entrepreneurs interested in cooperating with other entrepreneurs, with some "skin in the game" in others' projects. Seventh, the variable *Creator indirect experience* (indicating for how long the entrepreneur has had an active profile on Kickstarter) has a positive impact on the likelihood of success of the focal project. Eighth, when we focus on sentiment, we find that projects with a more positive *Sentiment of risk description* have a higher probability of success. This is expected because positive sentiment implies that a project's risks have been carefully foreseen and a contingency plan has been prepared (positive words). Finally, backers favor projects with higher values of *Project description length*.

We then test for herding behavior in columns 4 and 5. Here we check whether overall project success depends on the early funding of the project. Our main independent variable of interest is *Time to reach 10% goal*. This variable shows a strong negative impact on project success. If early backers show considerable interest, the time it takes to reach 10% of the funding goal falls, and this increases the probability of success.

These results are suggestive of herding behavior, although a finer test will require studying how late backers react to this interest of early backers. It is noteworthy that, in general, quality signal variables are less significant in column 5. This result is compatible with herding behavior implying that, as time goes by, the initial signals of entrepreneur quality become less important, and the decisions of late backers rely mainly on those of early backers and less on signals of project quality.

Nevertheless, the result shown in Table 1.3, using project success as the dependent variable, aggregate the behavior of early and late backers and does not allow us to clearly separate their behavior or to disentangle the signaling and herding behavior effects and the way they interact. This is further investigated in the next two sections.

1.4.3. Early backers' response to signaling

Table 1.4 shows the results of testing Hypothesis 1. Here we repeat the study conducted in Table 1.3, but we substitute *Success* with *Time to reach 10% goal* as our dependent variable to capture the behavior of early backers. The results show that early backers' interest is based on the previously analyzed quality signals: (i) those of project preparedness (*Project funding goal*, *Duration*, *Number of webs*, and *Quick update*); (ii) creator quality (*Created projects, Backed projects, Creator indirect experience*), and (iii) subjective quality of the project's description (*Sentiment of risk description*, *Project description length*). Hence, the behavior of early backers seems rational and objective. Their decisions do not appear to be the result of "love money" from friends or family, fads, or whims since early backers react strongly to the public information disclosed by the creator at the beginning of the campaign. We also confirm the previous results by Colombo et al. (2015), who found that the creator's previous involvement in funding reward-based crowdfunding campaigns launched by other entrepreneurs has a positive effect on early backers' funding behavior (as reflected in the negative sign for the *Backed projects* variable). The intuition is that participation in other projects generates social capital and trust among backers.

Overall, our results conform to Hypothesis 1 and show that early backers base their funding decisions on quality signals and, hence, do not create irrational trends. Thus, following early backers could be positive for later investors. Whether this herding really takes place and how late investors aggregate the quality signals and the information on early backers' interest information is analyzed next.

1.4.4. Late backers' decisions

To test our second hypothesis, we start by constructing a set of variables that measure the time it takes for the project to reach additional funding levels. Specifically, *Time a% to b%* measures the proportion in the number of days out of the total duration of the campaign that the project takes to increase funding from a% to b% (b>a). If our hypothesis is correct, we expect to see that, as we move toward higher intervals, quality signals have less impact on the changing dependent variables, and the behavior of earlier backers becomes a more important determinant of late backers' interest.

The results are presented in Table 1.5 (Panel A and B). In Table 1.5, Panel A we only incorporate quality signals and, as expected, we find that ex-ante quality variables lose their significance as determinants of late backers' decisions. Interestingly, the signaling effect starts to disappear quite quickly and exponentially since the results for the 10 to 15% range are far more similar to the 15 to 20% and even to the 60 to 80% range than to the results of early backers (0-10%) in Table 1.4 (also reported in column 1 of Table 1.5, Panel A). Moreover, it is important to note the existence of four variables that maintain their effect and significance, namely, *Project funding goal*, *Duration* (beyond 50% of funding), *Quick Update*, and *Project description length*. Interestingly, these variables are rather salient and seem particularly informative on project quality and its description. Just as we found in Table 1.3 (Model 5), variables indicating entrepreneur quality become insignificant as the campaign evolves.

In Table 1.5, Panel B we additionally incorporate the behavior of early backers as an independent variable that can be observed by late backers. This variable (*Time to reach 10% goal*) is highly significant: the shorter the time to reach the initial 10% of the funding goal, the lower the time to complete the intervals (10-15%), (15-20%), (20-40%), (40-60%), and (60-80%). This is clear evidence of herding behavior, which confirms Hypothesis 2. Moreover, almost all the quality variables that are the basis of signaling behavior, according to Table 1.4 become insignificant (the only ones that remain significant are *Duration* and *Quick Update*).

Hence, herding is shown to substitute signaling as the driving force behind the late backers. According to the results, a one standard deviation increase in the time it takes to reach the first 10% funding goal increases the time it takes to obtain an extra 5% of funding by about 0.44, and the time it takes to obtain the following extra 5% thresholds by about 0.42.

It is interesting to compare the coefficients for the quality signals in Table 1.5 Panel A and Panel B. To the extent that early backers rely on quality signals to decide whether to fund, the omission (inclusion) of the *Time to reach 10% goal* explanatory variable in Table 1.5 Panel A (Panel B) can generate a bias in the coefficients. We would expect an amplification bias, generating overestimation of these coefficients, in Table 1.5, Panel A, where it is not included, and a collider bias, generating underestimation, in Table 1.5, Panel B, where the additional explanatory variable is included. Therefore, the real impact of the quality signals on the decision of the late backers is expected to be in between that from Table 1.5, Panel A, and Table 1.5, Panel B.

Together all these results indicate that most of the funding in a reward-based crowdfunding campaign comes from backers that do not pay much attention to quality signals but rely instead on the observable behavior of the early backers to make their decisions. Therefore, the reward-based crowdfunding market seems characterized by the existence of a small number of backers that analyze signals and a large number of backers who follow them. Because of this configuration, this market can only work efficiently if, first, backers with a greater ability to interpret signals self-select to be early backers and make decisions based on the quality signals provided by the creator; and, second, backers of lower ability choose to come later and herd on previous backers' behavior.

1.5. Robustness Checks

We conducted several robustness checks on our results. All of them are available in the appendix of the chapter.

1.5.1. Controls for project size

Herding effects should be more important for projects that need a larger number of backers. To confirm the importance of herding for larger projects, we repeated the analysis keeping only large projects and dropping all projects with a funding goal below \$5,000 (whose funding goal can sometimes be obtained with a single pledge). Our results are similar to those found for the whole sample (see Table 1.A1, panels A to D in the appendix).

1.5.2. Alternative definitions of success and estimation methods

To control for measurement errors, we used alternative definitions of a project's success, such as (i) the number of days required to reach 100% of the funding goal and (ii) success level, assigning values 1, 2, 3 or 4 depending on the funding level achieved. We found consistent results. Moreover, we ran Linear Probability Regressions to confirm the results from the Logistic Regressions and we found similar results (See Table 1.A2, panels A to C in the appendix).

1.5.3. Seasonal and daily effects

We tested for seasonal effects, in particular for the Christmas period, and for possible day-of-the-week and weekend anomalies as determinants of a project's success and early and late backer behavior. Weekend effects are marginally significant but they do not change any of our main results.[7] (See Table 1.A3, panels A and B in the appendix).

1.5.4. Identification of early and late backers

We also repeated the entire analysis changing the threshold we used to classify investors as early and late backers. We did so by changing the initial funding percentage reached. The alternative thresholds used are 5%, 15%, and 20% of the target funding. The best results are obtained for 5% and (our original) 10%, and are weaker for 15% and 20% indicating that the percentage of informed backers who self-select into early backers and use quality signals is not high, which makes the herding behavior of late backers even more important to determine projects' success (See Table 1.A4, panels A to C in the appendix).

1.5.5. Additional measures of initial backers' interest

We used the *Number of comments to reach 10% goal* as a complement to the *Time to reach 10% goal* variable. This variable indicates unresolved issues that need to be solved, and/or uncertainty in early backers' assessment of the project. Therefore, if there is herding, it should have a negative impact on late backers' decision to pledge funds. The results indicate that every comment made before the project obtains the first 10% of its target funding is expected to subtract over 0.3 from the log odds of the project reaching the funding goal (See Table 1.A5, panels A and B in the appendix).

---

[7] According to the data, most of the projects are launched during labor days. The number of projects launched on weekends is much lower. However, we only appreciate a marginally significant and negative effect on the probability of success and on the interest of early backers when a project is launched on a weekend.

1.5.6. Controls for endogeneity

We want to rule out the possibility that our results for the impact of early adopters' interest on late backers' interest are not capturing herding and are simply the result of the correlation between early adopters' interest and omitted variables related to quality signals.

We think this is unlikely, given that *Time to reach 10% goal* remains a very significant variable even when we control for all available quality signals. However, it is still possible that some quality signals cannot be quantified, and we are unable to measure them. These signals could be simultaneously causing a lower *Time to reach 10% goal* and a higher interest by late backers, thus generating a spurious correlation between these two variables. We conducted several robustness checks to dismiss this possibility.

First, we performed an instrumental variable analysis (IV) due to the correlation between the variable *Time to reach 10% goal* and some of the quality signals. We need an instrument that affects the time to reach 10% but it is independent of project quality signals. We use a dummy that captures the existence of population aggregation effects for each of the product categories[8]. This dummy indicates which type of products are more interesting for the user when more people are using them (network effects at category level). In this case, when this dummy is equal to 1, there is a higher likelihood of herding effects for any given level of product quality. To test the validity of our instrument, we run a first-stage estimation that shows a strong impact of the instrument on the instrumented variable (F-statistic 16.73, see column 1 in Table 1.6, panel A).

[8] The dummy aggregation effect has a value of 1 for the categories in which the aggregation of backers have a positive effect on the project. These categories are Art, Comics, Design, Fashion, Film/Video, Games, Photography and Theater.

Also, the exclusion restriction is satisfied since the instrument explains late backers' interest only through the early backers' interest. In other words, projects that can have population aggregation effects do not have a direct impact on funding success because of this population aggregation effect, but they need to have generated previous backers' interest beforehand. In the second stage estimation, we find that the results with the instrumented variable are similar to our initial results (see columns 2 to 6 in Table 1.A6, panel A).

Secondly, we tried to measure the separate impact of public and private information on the behavior of both early and late adopters. Our basic assumption is that early backers have extra information not available to late backers and/or are able to analyze the existing information more precisely than late backers. Therefore, ideally, we seek to measure the extent and impact of the extra information that late backers do not have themselves and can only infer from the behavior of early backers. To do this, we reran our tests formally separating any information coming from, on the one hand, publicly available quality signals (which may drive both early and late backers' behavior simultaneously) and, on the other hand, any other information that is orthogonal to this quality signals but is captured by the rate of adoption of the early backers[9]. The results are consistent with our hypotheses and show that the behavior of late backers is, at least, partly driven by the behavior of early backers that conveys extra orthogonal information that cannot be extracted from the publicly available quality signals (see Table 1.B6, panel A in the appendix).

---

[9] We specifically used the following procedure. First, we computed the error term from specification (1) in Table 4 where the behavior of the early backers (measured with the variable Time to reach 10% goal) is modeled as a function of the observed quality signals of the project and the entrepreneur. The error term from this regression represents the part of the behavior of the early backers that cannot be explained by the publicly available quality signals, and it is orthogonal to the information included in those variables. We then re-estimated specification (1) from Table 5b but, rather than introducing the original variable Time to reach 10% goal, we used this error term.

Finally, as an alternative control for potential endogeneity, we reran our tests on a specifically matched subsample using the propensity score matching (PSM) technique, both with and without replacement. We construct two samples that are similar in observable quality but differ in their interest to early backers. First, we select a treated sample, consisting of projects that attract strong interest from early backers, which we define as projects that receive 10% or more of their funding goal in less than 1% of their campaign time. We then create a control sample by matching each of these observations with an observation from the remaining projects in the initial sample (i.e. an observation that did not attract strong interest from early backers) considered the closest neighbor observation in terms of observable characteristics (quality signals, categories FE, and ex-ante information). Therefore, in these estimations, the *Treated* variable is a dummy taking the value 1 when it takes less than 1% of the campaign time to reach at least 10% of the funding goal. Since the control sample is specifically selected to be similar in quality signals, this *Treated* variable should capture a pure herding effect. The re-estimation of the main results using the newly matched sample obtained from the PSM procedure shows consistent results (see Figure 1.A1, and Table 1.A6, panel C in the appendix).

1.6. Discussion, Conclusions, and Future Research

Crowdfunding is a novel source of funding for start-ups that also provides useful marketing information to entrepreneurs (Viotto de la Cruz, 2018). Just like other forms of early funding, such as angel financing or venture capital, the crowdfunding market has high levels of information asymmetry. However, in crowdfunding, these problems are compounded because unknown creators with little formal experience (e.g., Davis et al., 2017; Lin and Boh, 2019) are trying to raise funds from a "crowd" of small, dispersed, and unaccredited investors. In this paper, we focus on reward-based crowdfunding, where backers fund projects in exchange for a reward, which usually consists of one unit of the service or product being funded.

1.6.1. Theoretical and empirical contributions

We contribute to the ample literature that studies how signaling and herding can help solve or aggravate the asymmetric information problems faced in financial markets. The focus on the reward-based crowdfunding market and our detailed dataset offers the opportunity to study these effects in a setting with extreme information asymmetries. Moreover, previous empirical papers on reward-based crowdfunding have focused separately on identifying signaling or herding effects. However, to the best of our knowledge, the literature has not yet explained how these two effects interplay and their impact on the project selection and funding outcome. This is important because the efficiency of final funding outcomes will depend on the strength of herding relative to signaling and on whether herding pursues valuable information or simply reinforces initial whims and fads.

We specifically study how signaling and herding substitute and complement each other as the crowdfunding campaign evolves. Our contention is that backers self-select into early and late backers depending on their ability to evaluate projects. Therefore, early backers analyze quality signals, while late backers, with a lower ability to evaluate signals, can have access to this information by herding on the behavior of the early backers. This is reflected in the dynamics observed as the campaign progresses, with signaling dominating the initial phase but being substituted by herding in the final phase. Moreover, we show using different approaches that our results are not driven by endogeneity concerns connected to omitted variables problems on the way we measure projects' quality. Such omitted variables would cause a spurious connection between early backers' interests and the behavior of late backers (*herding behavior*).

Crucially, although over time, herding substitutes signaling, considering the whole campaign, and adopting an intertemporal approach, the quality of initial signaling complements and reinforces the later herding behavior. Interestingly, according to our findings, only a small fraction of the overall funding comes from backers that pay attention to quality signals (around 10% of funding). Therefore, the herding mechanism is also indispensable to induce the majority of late-low-evaluation-ability backers to participate in this highly uncertain market, and it substantially increases the funding accessible to creators without reducing the efficiency of funding decisions.

1.6.2. Implications for entrepreneurs and regulators

Our results can help regulators and entrepreneurs make better decisions.

For practitioners, our research implies that it makes sense for creators (i) to invest ex-ante to develop quality signals that show high levels of preparedness for the project, and (ii) to design their projects to cater to more experienced or better-informed backers, rather than to the average backer. These more sophisticated backers will be more likely to be actively looking for projects and base their financing decisions on quality signals (Allison et al., 2017). Moreover, by attracting well-informed early funders, entrepreneurs can then benefit from the herding behavior of the "followers" or late backers with less information. All in all, success in crowdfunding can also extend to posterior success in subsequent funding efforts (Roma et al., 2017; Yu et al., 2017).

Regarding the regulation of crowdfunding, some authors (Griffin, 2013; Bradford, 2012; Hazen, 2012) make a case for reinforcement of oversight and investor protection in crowdfunding, pointing out that these markets may suffer from irrational herding behavior, which increases fears that unscrupulous creators may take advantage of uninformed backers pursuing the latest fad. However, other authors (Gutiérrez and Saez, 2018) point out that the particular characteristics and incentives of creators in this market already protect investors from fraud to a large extent. We prove the existence of rational herding behavior, which indicates that, although a higher proportion of backers might lack the ability to evaluate quality, they achieve protection from fraud by following informed early backers, who act on the basis of observable quality signals and "put their money where their mouth is". Hence, we advocate for a policy of minimum intervention in the regulation of such markets.

These results can be extrapolated to other financial markets where there exist high informational asymmetries, such as equity crowdfunding. In these markets, and particularly in some countries such as the U.S., there exist limitations to investors, so they have to be accredited and comply with KYC requirements. Considering our results that show that information provided by the funding decisions of early backers is reliable in markets with high information asymmetries and heterogenous investors profiles, we expect that information coming from early investors in equity crowdfunding is even more reliable. Therefore, regulation in equity crowdfunding markets can be reconsidered, such as reducing accreditation requirements for small investors, which is already the case in some countries in the European Union.

1.6.3. Limitations and future research

Our study has some limitations that open the possibility for future research.

First, more research is needed before our result can be generalized to other platforms and/or other types of crowdfunding like equity crowdfunding or debt crowdfunding. Differences in investor profiles and the differential effects on project success of similar explanatory factors across platforms (Chan et al., 2018; Dushnitsky and Fitza, 2018) could result in a different relationship between signaling and herding. However, differently from Kickstarter and reward-based crowdfunding, many of these platforms require to comply with KYC requirements and/or being accredited before providing funds given the higher postcampaign failure rates of equity-crowdfunded firms versus nonequity-crowdfunded ones (Walthoff-Borm et al., 2016). Hence, this leads us to expect that information coming from early backers' funding decisions in these markets is even more reliable.

A second research avenue is connected to the analysis of factors that may accelerate the herding behavior that starts when projects attract early backers that are more able to evaluate the quality of the new project. Persuasion and legitimacy are expected to be two key elements to facilitate such a process (Chen et al., 2009; Chan et al., 2020).

A third avenue is related to the connection between crowdfunding as an initial financing mechanism of start-up projects and other forms of financing in later stages, such as venture capital and/or Initial Coin Offerings (ICOs) and its derivatives using blockchain technology. The existence of such a connection, leading to the setting up of successful ventures when the initial reward-based crowdfunding has been successful, has been established in the literature (Greenberg and Mollick, 2017; Hornuf et al. 2018; Roma et al. 2017; Roma et al. 2021). In this context, it may be worth exploring which are the characteristics of the informational cascades generated in reward-based crowdfunding that, in the medium-term, can lead to successful venture capital and/or fintech financing.

Finally, another research avenue would explore the potential strategic actions of the competitors, who may try to generate negative herding behavior over a creator's project through their own funding decisions as well as the comments made on the platform. Such stigmatization attempts (which can generate "negative crowds") may have very negative consequences, not only for the success of a project but also for future projects launched by a given creator.

1.6.4. Conclusions

The novelty of our study is to provide evidence on how signaling and herding behavior effects reinforce each other in the reward-based crowdfunding market. This is a way to increase the size and attractiveness of the market, with both early-informed and late-uninformed backers benefiting from each other's decisions. Late-uninformed backers benefit from herding on the information analysis that early backers make of the quality signals provided by the creator. Early-informed backers benefit from the funding provided by late backers to make the campaign succeed. This has important implications for market participants and regulators and indicates the need for more research on the dynamics of investor behavior in crowdfunding markets.

References

Agrawal, A., Catalini, C., & Goldfarb, A. (2013). Some simple economics of crowdfunding. *Innovation Policy and the Economy, 14*, 63–97.

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.

Allison, T. H., Davis, B. C., Short, J. C., & Webb, J. W. (2015). Crowdfunding in a Prosocial Microlending Environment: Examining the Role of Intrinsic Versus Extrinsic Cues. *Entrepreneurship Theory and Practice, 39*, 53–73.

Allison, T. H., Davis, B. C., Webb, J. W., & Short, J. C. (2017). Persuasion in crowdfunding: An elaboration likelihood model of crowdfunding performance. *Journal of Business Venturing, 32*, 707–725.

Banerjee, A. V. (1992). A Simple Model of Herd Behavior. *The Quarterly Journal of Economics, 107*(3), 797–819.

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

Bikhchandani, S., Hirshleifer, D., & Welch, I., (1992). A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy, 100*(5), 992–1026

Block, J., Hornuf, L., & Moritz, A. (2018). Which Updates During an Equity Crowdfunding Campaign Increase Crowd Participation? *Small Business Economics, 50*(1), 3–27.

Bradford, C. S. (2012). Crowdfunding and the Federal Securities Laws. Columbia Business Law Review, 1. *SSRN*

Cambridge University Centre for Alternative Finance. Global Alternative Finance Market Benchmarking Report 2021.

Cardon, M. S., Wincent, J., Singh, J., & Drnovsek, M. (2009). The nature and experience of entrepreneurial passion. *Academy of Management Review, 34*(3), 511–532.

Chan, C. S. R., Parhankangas, A., Sahaym, A., & Oo, P. (2020). Bellwether and the herd? Unpacking the u-shaped relationship between prior funding and subsequent contributions in reward-based crowdfunding. *Journal of Business Venturing, 35*(2) (in press).

Chan, C.R., Park, H.D., Patel, P., & Gomulya, D. (2018). Reward-based crowdfunding success: decomposition of the project, product category, entrepreneur, and location effects. *Venture Capital, 20*(3), 285–307.

Chen, X. P., Yao, X., & Kotha, S. (2009). Entrepreneur passion and preparedness in business plan presentations: a persuasion analysis of venture capitalists' funding decisions. *Academy of Management Journal, 52*(1), 199–214.

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

Cosh, A., Cumming, D., & Hughes, A. (2009). Outside Entrepreneurial Capital. *The Economic Journal, 119*, 1494–1533.

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

Croson, R., & Shang, J. (2008). The impact of downward social information on contribution decisions. *Experimental Economics, 11*(3), 221–233.

Davis, B. C., Hmieleski, K. M., Webb, J. W., & Coombs, J. E. (2017). Funders' positive affective reactions to entrepreneurs' crowdfunding pitches: The influence of perceived product creativity and entrepreneurial passion. *Journal of Business Venturing, 32*, 90–106.

Deichmann, D., Gillier, T., & Tonellato, M. (2021). Getting on board with new ideas: An analysis of idea commitments on a crowdsourcing platform. *Research Policy*, 50 (9), 104320

Devenow, A., & Welch, I. (1996). Rational herding in financial economics. *European Economic Review, 40*(3), 603–615.

Dushnitsky, G., & Fitza, M.A. (2018). Are we missing the platforms for the crowd? Comparing investment drivers across multiple crowdfunding platforms. *Journal of Business Venturing Insights, 10,* 1–10.

Eiteneyer, N., Bendig, D., & Brettel, M. (2019). Social capital and the digital crowd: Involving backers to promote new product innovativeness. *Research Policy*, 48(8), 103744.

Gafni, H., Marom, D., & Sade, O. (2019). Are the life and death of an early-stage venture indeed in the power of the tongue? Lessons from online crowdfunding pitches. *Strategic Entrepreneurship Journal, 1,* 3–23.

Ganatra, J. H. (2016). When a Kickstarter stops: exploring failures and regulatory frameworks for the rewards-based crowdfunding industry. *Rutgers University Law Review 68,* 1425–1470.

Gangi, F., & Daniele, L. (2017). Remarkable funders: How early-late backers and mentors affect reward-based crowdfunding campaigns. *International Business Research, 10*(11), 58-69.

Greenberg, J., & Mollick, E.R. (2017). Activist Choice Homophily and the Crowdfunding of Female Founders. *Administrative Science Quarterly, 62*, 341-374.

Griffin, Z. (2013). Crowdfunding: Fleecing the American Masses. *Journal of Law, Technology, and the Internet, 4*(2), 375–410.

Gutiérrez, M., & Sáez, M. I. (2018). The promise of reward crowdfunding. *Corporate Governance: An International Review, 26*(5), 355–373.

Hazen, T. L. (2012). Crowdfunding or Fraudfunding? Social Networks and the Securities Laws – Why the Specially Tailored Exemption Must be Conditioned on Meaningful Disclosure. *North Carolina Law Review, 90*, 1735-1752.

Hellman, T., Mostipan, I., & Vulkan, N. (2019). Be careful what you ask for: Fundraising strategies in equity crowdfunding. *National Bureau of Economic Research Working Paper Series*. WP2019-12, Said Business School.

Hornuf, L., Schmitt, M. & Stenzhorn, E. (2018). Equity crowdfunding in Germany and the United Kingdom: Follow-up funding and firm failure. *Corporate Governance: An International Review, 26*(5), 331–35.

Huang, S., Pickernell, D., Battisti, M., and Nguyen, T. (2022). Signaling entrepreneurs' credibility and project quality for crowdfunding success: cases from the Kickstarter and Indiegogo environments. *Small Business Economics, 58*(4), 1801–1821.

Jiang, H., Wang, Z., Yang, L., Shen, J., & Hahn, J. (2021). How rewarding are your rewards? A value-based view of crowdfunding rewards and crowdfunding performance. *Entrepreneurship Theory and Practice, 45*(3), 562–599.

Kaminski, J. C. & Hopp, C. (2020). Predicting outcomes in crowdfunding campaigns with textual, visual, and linguistic signals. *Small Business Economics, 55*, 627–649.

Kuppuswamy, V., & Bayus, B. L. (2018a). Crowdfunding creative ideas: The dynamics of project backers. *Cumming D., Hornuf L. (Eds), The Economics of Crowdfunding, Palgrave Macmillan, Cham.,* 151–182.

Kuppuswamy, V., & Bayus, B. L. (2018b). A review of crowdfunding research and findings, in *Handbook of Research on New Product Development*. Cheltenham, UK: Edward Elgar Publishing, 361–373.

Kuppuswamy, V., & Mollick, E. (2014). After the Campaign: Outcomes of Crowdfunding. *UNC Kenan-Flagler Research Paper No 2376997*.

Li, E., & Martin, J. S. (2019). Capital Formation and Financial Intermediation: The Role of Entrepreneur Reputation Formation. *Journal of Corporate Finance, 59*, 185–201.

Lin, Y., & Boh, W. F. (2019). How Different are Crowdfunders? Examining Archetypes of Crowdfunders. *Journal of the Association for Information Science and Technology, 71*, 1357–1370.

Meoli, M., Rossi, A. & Vismara, S. (2022). Financial literacy and security-based crowdfunding. *Corporate Governance: An International Review,* 30 (1), 27–54.

Merton, R. K. (1957). Priorities in Scientific Discovery: A Chapter in the Sociology of Science. *American Sociological Review, 22*, 635–659.

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

Roma, P., Messeni Petruzzelli, A., & Perrone, G. (2017). From the crowd to the market: The role of reward-based crowdfunding performance in attracting professional investors. *Research Policy*, *46*(9).

Roma, P., Vasi, M., & Kolympiris, C. (2021). On the signaling effect of reward-based crowdfunding: (When) do later stage venture capitalists rely more on the crowd than their peers? *Research Policy*, *50*(6).

Schelling, T. (1978). Micromotives and Macrobehavior. *New York and London W.W.Norton.*

Schwienbacher, A., & Larralde, B. (2012). Crowdfunding of entrepreneurial ventures, in: Cumming, D. (Eds.), *The Oxford Handbook of Entrepreneurial Finance*. Oxford University Press, 369–391.

Sewaid, A., Garcia-Cestona, M., & Silaghi, F. (2021). Resolving information asymmetries in financing new product development: The case of reward-based crowdfunding. *Research Policy*, *50* (10), 104345

Simonsohn, U., & Ariely, D. (2008). When rational sellers face nonrational buyers: evidence from herding on eBay. *Management Science, 54*(9), 1624–1637.

Stigler, G. J., & Becker, G. S. (1977). De Gustibus Non Est Disputandum. *The American Economic Review, 67(*2), 76-90.

Viotto da Cruz, J. (2018). Beyond financing: crowdfunding as an informational mechanism. *Journal of Business and Venturing, 33,* 371–393.

Vismara, S. (2018). Information cascades among investors in equity crowdfunding. *Entrepreneurship Theory and Practice, 42*(3), 467-497.

Walthoff-Borm, X., Vanacker T., & Collewaert, V. (2018). Equity crowdfunding, shareholder structures, and firm performance. *Corporate Governance: An International Review, 26(5)*, 314–330.

Wang, W., Mahmood, A., Sismeiro, C., & Vulkan, N. (2019). The evolution of equity crowdfunding: Insights from co-investments of angels and the crowd. *Research Policy*, *48*(8).

Yu, S., Johnson, S., Lai, C., Cricelli, A., & Fleming, L. (2017). Crowdfunding and regional entrepreneurial investment: an application of the CrowdBerkeley database. *Research Policy*, *46*(10), 1723–1737.

Zheng, H., Li, D., Wu, J., & Yun, X. (2014). The Role of Multidimensional Social Capital in Crowdfunding: A Comparative Study in China and US. *Information and Management, 51*, 488–496.

Tables and Figures

Figure 1.1: Theoretical model

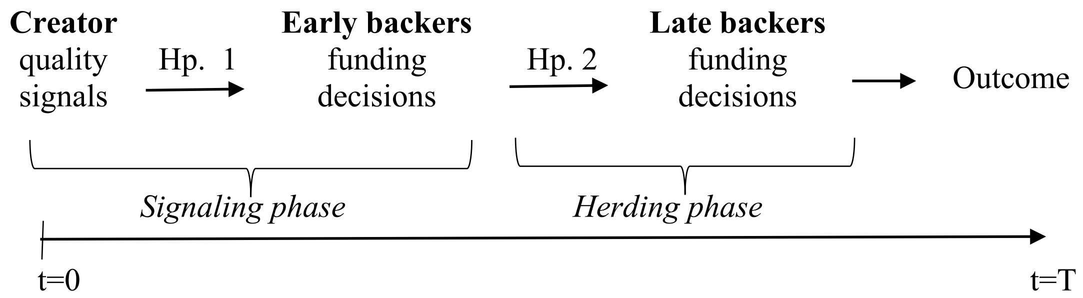


Figure 1.2: Distribution of level of funding reached (<100%= Failure; >100%=Success)

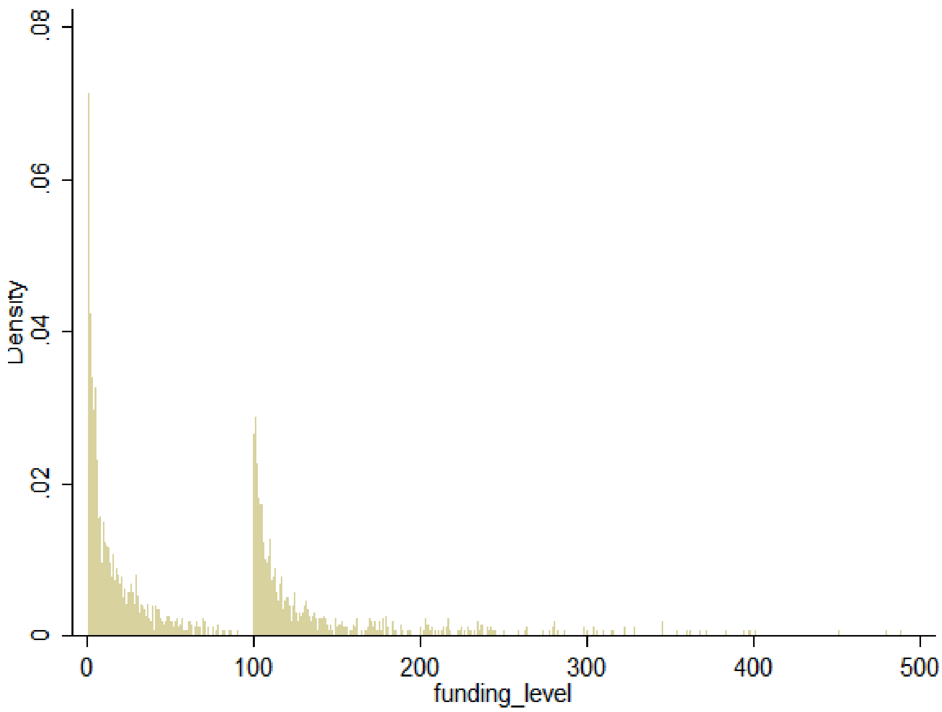


Figure 1.3: Daily pledges obtained by projects that succeed and projects that fail
(Zoom on 0% to 25% funding level reached)

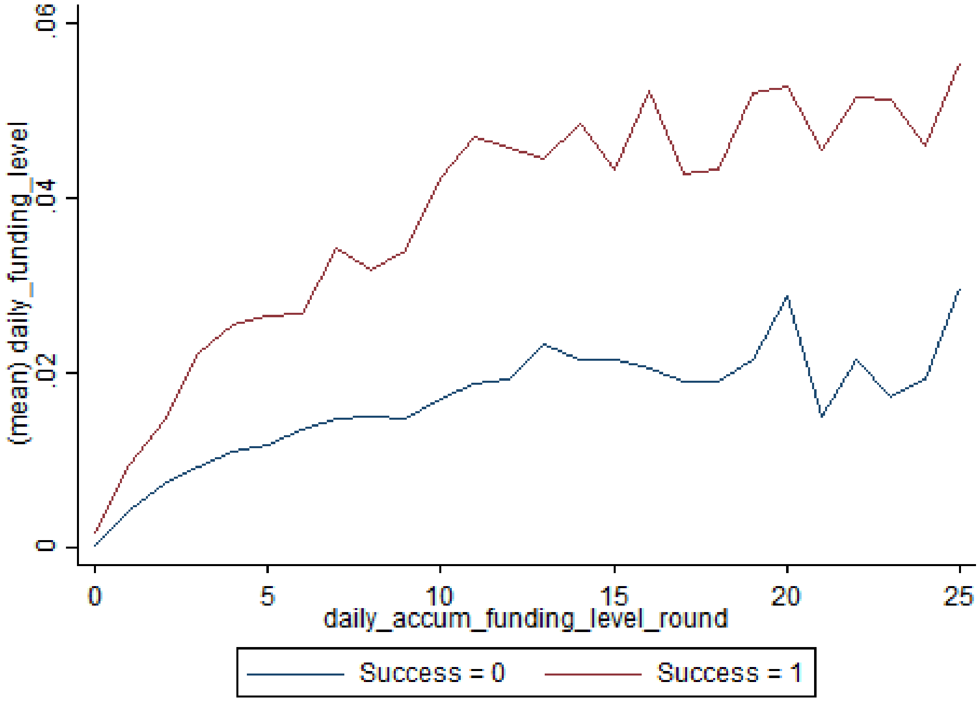


Table 1.1: Summary statistics for all Kickstarter projects and successful and unsuccessful projects

	Overall sample				Successful projects (Success=1)				Unsuccessful projects (Success=0)			
	Mean	Std. dev	Min	max	mean	Std. dev	min	max	mean	Std. dev	min	max
Success	0.40	0.49	0.0	1.0	1.00	0.00	1.0	1.0	0.00	0.00	0.0	0.0
Project funding goal	8.56	1.74	0.0	18.4	8.04	1.69	0.0	12.4	8.91	1.68	0.0	18.4
Duration	32.52	12.61	0.0	60.0	30.50	11.74	1.0	60.0	33.88	12.99	0.0	60.0
Number of webs	1.57	1.50	0.0	6.0	1.87	1.49	0.0	6.0	1.36	1.46	0.0	6.0
Quick update	0.04	0.20	0.0	1.0	0.06	0.24	0.0	1.0	0.03	0.16	0.0	1.0
Created projects	1.93	3.39	1.0	78.0	2.65	4.57	1.0	78.0	1.44	2.14	1.0	78.0
Backed projects	4.19	16.15	0.0	285.0	7.80	22.21	0.0	285.0	1.75	9.44	0.0	207.0
Creator indirect experience	1.36	1.97	0.0	8.6	1.90	2.25	0.0	8.6	1.00	1.67	0.0	8.3
Sentiment of risk description	2.21	1.70	0.2	19.0	2.35	1.83	0.2	19.0	2.11	1.61	0.2	18.0
Project description length	5.12	1.33	0.0	8.2	5.45	1.24	0.7	8.2	4.94	1.36	0.0	8.1
Time to reach 10% goal	0.48	0.47	0.0	1.0	0.03	0.10	0.0	0.9	0.77	0.38	0.0	1.0
Number of comments to reach 10% goal	0.18	0.91	0.0	19.0	0.05	0.31	0.0	4.0	0.27	1.15	0.0	19.0
N	3500				1405				2095			

Table 1.2: Matrix of correlations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Success	1.000											
(2) Project funding goal	–0.245	1.000										
(3) Duration	–0.133	0.261	1.000									
(4) Number of webs	0.164	–0.012	–0.022	1.000								
(5) Quick update	0.089	–0.008	–0.055	0.064	1.000							
(6) Created projects	0.175	–0.175	–0.164	0.088	0.049	1.000						
(7) Backed projects	0.184	–0.130	–0.120	0.125	0.048	0.379	1.000					
(8) Creator indirect experience	0.223	–0.083	–0.110	0.171	0.052	0.306	0.409	1.000				
(9) Sentiment of risk description	0.070	0.075	–0.003	0.082	0.018	–0.014	0.010	–0.004	1.000			
(10) Project description length	0.189	0.128	-0.012	0.138	0.050	0.032	0.086	0.117	0.075	1.000		
(11) Time to reach 10% goal	–0.773	0.279	0.192	–0.177	–0.104	–0.187	–0.200	–0.241	–0.077	-0.196	1.000	
(12) Number of comments to reach 10% goal	–0.122	0.142	0.035	0.006	0.015	–0.041	–0.035	–0.036	0.023	0.049	0.117	1.000

Table 1.3: Signaling and herding as a determinant of project success

	(1) Success	(2) Success	(3) Success	(4) Success	(5) Success
Time to reach 10% goal				-7.396***	-7.130***
				(0.336)	(0.354)
Project funding goal	–0.313***	–0.293***	–0.367***		-0.185***
	(0.025)	(0.026)	(0.028)		(0.042)
Duration	–0.013***	–0.009***	–0.007**		0.015***
	(0.003)	(0.003)	(0.003)		(0.005)
Number of webs	0.229***	0.181***	0.137***		0.091**
	(0.025)	(0.026)	(0.026)		(0.041)
Quick update	0.820***	0.772***	0.735***		0.295
	(0.186)	(0.191)	(0.194)		(0.261)
Created projects		0.065***	0.076***		0.031
		(0.020)	(0.021)		(0.021)
Backed projects		0.014***	0.010***		0.003
		(0.004)	(0.004)		(0.004)
Creator indirect experience		0.132***	0.126***		0.036
		(0.022)	(0.022)		(0.032)
Sentiment of risk description			0.100***		0.066*
			(0.022)		(0.035)
Project description length			0.396***		0.193***
			(0.037)		(0.044)
Category Controls	Yes	Yes	Yes	Yes	Yes
Constant	2.151***	1.562***	-0.040***	1.573***	1.021***
	(0.228)	(0.242)	(0.282)	(0.195)	(0.391)
Observations	3,500	3,479	3,479	3,923	3,479
chi2	455.4	579.3	740.4	3114	2761
P	0	0	0	0	0
McFadden's adjusted R^2	0.0966	0.124	0.158	0.608	0.589

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Category control for all models. Logistic regression is used. All variables are defined in the main text.

Table 1.4: Early backers' response to signaling

	(1)	(2)	(3)
	Time to reach 10% goal		
Project funding goal	0.067***	0.064***	0.072***
	(0.005)	(0.005)	(0.005)
Duration	0.004***	0.004***	0.003***
	(0.001)	(0.001)	(0.001)
Number of webs	-0.050***	-0.039***	-0.030***
	(0.005)	(0.005)	(0.005)
Quick update	-0.179***	-0.163***	-0.150***
	(0.037)	(0.036)	(0.036)
Created projects		-0.006***	-0.007***
		(0.002)	(0.002)
Backed projects		-0.002***	-0.002***
		(0.001)	(0.001)
Creator indirect experience		-0.032***	-0.029***
		(0.004)	(0.004)
Sentiment of risk description			-0.019***
			(0.004)
Project description length			-0.065***
			(0.006)
Category Controls	Yes	Yes	Yes
Constant	-0.147***	-0.037	0.254***
	(0.044)	(0.045)	(0.050)
Observations	3,500	3,479	3,479
P	0	0	0
R^2	0.161	0.193	0.229

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Category control for all models. OLS regression. All variables are defined in the main text.

Table 1.5: Late backers' response to signaling and early backers' behavior

Panel A. Late backers' response to quality signals

	Time to move from one % to the next					
	(0-10%)	(10→15%)	(15→20%)	(20→40%)	(40→60%)	(60→80%)
Project funding goal	0.072***	0.031***	0.032***	0.029***	0.028***	0.029***
	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
Duration	0.003***	-0.000	-0.001	-0.001	-0.001**	-0.001**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number of webs	-0.030***	-0.007	-0.007	-0.005	-0.004	-0.007
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Quick Update	-0.150***	-0.143***	-0.143***	-0.152***	-0.106***	-0.108***
	(0.036)	(0.035)	(0.035)	(0.036)	(0.036)	(0.037)
Created projects	-0.007***	-0.005**	-0.005**	-0.008***	-0.005**	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Backed projects	-0.002***	-0.001	-0.000	-0.001	-0.001	-0.000
	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Creator indirect experience	-0.029***	-0.014***	-0.013***	-0.006	-0.006	-0.006
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Sentiment of risk description	-0.019***	-0.003	-0.002	0.001	0.003	0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Project description length	-0.065***	-0.025***	-0.025***	-0.017***	-0.016***	-0.017***
	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
Category Controls	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.254***	0.177***	0.192***	0.231***	0.221***	0.217***
	(0.050)	(0.048)	(0.049)	(0.050)	(0.051)	(0.051)
Observations	3,479	3,479	3,479	3,479	3,479	3,479
R^2	0	0.057	0.054	0.043	0.031	0.029
P	0.229	0	0	0	0	0

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1 Category control for all models. OLS regression. All variables are defined in the main text.

Panel B. Late backers' response to early backers' behavior and quality signals					
	Time to move from one % to the next				
	(10→15%)	(15→20%)	(20→40%)	(40→60%)	(60→80%)
Time to reach 10% goal	0.442***	0.420***	0.252***	0.289***	0.316***
	(0.015)	(0.015)	(0.017)	(0.017)	(0.017)
Project funding goal	-0.000	0.002	0.011**	0.007	0.007
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
Duration	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number of webs	0.006	0.006	0.002	0.004	0.003
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
Quick Update	-0.078**	-0.080**	-0.114***	-0.063*	-0.061*
	(0.031)	(0.032)	(0.035)	(0.035)	(0.035)
Created projects	-0.002	-0.003	-0.006***	-0.003	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Backed projects	0.000	0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Creator indirect experience	-0.001	-0.001	0.002	0.002	0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Sentiment of risk description	0.005	0.006	0.006	0.009**	0.008*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Project description length	0.004	0.002	-0.001	0.003	0.003
	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
Category Controls	Yes	Yes	Yes	Yes	Yes
Constant	0.065	0.085*	0.168***	0.148***	0.137***
	(0.043)	(0.044)	(0.049)	(0.049)	(0.049)
Observations	3,479	3,479	3,479	3,479	3,479
R^2	0.253	0.226	0.101	0.108	0.120
P	0	0	0	0	0

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1 Category control for all models. OLS regression. All variables are defined in the main text.

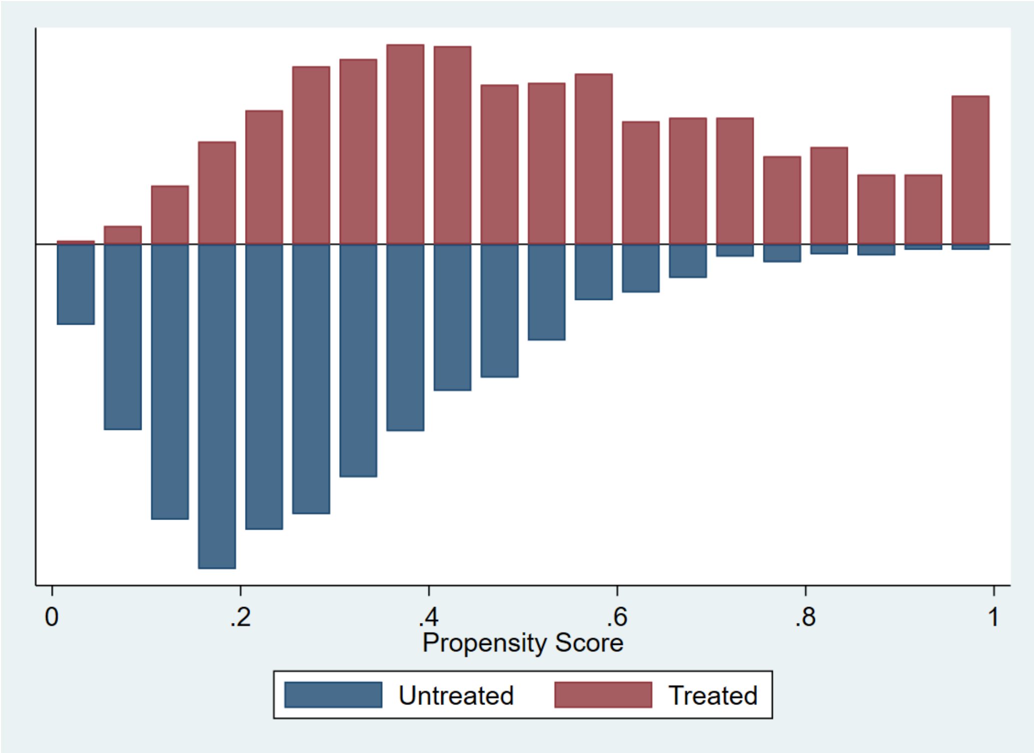
Table 1.6: Instrumental Variable analysis of late backers' response to early backers' behavior (instrumented by dummy on Aggregation Effect).

	(1)	(2)	(3)	(4)	(5)	(6)
	first	second	second	second	second	second
	Time to reach	Time to move from one % to the next				
VARIABLES	10% goal	(10→15%)	(15→20%)	(20→40%)	(40→60%)	(60→80%)
Dummy Aggregation Effect	-0.060***					
(*instrument*)	(0.015)					
Time to reach 10% goal		0.957***	0.891***	0.601**	0.611**	0.638**
(*instrumented*)		(0.244)	(0.243)	(0.252)	(0.249)	(0.249)
Project funding goal	0.069***	-0.037**	-0.032*	-0.016	-0.017	-0.017
	(0.004)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Duration	0.003***	-0.004***	-0.004***	-0.003**	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number of webs	-0.032***	0.023**	0.022**	0.014	0.015	0.014
	(0.005)	(0.009)	(0.009)	(0.010)	(0.009)	(0.009)
Quick Update	-0.155***	0.002	-0.010	-0.067	-0.020	-0.015
	(0.036)	(0.053)	(0.053)	(0.055)	(0.054)	(0.054)
Created projects	-0.007***	0.001	0.000	-0.005	-0.001	0.001
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Backed projects	-0.001***	0.001	0.001	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Creator indirect experience	-0.031***	0.015*	0.014	0.012	0.012	0.013
	(0.004)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Sentiment risk description	-0.020***	0.015**	0.015**	0.012*	0.015**	0.014**
	(0.004)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
Project description length	-0.066***	0.038**	0.033*	0.022	0.024	0.024
	(0.006)	(0.017)	(0.017)	(0.018)	(0.017)	(0.017)
Categories Control	No	No	No	No	No	No
Constant	0.320***	-0.056	-0.023	0.105	0.091	0.074
	(0.047)	(0.081)	(0.081)	(0.084)	(0.083)	(0.083)
Observations	3,479	3,479	3,479	3,479	3,479	3,479
R^2	0.211				0.002	0.016
F test	0	0	0	0	0	0
chi2	.	171.5	167.3	126.5	91.41	90.52
McFadden's adjusted R^2	0.209	.	.	.	-0.000890	0.0136

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Categories not controlled for any model due to collinearity with the instrument. OLS regression similar to Table 5b but instrumenting "Time to reach 10% goal" with instrumental variable "Dummy Aggregation Effect". The instrument is not weak (exclusion restriction is met) since it has a significant effect on instrumented variable (see column 1), F-test of the instrument in the first stage is 16.73, significantly high, and relatively high R-squared. All variables are defined in the main text.

Appendix 1

Figure 1.A1: Sample matching after Propensity Score Matching



Notes: The figure displays a distribution of treated and untreated observations among their propensity score. Treated observations are those that have a significantly high investors' interest, hence those projects that receive 10% of the funding goal in less than 1% of the time. We match based on all the control variables as well as the category the project belongs to. We match on replacement, and we require each observation of the projects affected by the significantly high investors' interest to be matched to the closest neighbor. The figure shows an equilibrated enough distribution of observations, achieving similarity between the two groups in terms of their project attributes.

Table 1.A1: Controls for project size.

Panel A. Equivalent to Tables 1.3 and 1.4 in text				
	(1)	(2)	(3)	(4)
Equivalent original tables:	(T1.3)	(T1.3)	(T1.3)	(T1.4)
	Success	Success	Success	Time to reach 10% goal
Time to reach 10% goal		-7.779***	-7.742***	
		(0.507)	(0.553)	
Number of comments to reach 10% goal		-0.332**	-0.329**	
		(0.138)	(0.140)	
Project funding goal	-0.577***		-0.147	0.099***
	(0.068)		(0.112)	(0.010)
Duration	-0.013***		0.017**	0.004***
	(0.005)		(0.009)	(0.001)
Number of webs	0.157***		0.080	-0.038***
	(0.036)		(0.057)	(0.006)
Quick Update	0.986***		0.391	-0.210***
	(0.247)		(0.348)	(0.047)
Created projects	0.034		0.003	-0.007
	(0.032)		(0.027)	(0.004)
Backed projects	0.013**		0.003	-0.003***
	(0.006)		(0.006)	(0.001)
Creator indirect experience	0.104***		-0.004	-0.023***
	(0.031)		(0.047)	(0.006)
Sentiment of risk description	0.106***		0.069	-0.019***
	(0.029)		(0.048)	(0.005)
Project description length	0.400***		0.172***	-0.062***
	(0.050)		(0.059)	(0.007)
Category Controls	Yes	Yes	Yes	Yes
Constant	2.072***	1.612***	1.194	0.001
	(0.723)	(0.459)	(1.164)	(0.113)
Observations	1,957	2,262	1,957	1,957
F test	0	0	0	0
chi2	375.6	1734	1551	.
p	0	0	0	0
R^2				0.228
McFadden's adjusted R^2	0.153	0.648	0.631	.

Notes: Project size restriction: drop those projects with a project goal lower than $5000. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Category control for all models.

Panel B. Equivalent to Table 1.5, Panel A in text					
	(1)	(2)	(3)	(4)	(5)
	Time to move from one % to the next				
VARIABLES	(10→15%)	(15→20%)	(20→40%)	(40→60%)	(60→80%)
Project funding goal	0.051***	0.048***	0.030***	0.030***	0.034***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Duration	-0.000	-0.000	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number of webs	-0.006	-0.007	-0.002	0.000	-0.000
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Quick Update	-0.158***	-0.161***	-0.170***	-0.108**	-0.124**
	(0.048)	(0.048)	(0.049)	(0.050)	(0.050)
Created projects	-0.006	-0.006	-0.010**	-0.005	0.001
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Backed projects	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Creator indirect experience	-0.009	-0.009	0.000	-0.003	-0.006
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Sentiment of risk description	0.001	0.000	0.002	0.004	0.001
	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
Project description length	-0.019***	-0.017**	-0.014*	-0.012	-0.012
	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)
Category Controls	Yes	Yes	Yes	Yes	Yes
Constant	-0.027	0.001	0.235**	0.202*	0.175
	(0.116)	(0.117)	(0.120)	(0.121)	(0.122)
Observations	1,957	1,957	1,957	1,957	1,957
F test	0	3.05e-10	3.71e-05	0.00345	0.00560
p	0	3.05e-10	3.71e-05	0.00345	0.00560
R^2	0.049	0.046	0.030	0.023	0.022

Notes: Project size restriction: drop those projects with a project goal lower than $5000. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Category control for all models.

Panel C. Equivalent to Table 1.5, Panel B in text					
	(1)	(2)	(3)	(4)	(5)
	Time to move from one % to the next				
VARIABLES	(10→15%)	(15→20%)	(20→40%)	(40→60%)	(60→80%)
Time to reach 10% goal	0.417***	0.386***	0.207***	0.253***	0.284***
	(0.021)	(0.022)	(0.024)	(0.024)	(0.024)
Number of comments to reach 10% goal	0.013	0.013*	0.014	0.010	0.011
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Project funding goal	0.008	0.008	0.008	0.005	0.005
	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)
Duration	-0.002***	-0.002***	-0.002**	-0.002**	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number of webs	0.010	0.008	0.005	0.010	0.010
	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
Quick Update	-0.071	-0.080*	-0.127***	-0.055	-0.064
	(0.044)	(0.045)	(0.049)	(0.049)	(0.049)
Created projects	-0.002	-0.003	-0.008*	-0.003	0.003
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
Backed projects	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Creator indirect experience	0.001	0.001	0.005	0.003	0.000
	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
Sentiment of risk description	0.009*	0.008	0.006	0.009	0.006
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Project description length	0.006	0.006	-0.002	0.003	0.005
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Category Controls	Yes	Yes	Yes	Yes	Yes
Constant	-0.016	0.013	0.248**	0.211*	0.185
	(0.106)	(0.109)	(0.117)	(0.117)	(0.118)
Observations	1,957	1,957	1,957	1,957	1,957
F test	0	0	0	0	0
p	0	0	0	0	0
R^2	0.209	0.182	0.070	0.080	0.093

Notes: Project size restriction: drop those projects with a project goal lower than $5000. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Category control for all models.

Panel D. Equivalent to Table 1.5, Panel C in text					
	(1)	(2)	(3)	(4)	(5)
	Time to move from one % to the next				
VARIABLES	(10→15%)	(15→20%)	(20→40%)	(40→60%)	(60→80%)
Time to reach 10% goal	0.420***				
	(0.021)				
Time to reach 15% goal		0.446***			
		(0.022)			
Time to reach 20% goal			0.299***		
			(0.024)		
Time to reach 40% goal				0.393***	
				(0.027)	
Time to reach 60% goal					0.446***
					(0.030)
Project funding goal	0.009	0.005	0.004	0.004	0.011
	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)
Duration	-0.002***	-0.002***	-0.002**	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number of webs	0.010	0.006	0.005	0.008	0.005
	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)
Quick Update	-0.070	-0.050	-0.090*	0.002	-0.016
	(0.044)	(0.044)	(0.048)	(0.048)	(0.048)
Created projects	-0.003	-0.002	-0.007	0.002	0.007
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
Backed projects	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Creator indirect experience	0.001	0.003	0.007	0.003	-0.001
	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
Sentiment of risk description	0.009*	0.008*	0.008	0.010*	0.007
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Project description length	0.007	0.007	0.000	0.005	0.004
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Category Controls	Yes	Yes	Yes	Yes	Yes
Constant	-0.028	-0.014	0.202*	0.048	-0.068
	(0.106)	(0.106)	(0.115)	(0.115)	(0.116)
Observations	1,957	1,957	1,957	1,957	1,957
F test	0	0	0	0	0
p	0	0	0	0	0
R^2	0.208	0.216	0.101	0.120	0.123

Notes: Project size restriction: drop those projects with a project goal lower than $5000. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Category control for all models.

Table 1.A2: Alternative definitions of success and estimation methods.

Panel A. Linear Probability Model for Table 1.3 in text			
	(1)	(2)	(3)
VARIABLES		Success	
Time to reach 10% goal		-0.795***	-0.768***
		(0.010)	(0.013)
Number of comments to reach 10% goal		-0.008*	-0.013**
		(0.004)	(0.006)
Project funding goal	-0.068***		-0.012***
	(0.005)		(0.004)
Duration	-0.002***		0.001**
	(0.001)		(0.000)
Number of webs	0.029***		0.005
	(0.005)		(0.004)
Quick update	0.151***		0.035
	(0.038)		(0.026)
Created projects	0.009***		0.004**
	(0.002)		(0.002)
Backed projects	0.002***		0.000
	(0.001)		(0.000)
Creator indirect experience	0.029***		0.006*
	(0.004)		(0.003)
Sentiment of risk description	0.020***		0.006*
	(0.004)		(0.003)
Project description length	0.067***		0.017***
	(0.006)		(0.004)
Category Controls	Yes	Yes	Yes
Constant	0.525***	0.787***	0.713***
	(0.053)	(0.018)	(0.037)
Observations	3,479	3,923	3,479
F test	0	0	0
p	0	0	0
R^2	0.189	0.629	0.610

Notes: LPM regression. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Panel B. Ordered Logistic Regression for Four Levels of Success, for Table 1.3 in text			
	(1)	(2)	(3)
VARIABLES		Level_of_success	
Time to reach 10% goal		-8.443***	-9.454***
		(0.280)	(0.334)
Number of comments to reach 10% goal		-0.105*	-0.133*
		(0.058)	(0.080)
Project funding goal	-0.404***		-0.213***
	(0.024)		(0.032)
Duration	-0.007**		0.014***
	(0.003)		(0.004)
Number of webs	0.153***		0.090***
	(0.023)		(0.030)
Quick update	0.818***		0.364*
	(0.174)		(0.212)
Created projects	0.073***		0.042**
	(0.018)		(0.018)
Backed projects	0.009***		0.001
	(0.003)		(0.003)
Creator indirect experience	0.131***		0.054**
	(0.019)		(0.024)
Sentiment of risk description	0.086***		0.025
	(0.019)		(0.025)
Project description length	0.372***		0.205***
	(0.031)		(0.036)
Category Controls	Yes	Yes	Yes
/cut1	-1.261***	-6.471***	-6.406***
	(0.247)	(0.298)	(0.421)
/cut2	-0.439*	-1.914***	-1.585***
	(0.246)	(0.147)	(0.310)
/cut3	0.307	-0.646***	-0.243
	(0.247)	(0.142)	(0.309)
Observations	3,479	3,923	3,479
F test	0	0	0
p	0	0	0
chi2	1022	4893	4814
McFadden's adjusted R^2	0.115	0.496	0.541

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Ordered Logistic Regression. Level_of_success = 4, 3 if funding_level<=1.1, 2 if funding_level<=0.99, 1 if funding_level<=0.1

Panel C. Other measures of success: Absolut number of days to reach 100% funding goal.				
	(1)	(2)	(3)	(4)
Equivalent table in text:		Table 1.3		Table 1.4
	# days to reach 100% funding goal			# days to reach 10% funding goal
Time to reach 10% goal		15.820***	10.421***	
		(0.514)	(0.366)	
Number of comments to reach 10% goal		0.856***	0.509***	
		(0.260)	(0.168)	
Project funding goal	1.510***		0.730***	2.312***
	(0.109)		(0.101)	(0.161)
Duration	0.868***		0.834***	0.664***
	(0.014)		(0.013)	(0.021)
Number of webs	-0.393***		-0.080	-1.088***
	(0.116)		(0.105)	(0.172)
Quick update	-5.378***		-3.820***	-4.456***
	(0.849)		(0.763)	(1.258)
Created projects	-0.185***		-0.107**	-0.036
	(0.056)		(0.050)	(0.082)
Backed projects	-0.040***		-0.023**	-0.045**
	(0.012)		(0.011)	(0.018)
Creator indirect experience	-0.534***		-0.222**	-0.825***
	(0.097)		(0.088)	(0.144)
Sentiment of risk description	-0.296***		-0.097	-0.681***
	(0.100)		(0.090)	(0.148)
Project description length	-0.881***		-0.210*	-2.320***
	(0.134)		(0.122)	(0.198)
Category Controls	Yes	Yes	Yes	Yes
Constant	-6.179***	16.444***	-8.576***	-6.879***
	(1.185)	(0.887)	(1.070)	(1.755)
Observations	3,479	3,501	3,479	3,479
F test	0	0	0	0
p	0	0	0	0
R^2	0.625	0.263	0.699	0.394

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Panel D. Other measures of success: Absolut number of days to reach 100% funding goal, for Table 1.5					
	(1)	(2)	(3)	(4)	(5)
	# days to move from one % to the next				
VARIABLES	(10→15%)	(15→20%)	(20→40%)	(40→60%)	(60→80%)
Project funding goal	2.386***	2.298***	2.110***	2.004***	1.988***
	(0.166)	(0.166)	(0.148)	(0.149)	(0.153)
Duration	0.660***	0.674***	0.747***	0.747***	0.738***
	(0.021)	(0.021)	(0.019)	(0.019)	(0.020)
Number of webs	-1.030***	-0.940***	-0.870***	-0.851***	-0.824***
	(0.177)	(0.177)	(0.158)	(0.159)	(0.163)
Quick update	-5.003***	-5.088***	-5.490***	-5.012***	-5.221***
	(1.298)	(1.294)	(1.159)	(1.166)	(1.196)
Created projects	-0.068	-0.093	-0.203***	-0.158**	-0.107
	(0.085)	(0.085)	(0.076)	(0.076)	(0.078)
Backed projects	-0.040**	-0.042**	-0.046***	-0.043***	-0.048***
	(0.019)	(0.018)	(0.017)	(0.017)	(0.017)
Creator indirect experience	-0.865***	-0.956***	-0.755***	-0.752***	-0.674***
	(0.149)	(0.148)	(0.133)	(0.133)	(0.137)
Sentiment of risk description	-0.674***	-0.699***	-0.611***	-0.620***	-0.661***
	(0.153)	(0.152)	(0.136)	(0.137)	(0.141)
Project description length	-2.414***	-2.371***	-1.954***	-1.892***	-1.900***
	(0.204)	(0.204)	(0.182)	(0.184)	(0.188)
Category Controls	Yes	Yes	Yes	Yes	Yes
Constant	-6.409***	-5.065***	-5.342***	-4.607***	-4.179**
	(1.812)	(1.805)	(1.617)	(1.627)	(1.668)
Observations	3,479	3,479	3,479	3,479	3,479
F test	0	0	0	0	0
p	0	0	0	0	0
R^2	0.380	0.385	0.455	0.444	0.422

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Panel E. Other measures of success: Absolut number of days to reach 100% funding goal, for Table 1.5					
	(1)	(2)	(3)	(4)	(5)
	# days to move from one % to the next				
VARIABLES	(10→15%)	(15→20%)	(20→40%)	(40→60%)	(60→80%)
# days to reach 10% goal	0.947***	0.908***	0.697***	0.677***	0.683***
	(0.007)	(0.008)	(0.010)	(0.011)	(0.011)
Number of comments to reach 10% goal	0.220*	0.316**	0.493***	0.419**	0.386**
	(0.113)	(0.133)	(0.166)	(0.175)	(0.183)
Project funding goal	0.181***	0.177**	0.465***	0.411***	0.382***
	(0.068)	(0.080)	(0.100)	(0.105)	(0.110)
Duration	0.030***	0.070***	0.284***	0.297***	0.284***
	(0.010)	(0.011)	(0.014)	(0.015)	(0.016)
Number of webs	-0.001	0.045	-0.115	-0.117	-0.084
	(0.070)	(0.083)	(0.104)	(0.109)	(0.114)
Quick update	-0.784	-1.045*	-2.388***	-1.999**	-2.180***
	(0.513)	(0.604)	(0.754)	(0.794)	(0.830)
Created projects	-0.032	-0.058	-0.174***	-0.130**	-0.079
	(0.034)	(0.040)	(0.049)	(0.052)	(0.054)
Backed projects	0.003	-0.001	-0.014	-0.012	-0.017
	(0.007)	(0.009)	(0.011)	(0.011)	(0.012)
Creator indirect experience	-0.081	-0.204***	-0.174**	-0.188**	-0.106
	(0.059)	(0.069)	(0.087)	(0.091)	(0.095)
Sentiment of risk description	-0.028	-0.080	-0.134	-0.157*	-0.195**
	(0.060)	(0.071)	(0.089)	(0.094)	(0.098)
Project description length	-0.220***	-0.270***	-0.346***	-0.329***	-0.322**
	(0.082)	(0.097)	(0.121)	(0.127)	(0.133)
Category Controls	Yes	Yes	Yes	Yes	Yes
Constant	0.215	1.336	-0.304	0.256	0.711
	(0.718)	(0.846)	(1.056)	(1.111)	(1.161)
Observations	3,479	3,479	3,479	3,479	3,479
F test	0	0	0	0	0
p	0	0	0	0	0
R^2	0.903	0.866	0.770	0.743	0.723

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Panel F. Other measures of success: Absolut number of days to reach 100% funding goal, for Table 1.5					
	(1)	(2)	(3)	(4)	(5)
		# days to move from one % to the next			
VARIABLES	(10→15%)	(15→20%)	(20→40%)	(40→60%)	(60→80%)
# days to reach 10% goal	0.948***				
	(0.007)				
# days to reach 15% goal		0.973***			
		(0.007)			
# days to reach 20% goal			0.822***		
			(0.009)		
# days to reach 40% goal				0.931***	
				(0.010)	
# days to reach 60% goal					0.985***
					(0.013)
Project funding goal	0.193***	0.114*	0.352***	0.221***	0.309***
	(0.068)	(0.065)	(0.080)	(0.081)	(0.095)
Duration	0.029***	-0.015	0.142***	-0.000	-0.086***
	(0.010)	(0.009)	(0.012)	(0.013)	(0.016)
Number of webs	0.002	-0.053	-0.178**	-0.200**	-0.272***
	(0.070)	(0.067)	(0.084)	(0.084)	(0.099)
Quick update	-0.776	0.347	-0.649	0.658	0.377
	(0.514)	(0.492)	(0.614)	(0.615)	(0.726)
Created projects	-0.033	-0.010	-0.101**	0.047	0.103**
	(0.034)	(0.032)	(0.040)	(0.040)	(0.047)
Backed projects	0.003	0.001	-0.008	0.001	-0.005
	(0.007)	(0.007)	(0.009)	(0.009)	(0.010)
Creator indirect experience	-0.083	-0.063	0.018	-0.029	0.005
	(0.059)	(0.056)	(0.070)	(0.070)	(0.083)
Sentiment of risk description	-0.028	-0.020	-0.036	-0.084	-0.209**
	(0.060)	(0.058)	(0.072)	(0.072)	(0.085)
Project description length	-0.213***	-0.287***	-0.324***	-0.381***	-0.647***
	(0.082)	(0.079)	(0.098)	(0.098)	(0.115)
Category Controls	Yes	Yes	Yes	Yes	Yes
Constant	0.115	1.521**	-0.258	0.992	1.809*
	(0.717)	(0.686)	(0.855)	(0.856)	(1.010)
Observations	3,479	3,479	3,479	3,479	3,479
F test	0	0	0	0	0
p	0	0	0	0	0
R^2	0.903	0.912	0.848	0.847	0.789

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 1.A3: Seasonal and daily effects

VARIABLES	(1) Success	(2) Success	(3) Success	(4) Success	(5) Success
D.week 47				-0.145	-0.141
				(0.127)	(0.127)
D.week 48				-0.163	-0.169
				(0.131)	(0.131)
D.week 49				-0.202	-0.222
				(0.218)	(0.218)
D.week 50				-0.363	-0.393*
				(0.225)	(0.226)
D.December		-0.074	-0.069	0.066	0.094
		(0.078)	(0.079)	(0.176)	(0.177)
D.Weekend	-0.158		-0.153		-0.169
	(0.113)		(0.114)		(0.114)
Project funding goal	-0.368***	-0.368***	-0.369***	-0.368***	-0.370***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Duration	-0.007**	-0.007**	-0.007**	-0.007**	-0.007**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Number of webs	0.137***	0.137***	0.137***	0.136***	0.136***
	(0.026)	(0.026)	(0.026)	(0.026)	(0.027)
Quick Update	0.730***	0.736***	0.731***	0.733***	0.727***
	(0.194)	(0.194)	(0.194)	(0.194)	(0.194)
Created projects	0.076***	0.076***	0.076***	0.076***	0.076***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Backed projects	0.010***	0.010***	0.010***	0.010***	0.010***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Creator indirect experience	0.127***	0.126***	0.127***	0.126***	0.127***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Sentiment of risk description	0.100***	0.100***	0.099***	0.099***	0.099***
	(0.022)	(0.022)	(0.022)	(0.023)	(0.023)
Project description length	0.397***	0.395***	0.396***	0.395***	0.396***
	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)
Category Controls	Yes	Yes	Yes	Yes	Yes
Constant	-0.009	0.002	0.029	0.114	0.146
	(0.283)	(0.286)	(0.286)	(0.297)	(0.298)
Observations	3,479	3,479	3,479	3,479	3,479
F test	0	0	0	0	0
chi2	742.4	741.3	743.2	745.0	747.2
p	0	0	0	0	0
McFadden's adjusted R^2	0.158	0.158	0.158	0.159	0.159

Notes: Weekend, week, and month dummies are incorporated in the regression. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 1.A4: Identification of early and late backers.

Panel A. Threshold at 5% funding goal level						
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Success	Success	days_5	int10_15	int15_20	int20_40
Time to reach 5% goal	-8.752***	-8.462***		0.365***	0.353***	0.198***
	(0.621)	(0.640)		(0.015)	(0.016)	(0.017)
Number of comments to reach 10% goal	-0.832***	-0.710***		0.024***	0.025***	0.019**
	(0.129)	(0.127)		(0.007)	(0.007)	(0.008)
Project funding goal		-0.206***	0.068***	0.005	0.006	0.014***
		(0.039)	(0.005)	(0.004)	(0.004)	(0.005)
Duration		0.007	0.003***	-0.002***	-0.002***	-0.001**
		(0.005)	(0.001)	(0.001)	(0.001)	(0.001)
Number of webs		0.074**	-0.036***	0.006	0.006	0.002
		(0.036)	(0.005)	(0.004)	(0.004)	(0.005)
Quick update		0.469*	-0.139***	-0.093***	-0.095***	-0.124***
		(0.246)	(0.036)	(0.032)	(0.033)	(0.035)
Created projects		0.038*	-0.005**	-0.003	-0.004*	-0.007***
		(0.021)	(0.002)	(0.002)	(0.002)	(0.002)
Backed projects		0.004	-0.001**	-0.000	-0.000	-0.000
		(0.004)	(0.001)	(0.000)	(0.000)	(0.001)
Creator indirect experience		0.064**	-0.026***	-0.004	-0.003	-0.000
		(0.029)	(0.004)	(0.004)	(0.004)	(0.004)
Sentiment of risk description		0.073**	-0.019***	0.004	0.004	0.005
		(0.031)	(0.004)	(0.004)	(0.004)	(0.004)
Project description length		0.174***	-0.071***	0.000	-0.001	-0.004
		(0.041)	(0.006)	(0.005)	(0.005)	(0.006)
Category Controls	Yes	Yes	Yes	Yes	Yes	Yes
Constant	1.227***	1.030***	0.258***	0.095**	0.113**	0.190***
	(0.178)	(0.365)	(0.050)	(0.045)	(0.046)	(0.050)
Observations	3,923	3,479	3,479	3,479	3,479	3,479
F test	0	0	0	0	0	0
p	0	0	0	0	0	0
R^2			0.222	0.195	0.180	0.081
McFadden's adjusted R^2	0.541	0.520	.	.	.	.

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Panel B. Threshold at 15% funding goal level							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	Success	Success	days_15	int10_15	int15_20	int20_40	int40_60
Time to reach 15% goal	-7.177***	-7.028***		0.531***	0.457***	0.283***	0.317***
	(0.271)	(0.290)		(0.014)	(0.015)	(0.017)	(0.017)
Number of comments to	-0.133	-0.120		0.002	0.006	0.007	0.002
reach 10% goal	(0.129)	(0.130)		(0.006)	(0.007)	(0.008)	(0.008)
Project funding goal		-0.167***	0.070***	-0.006	-0.001	0.008*	0.005
		(0.046)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Duration		0.022***	0.003***	-0.002***	-0.002***	-0.002***	-0.002***
		(0.006)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
Number of webs		0.121***	-0.026***	0.007*	0.005	0.002	0.004
		(0.045)	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)
Quick update		-0.034	-0.187***	-0.044	-0.058*	-0.099***	-0.047
		(0.271)	(0.035)	(0.029)	(0.031)	(0.035)	(0.035)
Created projects		0.016	-0.009***	-0.000	-0.001	-0.006**	-0.002
		(0.020)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Backed projects		0.003	-0.001***	0.000	0.000	-0.000	-0.000
		(0.004)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Creator indirect experience		0.012	-0.032***	0.003	0.002	0.003	0.004
		(0.034)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)
Sentiment of risk description		0.054	-0.020***	0.007**	0.007*	0.006	0.009**
		(0.038)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)
Project description length		0.213***	-0.061***	0.008	0.003	-0.000	0.004
		(0.048)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)
Category Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	2.036***	1.150***	0.292***	0.023	0.061	0.152***	0.130***
	(0.222)	(0.432)	(0.049)	(0.041)	(0.044)	(0.049)	(0.049)
Observations	3,923	3,479	3,479	3,479	3,479	3,479	3,479
p	0	0	0	0	0	0	0
R^2			0.233	0.332	0.254	0.116	0.122
McFadden's adjusted R^2	0.662	0.646	.	.	.	.	.

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Panel B. Threshold at 20% funding goal level							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	Success	Success	days_20	int10_15	int15_20	int20_40	int40_60
Time to reach 20% goal	-7.178***	-7.085***		0.519***	0.524***	0.313***	0.338***
	(0.239)	(0.260)		(0.015)	(0.015)	(0.017)	(0.017)
Number of comments to	0.022	0.016		0.001	0.001	0.005	0.000
reach 10% goal	(0.117)	(0.127)		(0.007)	(0.007)	(0.008)	(0.008)
Project funding goal		-0.165***	0.068***	-0.004	-0.004	0.007	0.005
		(0.049)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Duration		0.025***	0.003***	-0.002***	-0.003***	-0.002***	-0.002***
		(0.006)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
Number of webs		0.140***	-0.023***	0.005	0.005	0.002	0.003
		(0.049)	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)
Quick update		-0.294	-0.205***	-0.037	-0.036	-0.088**	-0.037
		(0.284)	(0.035)	(0.030)	(0.030)	(0.035)	(0.035)
Created projects		-0.001	-0.010***	0.000	0.000	-0.005**	-0.001
		(0.018)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Backed projects		0.003	-0.001***	0.000	0.000	-0.000	-0.000
		(0.005)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Creator indirect experience		-0.010	-0.032***	0.003	0.004	0.004	0.004
		(0.038)	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)
Sentiment of risk description		0.049	-0.019***	0.007*	0.008**	0.007*	0.010**
		(0.041)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)
Project description length		0.230***	-0.059***	0.005	0.005	0.001	0.004
		(0.052)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
Category Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	2.557***	1.577***	0.337***	0.003	0.016	0.128***	0.108**
	(0.252)	(0.472)	(0.048)	(0.042)	(0.042)	(0.049)	(0.048)
Observations	3,923	3,479	3,479	3,479	3,479	3,479	3,479
p	0	0	0	0	0	0	0
R^2			0.234	0.312	0.308	0.130	0.131
McFadden's adjusted R^2	0.698	0.686	.	.	.	.	.

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 1.A5: Additional measures of initial backers' interest (number of comments)

Panel A.		
	(1)	(2)
	Success	Success
Time to reach 10% goal	-7.274***	-7.034***
	(0.329)	(0.349)
Number of comments to reach 10% goal	-0.345***	-0.292**
	(0.124)	(0.123)
Project funding goal		-0.172***
		(0.042)
Duration		0.015***
		(0.005)
Number of webs		0.092**
		(0.041)
Quick update		0.310
		(0.264)
Created projects		0.029
		(0.020)
Backed projects		0.003
		(0.004)
Creator indirect experience		0.037
		(0.032)
Sentiment of risk description		0.063*
		(0.035)
Project description length		0.192***
		(0.044)
Category Controls	Yes	Yes
Constant	1.592***	0.945**
	(0.196)	(0.392)
Observations	3,923	3,479
chi2	3124	2768
P	0	0
McFadden's adjusted R^2	0.610	0.590

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Category control for all models. Logistic regression is used. All variables are defined in the main text.

Panel B.

	(1)	(2)	(3)	(4)	(5)
		Time to move from one % to the next			
	(10→15%)	(15→20%)	(20→40%)	(40→60%)	(60→80%)
Time to reach 10% goal	0.435***	0.415***	0.247***	0.285***	0.312***
	(0.015)	(0.015)	(0.017)	(0.017)	(0.017)
Number of comments to reach 10% goal	0.010	0.011	0.011	0.006	0.007
	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)
Project funding goal	-0.000	0.001	0.010**	0.007	0.006
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
Duration	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number of webs	0.006	0.006	0.002	0.004	0.003
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
Quick Update	-0.078**	-0.081**	-0.115***	-0.063*	-0.061*
	(0.031)	(0.032)	(0.035)	(0.035)	(0.035)
Created projects	-0.002	-0.002	-0.006***	-0.003	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Backed projects	0.000	0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Creator indirect experience	-0.001	-0.001	0.002	0.002	0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Sentiment of risk description	0.005	0.006	0.006	0.009**	0.008*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Project description length	0.003	0.002	-0.001	0.003	0.003
	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
Category Controls	Yes	Yes	Yes	Yes	Yes
Constant	0.072*	0.092**	0.174***	0.152***	0.141***
	(0.043)	(0.045)	(0.049)	(0.049)	(0.049)
Observations	3,479	3,479	3,479	3,479	3,479
R^2	0.249	0.225	0.101	0.107	0.119
P	0	0	0	0	0

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1 Category control for all models. OLS regression. All variables are defined in the main text.

Table 1.A6: Controls for endogeneity.

Panel A. Instrumental Variable analysis of late backers' response to early backers' behavior (Instrumented by dummy on Aggregation Effect)						
	(1)	(2)	(3)	(4)	(5)	(6)
IV regression stage:	first	second	second	second	second	second
	Time to reach	Time to move from one % to the next				
VARIABLES	10% goal	(10→15%)	(15→20%)	(20→40%)	(40→60%)	(60→80%)
Dummy Aggregation Effect (*instrument*)	-0.060***					
	(0.015)					
Time to reach 10% goal (*instrumented*)		0.957***	0.891***	0.601**	0.611**	0.638**
		(0.244)	(0.243)	(0.252)	(0.249)	(0.249)
Project funding goal	0.069***	-0.037**	-0.032*	-0.016	-0.017	-0.017
	(0.004)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Duration	0.003***	-0.004***	-0.004***	-0.003**	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number of webs	-0.032***	0.023**	0.022**	0.014	0.015	0.014
	(0.005)	(0.009)	(0.009)	(0.010)	(0.009)	(0.009)
Quick Update	-0.155***	0.002	-0.010	-0.067	-0.020	-0.015
	(0.036)	(0.053)	(0.053)	(0.055)	(0.054)	(0.054)
Created projects	-0.007***	0.001	0.000	-0.005	-0.001	0.001
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Backed projects	-0.001***	0.001	0.001	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Creator indirect experience	-0.031***	0.015*	0.014	0.012	0.012	0.013
	(0.004)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Sentiment of risk description	-0.020***	0.015**	0.015**	0.012*	0.015**	0.014**
	(0.004)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
Project description length	-0.066***	0.038**	0.033*	0.022	0.024	0.024
	(0.006)	(0.017)	(0.017)	(0.018)	(0.017)	(0.017)
Categories Control	No	No	No	No	No	No
Constant	0.320***	-0.056	-0.023	0.105	0.091	0.074
	(0.047)	(0.081)	(0.081)	(0.084)	(0.083)	(0.083)
Observations	3,479	3,479	3,479	3,479	3,479	3,479
R^2	0.211				0.002	0.016
F test	0	0	0	0	0	0
chi2	.	171.5	167.3	126.5	91.41	90.52
p	0	0	0	0	0	0
McFadden's adjusted R^2	0.209	.	.	.	-0.000890	0.0136

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Categories not controlled for any model due to collinearity with the instrument. OLS regression similar to Table 5b but instrumenting "Time to reach 10% goal" with instrumental variable "Dummy Aggregation Effect". The instrument is not weak (exclusion restriction is met) since it has a significant effect on instrumented variable (see column 1), F-test of the instrument in the first stage is 16.73 significantly high, and a relatively high R-squared. All variables are defined in the main text.

Panel B. Error term: Extra information added by followers. For Table 1.5, panel B.					
	(1)	(2)	(3)	(4)	(5)
	Time to move from one % to the next				
VARIABLES	(10→15%)	(15→20%)	(20→40%)	(40→60%)	(60→80%)
ehat (Private orthogonal information)	0.435***	0.415***	0.247***	0.285***	0.312***
	(0.015)	(0.015)	(0.017)	(0.017)	(0.017)
Project funding goal	0.031***	0.031***	0.028***	0.027***	0.029***
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)
Duration	-0.000	-0.001*	-0.001	-0.001**	-0.001**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number of webs	-0.007*	-0.007	-0.005	-0.004	-0.007
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
Quick Update	-0.144***	-0.143***	-0.152***	-0.106***	-0.108***
	(0.031)	(0.032)	(0.035)	(0.035)	(0.035)
Created projects	-0.005**	-0.005***	-0.008***	-0.005**	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Backed projects	-0.001	-0.000	-0.001	-0.001	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Creator indirect experience	-0.014***	-0.013***	-0.006	-0.006	-0.006
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Sentiment of risk description	-0.003	-0.002	0.001	0.003	0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Project description length	-0.025***	-0.025***	-0.018***	-0.016***	-0.017***
	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
Category Controls	Yes	Yes	Yes	Yes	Yes
Constant	0.182***	0.197***	0.237***	0.224***	0.221***
	(0.043)	(0.044)	(0.049)	(0.049)	(0.049)
Observations	3,479	3,479	3,479	3,479	3,479
R2	0.249	0.225	0.101	0.107	0.119
F test	0	0	0	0	0
p	0	0	0	0	0

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. . Category control for all models. First, we computed the error term from specification (1) in Table 1.4 and keep the error term (ehat) representing the part of the behavior of the early backers that cannot be explained by the publicly available quality signals, and it is orthogonal to the information included in those variables. We then re-estimated specification (1) from Table 1.5, panel B, but, rather than introducing the original variable Time to reach 10% goal, we used this error term (ehat).

Panel C. PSM-Propensity Score Matching: Late backers' response to early backers' behavior.					
	(1)	(2)	(3)	(4)	(5)
	Time to move from one % to the next				
VARIABLES	(10→15%)	(15→20%)	(20→40%)	(40→60%)	(60→80%)
Dummy Investors' interest	-0.326***	-0.304***	-0.234***	-0.211***	-0.229***
(*treatment variable*)	(0.015)	(0.015)	(0.017)	(0.017)	(0.017)
Project funding goal	0.020***	0.021***	0.031***	0.028***	0.025***
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
Duration	-0.003***	-0.004***	-0.002***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number of webs	0.012**	0.011**	0.006	0.010*	0.008
	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
Quick Update	-0.136***	-0.138***	-0.147***	-0.095***	-0.074**
	(0.027)	(0.027)	(0.031)	(0.031)	(0.031)
Created projects	-0.001	-0.004**	-0.007***	-0.004*	-0.004
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Backed projects	0.000	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Creator indirect experience	-0.010***	-0.013***	-0.010**	-0.010**	-0.008*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Sentiment of risk description	-0.001	-0.000	-0.001	0.002	-0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Project description length	-0.019***	-0.024***	-0.028***	-0.025***	-0.024***
	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)
Categories Control	Yes	Yes	Yes	Yes	Yes
Constant	0.385***	0.402***	0.374***	0.353***	0.376***
	(0.046)	(0.048)	(0.053)	(0.054)	(0.054)
Observations	1,988	1,988	1,988	1,988	1,988
R^2	0.239	0.222	0.144	0.119	0.124
p	0	0	0	0	0
McFadden's adjusted R^2	.	.	.	.	.

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Category control for all models. OLS regression similar to Table 1.5, panel B after performing propensity score matching. We match on replacement and we require each treated observation to be matched to the closest neighbor control observation. PSM has also been performed matching on no replacement with similar results. All variables are defined in the main text.

Chapter 2

Tough love: a story of stigmatization in crowdfunding communities

2.1. Introduction

Entering a community is a tough job. Entrepreneurs entering a new community lack the legitimacy of belonging (Zimmerman & Zeitz, 2002) and struggle to acquire the resources that will help them to acquire resources to grow (Lounsbury & Glynn, 2001; Wry, Lounsbury, & Glynn, 2011). These newcomer entrepreneurs are usually considered foreigners and aliens to the group and can be strongly criticized and even expelled from the community when not adhering to the social norms (Stinchcombe, 1965). As foreigners coming to a new country, entrepreneurs entering a new community can eventually suffer from the stigma of being new. Stigmatization happens when a critical mass of actors in the community perceives an organization as “having values counter to their own” (Devers, Dewett, Mishina, & Belsito, 2009, p. 157). The negative attribution that members of a stakeholder group in a community make is about the “core essence” of the organization (Mishina & Devers, 2012, p. 205), where the organization is seen as “a dangerous deviant”, that is the embodiment of everything that the community considers wrong (Devers et al., 2009, p. 162). In this situation, such organizations are therefore subjected to a strong social sanctioning or a vilification process that results in organizational stigma (Wiesenfeld, Wurthmann, & Hambrick, 2008).

Key to the stigmatization to happen is that early claim makers, the valuators, need to hold a status that allows them to be a credible source of negative comments (Bitektine & Haack, 2015). Status in the community is a key characteristic of valuators (Devers et al., 2009; Mishina & Devers, 2012). Valuators are actors who indicate “what evil looks like, what shapes the devil can assume” (Erikson, 1962, p. 310); they utilize shaming strategies as means of social control to define, clarify, and enforce collectively-held values and norms within the community (Creed, Hudson, Okhuysen, & Smith-Crowe, 2014; Doern & Goss, 2014; Scheff, 1990). A predominant assumption in this literature is that valuators use their power with a social purpose, that is, to maintain order and control over community norms and values (Hudson & Okhuysen, 2009) and care about the norm institutionalization that would help the community to reproduce (Bitektain & Haack, 2015). Challenging this assumption, we focus on the valuator's private interests in the stigmatization to ask: What is the valuator's interest in the stigmatization process? Are valuators moved by private interest to confer stigma to a newcomer?

To answer these questions, we extend the literature on social valuation and stigma by studying the private incentives of valuators.

We analyze this issue in the context of an online community of reward-based crowdfunding. We rely on Kickstarter, which is one of the biggest platforms facilitating the practice of soliciting financial contributions from a large number of people. We argue that newcomers can be socially sanctioned when not adhering to the community norms. Social sanctioning and the ulterior stigma push the entrepreneurs out of the community since they are unable to raise funds for their projects. We focus on the study of the valuators in Kickstarter to find that beyond community-oriented interest in preserving social norms, valuators have private interests.

We find that valuators' work on preserving social norms, via the social sanctioning of the deviant newcomers increases the status of the valuators in their community. Surprisingly, we also find limits in the harshness of social sanctioning. The accumulation of negative comments from valuators to new entrepreneurs leads to a reduction in valuators' status and eventually to a reduction in the attractiveness of the community for other potential entrants.

Our theory and findings contribute to extending our understanding of stigmatization processes in entrepreneurship by exploring: (1) the process through which stigmatization can be shaped by valuators' status and (2) the limits in this sanctioning social behavior in which valuators can be sanctioned by the community for being too hard with newcomers.

Extant literature has focused on the characteristics of the stigmatization arguments in the form of shaming strategies (Creed et al., 2014) and the structural organizational characteristics (Doern & Goss, 2014) that mediate the stigmatization processes. This paper focuses on the valuators' influence on the stigmatization processes and its consequences for them as well as for the community they belong to.

We contribute to the literature on organization stigma and entrepreneurship in three ways. First, prior literature on stigma has underlined the importance of power and status in the vilification process, shaming, and ultimately the reproduction and change of institutional arrangements in a community (Erikson, 1962; Doern and Goss, 2014; Creed, 2014). This paper moves a step forward and stresses the importance of attending to the valuators' private interest of gaining power and status beyond the more communitarian interest of norm preservation. Second, we contribute to the literature by showing the recursive nature of legitimacy and illegitimacy (Gehman & Soublière, 2017; J. F. Soublière & Gehman, 2020; Wry et al., 2011) but also its fragility (Hampel & Tracey, 2017; Siltaoja et al. 2020; Phung et al. 2020). We offer a more nuanced understanding of the role of valuators in the stigmatization process. Whereas prior work has viewed norm enforcement and vilification as positive to the valuator and the community, we theorize and show that cumulative vilification can be detrimental for both valuators and communities. Third, our work sheds new light on the dynamics of entrepreneurial endeavors and the formation of online communities. Whereas prior work has theorized that emerging categories are initially cooperative spaces that might then become competitive as they mature (Cattani et al. 2017), our study illustrates how valuators' behavior might lift or rock all boats at any point in time.

2.2. Research Setting: A Crowdfunding Platform

2.2.1. Crowdfunding as community building

Launching a project for the first time is a tough job for entrepreneurs. In fact, obtaining the resources necessary to grow without a funding track record gets very difficult (Lounsbury & Glynn, 2001; Wry et al., 2011). This is even exacerbated when entrepreneurs try to obtain financing from traditional financing methods or institutions such as issuing bonds, equity or commercial bank lending, or venture capital firms and angel investors. Lately, more and more entrepreneurs use platforms-enabled financing methods to finance their projects. One of those methods is reward-based crowdfunding.

Reward-based crowdfunding is an innovative way of overcoming difficulties in early-stage funding such as information asymmetries (Cosh et al., 2009) and lack of the legitimacy of belonging in new communities (Zimmerman & Zeitz, 2002) through small contributions from a large set of “backers”; online platform investors that support the project with individual “pledges”. This method allows entrepreneurs to bypass traditional financial investors, and raise funds from large, online communities that meet on crowdfunding platforms (Schwienbacher and Larralde 2012; Belleflamme et al., 2014; Agrawal et al., 2015; Kuppuswamy and Bayus, 2018b). Furthermore, crowdfunding allows for an interactive development process. It provides valuable information and feedback on the products offered through the comments that online platforms and potential investors provide (Belleflamme et al., 2014; Agrawal et al., 2015; Kuppuswamy and Bayus, 2018b, Eiteneyer et al., 2019). Moreover, reward-based crowdfunding is particular in that the funding comes from customers. This implies that reward-based crowdfunding unites a funding campaign with a marketing campaign. Therefore, it offers a low-cost way to promote products in the market, test the product and its potential demand, and gather information from the reactions of customers. The combination of cheap financing, information gathering, and low-cost promotion makes this financing method particularly attractive for entrepreneurs (Rodríguez-Garnica, Gutiérrez & Tribó, 2022).

Kickstarter is one of the biggest platforms facilitating the practice of crowdfunding, and the principal reward-based crowdfunding platform. Remarkably, the communities emerging from the use of these online platforms like Kickstarter are becoming increasingly closed as a way to maintain their identity and provide clear feedback to those participants in these communities. Such participants approach platforms like Kickstarter not only to obtain resources but, even more importantly, to get cheap feedback from the community in which they participate. However, in these closed communities, incumbents want to avoid cheap feedback behavior, because frequently such type of discontinued feedback results in "scammed" projects (failed projects).

2.2.2. Backers as community guardians

Extant research depicts resource providers (called backers and superbackers, since they support or back up the projects presented by the entrepreneurs) as relatively independent audience members, which understates the fact that their legitimacy assessments are sensitive to social cues (Navis & Glynn, 2011) and often "exhibit herding behavior" (Rodríguez-Garnica, Gutiérrez & Tribó, 2022; Pontikes & Barnett, 2017, p. 140). This cannot be ignored on crowdfunding platforms like Kickstarter, which has brought together an "enormous global community" of millions of entrepreneurs and resource providers. Like all online communities (Fisher, 2019; Massa, 2017), crowdfunding platforms are characterized by a "collective flow of knowledge among community participants" (Faraj, Krogh, Monteiro, & Lakhani, 2016, p. 668), providing fertile ground for studying the dynamics of social valuation strategies. Moreover, since the crowd's legitimacy and illegitimacy criteria derive from a community logic (Fisher, Kuratko, Bloodgood, & Hornsby, 2017), gaining better insights into crowdfunding is important to redress the "default presumption" that resource access is governed by market logics (Clough, Fang, Vissa, & Wu, 2019, p. 247).

In previous studies, Rodríguez-Garnica, Gutiérrez & Tribó, 2022 show that experience and a track record of both activity, investments, and projects launched on Kickstarter are significant determinants for entrepreneurs obtaining its funding goal. Hence, entrepreneurs that lack experience and are new to the platform have less chance of achieving the funding outcome. From our understanding, these newcomer entrepreneurs can be considered foreigners and aliens to the group. Hence backers can act as guardians so that they strongly criticize and even expel new entrepreneurs from the community when not adhering to the social norms (Stinchcombe, 1965). This phenomenon named "stigmatization" is addressed in this paper.

2.2.3. Stigma of the newcomers' and the valuators' interest

Legitimacy is key for entrepreneurs to obtain resources and grow in a community. When entrepreneurs entering a community fail to abide by established community standards and norms, they get discredited. This discredit is not only attributable to malpractice but can also be associated with the organization per se, which gets stigmatized in this process. Stigma has been typically associated with core characteristics of organizations, like the association with ill products. However, stigma can also happen in relation to the temporal status of the organization, organizational failure, and newness in the community (Shepherd & Haynie, 2011; Sutton & Callahan, 1987). Especially, closed communities, like university fraternities or military services, tend to stigmatize newcomers until they are submitted to hazing rituals (Cimino, 2011), or they are assimilated in the process of building the boundaries of the community (Siltaoja et al., 2020).

Stigmatization of newcomers may emerge as an outcome of the incentives of focal agents in a community. Valuators are key agents in a community that hold a significant status and act as a reference to define the essence of what a community is (Devers et al., 2009; Mishina & Devers, 2012). Social status refers to a "socially constructed, intersubjectively agreed-upon, and accepted ordering or ranking of individuals, groups, organizations, or activities in a social system." (Washington & Zajac, 2005: 283). Once a newcomer enters a community, these valuators have incentives to react in a negative and defensive way (Bitektine & Haack, 2015). Two opposite and conflicting reasons explain this behavior. First, valuators legitimately care about the norms and rules that define the essence of the community and might patrol the boundaries of the community in which they have a status, and approach newcomers in a conservative/negative way because newcomers may change the essence of the community (Hudson & Okhuysen, 2009). This is the community-oriented interest view. Second, valuators may only be interested in maintaining their status in their community and eliminating newcomers *per se* as a way to avoid any threat to their status and financial benefits (private interest view). There might be a conflict between both views, which defines an agency problem, given that valuators just following their own private interests end up damaging the survival of the community in which they have a ~~high~~ status. As we further develop, the key element to disentangle whether valuators pursue just their own interests or they also internalize the social interests of the community is whether the negative comments on newcomers are persistent and systematic or not.

Despite increasing research on crowdfunding (Short, Ketchen, McKenny, Allison, & Ireland, 2017; J.-F. Soublière & Gehman, 2020), we are the first to focus on the dynamics of interactions between the newcomers' entrepreneurs and the valuators and the stigmatization dynamics that can make entrepreneurs be banned from the online community and, thus, from an important source of capital. Remarkably, we are going to adopt a broad perspective, analyzing not only the consequence of social sanctioning to the newcomers' entrepreneurs but also to the valuators as well as Kickstarter's communities. The Kickstarter platform is ideal for investigating such relationships and interactions among participants, given that successes and failures are observable[10]. Also, the sample generated does not suffer from the left censoring and survival bias that frequently plagues other studies (Aldrich & Fiol, 1994).

[10]Success in "all-or-nothing" models of crowdfunding (Kickstarter) is reached when a project collects capital equal to or greater than the campaign's target financing amount.

2.3. Hypothesis Development: Private and Community Interest in a Model of Social Valuation of Entrepreneurs Entering a New Community

2.3.1. A valuator's status and stigmatization activity

Entrepreneurs are increasingly using online crowdfunding platforms to promote their ventures and get access to non-traditional funding. In this type of financing, the support from the audience—i.e., the crowd—to the venture is crucial as the entrepreneur has to solicit, through an online platform, financial contributions from a large number of individuals who become resource providers. Though relatively independent, individuals are sensitive to information cues given the complexity of the process of, first, scanning the business environment for relevant information and then interpreting it. Audience members, then, closely watch the words, actions, and opinions of key social valuators.

In crowdfunding platforms like Kickstarter, where there is incomplete information to assess the potential success of individual projects, resource providers—i.e., backers—base their investment decisions on their own assessment of the legitimacy of the projects. In guiding this assessment, backers react to what other backers do and say (Rodríguez-Garnica, Gutiérrez & Tribó, 2022). Of primary importance in this process of collective construction of legitimacy is the role played by the power of visibility and power of influence to measure social impact from backers.

In the first case, superbackers (frequent resource providers), have visibility power through their comments being tagged as "superbacked". Hence, they can be seen as experts in certain domains—or entrepreneurial communities—of the platform. Superbackers might follow a strong community logic and seek to protect and advance their community. For this reason, they typically penalize entrepreneurs that are: (i) new to their community; (ii) have no previous experience, and (iii) have not shown previously to conform with the rule.

In the second case, another way potential investors do categorize and classify information coming from earlier backers is through their status which measures their power of influence. Social status refers to a "socially constructed, intersubjectively agreed-upon, and accepted ordering or ranking of individuals, groups, organizations, or activities in a social system." (Washington & Zajac, 2005: 283). In this study, we define it as the capacity to influence and deviate resource allocation decisions among community members. We measure social status with a proxy that measures the historical (power of) decisiveness of the backer's negative comments to steer the future outcome of the project; in other words, to make the project fail with negative comments. Given a valuator's social valuation, we consider it decisive whenever a valuator adopts a negative stance, and the project fails to attract sufficient funding. Therefore, we will focus on negative comments in this study, since we are analyzing the stigmatization phenomenon. The more negative comments made on projects that end up failing, the higher decisiveness/influence power the backer has within the community (status). New backers in Kickstarter have neither influence power, nor status (value of *zero*).

Within the last decade, most studies on crowdfunding have widely analyzed project-related and entrepreneur-related determinants of funding success. However, not many studies have focused on backers' characteristics which can be categorized as backer-related factors. Backer-related factors are those associated with the people who back the projects; backers. Some of these factors include backers' Positive Affective Reactions to projects (Davis, Hmieleski, Webb, et al., 2017), geography factors measuring the same location of the backer and product backed (Lelo de Larrea, Altin &

Singh, 2019) have positive effects on success. According to another study, backers with higher experience in both backing or launching a project and a higher ratio of previously successful projects backed have a positive effect on/tend to invest higher amounts of funding on projects (Mahmood, Luffarelli & Mukesh, 2019). However, none of these studies have yet analyzed the backers' social/influential characteristics such as status or leadership ex-ante characteristics, and their potential effect on the funding campaign.

Additionally, there are some studies analyzing the sentiment and quantity of the comments (Jiang, Han, Xu, et al. 2020; Wang, Li, Liang, et al., 2018; Courtney, Dutta & Li, 2017) showing a positive effect on funding success. Also, more related to this paper, Butticè, Colombo & Wright (2017) study the social capital generated by serial entrepreneurs using the crowdfunding platform. They analyze the number of comments entrepreneurs historically have received on previously launched projects as a proxy for social capital generated. However, we are not yet aware of a study that gathers and analyses the backers' historical number of comments and their sentiment posted on previously backed projects, and its effect on future projects they comment on and back on.

Acting as a social valuator, superbackers' and high-status backers' opinions about a particular entrepreneurial endeavor are critical for entrepreneurs, particularly newcomers, to be successful in raising the required funds. Negative opinions about an endeavor can, thus, be spread quickly through the community, generating among its members a generalized vilification or negative perception of the new entrepreneur or newcomer and the project, leading to its financial collapse. Therefore, we propose as a baseline hypothesis:

Baseline Hypothesis (BH): *A valuator with high status that conducts social (negative) sanctioning of a new entrepreneur's project, reduces the likelihood of a new entrepreneur raising funding and being successful in his/her project.*

2.3.2. A valuators' private and social interests in social sanctioning

Stigmatization is a process through which valuators engage in audience-minded sensemaking. In this sensemaking process, valuators rely on their social and psychological context to come up with judgments on new entrepreneurs. Although this sensemaking process is typically driven by community norms, valuators are also influenced by different human biases: their own interests as well as biases and those they anticipate in their audiences (Wiesenfeld et al., 2008).

One of these personal biases, which creates agency problems, is valuators aim to gain influence in the community even at the expense of risking its very survival. Research on impression formation has long acknowledged that negative information has a more extreme influence on perceptions than positive information (e.g., Kanouse and Hanson 1972: Peeters and Czapinski 1990). So, audiences are more likely to follow the opinions and actions of valuators that assess entrepreneurs' ventures as inappropriate. This process ends up increasing the influence and status of valuators within their community. Overall, and taking into consideration the baseline hypothesis, valuators have incentives to provide negative information on newcomer entrepreneurs' projects as a way to improve their status within the community.

Hypothesis 1: *A valuator's social sanctioning of a new entrepreneur's project has a positive impact on a valuator's status (decisiveness) in the community.*

2.3.3. The limits to social sanctioning

Although community members become more influenced by a negative assessment than by a positive one, they may interpret a valuator's persistent social sanctioning and their extreme vilification of newcomers as biased behavior to protect private benefits. These are connected to maintaining a valuator's status, rather than an attempt at maintaining and reinforcing community norms. Therefore, when these concerns about the valuators become more salient, their opinions will generate less adhesion by the audience, so that valuators will lose status in their community and, with that, the capacity to influence resource allocation decisions among community members. Valuators significantly increasing the use of stigmatization on new entrepreneurs through the use of persistent negative comments in the community (*cumulative negative social sanctioning*) can have severe consequences for the valuators' individual interests. Therefore, we propose:

> ***Hypothesis 2:*** *A valuator's cumulative social sanctioning reduces the valuator's status (decisiveness) in the community.*

2.3.4. A valuator's cumulative sanctioning and its effect on the attractiveness of the community

At the same time, this persistent social sanctioning will signal to potential new entrepreneurs that other reasons than efficiency considerations are guiding the community's support of entrepreneurial endeavors. As a consequence, good entrepreneurs will not find this community attractive for developing their projects. Only entrepreneurs with inferior projects will apply to this community. The likely failure of these latter projects in getting funds or fulfilling the audience's expectations will create a generalized perception of negativity about the quality of the community's projects, making such a community less attractive for future project applicants. Therefore, we propose:

> ***Hypothesis 3:*** *A valuator's cumulative social sanctioning reduces the attractiveness of the community for potential new entrepreneurs.*

2.4. Data and Methods

2.4.1. Data Collection

We empirically analyze a total of 40,841 projects initiated on Kickstarter between 2013 and 2016. A total of 32,184 entrepreneurs, from which 30,813 are newcomers (95.74%), contributed to 152 different subcategories or tags, classified into 15 broad categories or industries (as defined by Kickstarter), including art, comics, crafts, dance, design, fashion, film-video, food, games, journalism, music, photography, publishing, technology and theatre. For each project, we gathered data on the entrepreneur's experience, financing goals, total funding achieved, start and end of campaign dates, duration of campaign, number, and dates of updates, comments made on the project campaign, and on the project updates and category. Beyond contextual data, for each funding initiative, we compiled all conversations taking place between the creator and the community of investors. We collected a total of 358,791 comments made by 192,712 valuators. For every comment, we computed the sentiments and emotions elicited by the valuator using the AFINN dictionary (Nielsen, 2011). The AFINN lexicon is a list of English terms manually rated for valence with an integer between -5 (negative) and +5 (positive) by Finn Årup Nielsen between 2009 and 2011.

Different from other dictionaries classifying words as negative or positive, AFINN provides a range between 1 to 5 to measure emotion intensity. We measure the sentiment of the comments provided by backers by summing the net words integer attributed through the AFINN dictionary. Specifically, a total of 35,768 (9.97%) comments are classified as negative comments according to the AFINN dictionary, while 313,516 are classified as positive and 9,507 as neutral comments. We have eliminated those comments made after the campaign is finished, and those made 2 days before it finishes, to reduce endogeneity issues linked to reverse causality (i.e. funding outcome within the last days of the campaign driving the comments).

2.4.2. Variables

Dependent Variables

Backer's decisiveness: Continuous variable. It measures the influence of the backer in the community and the project, how decisive her/his historical negative comments have been over the total number of negative comments the backer has provided, in percentage points, so that the project succeeds or fails[11]. Making a negative comment on a project and the project failing on Kickstarter afterward adds one unit to both the numerator and denominator. On the contrary, making a negative non-decisive comment on a project so that the project is successful when the campaign closes adds only one unit to the denominator, but not to the numerator. No positive comments are considered in this variable. In other words, we use it as a proxy for the status of a valuator, measured as his/her ability to steer the outcome of the project (funded/no funded).

[11] Success in "all-or-nothing" models of crowdfunding (such as in this case in Kickstarter) is reached when a project collects capital equal to or greater than the campaign's target financing amount.

Number of newcomers to the category after N days: Number of newcomers (new entrepreneurs) that will come for the first time to the category during the next N days. N is a time window of 30, 60, 90, 180, or 360 days.

Independent Variables

Backer's Social Sanctioning in Project: Dynamic sum of the total number of negative comments (using AFINN dictionary) the backer has provided to the project until today.

Backer's Social Sanctioning Cumulative: Dynamic sum of the total number of negative comments the backer has historically provided to all projects in which commented until today.

Previous Backer's Social Sanctioning Cumulative: Dynamic sum of the total number of negative comments the backer has historically provided to all the previous projects in which commented until the one it is being studied to fund.

Social Neg Sanctioning in Category 30_days before: Sum of the total number of negative comments that backers from the same category have provided to all projects in the category for the last (-T) days.

Control Variables

Controls for backers' characteristics and actions:

Previous backer's decisiveness: Same variable as the previous one measuring backer´s negative status, but measured before the start of the project campaign in which the backer provides a comment.

Previous newcomers to category 30_days before: Number of newcomers (new entrepreneurs) that have come for the first time to the category in the previous (-T) days.

Backer´s Total Interaction count: Dynamic sum of the total number of both positive and negative comments the backer has historically provided to all projects in which commented until today.

Backers´ Total Interaction count in Category 30_daysbefore: Sum of the total number of both positive and negative comments that backers from the same category have provided to all projects in the category for the last -T days.

Dummy Superbacker: A backer and his/her comment is tagged as "superbacker" when the backer has historically funded projects more than 10 times. This gives the backer's comment more visibility and power to influence the result of the project's campaign.

Time since last neg comment: The time (measured in number of days) since the last negative comment was posted by the backer.

Number of changes of category: Number of times the backer has commented and funded a project with a different category than the previous one he/she funded. It stands for how much a backer diversifies by funding and commenting on different categories as compared to backers that specialize and only back and comment on projects within the same category classification.

Controls for entrepreneur and project's characteristics:

Dummy creator experience: Dummy receives the value of 1 when the entrepreneur of the project that the backer funded and commented on has experience in Kickstarter and has launched at least one other project in Kickstarter before this one. Therefore, it receives a value of 0 when the entrepreneur does not have experience in Kickstarter; when this project is the first launched project in the platform.

Number of created projects: Dummy that receives the value of 1 when the entrepreneur of the project that the backer funded and commented on has experience in Kickstarter and has launched at least one other project on Kickstarter before this one.

Dummy quick update: Dummy that receives the value of 1 when the entrepreneur posts a project update within the first 24 hours once the project has been launched. According to the literature, this is a very good quality signal of a project and entrepreneur.

Category of the project controls: The category or industry in which the project is classified when the project is launched on Kickstarter. Right now there exists a total of 15 categories to classify projects. Each category has different success/failure rates.

2.4.3. Empirical methodology

We contrast two models for the purposes of testing the hypotheses. The first one is aimed at determining how the behavior of valuators, particularly the negative social sanctioning, impacts their negative status (decisiveness) in the community. The second one is aimed at studying the factors affecting the attractiveness of the community, particularly the propensity of entrepreneurs (both newcomers and recurring entrepreneurs separately) to look for funding through Kickstarter. The following models are developed next.

Model 1. Evolution of a valuator's negative decisiveness

This model adopts a valuator's perspective. Loosely borrowing from game theory (Shapley value), we conceptualize the status of a valuator as his/her ability to steer the outcome of the project (funded/no funded).

We define the stance of a valuator as the negative valuation the valuator conveys regarding a given project. Given a valuator's social valuation, we consider it decisive whenever a valuator adopts a negative stance, and the project fails to attract sufficient funding. In this study, we will focus on negative comments as we are analyzing the stigmatization phenomenon.

We further assume that if the social valuation of a valuator is decisive "most of the time" then his/her status in the community is high. Conversely, if his/her social valuation is not decisive, then his/her status in the community is low.

Given these definitions, we observe the evolution of a valuator's social valuations, projects' funding outcomes, and a valuator's decisiveness across time.

We assimilate the problem as a counting process (number of decisive negative outcomes across time per valuator). We make use of a Fixed Effects (within) regression to compute the evolution of the negative decisiveness of valuators (measured in terms of negative comments' decisiveness) across time. The empirical model we use to examine the effects of social sanctioning on a valuator´s negative decisiveness is as follows:

$$\textbf{\textit{Valuator's negative decisiveness}}_{it} = \alpha_0 + \alpha_1\, \textit{Social Sanctioning}_{it-1} + \alpha_2\, \textit{Social Sanctioning\^{}2}_{it-1} + \alpha_3\, \textit{Previous Valuator's negative decisiveness}_{it-1} + \alpha_4\, \textit{Backer Controls}_{it-1} + \alpha_5\, \textit{Project Entrepreneurs Controls}_{it-1} + \textit{Backer FE}_i + \epsilon_{it} \qquad (1)$$

Subscript i indexed valuator or backer and t indexed time period (*day x project*), respectively. The variables used to conduct our analysis were explained in the previous sub-section.

Using the previous specification (1), Hypothesis 1 would be supported if the coefficient α_2 were positive. Note that the satisfaction of Hypothesis 1 also confirms our baseline hypothesis given the way we have built our variable of decisiveness.

Concerning Hypothesis 2, the support will be connected to the negative values of coefficient α_3, that is the cumulative (quadratic) negative comments that a valuator makes have a negative impact on a valuator's decisiveness.

Model 2. Social sanctioning and community attractiveness.

This model assumes that every entrepreneur launching a new project on Kickstarter is subject to social valuation by valuators affiliated with the community. Social valuation is captured by computing the positive and negative emotions elicited in each comment made by valuators in relation to the projects being considered for funding. Being socially and negatively sanctioned (stigmatized) by the community can have severe consequences on the funding campaign outcome when being new into a category or into the Kickstarter platform as a whole.

We assimilate creators' lifecycle on Kickstarter as a counting process (Andersen et al., 1993). We adopt a linear regression model (OLS) to test the effect of social sanctioning and stigmatization on the attractiveness of the community/category of Kickstarter perceived by potential newcomers or new entrepreneurs in the mentioned category during the next T days.[12] The empirical model we use to examine the effects of social sanctioning on community attractiveness is as follows:

Number of newcomers to category after T days $= \beta_0 + \beta_1$ *Previous newcomers to category N days before* $+ \beta_2$ *Social Neg Sanctioning in Category N days before*$_{it-1}$ $+ \beta_3$ *Backers´ Total Interaction count in Category N days before*$_{it-1}$ + Category FE + ϵ_{it} (2)

Subscript i indexed community or category and t indexed time period, respectively. A category-specific component of the error term (η_i) was included to eliminate the unobservable category or community heterogeneity that could be correlated with the independent variables (e.g., number of negative comments).

Using the previous specification (2), Hypothesis 3 (attractiveness of the category) would be supported if the coefficient β_2 were negative.

In the OLS regression model the *output variable*: the number of newcomers within the next T=30, 60, 90, 180, or 360 days. The *Treatment variable*: cumulative sum of valuators' negative social valuation of the project within a previous N=30 days window.[13] The *control variables*: cumulative sum of comments made by valuators of projects within the same category during the previous 30 days window, community category, and previous valuator's negative status (please refer to previous sub-section for the definition of these variables).

12 As a robustness check, we also adopt a logistic regression model to test the attractiveness of Kickstarter as a whole for new entrepreneurs to repeat and return to Kickstarter with a new project for obtaining funds.

13 As a robustness check we use other N windows thresholds of 60 and 90 days before.

2.5. Results

2.5.1. Descriptive results

Considering all the projects launched on Kickstarter that have received comments during the live of the campaign[14], there have been launched 40,841 projects in Kickstarter from January 2013 to 2017. Among these projects, 30,813 were launched by newcomers (95.74%), while 1,371 were produced by repeating entrepreneurs with experience in Kickstarter. This result is consistent with different studies that have shown how Kickstarter is a good starting point for gathering funds for the first time, while then moving to other financing methods such as VCs, business angels, or traditional bank lending, among others (Roma, Messen Petruzzelli, & Perrone, 2017). An interesting phenomenon we can observe is that the tendency to provide negative comments has decreased during the last years (from 2013 to 2017), which can partially be explained by the detrimental effects on the community of stigmatization, that we are analyzing (see Figure 2.1). This evolution of a lower negativeness proportion of comments can be observed in most of the categories, though the proportion is much higher for the Games community (13% negativeness on comments) or industry as compared to the Dance community (4% negativeness on comments) (see Figure 2.2). Therefore, controlling projects by categories will be important to find the effect of persistent stigmatization on Kickstarter and community attractiveness in the short, medium, and long term.

Another interesting phenomenon can be seen in Figure 2.3, which shows the evolution of the ratio of negative comments over positive comments accumulated by

[14] As mentioned previously we have excluded comments that are provided after the project has finished, and also, for being conservatism, those provided during the last 2 days before the project campaign closes.

new entrepreneurs (or "newcomers") and repeating entrepreneurs (or experienced entrepreneurs). In this figure, we can observe that the tendency to receive a higher proportion of negative comments has decreased during the almost 5 years period analyzed for both types of entrepreneurs. Also, there is a much higher proportion in the negativeness of comments given to experienced entrepreneurs rather than to newcomers. However, this observation does not need to mean that financing is more easy to get among newcomer entrepreneurs given the lower negativeness of the comments received.[15]

Basic statistics are computed in Table 2.1. The level of decisiveness of backers is not particularly high (7.4%). Also, there is some stickiness in the negative status given the high correlation (47.2%) between the variables: *(i) Backer's negative status (decisiveness); (ii) Previous backer's negative status*. Finally, there is slight evidence of the downside effects on the backer's status of making persistent negative comments given the negative correlation (-0.9%) between (i) *Backer's negative status (decisiveness) and* (ii) *Social Neg Sanctioning Cumulative*.

2.5.2. Results of model 1: Evolution of valuator's negative decisiveness and boomerang effect.

Table 2.2 summarizes the results of fitting a Fixed-effects (within) regression on the evolution of the valuators' negative comments' decisiveness in projects to fail. We observe that negative social sanctioning by the valuator in the project and in previous projects has a statistically significant effect on increasing his decisiveness (α_1 = 0.992 and α_3 = 0.201, $p<0.01$ on the full model, column 2), therein providing support for

[15] This analysis has been performed in Robustness Check R2.1 and R2.2.

Hypothesis 1 (social sanctioning improves valuator's status (decisiveness) in the community).

We also find that cumulative or historical negative social sanctioning (quadratic transformation) in the project and previous projects reduce the valuator's negative status or ability to be decisive in the negative outcome of a project (coefficient $\alpha_2 = -0.057$ and $\alpha_4 = -0.003$, $p<0.01$, see full model, column 2 in Table 2.2: Panel A). In this regard, we find support for Hypothesis 2, claiming that cumulative and persistent negative social sanctioning reduces the valuator's (negative) status (decisiveness) in the community. Therefore, there exists a significant loss in evaluators' decisiveness (status) after an accumulation of social sanctioning, confirming Hypothesis 2.

Furthermore, the same results are obtained when no controlling for previous backers´ decisiveness (columns 3 and 4) and when no clustering of the standard errors by backer (column 1). The same results are obtained when doing a First Differences analysis (Table 2.2, Panel B), and when testing for each independent variable added separately as well as using random effects regression as well (Table 2.2, Panel C).

Furthermore, in this regression, the result that a higher number of negative social sanctioning increases the valuator's negative status can be directly interpreted in terms of increases in the failure probability of the project receiving negative comments. This result provides support for the baseline proposition.[16]

[16] An extra robustness check has also been performed using a logistic regression on the probability of the project success or failure to support the baseline proposition (see the robustness section).

2.5.3. Results of model 2: Social sanctioning and community attractiveness.

Table 2.3 outlines the results of fitting a linear regression model (OLS) on the number of newcomers launching a new project on Kickstarter for the first time in different time horizons (from 30 to 360 days) as a function of the number of valuators' historical total and negative social sanctioning/valuation, past successes and community that belongs to.

The more interaction there exists in the community within the last 30 days[17], the more attractive the community for newcomers in the coming months is (values for *Backer´s Total Interaction count in Category 30_days before* range from β_3=0.019 to β_3=0.421, p<0.01). On average, for every 2 to 3 comments a category tracks in the last 30 days, one additional newcomer will come to the category in the next 360 days.

However, a higher number of negative comments a community provides in the last 30 days, they have a negative significant impact on the number of newcomers (attractiveness) that will join the community financing platform for the coming months (β_2 ranges from -0.104 to -2.356, p<0.01). This means that among that comments, for every negative comment tracked in the category in the last 30 days, at least 2 newcomers won´t come to the category in the next 360 days (see last column). In other words, taking an average project among all the categories, the average negative comments a category accumulates in a month have a negative impact of 45.5% fewer newcomers that will come to the category within the next year.[18]

[17] Other thresholds such as 15, 45, or 60 days before have been applied obtaining similar results (available upon request).
[18] This is the result of (151.34 x 0.236 / 78.424 = 45.5%). See Table 1 for the mean values of num_newcomers_30_days, and neg_comments_30_daysbefore.

Furthermore, the same regression has been tested not controlling for the previous number of newcomers received in the category in the previous 30 days and the results hold (see Table 2.3, Panel B).

These results provide support for the impact of the accumulation of valuators' (negative) social sanctioning (vilification) reducing the attractiveness of the community for potential new entrepreneurs (newcomers) not only in the short term (30 or 60 days) but also in the long term (180 and 360 days), which confirms Hypothesis 3.[19]

Results change when clustering the standard errors by category. Only these results have an effect on the long-term attractiveness of the category, not on the short-term (see Table 2.3, Panel C).

2.6. Robustness Checks

2.6.1. Determinants of success

As a robustness check to test the baseline hypothesis, we conduct a logistic regression on the probability of the project campaign's success or failure, similar to Mollick (2014) (see Table 2.A1 in appendix). We have found that negative social sanctioning reduces the probability of a new entrepreneur being successful (see Model 4). This result conforms to the baseline hypothesis. Besides, we have found that positive comments increase the probability of newcomers being successful. Finally, we have found that experienced entrepreneurs do not suffer stigmatization from negative social sanctioning (see Models 1-2).

[19] An extra robustness check has also been performed using a logistic regression on the probability of a newcomer returning to Kickstarter with a new project to get financing based on community interaction with the project. A negative coefficient is found once we controlling for past success in the first project. This provides additional support for Hypothesis 3 (see Robustness check section).

2.6.2. Newcomers in previous or post T days.

We replicated our main analyses on Hypothesis 3, reported in Table 2.3, for different T windows/thresholds. In particular, different windows such as 15, 30, 60, 180, and 360 days have been checked and the results remain robust, confirming our Hypothesis 3.

2.6.3. Returning to Kickstarter.

As an extra test to complete Hypothesis 3, we model creators' likelihood to keep launching new initiatives in the future, controlling for factors such as experience, category, past successes, social sanctioning exerted by valuators, and the status of the valuator (see Table 2.A2 in appendix). With this test, we aim to understand not only the attractiveness of the category for newcomers (Hypothesis 3) but also the attractiveness of Kickstarter for entrepreneurs that have already used the crowdfunding platform. Hence, their probability to return and use Kickstarter again as a financing mechanism.

In this test, we use a logistic regression model. The *output variable* is a dummy variable that measures entrepreneurs returning to Kickstarter or not for funding a new project.

We have found that negative social sanctioning reduces the attractiveness of the platform to experienced entrepreneurs (β_1 = -0.032, p<0.01 in Model 2). Moreover, when we separate between entrepreneur whose last projects has been a failure (Success=0) from those whose last project has been a success, we have found a larger negative effect of social sanctioning in the unsuccessful projects (β_1 = -0.187, p<0.01 in

Model 3) in comparison to the successful (β_1 = -0.028, p<0.01 in Model 4). Therefore, the more negative social sanctioning entrepreneurs have received, particularly after a failed previous experience, the less attractive Kickstarter is as an early-financing platform.

In this line, we have analyzed the effect of negative sanctioning among those returning entrepreneurs. We have found that the variable of negative sanctioning does not have an impact on the status of the backers once we focus on repeating entrepreneurs. Hence, we can conclude that although social sanctioning detracts new entrepreneurs from returning to Kickstarter, if they finally decide to return, the eventual new negative comments of backers will not have an impact on the probability of these returning projects failing.

Then, results show that repeating or experienced entrepreneurs are not affected by the number of negative comments or stigma posted in their project´s funding campaign. Even if they tend to receive on average a higher proportion of negative comments, these negative comments do not have any impact on the funding outcome.

Therefore, it seems that stigmatization only affects newcomers that are not known in the community, but once you are in the community, only the characteristics of your project, rather than being stigmatized, are going to determine whether your project ends up being successful or not.

2.6.4. Stigmatizers versus valuators

As an additional analysis, we restrict the sample to those backers that tend to stigmatize in comparison to those that only provide a negative comment when they do not like the project. To isolate those backers that have a stigmatization purpose, we use different thresholds based on the number of historical negative comments each backer provides (using the variable "Social Neg Sanctioning Cumulative". Therefore, first column uses the last quartile of the sample of backers that provide the highest number of negative comments. In the next column (Model 3), the last decile of the sample is used. Finally, in columns (4) and (5), the sample of backers is separated between those that have provided more than 5 and 9 (respectively) negative comments. The sample becoming more and more restrictive, identifies better stigmatizers instead of valuators. Results confirm that negative sanctioning increases the negative status, which supports hypothesis 1. Moreover, such positive coefficients become lower as we move to the last columns of the most intense stigmatizers (rather than valuators), which conforms to Hypothesis 2.

2.6.5. High frequency or persistent stigmatizers

We separate among backers that have historically provided a lot of negative comments from others has not provided so many negative comments, controlling for the existing or previous status of the backer. Results show that having the same historical status or influence power, those that have historically provided a higher number of negative comments on Kickstarter, show significant decreases in their future status. Moreover, such decreases become larger (more negative) as valuators have historically provided more negative comments (they will have a lower influence). Such robustness analysis conforms to Hypothesis 2.

2.6.6. Other measures of the social impact of stigmatizers

As pointed out before, social impact can be measured in terms of visibility and influence (see Table 2.A3 in the appendix). We are measuring social impact through the visibility of the backer so that her or his comment can have a higher impact on the final outcome of the funding campaign. Visibility can be proxied through the dummy variable "Superbacker" which labels those backers funding a high number of projects.[20] Results (see Table 2.A3 in appendix) show that the interaction of negative sanctioning and superbacker is negative, meaning that backers with visibility end up losing influence if they are involved in negative sanctioning. Therefore, hypothesis 2 is supported, since once a backer has enough power of social impact or recognition (through higher visibility of their comments being tagged as "superbacker") more negative comments detriment its impact power.

2.7. Discussion and conclusions

This study aimed to provide a better understanding of the reasons that lead valuators to negatively sanction newcomers socially. Although past research has argued and shown that social sanctioning has the purpose of defending the norms of the community, we challenged this assumption by focusing on valuators' private interests. We suggested and found empirical support, for the argument that valuators engage in stigmatizing behaviors towards newcomers to gain visibility and influence—in short, social status—in their communities. Remarkably, the results presented in this study showed that this self-interested behavior of valuators comes at a cost for them and for

[20] The requisite for being tagged as "Superbacker" is providing funds to at least 10 projects more than $10 pledge in each. This requisite is very low and hinders the quality of this proxy as backers' visibility.

their communities when they follow persistent and systematic vilification of newcomers' entrepreneurs.

With these findings, and the theoretical arguments that support them, our research makes three key contributions to the literature. We first contribute to a better understanding of the stigmatization processes by adding to the widely-accepted community-oriented motivations for social sanctioning (articulated in the preservation of social norms) a new motivation: the private interest of the valuators, operationalized in terms of their status within the community. Second, we extend the existing literature on the recursive and fragile nature of the legitimacy-illegitimacy relationship (Gehman & Soublière, 2017; J. F. Soublière & Gehman, 2020; Wry et al., 2011); Hampel & Tracey, 2017; Siltaoja et al. 2020; Phung et al. 2020). In particular, we offer not only a more nuanced understanding of the role of valuators in the stigmatization process but also of the effects of cumulative social sanctioning on the status of the valuator and the attractiveness of the community. Third, our work sheds new light on the dynamics of crowdfunding and its relation to the formation of online communities. We elaborate on these three contributions next.

2.7.1. Social and private interest in the stigmatization process

Prior literature on stigma has underlined the importance of power and status in the vilification process, shaming, and ultimately generating the reproduction and change of institutional arrangements in a community (Erikson, 1962; Doern and Goss, 2014; Creed, 2014). This paper stresses the importance of private interest as an incentive mechanism for gaining power and status beyond the more communitarian interest of norm preservation.

2.7.2. The fragility of the stigmatization process

Illegitimacy and stigma have often been presented as strong and enduring characteristics (Crocker, Major, & Steele, 1998; Devers et al., 2009). However, increasing research looks at how organizations and communities can challenge stigma and improve the legitimacy of the organization (Hampel & Tracey, 2017; Siltaoja et al. 2020) and even of the category (Phung et al. (2020). We contribute to the literature by showing the recursive nature of legitimacy and illegitimacy (Gehman & Soublière, 2017; Soublière & Gehman, 2020; Wry et al., 2011). By attending to the cumulative effects of vilification, we theorize about the importance of the recursive and cumulative stigma-conferring behavior and analyze the levels of acceptability of stigmatization by the community. We argue that communities have their own regulation mechanism in which members might act against the influence of high-status actors when seeing their community in danger. Whereas prior work has viewed norm enforcement and vilification as positive to the valuator and the community, we theorize and show that this effect cannot be sustained when there is cumulative vilification.

2.7.3. Cumulative stigmatization that lifts or rocks all boats

Entrepreneurial behaviors have been founded to have mutualistic outcomes. For instance, once funded, entrepreneurial actors on Kickstarter frequently act as backers of new entrepreneur initiatives. In doing so, they sometimes draw attention to other actors who have recently launched their own campaigns or may be struggling to meet their funding goals (Soublier & Gehman, 2019). Although such behaviors may not directly benefit focal actors, they nevertheless contribute to strengthening the entrepreneurial community in which they all coevolve (Soublier & Gehman, 2019). A legitimacy threshold (Soublier & Gehman, 2019: Zimmerman & Zeit, 2002; Fisher et al. 2016) has been defined as the order of magnitude by which endeavors succeed or fail at crossing such threshold of support. We argue that in the vilification and subsequent stigmatization, there is also a threshold by which valuators may fail to have social support from their own community.

Our findings make a case for revisiting the temporal sequencing between social control vs free-floating identification of powerful members of the community that might affect these mutualistic behaviors and, in the long term, the survival of the community. We argue that controlling behaviors might have not only consequences for the object of stigmatization—i.e., the entrepreneurs that are expelled from the community but also for the community in general. Social sanctioning may lift or rock all boats (Chen & Miller, 2014), depending on how this control is perceived by not only the incumbent actors in the category, but also by the potential newcomers.

Since success and failures are interconnected over time (Soublier & Gehman, 2019), the perception of too extreme controls of the category hinder the willingness of new entrepreneurs to enter the category. What drives entrepreneurs to enter a new category is then, not only the extent to which they will benefit from the reputation of the others (Reschke et al. 2017), but the perception of how flexible and open the group is to innovation and experimentation with new norms.

2.7.4. Contributions to the understanding of community formation, maintenance, and dissolution in crowdfunding platforms

This paper contributes to our understanding of crowdfunding (Short et al., 2017), as well. Crowdfunding is seen as an important venue for entrepreneurial action, but it is also a platform for failure and stigmatization. Further, the idea of a community that can be transformed into a social entity is also inherent in crowdfunding platforms (Clough et al., 2019, p. 241). We show that social norms but also the enforcement of the social norms matter by powerful actors patrolling the community to control membership and behavior. We respond to recent calls that look at valuation and how processes of legitimization and illegitimization can become even more polarized in digital platforms than ever before (Zavyalova, Pfarrer, & Reger, 2017). Moreover, negative social evaluations, such as stigma, illegitimacy, and social disapproval, are gaining particular importance (Bundy & Pfarrer, 2015; Zavyalova et al., 2017). We finally help to conceptualize processes of negative social evaluation empirically and conceptually in online platforms and how these platforms might have their own particular dynamics that may perfectly end up destroying the survival of the community.

2.7.5. Limitations and future research

The study also suggests several future research opportunities. For example, we have focused on studying the private interest of valuators, or backers of a project that provide a comment to the project. However, competitors or entrepreneurs with projects in the same category can have their own private interest in preserving the category from providing funds to competitors. Particularly, potential competitors of a crowdfunding project can collude with valuators to eliminate a project that may be a fierce competitor in the category. In a setting such as Kickstarter in which fund providers are very small non-professional investors, and entrepreneurs are generally unknown without a historic project track, information asymmetries are so big that fierce competition can be a problem for newcomers. Competitors might be interested in, for instance, avoiding other similar projects from collecting funds when their project's campaign is still active or will start in the coming weeks. Also, avoiding competitors to enter the category from their inception can be both efficient and effective. Thus, studying the effects of strategic social sanctioning from competitors on other entrepreneurs' success is a future research avenue that can provide more insights into this topic.

Furthermore, topics such as the contagious effect of negative comments and sanctioning herding behaviors have been assumed in this paper. A broader analysis could test how valuators in a funding campaign follow the crowd and post negative comments based on previous social sanctioning they observe, as well as analyzing the role of a leader that vilifies the project so that other valuators take action. Future research can provide more insights into this topic.

2.7.6. Conclusion

Valuators' opinions about a particular entrepreneurial endeavor are critical for entrepreneurs to be successful in raising the required funds. Our study has extended the stream of literature that explores the social sanctioning of new entrepreneurs and the subsequent stigmatization, by focusing on valuators' private interests, rather than valuators' general interest in defending the norms of the community. Following private interest, particularly continuously looking to gain visibility and influence to have a high social status in the community, can be detrimental to the community they belong to, and hence, to them as well.

Therefore, whereas prior work has viewed norm enforcement and vilification as positive to the valuator and the community, we theorize and show that this effect cannot be sustained when it is persistent. We hope our research will inspire future studies on the social sanctioning and stigmatization effects on both valuators, communities, and entrepreneurs, as well as a better strategic understanding of early financing and crowdfunding platforms.

References

Aldrich, H. E., & Fiol, C. M. (1994). Fools rush in? The institutional context of industry creation. *Academy of Management Review, 19*(4), 645–670.

Bitektine, A., & Haack, P. (2015). The “macro” and the “micro” of legitimacy: Toward a multilevel theory of the legitimacy process. *Academy of Management Review, 40*(1), 49-75.

Bundy, J., & Pfarrer, M. D. (2015). A burden of responsibility: The role of social approval at the onset of a crisis. *Academy of Management Review, 40*(3), 345–369.

Butticè, V., Colombo, M. G., & Wright, M. (2017). Serial Crowdfunding, Social Capital, and Project Success. *Entrepreneurship: Theory and Practice*, *41*(2), 183–207.

Cimino, A. (2011). The evolution of hazing: Motivational mechanisms and the abuse of newcomers. *Journal of Cognition and Culture, 11*(3-4), 241-267.

Clough, D. R., Fang, T. P., Vissa, B., & Wu, A. (2019). Turning Lead into Gold: How Do Entrepreneurs Mobilize Resources to Exploit Opportunities? *Academy of Management Annals, 13*(1), 240-271.

Creed, W. D. E., Hudson, B. A., Okhuysen, G. A., & Smith-Crowe, K. (2014). Swimming in a Sea of Shame: Incorporating Emotion into Explanations of Institutional Reproduction and Change. *Academy of Management Review, 39*(3), 275-301.

Cornelius, P. B., & Gokpinar, B. (2020). The role of customer investor involvement in crowdfunding success. *Management Science*, *66*(1), 452–472.

Courtney, C., Dutta, S., & Li, Y. (2017). Resolving Information Asymmetry: Signaling, Endorsement, and Crowdfunding Success. *Entrepreneurship: Theory and Practice*, *41*(2), 265–290.

Crocker, J., Major, B., & Steele, C. (1998). Social stigma. In D. U. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The Handbook of Social Psychology.* (pp. 504–553). New York: Random House.

Crowston, K., Allen, E. E., & Heckman, R. (2012). Using natural language processing technology for qualitative data analysis. *International Journal of Social Research Methodology, 15*(6), 523-543.

Davis, B. C., Hmieleski, K. M., Webb, J. W., & Coombs, J. E. (2017). Funders’ positive affective reactions to entrepreneurs’ crowdfunding pitches: The influence of perceived product creativity and entrepreneurial passion. *Journal of Business Venturing*, *32*(1), 90–106.

Devers, C. E., Dewett, T., Mishina, Y., & Belsito, C. A. (2009). A general theory of organizational stigma. *Organization Science, 20*(1), 154-171.

Doern, R., & Goss, D. (2014). The role of negative emotions in the social processes of entrepreneurship: Power rituals and shame-related appeasement behaviors. *Entrepreneurship Theory and Practice, 38*, 863-890.

Erikson, K. T. (1962). Notes on the sociology of deviance. *Social Problems, 9*(4), 307-314.

Faraj, S., Krogh, G. v., Monteiro, E., & Lakhani, K. R. (2016). Special Section Introduction—Online Community as Space for Knowledge Flows. *Information Systems Research, 27*(4), 668-684.

Fisher, G. (2019). Online Communities and Firm Advantages. *Academy of Management Review, 44*(2), 279-298.

Fisher, G., Kuratko, D. F., Bloodgood, J. M., & Hornsby, J. S. (2017). Legitimate to whom? The challenge of audience diversity and new venture legitimacy. *Journal of Business Venturing, 32*(1), 52-71.

Gehman, J., & Soublière, J. F. (2017). Cultural entrepreneurship: from making culture to cultural making. *Innovation: Organization and Management, 19*, 61-73.

Hudson, B. A., & Okhuysen, G. A. (2009). Not with a ten-foot pole: Core stigma, stigma transfer, and improbable persistence of men's bathhouses. *Organization Science, 20*(1), 134–153.

Jiang, C., Han, R., Xu, Q., & Liu, Y. (2020). The impact of soft information extracted from descriptive text on crowdfunding performance. *Electronic Commerce Research and Applications*, *43*.

Lelo de Larrea, G., Altin, M., & Singh, D. (2019). Determinants of success of restaurant crowdfunding. *International Journal of Hospitality Management*, *78*, 150–158.

Lounsbury, M., & Glynn, M. A. (2001). Cultural Entrepreneurship: Stories, Legitimacy, and the Acquisition of Resources. *Strategic Management Journal, 22*(6/7), 545-564.

Mahmood, A., Luffarelli, J., & Mukesh, M. (2019). What's in a logo? The impact of complex visual cues in equity crowdfunding. *Journal of Business Venturing*, *34*(1), 41–62.

Massa, F. G. (2017). Guardians of the Internet: Building and Sustaining the Anonymous Online Community. *Organization Studies, 38*(7), 959-988.

Mishina, Y., & Devers, C. E. (2012). On being bad: Why stigma is not the same as a bad reputation. *The Oxford handbook of corporate reputation*, 201-220.

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

Navis, C., & Glynn, M. A. (2011). Legitimate Distinctiveness and The Entrepreneurial Identity: Influence on Investor Judgments of New Venture Plausibility. *Academy of Management Review, 36*(3), 479-499.

Nielsen, F. Å. (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. arXiv preprint arXiv:1103.2903.

Pandey, S., & Pandey, S. K. (2019). Applying natural language processing capabilities in computerized textual analysis to measure organizational culture. *Organizational Research Methods, 22*(3), 765-797.

Pontikes, E. G., & Barnett, W. P. (2017). The non-consensus entrepreneur: Organizational responses to vital events. *Administrative Science Quarterly, 62*(1), 140-178.

Rodríguez-Garnica, G., Gutiérrez, M., & Tribó, J.A., (2022). Signaling and Herding in Reward-Based Crowdfunding: Complements or Substitutes? *Working paper*.

Roma, P., Messeni Petruzzelli, A., & Perrone, G. (2017). From the crowd to the market: The role of reward-based crowdfunding performance in attracting professional investors. *Research Policy*, *46*(9), 1606–1628.

Scheff, T. J. (1990). Socialization of Emotion: Pride and Shame as Causal Agents. Albany: State University of New York Press.

Shepherd, D. A., & Haynie, J. M. (2011). Venture failure, stigma, and impression management: A self-verification, self-determination view. *Strategic Entrepreneurship Journal*, *5*(2), 178–197.

Short, J. C., Ketchen, D. J., McKenny, A. F., Allison, T. H., & Ireland, R. D. (2017). Research on Crowdfunding: Reviewing the (Very Recent) past and Celebrating the Present. *Entrepreneurship Theory and Practice, 41*(2), 149-160.

Siltaoja, M. E., Lähdesmäki, M., Granqvist, N., Kurki, S., Puska, P., & Luomala, H. (2020). The dynamics of (de)stigmatisation : boundary construction in the nascent category of organic farming. *Organization Studies, Online first*.

Soublière, J. F., & Gehman, J. (2020). The legitimacy threshold revisited: How prior successes and failures spill over to other endeavors on Kickstarter. *Academy of Management Journal, 63*(2), 472-502.

Sutton, R. I., & Callahan, A. L. (1987). The Stigma of Bankruptcy: Spoiled Organizational Image and Its Management. *Academy of Management Journal, 30*, 405-436.

Wachs, J., & Vedres, B. (2021). Does crowdfunding really foster innovation? Evidence from the board game industry. *Technological Forecasting and Social Change*, *168*.

Wang, N., Li, Q., Liang, H., Ye, T., & Ge, S. (2018). Understanding the importance of interaction between creators and backers in crowdfunding success. *Electronic Commerce Research and Applications*, *27*, 106–117.

Wiesenfeld, B. M., Wurthmann, K. A., & Hambrick, D. C. (2008). The stigmatization and devaluation of elites associated with corporate failures: A process model. *Academy of Management Review, 33*(1), 231–251.

Wry, T., Lounsbury, M., & Glynn, M. A. (2011). Legitimating Nascent Collective Identities: Coordinating Cultural Entrepreneurship. *Organization Science, 22*(2), 449-463.

Xiao, S., & Yue, Q. (2018). Investors' inertia behavior and their repeated decision-making in online reward-based crowdfunding market. *Decision Support Systems*, *111*, 101–112.

Zavyalova, A., Pfarrer, M. D., & Reger, R. K. (2017). Celebrity and Infamy? The Consequences of Media Narratives About Organizational Identity. *Academy of Management Review, 42*(3), 461-480.

Zimmerman, M. A., & Zeitz, G. J. (2002). Beyond Survival: Achieving New Venture Growth by Building Legitimacy. *Academy of Management Review, 27*(414-431).

Tables and Figures

Figure 2.1: Evolution % negativeness of comments in Kickstarter

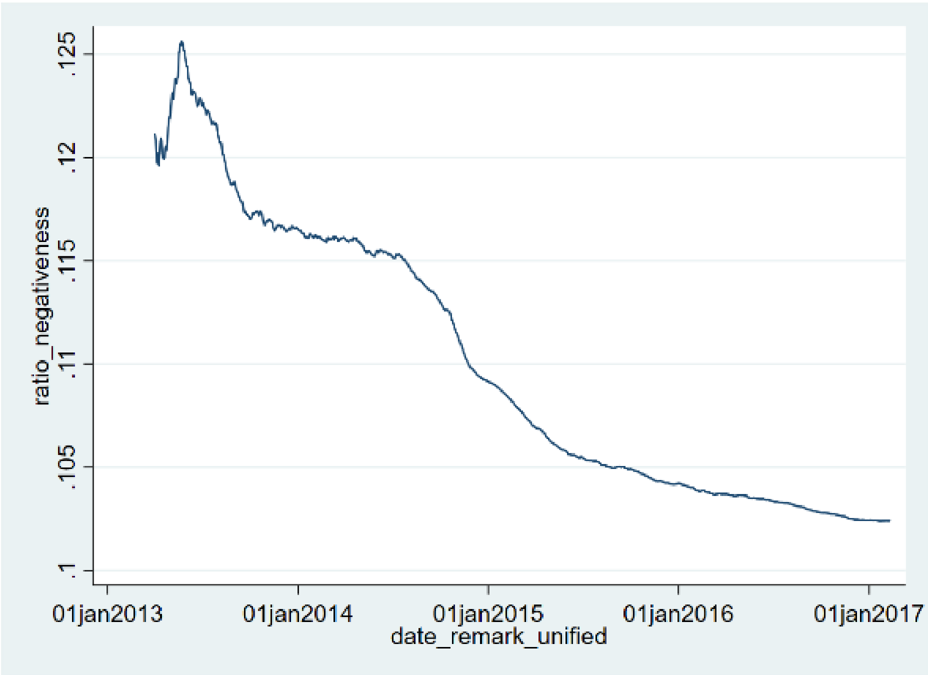


Figure 2.2: Evolution % negativeness of comments by category.

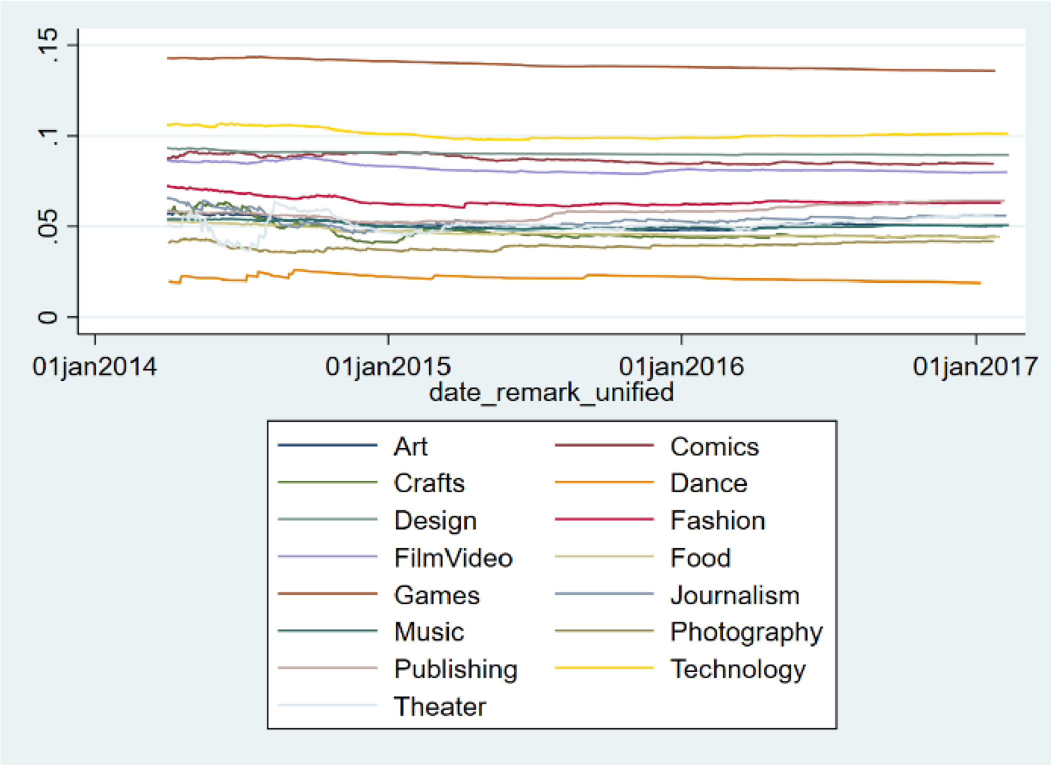


Figure 2.3: Ratio of negativeness of comments received by newcomers and repeating entrepreneurs.

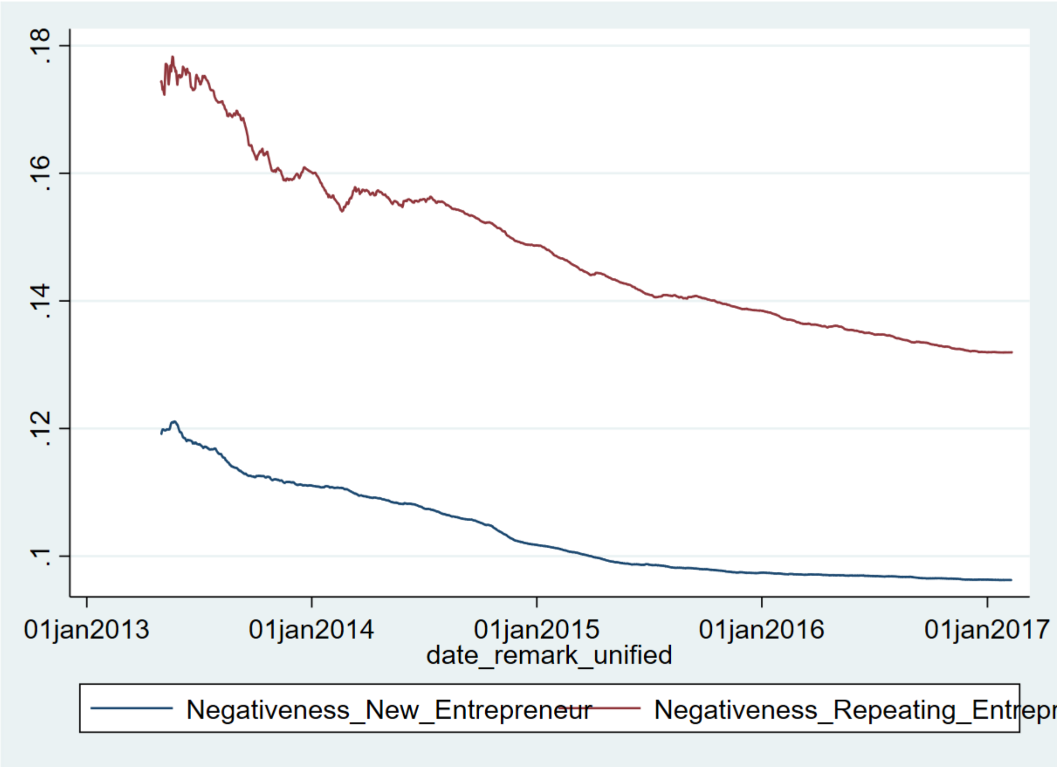


Table 2.1: Descriptive statistics and matrix of correlations

Panel A. Descriptive statistics and matrix of correlations (for H1, H2)

Variable	Obs	Mean	Std. Dev.	Min	Max	(1)	(2)	(3)	(4)	(5)
(1) Backer’s negative status (decisiveness)	44484	7.359	24.574	0	100	1.000				
(2) Previous backer’s negative status	44484	2.491	13.704	0	100	0.472	1.000			
(3) Social Neg Sanctioning in Project	44484	1.079	1.14	0	20	-0.032	-0.101	1.000		
(4) Social Neg Sanctioning Cumulative	44484	2.062	4.445	0	61	-0.009	0.103	0.209	1.000	
(5) Backer´s Total Interaction count	44484	11.769	31.555	0	807	0.009	0.114	0.103	0.837	1.000

Panel B. Descriptive statistics and matrix of correlations (for H3)

Variables	Obs	Mean	Std.Dev.	Min	Max	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) num_newcomers_30_days	142877	78.424	42.733	0	244	1.000							
(2) num_newcomers_60_days	142877	148.296	79.886	0	438	0.959	1.000						
(3) num_newcomers_90_days	142877	215.968	115.799	0	590	0.915	0.978	1.000					
(4) num_newcomers_180_days	142877	409.415	223.946	0	1149	0.857	0.913	0.951	1.000				
(5) num_newcomers_360_days	142877	748.46	441.304	0	2007	0.774	0.824	0.865	0.943	1.000			
(6) prev_num_newcomers_30_days	142868	74.859	40.13	0	240	0.847	0.812	0.805	0.777	0.719	1.000		
(7) neg_comments_30_daysbefore	142874	151.343	257.455	0	1415	0.307	0.313	0.315	0.317	0.302	0.311	1.000	
(8) sum_comments_30_daysbefore	142874	1292.936	1802.598	0	9690	0.359	0.361	0.362	0.365	0.352	0.367	0.993	1.000

Table 2.2: Evolution of backer´s decisiveness across time

Panel A. Fixed Effects regression with clustered errors by backer.				
	(1)	(2)	(3)	(4)
VARIABLES		Backer's decisiveness		
Social Sanctioning in Project	0.922***	0.922***	0.718***	0.718***
	(0.003)	(0.178)	(0.002)	(0.124)
Social Sanctioning in Project^2	-0.057***	-0.057***	-0.043***	-0.043***
	(0.000)	(0.016)	(0.000)	(0.009)
Previous Social Sanctioning Cumulative	0.201***	0.201***	0.563***	0.563***
	(0.002)	(0.047)	(0.001)	(0.111)
Previous Social Sanctioning Cumulative^2	-0.003***	-0.003***	-0.010***	-0.010***
	(0.000)	(0.001)	(0.000)	(0.003)
Backer controls:				
Previous backer's decisiveness	0.679***	0.679***		
	(0.000)	(0.028)		
D_Superbacker	0.368***	0.368	-2.820***	-2.820
	(0.044)	(0.483)	(0.039)	(2.684)
Time_since_last_neg_comment	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Num_changes_Category	0.049***	0.049	0.147***	0.147
	(0.000)	(0.032)	(0.000)	(0.091)
Backer F.E.	Yes	Yes	Yes	Yes
Entrepreneur/Project controls:				
Number of created projects	-0.015***	-0.015	-0.003***	-0.003
	(0.001)	(0.011)	(0.001)	(0.018)
D_Quick update	-0.027***	-0.027***	-0.017***	-0.017**
	(0.001)	(0.008)	(0.000)	(0.007)
Category of the Project Controls	Yes	Yes	Yes	Yes
Constant	1.247***	1.247***	4.388***	4.388***
	(0.021)	(0.274)	(0.018)	(0.558)
Observations	8,315,324	8,315,324	21,715,723	21,715,723
Number of backer_ID	8,307	8,307	21,575	21,575
R-squared	0.987	0.987	0.989	0.989
p	0	0	0	1.57e-08

Notes: The data are at the backer-project-day level from 2013 to 2017. Standard errors in parentheses. Fixed effects are denoted in the table. Standard errors are clustered by Backer ID in columns 2 and 4. *** p<0.01, ** p<0.05, * p<0.1.

Panel B. First Differences analysis				
	(1)	(2)	(3)	(4)
VARIABLES	Dif. Backer's decisiveness			
Dif. Social Sanctioning in Project	0.082***	0.139***	0.306***	0.306***
	(0.001)	(0.001)	(0.002)	(0.002)
Dif. Social Sanctioning in Project^2		-0.016***	-0.018***	-0.018***
		(0.000)	(0.000)	(0.000)
Dif. Previous Social Sanctioning Cumulative			0.229***	0.229***
			(0.002)	(0.002)
Dif. Previous Social Sanctioning Cumulative^2			-0.003***	-0.003***
			(0.000)	(0.000)
Constant	0.000***	0.000***		0.000***
	(0.000)	(0.000)		(0.000)
Observations	21,957,293	21,957,293	21,957,293	21,957,293
R-squared	0.000	0.001	0.002	0.002
p	0	0	0	0

Notes: First Differences regression analysis. The data are at the backer-project-day level from 2013 to 2017. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Panel C. FE and RE regressions incorporating controls without clusters.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES			Backer's decisiveness			
Social Sanctioning in Project		0.942***	0.904***	0.935***	0.922***	0.921***
		(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Social Sanctioning in Project^2		-0.058***	-0.056***	-0.058***	-0.057***	-0.057***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Previous Social Sanctioning	0.126***	0.244***	0.208***	0.214***	0.201***	0.197***
Cumulative	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Previous Social Sanctioning	-0.001***	-0.003***	-0.004***	-0.004***	-0.003***	-0.003***
Cumulative^2	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Backer controls:						
Previous backer's decisiveness	0.681***	0.681***	0.679***	0.679***	0.679***	0.684***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
D.Superbacker			0.229***	0.284***	0.368***	0.540***
			(0.044)	(0.044)	(0.044)	(0.041)
Time_since_last_neg_comment			-0.000***	-0.000***	-0.000***	-0.000***
			(0.000)	(0.000)	(0.000)	(0.000)
Num_changes_Category			0.053***	0.053***	0.049***	0.049***
			(0.000)	(0.000)	(0.000)	(0.000)
Backer F.E. or R.E.	FE	FE	FE	FE	FE	RE
Entrepreneur/Project controls:						
Number of created projects				-0.020***	-0.015***	-0.015***
				(0.001)	(0.001)	(0.001)
D_Quick update				-0.029***	-0.027***	-0.027***
				(0.001)	(0.001)	(0.001)
Category of the project Controls	No	No	No	No	Yes	Yes
Constant	1.576***	1.174***	1.171***	1.166***	1.247***	1.928***
	(0.003)	(0.003)	(0.010)	(0.010)	(0.021)	(0.050)
Observations	8,329,417	8,329,417	8,329,344	8,317,303	8,315,324	8,315,324
Number of backer_ID	8,334	8,334	8,334	8,310	8,307	8,307
R-squared / chi2	0.474	0.479	0.480	0.484	0.486	8.085e+06

Notes: The data are at the backer-project-day level from 2013 to 2017. Standard errors in parentheses. Fixed and random effects are denoted in the table. *** p<0.01, ** p<0.05, * p<0.1

Table 2.3: Attractiveness of the category

Panel A. Category Fixed Effects regression without clusters					
	(1)	(2)	(3)	(4)	(5)
	Number of newcomers to the category after x days				
VARIABLES	30_days	60_days	90_days	180_days	360_days
Previous newcomers to category	0.628***	0.999***	1.372***	2.055***	2.619***
30_days before	(0.002)	(0.005)	(0.007)	(0.014)	(0.029)
Social Neg Sanctioning in Category	-0.104***	-0.122***	-0.142***	-0.699***	-2.356***
30_days before	(0.002)	(0.005)	(0.007)	(0.013)	(0.027)
Backer´s Total Interaction count in	0.019***	0.026***	0.034***	0.135***	0.421***
Category 30_daysbefore	(0.000)	(0.001)	(0.001)	(0.002)	(0.004)
Category F.E.	Yes	Yes	Yes	Yes	Yes
Cluster errors (Categories)	No	No	No	No	No
Constant	22.550***	57.890***	91.084***	187.269***	364.210***
	(0.160)	(0.322)	(0.468)	(0.923)	(1.947)
Observations	142,866	142,866	142,866	142,866	142,866
R-squared	0.755	0.717	0.715	0.704	0.661
P	0	0	0	0	0

Panel B. Without controlling for previous newcomers into category					
	(1)	(2)	(3)	(4)	(5)
	Number of newcomers to the category after x days				
VARIABLES	30_days	60_days	90_days	180_days	360_days
Social Neg Sanctioning in Category	-0.321***	-0.467***	-0.615***	-1.408***	-3.259***
30_days before	(0.003)	(0.005)	(0.007)	(0.013)	(0.026)
Backer´s Total Interaction count in	0.056***	0.085***	0.115***	0.256***	0.575***
Category 30_daysbefore	(0.000)	(0.001)	(0.001)	(0.002)	(0.004)
Category F.E.	Yes	Yes	Yes	Yes	Yes
Cluster errors (Categories)	No	No	No	No	No
Constant	54.587***	108.807***	161.003***	292.016***	497.743***
	(0.130)	(0.244)	(0.350)	(0.658)	(1.326)
Observations	142,872	142,872	142,872	142,872	142,872
R-squared	0.633	0.629	0.636	0.656	0.641
p	0	0	0	0	0

Panel C. Category Fixed Effects regression with clusters					
	(1)	(2)	(3)	(4)	(5)
	Number of newcomers to the category after x days				
VARIABLES	30_days	60_days	90_days	180_days	360_days
Previous newcomers to category	0.628***	0.999***	1.372***	2.055***	2.619***
30_days before	(0.038)	(0.075)	(0.110)	(0.190)	(0.840)
Social Neg Sanctioning in Category	-0.104	-0.122	-0.142	-0.699	-2.356*
30_days before	(0.117)	(0.212)	(0.300)	(0.600)	(1.282)
Backer´s Total Interaction count in	0.019	0.026	0.034	0.135	0.421*
Category 30_daysbefore	(0.017)	(0.032)	(0.046)	(0.093)	(0.204)
Category F.E.	Yes	Yes	Yes	Yes	Yes
Cluster errors (Categories)	Yes	Yes	Yes	Yes	Yes
Constant	22.550***	57.890***	91.084***	187.269***	364.210***
	(3.228)	(6.126)	(9.534)	(22.680)	(28.483)
Observations	142,866	142,866	142,866	142,866	142,866
R-squared	0.755	0.717	0.715	0.704	0.661
p	0	0	0	0	0

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Category Fixed Effects regression. All independent variables are constructed as the accumulated variable to measure in each category during a 30 days window. The data are at the category-day level from 2013 to 2017.

Appendix 2

Table 2.A1: Determinants of Success

	(1)	(2)	(3)	(4)
	Experienced-creator		1st time creator	
VARIABLES	Success	Success	Success	Success
Log_Funding Goal	-0.846***	-0.852***	-0.662***	-0.660***
	(0.132)	(0.133)	(0.016)	(0.016)
Duration of campaign	-0.058***	-0.057***	-0.010***	-0.010***
	(0.018)	(0.018)	(0.002)	(0.002)
Social Neg Sanctioning in Project		0.083		-0.276***
		(0.146)		(0.036)
Social Pos Sanctioning in Project	0.146***	0.140***	0.174***	0.194***
	(0.026)	(0.029)	(0.007)	(0.008)
Constant	3.668***	3.690***	6.642***	6.620***
	(1.047)	(1.047)	(0.155)	(0.155)
Observations	1,668	1,668	14,608	14,608
chi2	71.37	71.67	2856	2910
p	0	0	0	0
r2_p	0.266	0.267	0.164	0.167

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Logistic regression on the probability of the project to success (1) or fail (0).

Table 2.A2: Attractiveness of Kickstarter platform.

	(1)	(2)	(3)	(4)
VARIABLES	Return To Kickstarter *(Dummy: 0/1)*			
			Success=0	Success=1
Social Neg Sanctioning in Project	-0.035***	-0.032***	-0.187***	-0.028***
	(0.008)	(0.007)	(0.060)	(0.007)
Backer´s Total Interaction count in Project	0.007***	0.006***	0.053***	0.005***
	(0.001)	(0.001)	(0.007)	(0.001)
Success *(Dummy: 1/0)*		0.243***		
		(0.052)		
Constant	-2.408***	-2.599***	-2.793***	-2.351***
	(0.021)	(0.046)	(0.056)	(0.023)
Observations	32,183	32,183	7,097	25,086
chi2	85.08	108.2	66.38	65.22
p	0	0	0	0
r2_p	0.00449	0.00572	0.0184	0.00427

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Logistic regression on the probability of a newcomer returning and repeating again in Kickstarter. Treatment variables: The exposure of the entrepreneur to social sanctioning. Control by interactions received and whether the first project was successful or failed.

Table 2.A3: Impact of visibility of backer on its decisiveness of Kickstarter platform

VARIABLES	(1)	(2)	(3)	(4)
		Backers' decisiveness		
neg_status_PREVIOUS		84.355***		
		(0.586)		
D_SuperBacker	3.760***	-0.053	3.604***	4.687***
	(0.233)	(0.204)	(0.234)	(0.313)
neg_comments_backerstot_PROJECT			-0.685***	-0.506***
			(0.080)	(0.087)
Interaction:				
D_SuperBacker#				-1.150***
c.neg_comments_backerstot_PROJECT				(0.222)
Constant	6.200***	4.705***	6.975***	6.772***
	(0.108)	(0.094)	(0.141)	(0.147)
Observations	61,362	61,362	61,344	61,344
R-squared	0.004	0.255	0.005	0.006
p	0	0	0	0

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Chapter 3

Token use and exchange-related transactions’ informativeness in token markets. Are transactions being used strategically?

3.1. Introduction

Crypto-assets have gained significant attention in recent years, with many investors seeking to profit from their potential high returns. Alongside this, researchers have focused on developing theoretical models explaining the valuation of cryptocurrencies. Many of these models use traditional equity market models focusing on traditional factors while others introduce specific crypto-asset-related and blockchain functional factors to explain crypto-asset returns. However, there does not exist yet a comprehensive study that separates the public investors' behavior observed on the on-chain transactions based on the purpose of the transaction for a better understanding of the complexities of the relationship between the dynamics of investors' behavior and crypto-assets returns. This study is the first that separates real token-use-related transactions from speculative and trading transactions, as well as the first that analyzes the effects of investors' public behavior on the sentiment of the market, the network actions, and crypto-asset prices.

Wikipedia or a web browser would not make any sense without someone using and posting information on the website. TV commercials with no audience would neither make any sense. Nevertheless, this seems to be the current situation we observe in token markets. Most tokens are created for particular uses, but the ultimate use of some of them is mainly ignored and tokens are considered as merely speculative vehicles. Interestingly, in an efficient market, if a token is not used in the ecosystem it was created for, the value of the token should drop. And, the other way around, a higher intensity of use of a token in the ecosystem it was created for should increase the demand for the token, and therefore, the token value should also increase.

This paper is, to the best of my knowledge, the first to present a proxy for "token use". I create this proxy by identifying on-chain transactions and classifying them as either (i) centralized and decentralized exchange-related transactions (from now on: "CeDex transactions") or (ii) transactions for other use purposes (token use transactions). Some authors have analyzed token adoption and network value by counting the number of wallets that own the token around the world. Nevertheless, investors can adopt and keep tokens in their wallets for purely speculative purposes with no intention of using them for their ultimate purpose.

I then use my proxy to answer important questions about token markets and crypto-assets not yet empirically addressed precisely because of the difficulty of identifying token-use transactions. In particular, I first test whether token use in its ecosystem is a significant determinant of observed token exchange prices or whether token price dynamics depend only on trading and speculation. Secondly, separating exchange-related transactions (CeDex-related transactions) from token use-related transactions. Furthermore, I identify transactions "to" and "from" exchange wallets (both centralized and decentralized exchanges) to better study market sentiment as well as momentum and feedback trading strategies, broadening our knowledge of investors' behavior in token markets. Specifically, I improve upon previous literature by analyzing market sentiment based on on-chain transactions rather than using NLP techniques on unstructured datasets of texts extracted from social media, news, and webpage searches such as Google trends. Finally, I also address the importance of feedback effects from exchange prices to token use value, and users' and investors' behavior in the token ecosystem.

The results of the paper show that, differently from what is expected in an efficient market, token users do not consider exchange token prices or transactions when they use the token for its ultimate purpose. Neither does the use of the token affect its market price. From the investors' side, results show that token investors do observe other investors exchange related transactions and prices and use this information to determine their trades. Therefore, I find evidence of the existence of feedback trading and momentum trading strategies in token markets. Moreover, investors' on-chain exchange-related transactions, that move their tokens "to" and "from" exchanges, do impact token prices. These transactions are costly to produce signals that are interpreted as direct, transparent, and reliable information that shows the sentiment of investors in the market. While some investors with higher levels of expertise could be using these signals to strategically move prices, other investors do make strategic use of the signals from these ones to make their investment decisions.

Precisely, in the paper I disentangle the two channels through which on-chain to and from exchange transactions have an impact on token prices: market sentiment revealed by these transactions and token supply and demand forces in the exchange. To disentangle both channels, I analyze the number and the volume of CeDex-related transactions. A higher number of transactions, controlling for the quantity transferred, relates to the market sentiment channel since it increases the visibility of the actions and the signal becomes more explicit. However, a higher volume or number of tokens transferred, controlling for the number of transactions, relates to changes to the supply and demand equilibrium of tokens in the exchanges. My results show that surprisingly the first channel (market sentiment) is statistically significant and coexists with the second channel (i.e. the amount or volume of tokens transferred has no impact on prices). Therefore, my interpretation of the results is that token prices are affected by market sentiment through investors' strategical use of on-chain transactions information as well as by token supply and demand forces in exchanges.

Even though some fraction of these transactions may be intended to manipulate and executed to pump or dump the market, on average the market reacts positively to these signals because they are costly (i.e. high required gas and transaction fees). The findings' implications do not relate to the intention of investors with their actions, but the effect of their actions on the token market, and how other investors do strategically use this publicly available on-chain information to trade the token, impacting token prices.

Overall, the use of these two new proxies of token use intensity and market sentiment opens a new avenue of research in the token and crypto-asset markets literature that can help academia answer unsolved questions about crypto-asset valuation, as well as investors' and users' behavior on token markets.

The structure of the chapter is as follows. Section 2 presents the theoretical background and develops the hypotheses. Section 3 describes the data and the methodology used. Section 4 presents the results. Section 5 conducts some robustness tests. Section 6 discusses the results obtained. The chapter finishes with some concluding remarks and some guidelines for future research.

3.2. Theoretical Background

Cryptocurrencies have gained significant attention in recent years, with many investors seeking to profit from their potential high returns. Alongside this, the role of traditional factors and specific blockchain-related factors in driving cryptocurrency returns has also been a topic of interest for researchers. An extensive literature has focused on developing theoretical models explaining the valuation of cryptocurrencies (Weber, 2016; Huberman, Leshno, and Moallemi, 2017; Chiu and Koeppl, 2017; Biais et al., 2018; Cong, He, and Li, 2018; Jermann, 2018; Abadi and Brunnermeier, 2018; Routledge and Zetlin-Jones, 2018; Saleh, 2018; Congand He, 2019; Cong, Li, and Wang, 2019; Sockin and Xiong, 2019; Schilling and Uhlig, 2019; Liu and Tsyvinski, 2021). Some papers argue that the evolution of cryptocurrency prices should follow a martingale, and thus cryptocurrency returns are not predictable (e.g., Schilling and Uhlig, 2019). Other papers argue that, in dynamic cryptocurrency valuation models, cryptocurrency returns could potentially be predicted by traditional factors (e.g., Cong, Li, and Wang 2019; Sockin and Xiong 2019; Shams, 2020; Liu and Tsyvinski, 2021). Many of these models use traditional equity market models focusing on factors such as momentum, book-to-market, investors' attention and sentiment, valuation ratios, and correlation with traditional asset markets such as commodities, stocks, currencies, and macroeconomic factors finding both the same and different results compared to traditional markets for each of the factors (e.g., Athey et al. 2016; Jermann 2018; Schilling and Uhlig 2019; Liu and Tsyvinski, 2021). Other models focus on cryptocurrency-specific factors such as production and network factors. The production factors analyzed refer to the cost of cryptocurrency production, i.e. the miners' problem concerning rewards and energy, production, and computing costs (e.g., Cong, He, and Li, 2018; Sockin and Xiong, 2019). However, some authors find no empirical evidence of returns being affected by cryptocurrency production factors (Liu and Tsyvinski, 2021). The network factors analyzed refer to cryptocurrency adoption and its positive effect on prices (e.g., Pagnotta and Buraschi 2018; Biais et al. 2018; Cong, Li, and Wang 2019). Empirical research has introduced proxies for measuring cryptocurrency adoption such as the number of users, active addresses, number of total transactions, and other proxies (e.g., Benedetti and Nikbakht, 2021; Liu and Tsyvinski, 2021; Lyandres, Palazzo, and Rabetti, 2020). Liu and Tsyvinski (2021) find that there is a significant time-series momentum phenomenon in cryptocurrency markets, while financial ratios are not significant in predicting cryptocurrency returns. Moreover, they find that cryptocurrency prices not only reflect current cryptocurrency adoption but also contain information about expected future network growth.

This paper contributes to a fast-growing literature on cryptocurrency valuation and investors' behavior through using public on-chain data. In previous studies, the network effect has been theoretically studied by Cong, Li, and Wang (2019) and Sockin and Xiong (2020), and then empirically quantified by Shams (2020) through textual analysis of the comments on different social media such as Reddit. However, to the best of my knowledge, this is the first paper that uses wallet-level on-chain data to identify wallets and measure market sentiment, demand, and the network effect and its effect on the returns of a large cross-section of tokens. The first authors that used wallet-level blockchain data and identified and classified transactions studied market manipulation and illegal activities on the Bitcoin blockchain (Foley, Karlsen, and Putnin (2019); Griffin and Shams (2019); and Makarov and Schoar (2021)).

Few empirical studies have studied the impact of market sentiment on cryptocurrency market returns (e.g., Liu and Tsyvinski, 2021). Liu and Tsyvinski (2021) approximate market sentiment and speculative interest of investors using the information on investor attention through Google searches, social media, and news. However, in this study, I propose a more direct form of testing for momentum effects and measuring market sentiment and speculative interest of investors in token markets through using the information of public, on-chain transactions, through the identification of blockchain wallets and transactions level data. This is a better way to measure market sentiment since investors "*put their money where their mouth is*". In other words, on-chain transactions are costly actions since on-chain transactions require high amounts of gas and fees expenses, making them good enough signals to rely on. Interestingly, this belief, together with on-chain transactions being public, transparent, and irreversible, can also be misused by investors to manipulate the market based on their interests provoking price changes. A number of studies have studied price manipulation of Bitcoin and other cryptocurrencies (e.g.: Cheah and Fry (2015); Corbet, Larkin, Lucey, and Yarovaya (2018); Baur, and Hong (2018); Gandal et al. (2018); Hamrick et al. (2018); and Li, Shin, and Wang (2019); Griffin and Shams (2020)). In this paper, I also contribute to the literature on price manipulation in token markets in general, rather than just on specific crypto-assets such as Bitcoin or other cryptocurrencies.

Therefore, different from previous studies that use a global cryptocurrency index, a set of cryptocurrencies, or only ICO tokens (such as Howell, Niessner, and Yermack (2018), Benedetti and Kostovetsky (2018), Li and Mann (2018), Lee, Li, and Shin (2019), Liu et al. (2020), and Makarov and Schoar (2021) among others), this paper uses a very comprehensive database that gathers on-chain transaction data on most Ethereum-based tokens that are listed and traded on a sufficient number of centralized and decentralized exchanges (similar to Shams, 2020). Consequently, in this paper, I study the price dynamics, token use intensity, and investors' behavior in the secondary market (market sentiment) for a large set of cryptocurrencies and tokens, not limited to only ICO or STO tokens or just in the primary market.

We show that cryptocurrency returns are higher when speculative or trading interests increase, and also token prices contain information about future expected speculative growth. However, we do not predict neither higher token returns when token use intensity increases, nor cumulative future cryptocurrency token use growth over different time horizons using current token market returns. These results are important since they show that token use intensity does not create value in token markets, but rather, it seems that prices are moved by speculative investors´ interest. Furthermore, prices positively predict and anticipate cumulative speculative-related transaction growth or speculative investors´ interest, but cannot predict cumulative token use intensity growth.

3.2.1. Network effect and market sentiment through exchange-related transactions.

Among the specific cryptocurrency factors studied by scholar, token adoption and community building is considered key sources of cryptocurrencies' value. Demand or buying pressures for holding cryptocurrencies can have an amplified effect on cryptocurrencies. This is due to the network effect because they may be interpreted as a sign of "adoption" by the community, not only from investors, but also from users and developers demanding the token (Gandal and Halaburda (2016); Li and Mann (2018); Cong, Li, and Wang (2020); Sockin and Xiong (2020); Shams (2020)). Therefore, as pointed out in previous studies, returns can be explained by investor demand (e.g. Barberis and Shleifer (2003) and Barberis, Shleifer, and Wurgler (2005)). Regarding token adoption and the network effects, a number of authors studied these effects, and some used the number of active users on the Ethereum network, finding a positive relationship with the value of Ethereum (Cong et al. (2018); Li and Mann (2018); Sockin and Xiong (2018); Lyandres, Palazzo, and Rabetti (2022); Benedetti and Nikbakht (2021)).

In this paper, I study how investors do observe and use strategically the information coming from other investors' exchange-related transactions and previous prices to make trading decisions. Understanding their behavior is important in order to predict future price variations. There exist two potential channels that affect token prices or returns through token demand and the network effects. Both channels work in the same direction. However, separating them can provide us with more information on how the network effects work in token markets.

One potential channel is through the visibility of exchange-related transactions, while the other channel is through the supply and demand forces of tokens at the exchanges; through the amount and value of tokens transferred on these transactions.

The first channel is more related to the demand forces and network effects, and it also proxies the market sentiment. Overall, the literature suggests that sentiment and network effects play important roles in driving returns for Ethereum-based assets. Alongside this, the role of sentiment in driving cryptocurrency returns has also been a topic of interest for researchers. This literature extends beyond assets such as Bitcoin or Ethereum, and there is a growing interest in studying the relationship between Ethereum-based token returns and sentiment in the market as well.

Several studies have explored the impact of news on cryptocurrency returns, volume, and volatility. However, results show that irrespective of a positive or negative sentiment of the news related to Bitcoin, the enthusiasm alone affects positively its returns (Rognone, Hyde, and Zhang, 2020). Similarly, news media's different types of sentiment have medium and long-term effects on Bitcoin volatility (Sapkota, 2019). Then, the relationship between sentiment and cryptocurrency returns appears to be complex and multifaceted. Further research is needed to fully understand the relationship between market sentiment and cryptocurrency returns.

3.2.2. Utility tokens and the use of tokens

Tokens are digital units of accounts issued on a blockchain-based platform. There exists a variety of types of tokens such as cryptocurrencies, security tokens, tokenized assets and asset-backed tokens, and utility tokens (Benedetti and Rodriguez-Garnica, 2023). In particular, utility tokens are digital units that serve as the only accepted means of payment for consumptive rights to services and products provided through a blockchain-based platform (Benedetti, Caceres, and Abarzúa, 2023). Many of these tokens, and in particular the utility tokens have been created on the Ethereum Blockchain, using standardized contracts such as ERC-20 smart contracts. An example of an existing and successful token is the Basic Attention Token (BAT). This token is created with the purpose of being used in its own token ecosystem. This token serves as a means of payment for advertising and services of website users on the web browser "Brave" to the sites and advertisers they visit rather than being the advertisers being the payers to the websites. This is only an example of the utility and use of tokens in a token ecosystem, however, there are many more.

This paper is the first paper to proxy for token use in the token ecosystem and analyze its effects on prices, as well as how token users behave depending on token prices. In an efficient market, the higher level of utility and usage of a token, the higher the demand for the token, and the value of the token should increase. Prices should reflect the pressure for an increment on the intensity of use of the token.

3.2.3. Hypothesis development

In token markets, we have the possibility to analyze the public on-chain transactions to better understand investors´ behavior in a direct and trustful manner. All transactions occurring on-chain are publicly available and disclosed, transparent, and traceable. Moreover, I am able to separate and proxy token use-related transactions as well as exchange-related transactions. This allows me to develop several testable hypotheses that can shed light on the process of token price dynamics.

In general, prices are moved through supply and demand. A higher demand for a product, ceteris paribus on other factors, should make pressure on prices to rise. The same should occur in token markets, a higher token adoption among the broader public, and the higher the intensity of token use in the ecosystem it was created for should increase the token's utility. Hence, a higher utility of the token should increase its price/value. If a token's price does not increase with token utility, network's value, and token use intensity, there is no reason to hold the token other than for speculative purposes. This should make the token less interesting and eventually will put downward pressure on its price. Therefore, my first testable hypothesis refers to the impact of token use transactions on token prices, and it is stated as follows:

Hypothesis 1*: An increase in token use transactions increases token prices.*

Interestingly, securities prices can also be moved, not because demand or supply for products change, but because of the information conveyed by investors' behavior. For example, if it was known that Warren Buffet had sold his stake in a particular stock, other investors would be likely to imitate him and sell it as well, which in turn would make the price decrease. More generally, a higher interest of investors in a particular stock and a broader acquisition of stock by more investors have an impact on prices. In token markets, information disclosed through transactions should also impact prices through investors observing other investors' behavior and taking action.

Additionally, in token markets, we have the particularity of being able to observe the information of each transaction. Identifying the wallet that makes each transaction with an exchange-related wallet is possible. Therefore, we can observe the number of transactions and tokens moved to exchanges and from exchanges. This is similar to observing the order book of the exchange (in terms of learning about investors' behavior and sentiment of the market). Unless the investor indirectly knows a buyer and agrees on the terms, the only way of liquidating his/her token investment is through the use of an exchange. To trade crypto-assets in any centralized or decentralized exchange (from now on CeDex exchange) the tokens must first be moved from the private on-chain investor's wallet (also known as a cold wallet) to the exchange's off-chain wallet. On the other hand, investors might want to move their acquired tokens on exchanges to their on-chain wallets. Doing this is costly but investors will do it for two reasons. The first reason has to do with security concerns. Token holders, even if they intend to trade in the future, prefer to keep their own tokens instead of relying on second-party custody services. In fact, decentralization supporters' claim is "not your keys, not your coins". The second reason may be strategic, even though Ethereum blockchain transactions are still very costly, investors might be interested in sending intentional and strategic signals to the market to make prices fluctuate. Some authors have found pump-and-dump schemes in cryptocurrency markets through significant wealth transfers between insiders and outsiders (Li, Shin, and Wang, 2023).

Whatever the reason, token market investors can directly observe on-chain token movements to and from market exchanges and this can be used as an indicator of investors' intention on the token investments and therefore, the market sentiment on the token.

”From” exchange transactions, with investors moving their tokens from an exchange to their own private on-chain wallet, stands for positive market sentiment. Investors take long positions (known as “HODL” in the token market) on the token, since they think the asset is undervalued and prices should eventually rise. An additional channel going in the same direction is that after a significant proportion of “from” exchange transactions, fewer tokens will be available in the exchange pool. Tokens would be moved to cold wallets, reducing the supply of tokens offered on exchanges for trading. This behavior then would put downward pressure on token prices in the exchanges.

On the other hand, investors moving their tokens from their private wallets to exchange wallets show a negative market sentiment. In other words, “to” exchange transactions show that investors are willing to sell the token. These investors may have liquidity needs but they may also think the token might be overvalued and wish to sell it for speculative purposes. Again, another channel going in the same direction is that the more tokens brought to the exchange, the more supply there exists in the exchange’s tokens pool. Therefore, a higher supply of tokens available on exchanges for being sold puts downward pressure on prices.

Recall that moving tokens to or from an exchange is a costly on-chain transaction because of blockchain transaction fees (i.e. gas used) that fluctuate and can reach very large costs, sometimes larger than the value of tokens moved. Therefore, according to the signaling theory, the more costly the signal is, the higher the quality of the information disclosed. Hence, signals produced and observed through the use of on-chain transactions can have high enough quality to be considered by other investors and the market, moving market sentiment and generating network effects.

All these expected effects are summarized in my second testable hypothesis and sub-hypotheses 2A and 2B:

Hypothesis 2*: Exchange-related transactions impact token prices.*

The first channel (market sentiment channel) is tested through the following sub-hypothesis 2A. A higher number of transactions, controlling for the quantity transferred, relates to the market sentiment channel since it increases the visibility of the actions and the signal becomes more explicit.

Hypothesis 2A*: An increase in the proportion of the number of "to" exchange-related transactions over the number of "from" exchange-related transactions have a negative impact on token prices. This channel is explained by the visibility of on-chain signals showing a negative market sentiment.*

Additionally, in order to disentangle the two potential channels (market sentiment channel, and supply and demand channel) through which on-chain "to" and "from" exchange transactions have an impact on token prices, I analyze and compare not only the number of exchange-related transactions (as proposed on Hypothesis 2A), but also the number of tokens transferred in these transactions; the volume of tokens transferred. If prices would be affected by token supply and demand, the proportion of amount transferred would have an impact on the supply and demand equilibrium of tokens in the exchanges, and hence on token prices.

Hypothesis 2B: *An increase in the proportion of the volume or number of tokens transferred "to" exchanges over the number of tokens transferred "from" exchanges have a negative impact on token prices. This channel is explained by a higher supply pressure on exchanges.*

Momentum trading and feedback trading strategies are implemented by investors and high-frequency traders (intraday investors). These strategies are based on historical price movements. Based on previous price fluctuations, investors will decide to buy or sell a token, therefore observing price movements is key for these investors. However, token users should not be considered investors but consumers or agents using the token for obtaining a good or service in a platform or ecosystem intended for this purpose. Therefore, while I expect token investors to be affected by historical prices, I do not expect token users to be affected by exchange price movements, since their intention is not speculative or for profit. We might think that if tokens change their price, consumers might deviate their intention. However, unless prices suffer a huge variation, token users are aware of the volatility and fluctuation of token prices beforehand, therefore their behavior should not be affected by these fluctuations. Hence, a price increase or drop should not affect the number of tokens that are used in the ecosystem in the following days.

Hypothesis 3: *Investors in token exchanges will make use of the information contained in previous price movements (i.e. they will engage in feedback strategies), while token users will not.*

3.3. Data and Methods

3.3.1. Data

Studying this research question requires an extensive database on both prices, transactions, and wallet identities in the growing token market. I build a unique and very complete database of tokens based on the Ethereum blockchain by crawling data on token prices from coinmarketcap.com and on transactions from etherscan.io.

In particular, I gather data on daily information on prices and transactions for all existing, successful, and listed Ethereum blockchain-based projects launched since the Ethereum blockchain inception and up to June 2021. The final sample of all Ethereum-based tokens that are listed on CoinMarketCap.com includes daily information on transactions and prices for a total of 513 projects[21].

This section describes the sources used for obtaining the comprehensive database, the strategy employed to identify and separate the different types of transactions, the variables used, and its descriptive statistics.

3.3.2. Transactions data on the Ethereum blockchain

I obtain data on daily token transactions taking place between different wallets on the Ethereum blockchain from etherscan.io. Etherscan is a leading block explorer that provides information on Ethereum Blockchain blocks and on every transaction from Ethereum-based tokens. For each token contract hash address (42-character string), I crawl information for all transactions taking place. This information includes the

[21] This sample has suffered some cleaning process since not all tokens hosted in CoinMarketCap (1115) are hosted in Ethereum blockchain nor its hash address has successfully been obtained. Some tokens do not show its hash address in CoinMarketCap, some others have not being matched with the Ethereum information, and some others have more than one hash address.

transaction address, sender address, receiver address (all of which are 66-character strings), transaction time (e.g., June-28-2021 05:36:08 PM +UTC), and quantity of tokens transferred.

3.3.3. Centralized and decentralized exchanges identification

I manually identify centralized and decentralized exchange (CEx [22]and DEx[23]) wallet hash addresses (42-character string) that send or receive tokens. Therefore, I can separate those transactions in which the receiver address or the sender address is a centralized or decentralized exchange (CeDex: CEx and DEx) from those transactions in which non-exchange wallets (non-CeDex) interact among them. The identification process is similar to the one used by Foley, Karlsen, and Putniņš (2019) in the well-cited paper "Sex, drugs, and Bitcoin: How much illegal activity is financed through cryptocurrencies?".

3.3.4. Secondary market daily prices and other token market data

Secondary market daily prices and volume for all tokens and those ICOs that have been successful and are listed in a secondary market are gathered from coinmarketcap.com. This is the most comprehensive and credible source of trading data for digital assets, with indices featured in the data feeds provided by NASDAQ, Bloomberg Terminal, Thomson Reuters, and others. CoinMarketCap aggregates daily data from those public exchanges that charge trading fees. This is relevant since

---

[22] Examples of Centralized Exchanges we identify are Coinbase, Binance, Kraken, Gemini, GDAX, Huobi Global, Bithumb, Bitfinex, Bitstamp, Bittrex, KuCoin, FTX, Poloniex, bitFlyer, Celsius Network, BlockFi, youhodler, CEX.io, OKEx, Mercatox, and others.

[23] Examples of Decentralized Exchanges we identify are Uniswap (V2), Uphold, 0x Protocol, Venus, Tokenlon, Bisq, AirSwap, Blocknet, Barterdex, Sushiswap, Compound, BurgerSwap, Curve Finance, 1inch Exchange, PancakeSwap, Paraswap, Stellar Decentralized Exchange, and others.

exchanges without fees permit issuers or other stakeholders to generate false volume, with traders (or their bots) trading back and forth with themselves.

3.3.5. Variables and descriptive statistics

Descriptive statistics and correlation matrix are shown in Table 3.1, Panels A and B.

Dependent Variable

Daily returns: Daily return of the token. It is calculated as the daily closing price over the daily opening price of the token. It is important to clarify that, differently from traditional markets, the token market "never sleeps", it never closes so the market "opens" and "closes" at 00:00:00 +UTC. This is the reference time zone used by CoinMarketCap and most of the token applications to construct the variables on token-related daily information.

Independent Variables

- Daily Ethereum returns: Daily return on Ether, is considered as the industry reference on market returns. Previous papers such as Hu, Parlour, and Rajan (2018) show that individual cryptocurrency returns correlate with Bitcoin returns, and more specifically Ethereum blockchain-based tokens correlate to Ether returns, considering it a more direct reference for controlling for the industry (Reference, 202X).

- Previous day returns: Previous daily return of the token analyzed.

- Volume: Total daily token trading volume in all exchanges (off-chain) that CoinMarketCap.com gathers information from (i.e. public exchanges that charge trading fees).

- Volume 7 day: Mean volume of token trading in exchanges (off-chain) during the previous 7 days.

- Return volatility 7 day: Volatility of token returns during the previous 7 days.

- Amihud ratio 7 day: Proxy that measures the illiquidity ratio of the token during the previous 7 days.

- Number of transactions: Total number of daily transactions that occurred on-chain in the Ethereum blockchain

- Number of CeDex transactions: Total number of daily transactions that occurred on-chain in the Ethereum blockchain and in which the receiver or the sender is a Centralized or Decentralized Exchange wallet.

- Number of Non-CeDex transactions: Total number of daily transactions that occurred on-chain in the Ethereum blockchain and in which neither the receiver nor the sender is a Centralized or Decentralized Exchange wallet.

- Negative Market Sentiment: Proxy measuring negative market sentiment that is constructed as the Ratio of CeDex "to" transactions over CeDex total transactions, or ratio of transactions in which tokens are moved from a wallet to a CeDex wallet over total transactions that occur in a day.

- Exchange supply and demand ratio: Ratio that measures the ratio of the volume of supply and demand of tokens on the transactions with exchanges. The ratio is constructed as the proportion of total number of tokens transferred "to" CeDex over the total number of tokens transferred "from" and "to" CeDex. Therefore, this is the ratio of the volume of tokens in which tokens are moved from a wallet to a CeDex wallet over the total number of tokens transferred with exchange wallets in a day.

3.4. Results

3.4.1. Token use, CeDex transactions, and token value.

Determining whether a token is correctly valued or is overvalued or undervalued in the token market is a question that many researchers try to answer. In order to better solve this question is important to understand how value is generated and varies in token markets. According to Hypothesis 1, higher token adoption and use in their ecosystem should generate value. Moreover, according to Hypothesis 2 market information on exchange-related transactions should also be incorporated into prices.

In Table 3.2, I start by testing whether the total number of global on-chain transactions on the public blockchain has an impact on future prices, and hence on token returns[24]. More interestingly, I then separate those transactions between two private wallets from those transactions between a private wallet and a Centralized or Decentralized Exchange (CeDex) wallet. Isolating the transactions of which CeDex is a part allows me to create a proxy for "token use" transactions on the ecosystem it was created for. Hence, this proxy can help to contribute to the interesting question of whether "token use" intensity impacts token prices. In other words, I can test whether the use of tokens in the ecosystem, compared to liquidity or speculative market transactions, increase the value of the token and the business. To the best of my knowledge, this is the first paper that quantitatively measures token use, differently from token adoption measured in the existing literature.

[24] I use total daily transactions for each token as the main independent variable. Theoretically, the same analysis could be conducted using the total daily volume or value (in Ethers) transacted, which apparently would make more sense. However, in practice, these alternatives are worst because the valuation of transactions is not objective and requires choosing an arbitrary price. We cannot use the price at the time of the transaction since we do not have this information. We can only use mid-price, closing price or opening price, or a derivative of them. But, this question is difficult to solve in token markets because of their high volatility and continuous pricing, with opening and closing of the market stablished at UTC time on CoinMarketCap webpage.

I also control for Ethereum returns, previous-day token returns, and a proxy for liquidity in the market (Volume). Ethereum return is broadly used in the literature as a proxy for industry returns. It is highly correlated to all tokens as it is the tool and token used for creating other tokens through "ERC20" smart contracts. The volume of global exchanges trading is the proxy for token liquidity, but I also use other proxies such as the Amihud illiquidity ratio and the average return volatility, all measured as the mean of the previous 7 days. Digital assets markets are continuous markets. Therefore opening and closing prices are determined at +UTC time. On these regressions I do not use the midprice ([Open+Close]/2), but the earliest ("open") and latest ("close") price at UTC time.

A fixed effects regression analysis with clustering of the standard errors is performed in Table 3.2 for testing for Hypothesis 1. We control for token fixed effects. All the control variables are significant and have the expected effect on token returns. However, a higher number or an increase in the number of token-use transactions does not have a significant impact on token returns or value (see column 2) nor does a higher number of tokens used (see column 5). Furthermore, considering the total number of transactions, increasing the number of total on-chain transactions a token has during a day does not have a significant effect on token value (see columns 1 and 4, Table 3.2). Therefore, Hypothesis 1 is rejected, meaning that changes in the number of token-use transactions do not impact token prices.

Hence, finding that token value does not depend on utility measures raises questions about the efficiency of token markets and token value. This is contrary to what is expected in an efficient market, where the higher the use and the utility of a token, would create higher demand for the token as well as prices to rise. A token with higher intensity of its use and utility should have a higher value. The next question that we should ask is whether information about exchange-related transactions impact on token value, as stated in Hypothesis 2.

In an initial approximation to this question and confirmation of Hypothesis 2, columns 3 and 6 (Table 3.2) show that, after controlling for industry returns and the previous day's number of CeDex-related transactions, an increase in the number of CeDex-related transactions increases the token returns of the day. Therefore, hypothesis 2 is confirmed, hence, exchange-related transactions impact token prices.

As robustness tests, the same analysis using the Random Effects GLS regression model (see Table 3.A1 in the appendix) is performed obtaining the same results. Furthermore, an OLS regression without controlling for Fixed Effects (see Table 3.A2 in appendix) has been tested to check whether token characteristics (i.e. tokens that tend to be used more as compared to others that are not commonly used on a daily basis) are taking all the variation and driving the results to insignificant coefficients. Results remain robust.

3.4.2. Market sentiment and token value

To investigate the mechanism that produces this impact of on-chain exchange-related transactions on prices I measure the market and investors' sentiment by further classifying the CeDex-related transactions into two groups. In particular, I separate those transactions in which a CeDex wallet sends tokens to a private wallet ("from" CeDex transactions) and those in that a private wallet sends tokens to a CeDex wallet ("to" CeDex transactions). The first group of transactions is related to investors that own tokens that were most likely bought on the exchange and are interested in moving their stock of tokens from the exchange's private wallet (off-chain) to their own personal private on-chain wallet (known as "cold wallet"). This is a common action in crypto markets for both long-term investors and decentralization supporters. The reason is that investors who buy tokens in an exchange do not truly own the custody of the token until they move it to their private on-chain wallet. While the tokens are in the exchange's private wallet, the exchange owns the custody and must be trusted to keep it safe. However, since token markets and many digital assets investors intend to be completely decentralized, they claim the best custody is the investor itself. On the other hand, the later group of transactions ("to" CeDex transactions) are related to investors that own tokens and are interested in exchanging them, liquidating them, or taking on a short position within an exchange. It is important to clarify the exchange of tokens without using an exchange, though possible is very difficult. Peer-to-peer exchanges can only happen between parties that know each other and agree previously on prices and conditions. Therefore, most of the listed token exchanges occur using centralized and decentralized exchanges.

Results in all columns 1 to 4 (see Table 3.3, Panel A) show that negative market sentiment (proxied by the ratio of CeDex "to" transactions compared to total CeDex-related transactions) leads to lower prices and returns. In other words, the proportion of investors who go to the market to sell or short-sell has a negative impact on prices. Therefore, Hypothesis 2A is confirmed. This first channel states that investors that predict that tokens are overvalued (undervalued) and prices should decrease (increase) during the day, tend to move their positions to (from) an exchange in order to sell (after buying it) before the price drops (rises), showing a signal that other investors can observe. Therefore, these investors generate network effects from showing publicly (through on-chain transactions) their behavior, hence the market sentiment. Again, it is important to clarify that moving your tokens from your public Ethereum wallet to an exchange wallet is time-consuming and very costly (i.e. high gas prices), hence most investors keep their tokens on cold wallets for long periods of time. Therefore, since "to" and "from" exchange transactions are costly, they are very good market sentiment signals. Observing these transactions can proxy for market sentiment and other investors can observe the information on market sentiment and quickly buy or sell depending on the market sentiment predicted through the costly signals (network effect). The same analysis has been performed using Random Effects GLS regression and results remain robust (see Table 3.A3 in the appendix).

According to the other stated complementary channel tested in Hypothesis 2B, as more tokens are available for selling (buying) in the exchange, there is more downward (upward) pressure on prices. Results show that observing the proportion of the number of tokens transferred to exchanges (the Exchange supply and demand ratio) has an impact on token returns (see Table 3.3, Panel B). Therefore, Hypothesis 2B is confirmed. Supply and demand forces of tokens have an impact on prices.

Together, all these results tested in H2, H2A, and H2B show that the channels that make prices to move are not only the supply and demand pressures in the exchanges (H2B), but also the visibility of these transactions, which shows the sentiment of the market on the token, enlarging and producing herding and network effects (H2A).

In order to know which channel is more important, we perform the same analysis using both ratios (on the number of transactions and on the number of tokens transferred). Results show that both ratios have an impact (see columns 1 and 2 in Table 3.3, Panel C) on token prices, however after standardizing the betas, the second channel (supply and demand forces in exchanges) has the largest impact (-0.015*** ass compared to -0.007***) on token prices (see columns 3 and 4 in Table 3.3, Panel C). It is interesting to note that still, visibility of on-chain transactions is important in determining token prices, hence this mechanism or channel is enlarging the effect of token transactions on prices. This raises the question of whether the token market can be manipulated by investors moving their tokens to and from CeDex exchanges in order to move prices. Market players could be using this strategy to show real or artificial market sentiment, hence, creating pump-and-dump schemes. However, as explained above, these movements are costly, which should prevent this type of market manipulation[25].

All in all, Hypothesis 2 is supported since exchange-related transactions do impact prices. Particularly, the effect is enlarged since the visibility of the signal of exchange-related transactions is used by investors as a proxy for market sentiment.

3.4.3. Momentum and feedback trading.

Feedback trading can be categorized as a special case of herding. Feedback trading occurs when a set of investors during a period of time trade in the same direction. Investors would agree on trading in the same direction when lag returns, or other variables such as variation of firm characteristics, decisions of previous traders, or earnings momentum of the stock are observed as the common signal for taking action on trading. Therefore, feedback trading involves a correlation between present and lag returns (Sentana and Wadhwani, 1992). "Positive feedback trading" occurs when the correlation is positive and it is often referred to as momentum trading. Specifically, positive-feedback trading is a trading strategy followed by investors who buy securities

[25] It is not the purpose of the paper to address this question, but to analyze the effect of on-chain transactions on token prices. Therefore, this analysis on the intentionality of using or not transactions to manipulate the market will be performed in a future research.

after price increases and sell after price declines. In many models though, feedback trading is tested for in intraday settings.

The results in Table 3.4, Panel A (columns 1 and 2), and Panel B (columns 1 and 2) show evidence consistent with momentum strategies in token exchange markets. In particular, we can see that higher price increases or decreases translate into higher trade volume on exchanges on the next day.

Additionally, in columns 3 and 4 (on both Panel A and B) we can observe that higher price movements also affect positively the CeDex number of transactions on the following day.

However, in columns 5 and 6, the results do not show a significant effect. Previous-day token extreme returns (both positive or negative extreme returns) do not have an impact on the total number of token use transactions (non-CeDex transactions). This might be indicative of "real token use", and confirm that my proxy is a good measure of real token use intensity, since extreme returns (positive or negative) do not have an impact on the token use intensity on the following day.

Therefore, all these results together confirm Hypothesis 3. Investors in token exchanges make use of the information contained in previous price movements (i.e. they engage in feedback strategies) while token users do not make use of the information contained in previous price movements before using the token.

3.4.4. Are investors taking positive or negative feedback trading strategies?

Unfortunately and differently from most feedback trading literature, I can't observe order books and flows, and I do not know whether the volume of transactions increases more on the buying orders part or the seller part (Sentana and Wadhwani, 1992). Nevertheless, we can approximate that type of analysis using CeDex "from and to" transactions and ratios.

In column 1 (see Table 3.4 Panel A and B) we observe that in general, as |absolute returns| increase (i.e. as more extreme prices are reached), there are less transactions to CeDex (fewer investors going short/trade), and more transactions from CeDex to public Ethereum wallets or cold wallets ("HODL", investors on long positions). The higher the losses or the higher the gains on the previous day, the more investors do "HODL" (meaning hold the asset for the long term, used in the blockchain argot), hence they do not want to sell and prefer to maintain for the long-term.

In columns 1 to 4 (for both Panel A and B), we can see how classical models of positive feedback trading do not seem to apply here. In fact, instead of past positive returns encouraging investors to buy, we see opposite results, indicating negative feedback strategies. When positive returns are obtained on the previous days, investors prefer to trade (i.e. going short, moving tokens to centralized or decentralized exchanges) on the following day. However, after negative returns on the previous day, investors prefer to "HODL" and maintain their money in cold wallets until "good times" arrive. Existing literature points out the existence of feedback trading strategy on good times, but not that much on "bad times". Therefore, we show a new setting (token markets) in which these strategies do occur.

The strategy we observe here is that investors prefer to trade (i.e. short-sell or move tokens to CeDex) after a period of positive returns. On the other hand, investors do "HODL" and deposit their money in cold wallets after "long periods" of losses (see columns 2 and 4 of Table 3.4, Panel B). Furthermore, the higher the benefits or losses on the previous day, the fewer investors trade or move tokens to CeDex and the more investors do HODL. The other way around, the more insignificant, small benefits or losses the previous days, the more investors do trade or move coins to CeDex, and the fewer investors do HODL or move to cold wallets.

Considering this as a moderation effect, it can be confounding. When returns are positive but small, investors prefer to move tokens to the CeDex wallet in order to trade. However, when returns are negative and have a large drop, investors prefer to keep their coins in cold wallets and keep them for the long term, they hope that things will recover in the future, or they took advantage to buy at the deep (i.e. shark investors that "buy the dip") and now keep the tokens for the long term.

Therefore, when we take into account the possible moderation effect (see Table 3.5 and Figure 3.1), the higher the losses (blue line) on the previous day, the more investors do have long positions (i.e. "HODL" on the blockchain argot) on the public Ethereum blockchain compared to those investors moving their tokens to the exchange to sell them on the next day. However, it does not look like the gains or positive returns increment (red line) affects the ratio of tokens moved to exchanges compared to those moved to Ethereum private wallets.

3.5. Conclusion

This study opens a new stream of literature in the token and crypto-asset markets through the introduction of a proxy to measure real token use intensity and measuring the impact of token prices on both token use and token exchange and trade, hence on investors' behavior and strategic use of information.

Among the implications for future studies, this paper includes a new proxy to measure token use transactions that can be used for further analysis of the utility of crypto assets. Regarding the implications of the results for investors show that token exchange-related transactions are a good signal to investors for price predictions, doesn´t matter if the market is being manipulated or not, they can get benefit by considering the market sentiment proxy (the evolution of exchange related transactions) used in this paper for better predicting. Remarkably, token users do not observe token price or transactions to use the token, neither the use of the token affects the value. On the other way around, token investors do observe other investors exchange related transactions and prices to trade the token, while their trades do impact the token's value. Definitely, understanding investors' behavior and sentiment in crypto markets can help us better estimate price movements both in the short and medium term.

Therefore, whereas prior work has studied token adoption from the network point of view, in this paper real token use in the ecosystem it was created for is empirically studied for the first time through the introduction of a new proxy that measures token use intensity. This has helped not only separate token-use transactions but also token exchange-related transactions and better analyze market sentiment. I hope this research will inspire future studies on the token and crypto-assets markets.

References

Adhami, S., Giudici, G., & Martinazzi, S. (2018). Why Do Businesses Go Crypto? An Empirical Analysis of Initial Coin Offerings. *Journal of Economics and Business, 100*(C), 64−75.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets, 5*(1), 31-56.

Aspris, A., Foley, S., Svec, J., & Wang, L. (2021). Decentralized exchanges: The "wild west" of cryptocurrency trading. *International Review of Financial Analysis, 77*. 101845.

Baur, D. G., & Hong, K. (2018). Bitcoin: Medium of exchange or speculative assets? *Journal of International Financial Markets, Institutions and Money, 54*, 177-189.

Benedetti, H., & Kostovetsky, L. (2021). Digital Tulips? Returns to investors in initial coin offerings. *Journal of Corporate Finance, 66.*

Benedetti, H., & Nikbakht, E. (2021). Returns and network growth of digital tokens after cross-listings. *Journal of Corporate Finance, 66*.

Benedetti, H., Caceres, C. & Abarzúa, L.Á. (2023). Utility Tokens. Baker, H.K., Benedetti, H., Nikbakht, E. and Smith, S.S. (Ed.) *The Emerald Handbook on Cryptoassets: Investment Opportunities and Challenges,* Emerald Publishing Limited, Bingley, 79-92.

Benedetti, H. & Rodríguez-Garnica, G. (2023). Tokenized Assets and Securities. Baker, H.K., Benedetti, H., Nikbakht, E. and Smith, S.S. (Ed.) *The Emerald Handbook on Cryptoassets: Investment Opportunities and Challenges,* Emerald Publishing Limited, Bingley, 107-121.

Bleher, J., & Dimpfl, T. (2019). Today I got a million, tomorrow, I don't know: On the predictability of cryptocurrencies by means of Google search volume. *International Review of Financial Analysis, 63*. 147-159.

Blémus, S., & Dominique G. (2020). Initial Crypto-asset Offerings (ICOs), Tokenization and Corporate Governance. *Capital Markets Law Journal, 15*(2), 191–223.

Boreiko, D., & Risteski, D. (2021). Serial and large investors in initial coin offerings. *Small Business Economics, 57*(2), 1053–1071.

Campino, J., Brochado, A., & Rosa, A. (2022). "Initial Coin Offerings (ICOs): Why Do They Succeed?" *Financial Innovation, 8*(17), 1−35.

Capponi, A., & Jia, R. (2021). The Adoption of Blockchain-based Decentralized Exchanges: A Market Microstructure Analysis of the Automated Market Maker. *SSRN Working Paper*

Cheah, E. T., & Fry, J. (2015). Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters, 130*, 32-36.

Chen, Y. (2018). Blockchain Tokens and the Potential Democratization of Entrepreneurship and Innovation. *Business Horizon*s, *61*(4), 567−575.

Cong, L. W., & He, Z. (2019). Blockchain Disruption and Smart Contracts. *Review of Financial Studies, 32*(5),1754–1797.

Cong, L. W., Karolyi, G.A., Tang, K. & Zhao, W. (2022). Value Premium, Network Adoption, and Factor Pricing of Crypto Assets 1. *SSRN*

Cong, L.W., Li, Y., & Wang, N. (2022). Token-based platform finance. *Journal of Financial Economics, 144*(3), 972-991.

Cong, L. W., Li, Y., & Wang, N. (2021). Tokenomics: Dynamic Adoption and Valuation. *Review of Financial Studies, 34*(3), 1105–1155.

Corbet, S., Larkin, C., Lucey, B., & Yarovaya, L. (2018). Datestamping the Bitcoin and Ethereum bubbles. *Finance Research Letters, 26*, 81-88.

Fahlenbrach, R., & Frattaroli, M. (2019). ICO investors. *Financial Markets and Portfolio Management, 35*, 1 - 59.

Foley, S., Karlsen, J. R., & Putnins, T. J. (2019). Sex, Drugs, and Bitcoin: How Much Illegal Activity Is Financed through Cryptocurrencies? *Review of Financial Studies, 32*(5), 1798–1853.

Gan, J.R., Tsoukalas G., & Netessine, S. (2021). Initial Coin Offerings, Speculation, and Asset Tokenization. *Management Science, 67*(2), 914−931.

Griffin, J. M., & Shams, A. (2020). Is Bitcoin Really Untethered? *The Journal of Finance, 75*(4), 1913–1964.

Howell, S. T., Niessner, M., & Yermack, D. (2020). Initial Coin Offerings: Financing Growth with Cryptocurrency Token Sales. *Review of Financial Studies, 33*(9), 3925−3974.

Hu, A. S., Parlour, C. A., & Rajan, U. (2019). Cryptocurrencies: Stylized facts on a new investible instrument. *Financial Management, 48*(4), 1049–1068.

Jiménez, I., Mora-Valencia, A., & Perote, J. (2022). Semi-nonparametric risk assessment with cryptocurrencies. *Research in International Business and Finance, 59.*

John, K., O'hara, M., & Saleh, F. (2022). Bitcoin and Beyond. *Annual Review of Financial Economics, 14*, Forthcoming.

Lambert, T., Liebau, D., & Roosenboom, P. (2021). Security Token Offerings. *Small Business Economics*, *12*, 1−27.

Lee, J., & Parlour, C.A. (2021). Consumers as Financiers: Consumer Surplus, Crowdfunding, and Initial Coin Offerings. *The Review of Financial Studies, 35*(3), 1105–1140

Li, T., Shin, D., and Wang, B. (2023). Cryptocurrency Pump-and-Dump Schemes. *SSRN Working Paper.*

Liu, Y., & Tsyvinski, A. (2021). Risks and returns of cryptocurrency. *Review of Financial Studies, 34*(6), 2689–2727.

Liu, Y., Tsyvinski, A., & Wu, X. (2022). Common Risk Factors in Cryptocurrency. *The Journal of Finance, 77*(2), 1133–1177.

Lyandres, E., Palazzo, D., Board, R., & Rabetti, D. (2019). Are Tokens Securities? An Anatomy of Initial Coin Offerings. *SSRN*

Lyandres, E., Palazzo, B., & Rabetti, D.(2022). Initial Coin Offering (ICO) Success and Post-ICO Performance. *Management Science, 68*(12).

Makarov, I., & Schoar, A. (2020). Trading and arbitrage in cryptocurrency markets. *Journal of Financial Economics, 135*(2), 293–319.

Makarov, I., & Schoar, A. (2021). Blockchain Analysis of the Bitcoin Market. *NBER*. Retrieved from: https://doi.org/10.3386/W29396

Makridis C.A., Fröwis M., Sridhar K., & Böhme R. (2023). The rise of decentralized cryptocurrency exchanges: Evaluating the role of airdrops and governance tokens. *Journal of Corporate Finance, 79*. 102358.

Momtaz, P. (2019). Token Sales and Initial Coin Offerings: Introduction. *Journal of Alternative Investments, 21*(4), 7–12.

Momtaz, P. (2021). Entrepreneurial Finance and Moral Hazard: Evidence from Token Offerings. *Journal of Business Venturing, 36*(5), 106001.

Momtaz, P., Rennertseder, K., & Schröder, H. (2019). Token Offerings: A Revolution in Corporate Finance? *Journal of Financial Transformation, 49*. 32−41.

Rognone, L., Hyde, S., & Zhang, S. (2020). News sentiment in the cryptocurrency market: An empirical comparison with Forex. *International Review of Financial Analysis, 69*, 101462.

Saleh, F. (2021). Blockchain without Waste: Proof-of-Stake. *The Review of Financial Studies, 34*(3), 1156–1190.

Sapkota, N. (2019). News-based sentiment and bitcoin volatility. *International Review of Financial Analysis, 82*, 102183.

Schär, F. (2021). Decentralized finance: on blockchain-and smart contract-based financial markets. *Federal Reserve Bank of St. Louis Review, 103*(2), 153–174.

Sentana, E. & Wadhwani, S.B. (1992). Feedback Traders and Stock Return Autocorrelations: Evidence from a Century of Daily Data. *Economic Journal, 102.* 415-425.

Shams, A. (2020). The Structure of Cryptocurrency Returns. *SSRN.*

Tapscott, D. (2020). Token Taxonomy: The Need for Open-Source Standards Around Digital Assets. Retrieved from: www.blockchainresearchinstitute.org

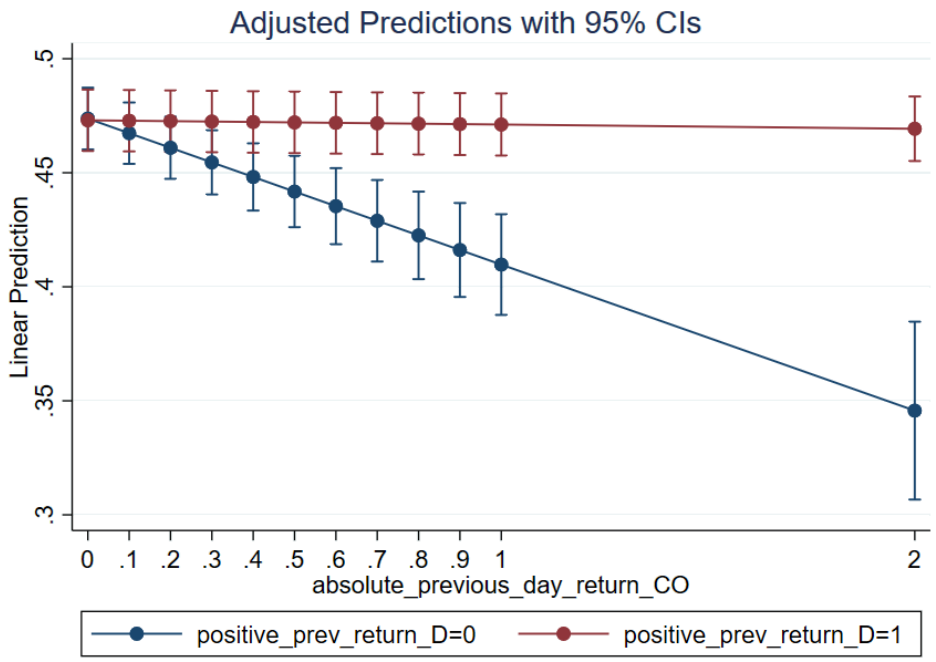
Urquhart, A., & Zhang, H. (2019). Is Bitcoin a hedge or safe haven for currencies? An intraday analysis. *International Review of Financial Analysis, 63*, 49-57

Yermack, D. (2015). Is Bitcoin a real currency? An economic appraisal. In *Handbook of Digital Currency*, ed. D.L. K. Chuen, 31–43. Amsterdam: Elsevier.

Yermack, D. (2017). Corporate Governance and Blockchains. *Review of Finance, 21*(1), 7–31.

Tables and Figures

Figure 3.1: Feedback trading strategies. Adjusted prediction of CeDex-related transactions based on previous returns.



Notes: In this table predicted at a 95% confidence interval the CeDex related transactions, relative to a negative market sentiment, measured as ratio_CeDex_to_transacitons, based on previous absolute returns. Positive returns are fitted in the red line and negative returns are shown in the blue line.

Table 3.1: Descriptive statistics and matrix of correlations

Panel A. Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
daily return CO	246974	.015	.604	-1	93.5
ethereumreturns_CO	246974	0	.05	-.423	.335
previous_daily_return_CO	246453	.015	.604	-1	93.5
Volume	246974	942424.85	5659613.8	0	6.981e+08
Volume 7day	243859	934254.84	4398643.5	0	3.173e+08
Amihud 7day	243859	.422	30.96	0	5305.426
RetVolatility_7day	243859	.14	.572	0	35.902
day_num_transactions	246974	86.212	641.096	1	93258
day_non CeDex_transactions	246974	55.987	606.582	0	93241
day_ CeDex_total_transactions	246974	30.225	126.558	0	15743
ratio CeDexto trans (sentiment)	206144	.447	.302	0	1

Panel B. Matrix of correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) daily_return_CO	1.000										
(2) ethereumreturns_CO	0.069	1.000									
(3) previous_daily_return_CO	-0.028	-0.005	1.000								
(4) Volume	0.012	0.009	0.007	1.000							
(5) L.Volume_7day	-0.006	0.001	-0.004	0.639	1.000						
(6) L.Amihud_7day	0.007	0.005	0.003	-0.002	-0.002	1.000					
(7) L. RetVolatility_7day	0.078	-0.000	0.405	-0.006	-0.006	0.009	1.000				
(8) day_num_transactions	-0.000	0.002	0.000	0.139	0.140	-0.001	-0.006	1.000			
(9) day_non CeDex_transactions	-0.001	0.001	-0.001	0.078	0.084	-0.001	-0.004	0.980	1.000		
(10) day_CeDex_total_transactions	0.004	0.006	0.004	0.329	0.303	-0.003	-0.012	0.375	0.185	1.000	
(11) ratio CeDexto trans (sentiment)	-0.017	0.008	0.000	-0.018	-0.026	-0.001	-0.004	0.020	0.024	-0.011	1.000

Table 3.2: Token use and CeDex transactions impact on token returns

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES			daily_return_CO			
day_num_transactions	0.001					
	(0.389)					
day_noncedex_transactions		-0.001				
		(0.328)				
day_cedex_total_transactions			0.007***			
			(0.001)			
day_totalvalue				-0.004		
				(0.133)		
day_noncedex_total_value					-0.005	
					(0.169)	
day_cedex_total_value						-0.003
						(0.231)
Controls:						
ethereumreturns_CO	0.068***	0.068***	0.068***	0.068***	0.069***	0.068***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
previous_daily_return_CO	-0.074***	-0.074***	-0.074***	-0.074***	-0.074***	-0.074***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Volume	0.027***	0.028***	0.026***	0.028***	0.028***	0.028***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
L.Volume_7day	-0.020***	-0.020***	-0.021***	-0.020***	-0.020***	-0.020***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
L.Amihud_7day	0.006***	0.006***	0.006***	0.006***	0.006***	0.006***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
L.RetVolatility_7day	0.027*	0.027*	0.027*	0.027*	0.027*	0.027*
	(0.078)	(0.078)	(0.079)	(0.079)	(0.076)	(0.079)
Token F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Observations	221,031	221,031	221,031	221,030	221,030	221,031
N clusters	512	512	512	512	512	512
R-squared	0.046	0.046	0.046	0.046	0.046	0.046
R-squared Adj.	0.0435	0.0435	0.0436	0.0436	0.0436	0.0436
p	0	0	0	0	0	0

Notes: Standardized beta coefficients. Robust standard errors in parentheses, clustered by token. Fixed Effects regression analysis. *** p<0.01, ** p<0.05, * p<0.1

Table 3.3: CeDex transactions channels impact on token returns

Panel A. Number of transactions				
	(1)	(2)	(3)	(4)
VARIABLES		daily_return_CO		
Negative Sentiment:				
ratio_cedexto_trans	-0.037***	-0.041***	-0.037***	-0.042***
	(0.005)	(0.007)	(0.005)	(0.007)
L.ratio_cedexto_trans		0.018*		0.019**
		(0.010)		(0.010)
day_cedex_total_transactions			0.000***	0.000***
			(0.000)	(0.000)
L.day_cedex_total_transactions				-0.000***
				(0.000)
Controls:				
ethereumreturns_CO	0.779***	0.774***	0.779***	0.773***
	(0.022)	(0.022)	(0.022)	(0.022)
previous_daily_return_CO	-0.062***	-0.063**	-0.062***	-0.063**
	(0.020)	(0.026)	(0.020)	(0.026)
Volume	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
L.Volume_7day	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
L.Amihud_7day	0.001	0.001	0.001	0.001
	(0.000)	(0.000)	(0.000)	(0.000)
L.RetVolatility_7day	0.021*	0.046	0.021*	0.046
	(0.011)	(0.035)	(0.011)	(0.035)
Token F.E.	Yes	Yes	Yes	Yes
Constant	0.025***	0.015**	0.024***	0.015**
	(0.002)	(0.006)	(0.002)	(0.006)
Observations	190,676	177,606	190,676	177,606
N_clusters	498	487	498	487
R-squared	0.038	0.047	0.038	0.047
R-squared Adj.	0.0353	0.0445	0.0353	0.0446
p	0	0	0	0

Panel B. Number of tokens transferred				
	(1)	(2)	(3)	(4)
VARIABLES		daily_return_CO		
Negative Sentiment:				
ratio_cedexto_value	-0.034***	-0.034***	-0.034***	-0.034***
	(0.004)	(0.004)	(0.004)	(0.004)
L.ratio_cedexto_value		0.008		0.008
		(0.006)		(0.006)
day_cedex_total_value			-0.000	-0.000***
			(0.000)	(0.000)
L.day_cedex_total_value				-0.000***
				(0.000)
Controls:				
ethereumreturns_CO	0.779***	0.774***	0.779***	0.774***
	(0.023)	(0.022)	(0.023)	(0.022)
previous_daily_return_CO	-0.061***	-0.063**	-0.061***	-0.063**
	(0.020)	(0.027)	(0.020)	(0.026)
Volume	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
L.Volume_7day	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
L.Amihud_7day	0.001	0.001	0.001	0.001
	(0.000)	(0.000)	(0.000)	(0.000)
L.RetVolatility_7day	0.021*	0.046	0.021*	0.046
	(0.011)	(0.035)	(0.011)	(0.035)
Token F.E.	Yes	Yes	Yes	Yes
Constant	0.024***	0.017***	0.024***	0.017***
	(0.002)	(0.005)	(0.002)	(0.005)
Observations	190,673	177,602	190,673	177,602
N_clusters	498	487	498	487
R-squared	0.038	0.047	0.038	0.047
R-squared Adj.	0.0353	0.0445	0.0354	0.0447
p	0	0	0	0

Panel C. Number of transactions and tokens transferred				
	(1)	(2)	(3)	(4)
VARIABLES		daily_return_CO		
Negative Sentiment:				
ratio_cedexto_trans	-0.013**	-0.013**	-0.007**	-0.007**
	(0.006)	(0.006)	(0.040)	(0.039)
day_cedex_total_transactions		0.000***		0.008***
		(0.000)		(0.001)
ratio_cedexto_value	-0.025***	-0.025***	-0.015***	-0.015***
	(0.004)	(0.004)	(0.000)	(0.000)
day_cedex_total_value		-0.000		-0.005
		(0.000)		(0.207)
Controls:				
ethereumreturns_CO	0.779***	0.779***	0.071***	0.071***
	(0.023)	(0.023)	(0.000)	(0.000)
previous_daily_return_CO	-0.061***	-0.062***	-0.059***	-0.059***
	(0.020)	(0.020)	(0.002)	(0.002)
Volume	0.000***	0.000***	0.030***	0.028***
	(0.000)	(0.000)	(0.000)	(0.000)
L.Volume_7day	-0.000***	-0.000***	-0.022***	-0.023***
	(0.000)	(0.000)	(0.000)	(0.000)
L.Amihud_7day	0.001	0.001	0.006	0.006
	(0.000)	(0.000)	(0.131)	(0.131)
L.RetVolatility_7day	0.021*	0.021*	0.018*	0.018*
	(0.011)	(0.011)	(0.053)	(0.054)
Token F.E.	Yes	Yes	Yes	Yes
Constant	0.026***	0.025***		
	(0.002)	(0.002)		
Observations	190,673	190,673	190,673	190,673
N_clusters	498	498	498	498
R-squared	0.038	0.038	0.038	0.038
R-squared Adj.	0.0353	0.0354	0.0353	0.0354
p	0	0	0	0

Notes: Fixed effects regression controlling for each token. In Panel A, B, and C (columns 1-2): robust standard errors in parentheses. In Panel C (columns 3-4): standardized beta coefficients, and p-values in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.4: Feedback strategies. Previous returns impact on CeDex and token use transactions

Panel A. Lagged previous absolute returns						
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Volume		CeDex_transactions		Token use transactions	
Abs_Previous_return	107,118.953***	160,390.130***	2.545***	3.638***	1.001	1.140
	(17,331.604)	(23,292.283)	(0.378)	(0.458)	(1.978)	(2.486)
L.Abs_Previous_return		108,646.090***		2.141***		0.868
		(23,130.371)		(0.455)		(2.469)
L2.Abs_Previous_return		102,990.424***		2.112***		0.370
		(24,951.136)		(0.491)		(2.660)
L3.Abs_Previous_return		105,257.779***		1.921***		1.104
		(25,047.434)		(0.493)		(2.686)
Constant	865,680.915***	862,350.635***	22.487***	22.934***	49.647***	54.356***
	(124,975.686)	(128,796.468)	(2.161)	(2.136)	(4.752)	(3.404)
Observations	246,453	199,936	246,453	199,936	246,453	199,936
Number of Tokens	521	509	521	509	521	509
chi2	38.20	99.95	45.36	114.7	0.256	0.536
p	6.39e-10	0	0	0	0.613	0.970

Panel B. Interaction dummy positive/negative previous return and absolute previous return						
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Volume		CeDex_transactions		Token use transactions	
D_Prev_Positive_return	232,434.757***	329,887.361***	3.081***	5.283***	3.687	1.913
	(21,323.049)	(24,466.536)	(0.477)	(0.533)	(2.441)	(2.737)
Abs_Previous_return		1478532.081***		39.767***		15.453
		(154,620.549)		(3.366)		(17.244)
Interaction (Dummy x Abs)		-1392667.899***		-37.758***		-14.859
		(155,176.960)		(3.378)		(17.315)
Constant	786,229.070***	651,666.228***	21.947***	18.548***	49.646***	48.384***
	(126,709.881)	(124,122.301)	(2.210)	(2.156)	(4.998)	(5.009)
Observations	238,549	238,028	238,549	238,028	238,549	238,028
Number of Tokens	521	521	521	521	521	521
chi2	118.8	224.0	41.82	189.8	2.282	0.984
p	0	0	1.00e-10	0	0.131	0.805

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Random-effects linear regression by GLS.

Table 3.5: Feedback trading strategies. Previous returns impact on the ratio of CeDex "to" transactions

VARIABLES	(1) ratio_CeDexto_trans	(2) ratio_CeDexto_trans
L.ratio_CeDexto_trans	0.253***	0.253***
	(0.002)	(0.002)
D_Prev_Positive_return	0.004**	0.004**
	(0.001)	(0.001)
Abs_Previous_return	-0.088***	
	(0.010)	
Previous_daily_return		0.088***
		(0.010)
Interactions:		
D_Prev_Positive_return x Abs_Previous_return	0.088***	
	(0.010)	
D_Prev_Positive_return x Previous_daily_return		-0.089***
		(0.010)
Constant	0.353***	0.353***
	(0.003)	(0.003)
Observations	176,642	176,642
Number of Tokens	499	499
chi2	12080	12080

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Random-effects linear regression by GLS.

Appendix 3

Table 3.A1: Robustness test. Token use and CeDex transactions impact on token returns. RE.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
			daily_return_CO			
day_nonCeDex_transactions	-0.000					
	(0.000)					
day_CeDex_total_transactions		0.000***				
		(0.000)				
day_num_transactions			0.000			
			(0.000)			
var_nonCeDextrans_7daymean				0.000		
				(0.000)		
var_tot_CeDextrans_7daymean					0.001***	
					(0.000)	
var_tot_trans_7daymean						0.000
						(0.000)
ethereumreturns_CO	0.752***	0.752***	0.752***	0.752***	0.756***	0.752***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.024)	(0.023)
previous_daily_return_CO	-0.068***	-0.068***	-0.068***	-0.068***	-0.068***	-0.068***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Volume	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
L.Volume_7day	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
L.Amihud_7day	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
L.RetVolatility_7day	0.064***	0.064***	0.064***	0.064***	0.063***	0.064***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Random Effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.017***	0.017***	0.017***	0.017***	0.018***	0.017***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
Observations	221,032	221,032	221,032	220,843	214,242	221,032
Number of Tokens(name2)	513	513	513	513	504	513

Notes: Non-standardized coefficients. Standard errors in parentheses.. Random-effects linear regression by GLS. *** p<0.01, ** p<0.05, * p<0.1

Table 3.A2: Robustness test. Token use and CeDex transactions impact on token returns. OLS.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES			daily_return_CO			
day_num_transactions	-0.000	-0.000				
	(0.000)	(0.000)				
L.day_num_transactions		0.000				
		(0.000)				
day_noncedex_transactions			-0.000	-0.000		
			(0.000)	(0.000)		
L.day_noncedex_transactions				0.000		
				(0.000)		
day_cedex_total_transactions					0.000*	0.000***
					(0.000)	(0.000)
L.day_cedex_total_transactions						-0.000***
						(0.000)
ethereumreturns_CO	0.752***	0.752***	0.752***	0.752***	0.752***	0.752***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
previous_daily_return_CO	-0.067***	-0.067***	-0.067***	-0.067***	-0.067***	-0.067***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Volume	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
L.Volume_7day	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
L.Amihud_7day	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
L.RetVolatility_7day	0.110***	0.110***	0.110***	0.110***	0.110***	0.110***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Token F.E.	No	No	No	No	No	No
Constant	-0.002	-0.002	-0.002	-0.002	-0.002*	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	221,032	221,032	221,032	221,032	221,032	221,032
R-squared	0.016	0.016	0.016	0.016	0.016	0.016
r2_a	0.0155	0.0155	0.0155	0.0155	0.0155	0.0156
p	0	0	0	0	0	0

Notes: Non-standardized coefficients. OLS Regression, no FE. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.A3: Robustness test. CeDex transactions channels impact on token returns. RE.

	(1)	(2)	(3)	(4)	(5)
VARIABLES			daily_return_CO		
Negative Sentiment:					
ratio_CeDexto_trans	-0.037***	-0.041***	-0.037***	-0.042***	-0.042***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
L.ratio_CeDexto_trans		0.018***		0.018***	0.019***
		(0.005)		(0.005)	(0.005)
day_CeDex_total_transactions			0.000***	0.000***	0.000***
			(0.000)	(0.000)	(0.000)
L.day_CeDex_total_transactions					-0.000***
					(0.000)
Controls:					
ethereumreturns_CO	0.776***	0.772***	0.776***	0.772***	0.772***
	(0.025)	(0.026)	(0.025)	(0.026)	(0.026)
previous_daily_return_CO	-0.062***	-0.064***	-0.062***	-0.064***	-0.064***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Volume	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
L.Volume_7day	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
L.Amihud_7day	0.001**	0.001***	0.001**	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
L.RetVolatility_7day	0.041***	0.060***	0.041***	0.060***	0.061***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Random Effects	Yes	Yes	Yes	Yes	Yes
Constant	0.042***	0.036***	0.041***	0.036***	0.035***
	(0.007)	(0.009)	(0.007)	(0.009)	(0.008)
Observations	190,681	177,612	190,681	177,612	177,612
Number of Tokens	503	493	503	493	493

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Random-effects linear regression by GLS.