Small Bus Econ
https://doi.org/10.1007/s11187-024-00933-z

RESEARCH ARTICLE

Signaling and herding in reward-based crowdfunding

Gabriel Rodríguez-Garnica ·
María Gutiérrez-Urtiaga · **Josep A. Tribo**

Accepted: 3 May 2024

Abstract This paper investigates how signaling and herding behavior interact in crowdfunding markets to give raise to an information cascade, even when there are no identifiable experts, which is the typical case in reward-based crowdfunding. Using daily funding data for on all the projects launched on Kickstarter during one month, we find that during the initial phase of the campaign, the funding decisions of a reduced number of early backers are based on information and 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. The results suggest that, even in the absence of identifiable experts, backers self-select into early or late backers depending on their ability to process the information, so that herding after signaling generates an information cascade that ameliorates asymmetric information problems. The findings are relevant for (i) creators, that will obtain better results by targeting their crowdfunding campaigns at better informed potential contributors, and (ii) regulators, that can expect backers' self-selection and herding to work together to protect uninformed backers from fraud and deception even when participation is not restricted.

Plain English Summary Information Cascades in Funding Markets without Experts: Our paper shows that in funding markets where there are no identifiable experts, such as reward-funding crowdfunding, a phenomenon of information cascades leading to efficient finance decisions may appear. In reward-based crowdfunding most backers are not experts in the products and services they are buying and, even those that may know more, cannot be identified by the rest of the backers. Thus, these funding markets are expected to produce inadequate funding decisions. However, we show that better-informed backers self-select into becoming early backers, while worst-informed backers prefer to be late backers. This results in a positive reinforcing information cascade that improves the quality of funding decisions. This result has implications for entrepreneurs, who should

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11187-024-00933-z.

G. Rodríguez-Garnica (✉)
Universidad Pontificia Comillas, School of Economics and Business Administration, Department of Finance, 28015 Madrid, Spain
e-mail: grgarnica@comillas.edu

M. Gutiérrez-Urtiaga
Department of Business Administration, Universidad Carlos III de Madrid and ECGI, Calle Madrid, 126, 28903 Getafe (Madrid), Spain
e-mail: maria.gutierrez@uc3m.es

J. A. Tribo
Stevens Institute of Technology School of Business, 1 Castle Point Terrace, Hoboken, NJ 07030, USA
e-mail: jtribo@stevens.edu

target their campaigns at better informed potential backers, for backers looking for attractive products to fund, and for regulators worried about consumer protection in reward-based platforms.

Keywords Reward-based crowdfunding · Herding behavior · Information cascades · Signaling · Observational learning · Wisdom of the crowd · Kickstarter

JEL Classification G30 · G32

1 Introduction

Creative ideas are typically generated by individuals with reduced financial resources and no access to banks, angel investors, or venture capital funds (Agrawal et al., 2015; Schwienbacher & Larralde, 2012). In this setting, 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 (Agrawal et al., 2015; Belleflamme et al., 2014; Kuppuswamy & Bayus, 2018b; Schwienbacher & Larralde, 2012) where the sense of belonging to a community is important (Goethner et al., 2021a). On these platforms, entrepreneurs are known as “creators” and the funds are provided by “contributors” or “backers”. Depending on the type of crowdfunding, in exchange for the funds, the backers receive either equity or debt securities (in equity and debt crowdfunding) or the commitment to receive the product or service once it is developed (in reward-based crowdfunding).

Although the particular characteristics of backers depend on the type of crowdfunding, many backers are usually small, first time contributors 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 by engaging in signaling and herding behavior.

First, signaling is used by creators to convince platforms and potential backers of the quality of their projects. Creators provide detailed information and project descriptions and also use costly signals to convince backers of the quality of their projects and backers react to information and signals. And these signals have been shown to determine to a large extent the probability of campaign success (Ahlers et al., 2015; Mollick, 2014). However, the effectiveness of signaling by the creator also depends on the ability of the target receiver of the signal to process the information (Vismara, 2018b). And, specifically, in reward-based crowdfunding -where backers are found among small, dispersed, potential consumers-, this ability is expected to be small, casting doubts on the efficiency of signaling.

Second, backers also pay attention to the funding decisions made by other backers and herd (i.e. late backers imitate early backers) as an alternative mechanism to reduce asymmetric information. Interestingly, this herding behavior may result either in information cascades, with each new backer having more information than the previous backers, giving rise to a “wisdom of the crowd” effect, or in irrational herding, with late backers following the whims of early ones.[1] The effectiveness of herding will also depend on the characteristics of the first receivers of the creators’ signals at the beginning of the campaign. In equity and debt crowdfunding, the signals are first analyzed by the platforms that give access to the platform to a small number of applicants (Kleinert & Vismara, 2024) and by reputed professional investors, and this can then trigger an information cascade (Vismara, 2018a, b), with poorly informed potential backers imitating the behavior of these experts. However, reward-based crowdfunding platforms accept almost all projects and there are no identifiable experts, which casts doubts on the possibility of information cascades in this market and raises concerns of irrational herding.

[1] Throughout the paper we use the term “herding” in a neutral fashion indicating that late backers’ decision to pledge is influenced by the funding decisions made by earlier backers. We use the term “irrational herding” when this influence is based on fashions and whims and the term “information cascade” when the influence is rational and due to information aggregation effects.

In this paper we show that in reward-based crowdfunding campaigns there is a dynamic relationship between signaling and herding that is consistent with an informational cascade, even in the absence of recognizable expert backers. Our results indicate that early backers do learn from the information and signals offered by the creator, and late bakers passively herd on early backers' behavior and benefit from their information, even though at the onset they were unidentifiable. We posit that this happens because potential backers can self-select into being early, and using their own information and quality signals and information provided by the creator, or being late, and imitating previous backers. Backers who do not have the ability to process this information prefer to come late and rely on herding, while backers that are better informed arrive early and make their decisions based on available information and processing the signals of the creator. Therefore, the creator signals are important for the better-informed potential contributors, who can become early backers, and then, other potential contributors simply herd, imitating the behaviour of early backers or ignoring the project if there are not enough early backers. This implies that *signaling and herding seem to substitute each other as the campaign evolves but they are effectively complementing each other to produce an information cascade.*

We use a unique dataset composed of the population of projects launched on Kickstarter from November 15, 2017 to December 20, 2017 (3,923 Kickstarter projects). This database contains daily information on funding, which is pivotal to study the dynamics of the signaling and the herding behavior along time. In particular, first, we prove that early backers (i.e. those that provide the first 10% of the funding goal) make their funding decisions relying heavily on information of projects' and entrepreneurs' characteristics as well as on costly signals, rather than acting on the bases of whims. Specifically, the time it takes to raise this first 10% of funds is significantly lower for projects of higher observed quality. Moreover, we show that the ultimate success of the funding campaign also depends on costly information variables. Second, we show that as the funding campaign progresses, late backers' interest in the project and their decision on whether to pledge funds (i.e. measured as the time it takes to move the funding from 10 to 20%, 20% to 30%, etc.) is less influenced by quality information and more by the behavior of early backers. Hence, their herding becomes more pronounced as more information from earlier backers accumulates (i.e. as the backers can observe how much time it took to raise the first 10%, 20%, etc.). Remarkably, the dependence of late backers on early backers' behavior is more pronounced when early backers act in a more informative way (i.e. there is less dispersion in the early pledges). Overall, the pattern of behavior we identify in this market is compatible with the dynamics of information cascades rather than irrational herding. Additionally, we show 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.

We contribute to the crowdfunding literature in general by showing how signaling and herding, as separate strategies to reduce asymmetric information, can complement each other and work to produce an information cascade even when there are no identifiable experts. We identify a self-selection mechanism among backers depending on their information processing abilities. Zhang and Liu (2012), Vismara (2016), Vismara (2018a) and Wang et al. (2019), among others, have already documented information cascades in crowdfunding, but they have mainly focused on equity and debt crowdfunding platforms, where signals can be targeted towards reputed professional investors and late backers can observe the decisions of these experts. We also contribute to the specific literature on reward-based crowdfunding. There is, in fact, a large empirical literature showing evidence of signaling and herding in reward-based crowdfunding. Our contribution to this literature is twofold. First, we show that the signals that have been shown to impact the probability of success (Mollick, 2014; Zheng et al., 2014) are being processed by early backers (5–10% of the total backers) with higher information processing ability but not by late backers. And, second, we provide evidence indicating that the herding behavior, previously documented by authors such as Colombo et al. (2015), Gangi and Daniele (2017), Kuppuswamy and Bayus (2018b) and Chan et al. (2020), happens after the signaling phase is over, represents most of the funding raised and is consistent with an information cascade rather than with irrational herding.

Our findings also have important policy implications for creators seeking funding, regulators trying to protect uninformed backers, and for our understanding of the crowdfunding market in general. Regarding entrepreneurs/creators, our results indicate that they will obtain better results from crowdfunding campaigns if they target their products and campaigns at better informed potential backers, which may include potential customers that have already used similar products or services. These types can interpret betters the signals of the entrepreneur and, if they chose to back the project, they will 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 crowdfunding because consumer protection is scarce in these markets (Bradford, 2012; Griffin, 2013; 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 crowdfunding markets. And second, the interplay between signaling and herding increases both the funding opportunities and the quality of decisions so that the crowdfunding market, in general, and the reward-based crowdfunding market, in particular, becomes more attractive. Rational uninformed backers would not participate in 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 paper is as follows. Section 2 presents the theoretical background and develops the hypotheses. Section 3 and 4 describe the data and the methodology used. Section 5 presents the results. Section 6 explains the robustness tests. Section 7 discusses the results obtained and their policy implications, presents limitations and guidelines for future research. Finally, Section 8 includes some brief concluding remarks.

2 Theoretical background, literature review, and hypothesis development

2.1 Asymmetric information in crowdfunding

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*". In this sense crowdfunding increases the opportunities for funding new ventures and it can "*democratize entrepreneurial finance by providing access to funding to underrepresented groups of potential entrepreneurs*" (Cumming et al., 2021).

In equity and debt crowdfunding the entrepreneurs must have set up a firm before launching the campaign to raise funds by selling standard financial contracts and the backers can be divided into accredited (typically venture capitalists, business angels and high net worth individuals) and non-accredited, and there are restrictions in the amounts that non-accredited investors can pledge.[2] Moreover, because security issuance needs to comply with the financial regulation of the country where the firm is incorporated, equity and debt crowdfunding platforms usually operate at the national level and filter applications from entrepreneurs that want to raise funds (according to Kleinert et al., 2022 these platforms reject as much as 90% of applicants).

In reward-based crowdfunding any would-be entrepreneur with a project at an early development stage can raise funding by pre-selling the product or service (Vismara, 2018b). Therefore, typically, creators in reward-based crowdfunding have little experience in moving products from their initial concept to the market and they are relatively unknown to the potential backers (Ganatra, 2016). Regarding backers, everyone is a potential backer and contributions are usually small because they relate to the cost of buying a number of units of the product or service. In fact, in reward-based crowdfunding platforms the majority of contributors are one-time backers who join the platform and pledge in the same day, while in equity and debt platforms there are many serial backers

[2] Crowdfunding is regulated by the JOBS act in the US. 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. Note that an investor is considered accredited if s/he had an annual income of at least $200,000 a year for the past two years (or household annual income of $300,000) with the expectation that it will continue; or a net worth of $1 million US or more, excluding the investor's primary residence.

(Agrawal et al., 2015). Finally, reward-based crowdfunding is dominated by a few very large platforms that operate internationally (Kickstarter & Indiegogo).

Three additional characteristics of reward-based crowdfunding can exacerbate these information problems. First, backers in reward-based crowdfunding are also future consumers. While this can make project selection easier it can also increase the likelihood of irrational herding behaviors if backers want to follow the latest fashions and fads that only bring utility because of their novelty, rather than from objective quality (Becker & Stigler, 1977). Second, early investors may try to manipulate later backers into herding making large contributions that they later can withdraw. Meoli and Vismara (2021) show that in equity crowdfunding there are frequent investment withdrawals (10.2%) and in Kickstarter this is much easier because, while a project is still live, backers can cancel their pledge anytime easily by clicking the "Cancel Pledge" link. Third, love money could appear at the beginning of the campaign, implying creators would receive funding at early stages from family and friends, rather than from informed backers (Abrams, 2017). This problem may even be compounded by the easiness of posterior withdrawals.

All of these characteristics point in the direction of uninformed and even irrational herding by reward-based crowdfunding backers (Bogost, 2012). However, the empirical evidence shows that reward-based crowdfunding is very successful with more than 257 thousand projects having been successfully funded to date only in Kickstarter, with a small percentage of funded projects failing to deliver (Mollick, 2015) and total global funding amounts very similar to those of equity crowdfunding.[3]Moreover, a large number of businesses successfully financed through reward-based crowdfunding platforms have become, at a later stage, high-growth ventures (Greenberg & Mollick, 2017; Kuppuswamy & Mollick, 2014, 2014; Schwienbacher & Larralde, 2012). As pointed out by Vismara (2018b), "*it is particularly important for the future of these markets to demonstrate signals' validity, as once receivers have received a signal and have used it to successfully make an informed choice, they are more likely to attend to similar signals in the future*". We therefore set out to explain this paradox.

2.2 Literature review

Crowdfunding, as a form of financing suffering from important information asymmetries, is an ideal setting to investigate both the impact of quality signals and herding behavior (Vismara, 2016). Moreover, in crowdfunding, the study of signaling is facilitated because the information and quality signals provided by the creators can be directly observed by researchers (Kleinert & Vismara, 2024); and herding is easy because platforms allow backers to observe previous funding with minimum costs (Vismara, 2018a). Here we will focus on the empirical literature on reward-based crowdfunding and its differences with respect to other types of crowdfunding.

The initial studies on reward-based crowdfunding 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). This line of research has provided important insights into the information that backers pay attention to (Allison et al., 2015; Block et al., 2018; Courtney et al., 2017; Davis et al., 2017; Deichmann et al., 2021; Gafni et al., 2019; Hellman et al., 2019). Thus, there is ample evidence indicating that signaling works in reward-based crowdfunding platforms, which seems to contradict the fact that most potential backers in reward-based crowdfunding are expected to have low information processing ability.

The more recent empirical studies have focused on the dynamics of crowdfunding and, specifically, on how the behavior of early backers affects late 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. Interestingly, herding may have either positive or negative information effects since it can be the consequence of information cascades or of irrational behavior. In an information cascade, people follow the crowd because, using Bayesian inferences, they assume that the collective wisdom of previous decision-makers must be more accurate or informed than their individual judgment, i.e. the informational content of the aggregate behavior

[3] According to The 2nd Global Alternative Finance Market Benchmarking Report by Cambridge University, in year 2020, the total volume of funds raised over the world in equity crowdfunding platforms was 1520 US million compared to 1250 Us million in reward-based crowdfunding platforms.

of the individuals that have already made their decisions is stronger than their individual judgment (Banerjee, 1992; Bikhchandani et al., 1992; Nofsinger & Sias, 1999). However, irrational herding occurs when individuals make decisions based solely on the observation of what others are doing, without necessarily having any information or judgment about the underlying reasons or merits of those choices. In irrational herding individuals blindly follow the actions of others without a proper assessment of risks and fundamentals. The reason for following is usually a fear of missing out or a desire to conform to social norms and not the result of information analysis (Becker & Stigler, 1977; Croson & Shang, 2008; Simonsohn & Ariely, 2008).

The first to document a positive correlation between daily lending amounts and the previously accumulated amounts were Zhang and Liu (2012) using data on a peer-to-peer lending web site (Prosper.com). After that, most papers focused on equity crowdfunding, studying how the behavior of expert and sophisticated investors is followed by later contributors, as an indication of positive information cascades. Specifically, Vismara (2018a) shows that, in equity crowdfunding, more contributions by sophisticated investors (whose names are disclosed on the platform) increase the attractiveness of the offer among early investors, who in turn attract late investors. Also, Signori and Vismara (2016) show that, for initial equity crowdfunding, none of the companies initially backed by professional investors failed in their campaigns. Moreover, Wang et al. (2019) also find that in equity crowdfunding, information flows from business angels to the crowd. Additional evidence of herding in different equity crowdfunding platforms is presented by Hornuf and Schwienbacher (2018), Vulkan et al. (2016), Goethner et al. (2021b). Because in these platforms expert backers can be identified, the influence of early backers on late backers is generally interpreted as evidence of information cascades. This idea also extends to some reward-based crowdfunding platforms, such as Ulule, where the platform provides information on the expertise of previous backers (Petit & Wirtz, 2022).

Herding behavior is also well documented in reward-based crowdfunding but, in this case, because there are no identifiable experts, the correlation that the different authors find between late backers' pledges and earlier backers' accumulated pledges does not necessarily correspond to an information cascade argument. Hence, the observed behavior may well be the result of irrational herding. Colombo et al. (2015) are the first to test for herding in reward-based crowdfunding, showing 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 and that, although late backers do not seem to pay attention to social capital, they follow the behavior of earlier backers. Kuppuswamy and Bayus (2018a) focus on the number of added 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. Gangi and Daniele (2017) also focus on the late part of the campaign and find that a higher number of backers at the beginning of the end of the campaign (i.e. a higher number of late-early backers) increase the probability of success. Interestingly, these authors use data from Italian crowdfunding platforms that allow backers to see the names of previous backers and particularly to know if a company has backed the campaign. They report that campaigns which have been supported by a company (a "mentor"), which can be interpreted as an identifiable expert, have a higher probability of success. Finally, Chan et at. (2020) use daily campaign data to show that the total pledged amount (to date) exhibits a U-shaped relationship with the daily pledged amount.

While all of these studies show the existence of herding in reward-based crowdfunding, some additional evidence indicates irrational behavior may be at play. Specifically, 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 for the backers. These extra sources of value for the backers may easily be driven by irrational impulses. Additionally, Allison et al. (2017) find that inexperienced first-time funders are likely to be influenced by subjective cues about group identity and pay little attention to objective information.

2.3 Hypothesis development

As we have already seen, almost all the informational characteristics of reward-based crowdfunding make this market a likely candidate for poor informational outcomes. Creators signals are targeted to the general crowd, which is expected to have poor information processing ability. And, moreover, there is evidence of herding behavior, which could lead to irrational herding given the absence of identifiable experts in most reward-based crowdfunding platforms: if followers passively mimic the behavior of no-better informed early backers the final result will be an uninformed, inefficient market (Croson & Shang, 2008; Simonsohn & Ariely, 2008).[4]

On the other hand, the reward-based crowdfunding market has been very successful and there is evidence indicating that creators that offer more information and quality signals increase the chances of success of their campaigns. To explain this apparent paradox, we hypothesize that there is a self-selection mechanism among backers that makes signaling and herding behaviors work together and generate an informational cascade even in the absence of identifiable experts.

This hypothesis is consistent with the idea that the "crowd" is not a homogeneous community, even if its heterogeneity cannot be observed by the researcher (Lin & Boh 2020; Goethner et al., 2021a). This heterogeneity implies 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 information and believes other backers to be better informed. Specifically, some backers will have more private information or better judgement, either because they have previous experience in evaluating projects on the platform or because they have 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 good project incentivizes these backers to provide early funds, which make this decision an informative signal for late 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.

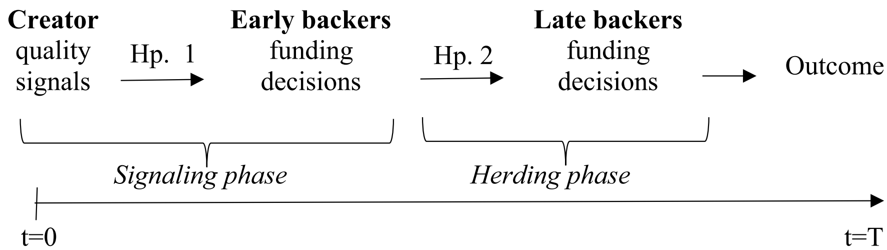
Regarding the choices of less informed backers, we expect that they choose to wait and see how many early backers are interested in and back the project. They either choose to become followers or refrain from backing the project depending on the strength of the early backers' interest in the project. 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 as we have explained in the development of Hypothesis 1. 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 project information and creator's signaling as the main determinant of the funding decisions of late backers.

The simultaneous validation of both hypotheses would be consistent with an information cascade, and allow us to disregard irrational herding. This is because two conditions are necessary for an information cascade to appear. First, the decisions made by early adopters should be based on useful information (Hypothesis 1). Second, in an information cascade the importance of individual information

[4] For example, Simonsohn and Ariely (2008) found that those bidders that behave irrationally in eBay, herd into auctions with many bids. However, these bidders do not realise that these auctions have historically obtained more bids because the starting price of the auction was very low. Hence, at very low prices, any consumer is willing to buy the product and bids for it. This produces a signal of interest that is not always as reliable as some irrational investors consider. There is also evidence of inefficient follower behavior in markets for music downloads (Salganik et al., 2006) and, even more closely related to our study, in equity crowdfunding, where amateur (but self-confident) funders are more likely to disregard objective quality signals and go with the crowd and end up investing in lower quality ventures (Stevenson et al., 2019).

Fig. 1 Theoretical model



will decrease as the choices of previous individuals accumulate (i.e. as previous information accumulates), while in irrational herding the very first individuals may have much more influence. In this sense, the validation of our second hypothesis, with results indicating that backers paying less and less attention to quality variables as information from previous backers accumulates, would be consistent with an information cascade.

The interplay between the hypotheses is summarized in Fig. 1, which shows how herding substitutes signaling as the campaign evolves but also how the outcome of the herding phase, and the generation of an informational cascade, depends critically on initial signaling in such a way that the quality of signaling complements herding decisions once we consider an intertemporal view. In this sense we can also say that signaling does not disappear but changes into rational herding. Early backers use the information directly provided by the creator and their previous knowledge, while late backers use the derivative signal and information provided by the behavior of early backers.

3 Data and methods

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 projects) to build a panel of data with daily information on the funding dynamics until February 13, 2018 when the last project campaign finalizes. 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, the success rate was 40%, while for the overall period of projects financed through Kickstarter (257,715 projects from April 28, 2009 until March 2024), the success rate is 41%. Note that we work with a sample composed of the population of all projects launched during the mentioned period, hence, reducing sample selection issues.

3.2 Variables

3.2.1 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.[5]

[5] The variable takes values between 0 and 1 and projects that never reach 10% of their funding goal are assigned a value of 1. We follow the same criterium when we define the variables *Time to move from one % funding threshold to the next*. In robustness test we do not apply this adjustment and the results remain robust.

To capture the behavior of late backers we use as dependent variables *Time to move from one % funding threshold to the next* that measure the time it takes to move from one percent of the funding goal threshold 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 on reward-based crowdfunding platforms before making their decisions. Herding behavior would be consistent with a positive relationship between the time to reach the 10% goal and the time to reach higher percentages by late backers.

3.2.2 Independent variables

We classify the information that should be relevant for backers into (i) observable project characteristics connected to their quality, (ii) entrepreneur characteristics connected to their abilities, and (iii) project description variables. Remarkably, some of the information on project and entrepreneur characteristics is also considered a quality signal in the Spence framework (i.e. costlier to produce when projects have low quality rather than high quality). This information is ex-ante because it is released and available for all potential investors when the project is launched.

First, regarding the observable project characteristics, we include in the analysis are *Funding goal*, *Duration*, *Number of webs* and *Quick updates*. We use *Quick updates* and *Number of webs* as our measures of project preparedness indicating that the entrepreneur has invested more time and effort to ensure that the project pitches conformed to standards for successful pitches as indicated by the platform. In the crowdfunding literature project preparedness is generally considered a signal of project quality (Chen et al., 2009; Huang et al., 2022; Mollick, 2014).

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. This variable is found relevant in studies like Hornuf and Schwienbacher (2016).

Duration The number of days during which the project's funding campaign will be active. Longer duration makes it more likely to reach the funding goal but may 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 Dummy taking value one if updates are provided within the first day. 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. This variable has been shown to be relevant in equity crowdfunding (Block et al., 2018).

Second, entrepreneurs experience is a particularly powerful signal in crowdfunding (Kleinert & Vismara, 2024). We proxy for this experience using the variables *Created projects*, *Backed projects* and *Creator indirect 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 of 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. This signal was found relevant by Kleinert et al. (2020).

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. This is a particularly relevant signal for Vismara (2006) and Belleflamme et al. (2014).

Our third and last set of variables includes those related to project's description.[6] These variables both capture project characteristics and are a signal of project quality as well. In particular, we define the following variables:

Project description length The natural logarithm of one plus the number of *"net"* words (after cleaning prepositions and conjunctions) used in the funding campaign description.

Sentiment of risk description Defined as (# 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.

Finally, we control for *Project category*. Specifically, Kickstarter's classifies projects into 15 industries according to their characteristics (design, games, technology, etc.). This is important because backers are likely to start their search by selecting a project category and also to control for industry dependent trends.

4 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:

$$\begin{aligned} \textit{Funding outcome}_i = {} & \alpha + \beta\ \textit{Project characteristics}_i \\ & + \gamma\ \textit{Entrepreneur characteristics}_i \\ & + \delta\ \textit{Project description}_i \\ & + \theta\ \textit{Early backers interest in the project}_i \\ & + \kappa\ \textit{Project category}_i + \varepsilon_i \end{aligned} \tag{1}$$

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.

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

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 variables ($\beta < 0$, $\gamma < 0$, $\delta < 0$) will indicate that early backers act as informed investors that rely on information and quality signals, and they are not simply following whims, fashions or fads.

To test Hypothesis 2, we will use alternative dependent variables, *Time to move from one % funding threshold to the next*, capturing the time it takes for late backers to move from one funding level to another. According to Hypothesis 2, quality signals should become less significant and early backers' interest more significant ($\theta > 0$) as we move to 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).

5 Results

5.1 Descriptive results

The summary statistics can be found in Table 1. The table shows that the mean values of the variables that measure the level of project preparedness directly connected to a project's quality (*Number of webs, Quick update*), as well as those reflecting entrepreneur's experience (*Created projects*, *Backed projects*, *Creator indirect experience*) are higher for projects that end up being successful.

These results are also confirmed in the correlation matrix, where the correlations of the previous variables with *Success* are positive and the correlations with *Time to reach 10% goal* are negative (see Table 2). Also notable is the significant negative correlation (-0.77) 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.

An interesting phenomenon can be seen in Fig. 2, which shows the evolution of funding relative to the funding goal. Successful projects and failures can be distinguished early on by looking at the daily funding they attract. The graph suggests that from the beginning of the campaign, daily pledges are on average larger for the group of projects that will succeed than for the group of projects that will fail. This difference is sustained along the campaign. The graph indicates that successful projects, not only start with higher funding levels (pledges), but also experience faster growth in funding over time due, which is consistent with the herding effect.

5.2 Early backers' response to information and signals

Table 3 shows the results of testing Hypothesis 1 using *Time to reach 10% goal* as the dependent variable to capture the behavior of early backers.[7] The results show that early backers' interest is based on information and quality signals regarding (i) project quality and preparedness (*Project funding goal*, *Duration*, *Number of webs*, and *Quick update*); (ii) creator's experience and preparedness (*Created projects, Backed projects, Creator indirect experience*), and (iii) 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.

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 variables and the information on early backers' interest information is analyzed next.

5.3 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 these variables, while the behavior of earlier backers becomes a more important determinant of late backers' interest.

The results are presented in Tables 4, 5, and 6. First, in Table 4 we only incorporate quality signals and, as expected, we find that ex-ante quality variables lose their significance as determinants of late backers' decisions (see Table 4, Model 2). Then, in the same Table 4, we 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, suggesting that the shorter the time to reach the initial 10% of the funding goal, the lower will be the time to complete the campaign (10–100%). This is clear evidence of herding behavior, which confirms Hypothesis 2. Moreover, most of the quality variables that are the basis of signaling behavior, as reflected in Tables 3 and 4 (Model 1), lose significance here, become insignificant or even change signs.

Furthermore, in Tables 5 and 6 we perform a similar analysis to that one in Table 4 (Models 2 and 3) but separating late backers into intervals. We test whether the herding behavior of late backers is relevant only at the beginning of the campaign (i.e. for the backers that provide funds from 10 to 15% and from 15 to 20%) or whether it continues to be relevant for the very late backers. Therefore, in Table 5 we only incorporate quality signals and, as expected (and similar to results in Table 4, Model 2), 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 3 (also reported in column 1 of Table 4 and 5). 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 have a clear informative signal component about project quality.

[7] In the Online Appendix (Table A1) we replicated these results using Success as dependent variable. We found significant results for the explanatory variables capturing the quality of the project and/or entrepreneur ($\beta > 0$, $\gamma > 0$, $\delta > 0$) indicate signaling effects. We also found significant results for early backers' interest in the project ($\theta > 0$) indicative of herding behavior.

Table 1 Summary statistics for all Kickstarter projects and for successful and unsuccessful projects

	Overall sample				Successful projects (Success = 1)				Unsuccessful projects (Success = 0)				Test on the Difference
	Mean	Std. dev	Min	max	mean	Std. dev	min	max	mean	Std. dev	min	max	mean
N	3500				1405				2095				Succ. Vs. Unsuccesful
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	-0.87***
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	-3.38***
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	0.51***
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	0.03***
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	1.21***
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	6.04***
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	0.89***
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	0.25***
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	0.51***
Time to reach 10% goal	0.48	0.46	0.0	1.0	0.04	0.11	0.0	0.9	0.77	0.37	0.0	1.0	-0.73***
Time to move from 10% goal to 15%	0.28	0.43	0.0	1.0	0.02	0.11	0.0	0.94	0.45	0.48	0.0	1.0	-0.42***
Time to move from 15% goal to 20%	0.29	0.43	0.0	1.0	0.03	0.06	0.0	0.75	0.46	0.48	0.0	1.0	-0.43***
Time to move from 20% goal to 40%	0.37	0.44	0.0	1.0	0.13	0.18	0.0	0.79	0.53	0.48	0.0	1.0	-0.40***
Time to move from 40% goal to 60%	0.36	0.44	0.0	1.0	0.13	0.17	0.0	1.0	0.52	0.49	0.0	1.0	-0.39***
Time to move from 60% goal to 80%	0.35	0.44	0.0	1.0	0.11	0.15	0.0	1.0	0.52	0.50	0.0	1.0	-0.41***

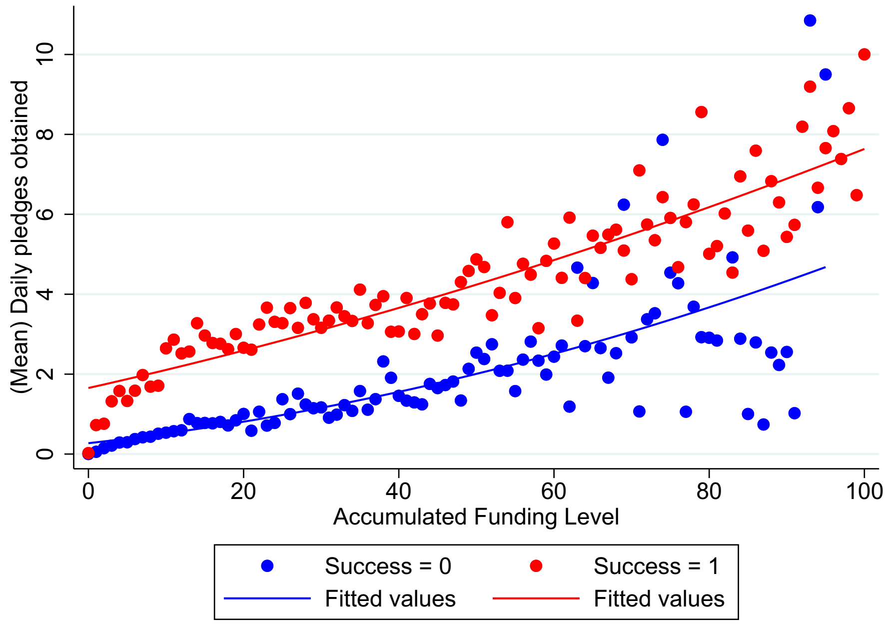
Pairwise differences of the means of successful versus unsuccessful projects are provided in the last column. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All variables are defined in the main text

Table 2 Variables correlation matrix

Matrix of correlations	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
(1) Success	1.000															
(2) Project funding goal	-0.246	1.000														
(3) Duration	-0.132	0.260	1.000													
(4) Number of webs	0.167	-0.012	-0.020	1.000												
(5) Quick update	0.090	-0.008	-0.054	0.065	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	-0.082	-0.110	0.171	0.052	0.306	0.409	1.000								
(9) Sentiment of risk description	0.071	0.073	-0.003	0.084	0.018	-0.014	0.010	-0.004	1.000							
(10) Project description length	0.190	0.126	-0.009	0.140	0.050	0.032	0.086	0.117	0.074	1.000						
(11) Time to reach 10% goal	-0.771	0.279	0.188	-0.179	-0.109	-0.189	-0.202	-0.241	-0.076	-0.195	1.000					
(12) Time to move from 10% goal to 15%	-0.486	0.141	0.059	-0.074	-0.096	-0.111	-0.110	-0.130	-0.014	-0.089	0.533	1.000				
(13) Time to move from 15% goal to 20%	-0.491	0.137	0.047	-0.070	-0.092	-0.110	-0.104	-0.124	-0.013	-0.085	0.504	0.934	1.000			
(14) Time to move from 20% goal to 40%	-0.453	0.126	0.042	-0.050	-0.097	-0.123	-0.098	-0.091	-0.003	-0.057	0.340	0.825	0.862	1.000		
(15) Time to move from 40% goal to 60%	-0.432	0.115	0.022	-0.044	-0.076	-0.089	-0.078	-0.076	0.006	-0.045	0.349	0.823	0.859	0.894	1.000	
(16) Time to move from 60% goal to 80%	-0.449	0.113	0.016	-0.048	-0.078	-0.073	-0.069	-0.069	0.003	-0.045	0.363	0.823	0.859	0.884	0.944	1.000

All variables are defined in the main text

Fig. 2 Daily pledges obtained by projects that succeed and projects that fail



To further investigate the impact of early backers on the behavior of late backers, in Table 6 we check whether the behavior of the early backers is relevant only at the beginning of the campaign (i.e. for the backers that provide funds from 10 to 15% and from 15 to 20%) or whether it continues to be relevant for the very late backers (i.e. the ones that provide funds from 60 to 80%). We observe the behavior of early backers is relevant for all subsequent investors, even as the quality signals become less and less significant for very late backers.[8]

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.

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, even in the absence of identifiable experts, seems characterized by the existence of a small number of backers that analyze signals and a large number of backers who follow them.

6 Robustness checks

We conducted several robustness checks to validate our findings. The results of most of these checks are reported in an additional Appendix which is available upon request.

6.1 Endogeneity issues

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'

[8] The changes of the estimated coefficients in the subsequent funding intervals of Table 6 are significant, as shown in the tests on differences (see online Appendix, Table A2). In order to better compare these un-nested models with different dependent variables we included statistical measures of information criteria such as AIC or BIC to assess model fit and complexity (see Table 6). These criteria provide a measure of model quality that balances goodness of fit against model complexity. Results show that both BIC and AIC values increases, indicating a worst model fit. This is consistent with early backers' behavior directly affecting to a larger extend the initial window of late backers' behavior rather than the last window of late backers' behavior.

Table 3 Early backers' response to quality signals

	(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
F	37.34	39.71	45.03
P	0	0	0
R^2	0.161	0.193	0.229

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

interest and omitted variables related to quality signals.

We think this is unlikely. Since we have the exact timing of every pledge made in every campaign, we can focus the analysis on within-campaign dynamics, controlling for all time-invariant unobserved heterogeneity across campaigns. Specifically, we show that late backers interest changes with changes in early backers' interests, when information variables are not changing and cannot be affecting the direct effect that we find between changes in these two variables (Arellano & Carrasco, 2003). Hence, we have changes of both variables of interest while the predetermined variables remain unchanged, which reduces potential endogeneity concerns.[9]

However, it is still possible that there are some quality signals that are obtained after the campaign starts (e.g. a project is reported in the news) and 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). We need an instrument that affects the time to reach 10% but it is independent of other quality information. We use a dummy that captures the existence of population aggregation or network effects for each of the product categories.[10] This dummy indicates which type of products are more interesting for the user when more people are using them as explained by Katz and Shapiro (1986, 1992). 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 (F-statistic 16.73) of the instrument on the instrumented variable (see column 1 in Table 7). Moreover, we believe the exclusion restriction is satisfied too because network effects should impact late backers specifically through their observation of the interest of early backer: products with population aggregation effects become more interesting for the late backers precisely because there is already a large number of early backers interested in them. 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 7).

[9] Notice that, for example, in Cumming, Meoli and Vismara (2019) the main interest is in knowing whether the setting up of a campaign with dual-class shares (i.e. where there is a maximum investment threshold beyond which no voting rights are granted for the additional shares to be issued) affects success, and the dual-class variable is also predetermined and does not change throughout the campaign.

[10] To construct the dummy, we asked 12 PhD students to go through the list of product categories and to report independently up to five categories which they thought had the largest potential aggregation effects. Based on their responses we created a dummy that takes the value 1 for the eight most voted categories: Art, Comics, Design, Fashion, Film/Video, Games, Photography and Theater.

Table 4 Early and late backers' response to quality signals and early backers' behavior

	(1) Time to reach 10% goal	(2) Time to move from 10% funding threshold to 100%	(3)
Time to reach 10% goal			-0.093***
			(0.018)
Project funding goal	0.072***	0.012**	0.018***
	(0.005)	(0.005)	(0.005)
Duration	0.003***	-0.002***	-0.001**
	(0.001)	(0.001)	(0.001)
Number of webs	-0.030***	0.006	0.003
	(0.005)	(0.005)	(0.005)
Quick Update	-0.150***	-0.176***	-0.191***
	(0.036)	(0.037)	(0.037)
Created projects	-0.007***	-0.009***	-0.010***
	(0.002)	(0.002)	(0.002)
Backed projects	-0.002***	-0.001	-0.001*
	(0.001)	(0.001)	(0.001)
Creator indirect experience	-0.029***	0.001	-0.001
	(0.004)	(0.004)	(0.004)
Sentiment of risk description	-0.019***	0.009**	0.007*
	(0.004)	(0.004)	(0.004)
Project description length	-0.065***	0.010*	0.004
	(0.006)	(0.006)	(0.006)
Category Controls	Yes	Yes	Yes
Constant	0.266***	0.435***	0.460***
	(0.049)	(0.051)	(0.051)
Observations	3,479	3,479	3,479
F	45.03	6.983	7.885
p	0	0	0
R^2	0.231	0.044	0.052

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1 Category control for all models. OLS regression. Column 1 tests the decisions of early backers (same as last column of Table 3). Column 2 tests all late backers' response to quality signals. Column 3 tests all late backers' response to both quality signals and to early backers' behavior. All variables are defined in the main text

Secondly, we have tried a second type of instrument following Cumming et al. (2019). In Table 8 we instrument *Time to reach 10% goal* using a mimicking variable computed as the mean value within project category, considering all projects in the category that were active during the week before the launching of the focal project.[11] After estimating a 2SLS model using this instrument, results are consistent to those of Table 6.

Thirdly, 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 information and quality signals (which may drive both

[11] To avoid multicollinearity problems and ensure the validity of the instruments, we have not included the category controls in the first-stage estimation to build the instrument.

Table 5 Late backers' response to quality signals

	Time to move from one % funding threshold 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
F	45.03	10.86	9.77	8.28	5.80	5.36
P	0	0	0	0	0	0
R^2	0.229	0.057	0.054	0.043	0.031	0.029

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

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. The results (Table A3 in the Online Appendix) 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.

Fourthly, 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 (ex-ante information and category fixed effects). 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 information and signals, this *Treated* variable should capture a pure herding effect.

Table 6 Late backers' response to quality signals and early backers' behavior

	(1)	(2)	(3)	(4)	(5)
	Relative time to move from one % funding threshold 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
F	60.13	51.41	21.98	22.19	24.21
P	0	0	0	0	0
R^2	0.253	0.226	0.101	0.108	0.120
AIC	2759.7	3016.4	3646.5	3690.1	3735.2
BIC	2913.6	3170.2	3800.4	3844.0	3889.1

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

The re-estimation of the main results using the new matched sample shows consistent results (Table A4 and Figure A1 in the Online Appendix).

Finally, as an alternative control for potential endogeneity, we also repeat the analysis for larger projects (funding goal above $5,000), smaller projects (funding goal below $5,000), projects with shorter campaigns (lower than 31 days) and longer campaigns (higher than 30 days). With these tests, we check whether herding or signaling is driven by entrepreneurs' decisions before the opening of the campaign. Our results for these subsamples are similar to those found for the whole sample (Table A5 in the Online Appendix).

6.2 Additional tests for information cascades versus irrational herding

We have argued that the simultaneous validation of hypothesis 1 and 2 indicates that the behavior we observe in reward-based crowdfunding is consistent with an information cascade rather than irrational herding. An additional difference between an information cascade and irrational herding is that

Table 7 Instrumental Variable analysis of Late backers' response to early backers' behavior (instrumented by dummy on Aggregation Effect)

	(1)	(2)	(3)	(4)	(5)	(6)
	First stage	Second stage				
	Time to reach 10% goal	Time to move from one % funding threshold to the next				
		(10-15%)	(15-20%)	(20-40%)	(40-60%)	(60-80%)
Dummy Aggregation Effect	-0.060***					
(*instrument*)	(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
F test	92.96					
Wald chi2		171.5	167.3	126.5	91.41	90.52
R^2	0.211	0.102	0.134	0.016	0.031	0.064

Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Categories not controlled for any model due to collinearity with instrument. OLS regression similar to Table 5 but instrumenting *Time to reach 10% goal* with instrumental variable *Dummy Aggregation Effect*. 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

the strength of information cascades should be correlated with the informational quality of the decisions made by early adopters, while the strength of irrational herding is not expected to be correlated to informational quality.

To investigate this difference, we measure the dispersion of initial contributions (i.e. the volatility of the mean daily funding obtained per backer, or average daily pledge provided by the backer) during the initial phase of the campaign.[12] More dispersion at the beginning of the funding campaign should imply less agreement

[12] In this case, we use the shortest possible period, the first 3 days of the campaign, to conserve more observations.

Table 8 Instrumental Variable analysis (Mimicking variable for "Time to reach 10% goal") of early and late backers' response to quality signals and early backers' behavior.

	(1)	(2)	(3)	(4)	(5)	(6)
	First stage	Second stage				
	Time to reach 10% goal	Time to move from one % funding threshold to the next				
		(10-15%)	(15-20%)	(20-40%)	(40-60%)	(60-80%)
Mv. Category Mean Time to reach 10% goal (*instrument*)	0.231***					
	(0.052)					
Time to reach 10% goal		0.772***	0.832***	0.730***	0.626***	0.543**
(*instrumented*)		(0.208)	(0.221)	(0.246)	(0.236)	(0.231)
Project funding goal	0.070***	-0.022	-0.026	-0.022	-0.016	-0.010
	(0.004)	(0.016)	(0.017)	(0.018)	(0.018)	(0.017)
Duration	0.003***	-0.005***	-0.006***	-0.005***	-0.006***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number of webs	-0.031***	0.015*	0.017**	0.017*	0.012	0.007
	(0.005)	(0.008)	(0.009)	(0.010)	(0.009)	(0.009)
Quick Update	-0.168***	-0.038	-0.025	-0.063	-0.040	-0.064
	(0.037)	(0.049)	(0.052)	(0.058)	(0.056)	(0.055)
Created projects	-0.007***	-0.001	-0.002	-0.006*	-0.003	-0.001
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Backed projects	-0.002***	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	-0.028***	0.007	0.009	0.013	0.010	0.009
	(0.004)	(0.007)	(0.008)	(0.009)	(0.008)	(0.008)
Sentiment of risk description	-0.019***	0.013**	0.014**	0.014**	0.014**	0.013**
	(0.004)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)
Project description length	-0.066***	0.028*	0.032**	0.034**	0.029*	0.023
	(0.006)	(0.015)	(0.016)	(0.017)	(0.017)	(0.016)
Category Controls	No	No	No	No	No	No
Constant	0.169***	0.036	0.048	0.129	0.153*	0.166**
	(0.052)	(0.071)	(0.075)	(0.084)	(0.080)	(0.079)
Observations	3,282	3,282	3,282	3,282	3,282	3,282
F	90.73					
chi2		232.9	207.6	143.1	122.5	120.4
R^2	0.217	0.199	0.123	0.045	0.039	0.100

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Categories not controlled for any model due to collinearity with instrument. OLS regression similar to Table 5 but instrumenting *Time to reach 10% goal* with instrumental variable *Mv. Category Mean Time to reach 10% goal* that stands for the category mean value of the variable *Time to reach 10% goal*, by taking the previous 7 days to the start of the project funding campaign. 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 19.36, significantly high, and relatively high R-squared. First-stage in column 1. Results remain robust when using the instrument (Second-stage in columns 2-6). All other variables are defined in the main text

among early backers and less information aggregation to sustain an information cascade. However, for a given amount of funds raised, it should not have an impact on irrational herding. The results in Table 9 show that the interaction between the dispersion of the average daily pledges of early backers and the herding of late backers has a significant negative coefficient, reducing late backers' herding, which is additional evidence in favor of the existence of an information cascade.

Table 9 Information cascades versus irrational herding

	(1)	(2)	(3)	(4)	(5)
	Time to move from one % funding threshold to the next				
	(10-15%)	(15-20%)	(20-40%)	(40-60%)	(60-80%)
Interaction:					
LogVolPledges(3 days) x Time to reach 10% goal	-0.009	-0.012	-0.028***	-0.025***	-0.021**
	(0.008)	(0.009)	(0.010)	(0.010)	(0.010)
LogVolPledges(3 days)	0.016***	0.018***	0.036***	0.031***	0.028***
	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)
Time to reach 10% goal	0.541***	0.526***	0.410***	0.429***	0.444***
	(0.028)	(0.029)	(0.032)	(0.032)	(0.032)
Project funding goal	-0.003	-0.001	0.007	0.003	0.002
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
Duration	-0.001**	-0.002***	-0.001**	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number of webs	0.005	0.005	0.003	0.004	0.003
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
Quick Update	-0.081***	-0.079**	-0.117***	-0.082**	-0.090**
	(0.031)	(0.032)	(0.035)	(0.036)	(0.036)
Created projects	-0.001	-0.002	-0.006**	-0.002	-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)	(0.001)	(0.001)
Creator indirect experience	0.000	0.000	0.003	0.004	0.005
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Sentiment of risk description	0.006*	0.006	0.005	0.008*	0.007*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Project description length	0.005	0.005	0.001	0.006	0.008
	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
Category Controls	Yes	Yes	Yes	Yes	Yes
Constant	-0.014	0.006	0.049	0.042	0.037
	(0.047)	(0.049)	(0.053)	(0.054)	(0.054)
Observations	3,462	3,462	3,462	3,462	3,462
F	54.83	47.04	20.95	20.93	22.63
P	0	0	0	0	0
R^2	0.293	0.263	0.137	0.137	0.146

Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$ Category control for all models. OLS regression. Variable *LogVolPledges(3 days)* stands for the logarithmic adjustment of the volatility of the mean daily pledges obtained during the first 3 days of the funding campaign, hence measuring the dispersion of the pledges at the beginning of the funding campaign for each project. All other variables are defined in the main text

6.3 Alternative definitions of backers' behavior, success and estimation methods

We have used alternative variables and methods to measure success and investors funding behavior.

Specifically, we have considered an alternative estimation model using duration thresholds during the life of the campaign. For this alternative model, we measure the interest of early and late backers as the percentage of funds achieved in the first 10% of campaign duration, avoiding distortions related to projects that never reach the different funding thresholds. Our main results remain robust (Table A6 in Online Appendix).

Moreover, to control for measurement errors, we have used alternative definitions of 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 (panels A and B, Table A7 in the Online Appendix). We found consistent results. Moreover, we ran Linear Probability Regressions to confirm the results from the Logistic Regressions and we found similar results (panel C, Table A7 in the Online Appendix).

6.4 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 (Table A8 in the Online Appendix).

6.5 Identification of early and late backers

We also repeated the entire analysis changing the threshold we used to classify investors as early or 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 (panels A to C in Table A9 in the Online Appendix). We also consider different jumps in the intervals measuring late backers' funding behavior, specifically we used 10% and 20% jumps (panels A and B in Table A10 in the Online Appendix), and the results remain robust.

We also re-run our estimations using as dependent variable the number of backers that provided funds to reach % target milestones. This is similar to the approach in Vismara (2018a), but using as threshold to separate early and late backers the relative time instead of a specific number of days (i.e., 5 days) because in reward-based crowdfunding there is a lot of heterogeneity in campaign durations. Results show that, the lower the time to reach 10% of the funding goal, the larger the number of investors that provided funds for the different percentages. This indicates the existence of herding by an increasing number of small backers (panels A and B in Table A11 in the Online Appendix).

6.6 Controls for truncation of variables and potential selection bias

Finally, we control for the potential distortions introduced by projects that do not reach the different funding thresholds. In our main estimations, to keep as many observations as possible, we arbitrarily assigned a value of 1 to the *Time to reach 10% goal* for projects that do not reach this goal or other subsequent goals. This is done for all variables that are computed as the ratio between the time it takes to reach a funding threshold and the overall duration of the campaign. As we explained in subSect. 5.2, in an alternative estimation model, we measure the interest of backers using duration threshold variables instead (Table A6 in Appendix). Here we discuss other alternative ways to tackle this problem.

First, we introduced in the main estimations a dummy capturing these projects and the results were not affected, indicating that these unsuccessful projects are not the drivers of our results (Table A13 in the Appendix). Secondly, we used a Tobit model to control for the truncation of these variables (Table A13 in the Appendix) and obtained similar results. And, in third place, we used a two-step Heckman model measuring the probability of a project never reaching the 10% funding goal (Table A14 in the Appendix).[13]

7 Discussion, future research and conclusions

Crowdfunding is a novel source of funding for startups. Just like other forms of early funding, it suffers from high levels of information asymmetry. However, in crowdfunding, these problems are compounded

[13] This estimation allows computing the inverse Mills ratio that we will include in the second step were we then exclude those projects that never reach the funding thresholds.

because unknown creators with limited experience (e.g. Davis et al., 2017; Lin & Boh, 2020) are trying to raise funds from a "crowd" of dispersed backers. In this paper, we focus on the type of crowdfunding where informational problems are potentially larger and there are no identifiable expert backers: reward-based crowdfunding. We provide an explanation for the apparent paradox of the remarkable success of this funding market despite its severe information problems.

Specifically, our database, with detailed daily funding information for 3,923 Kickstarter projects, allows us to investigate how the signaling and herding that previous authors have documented in crowdfunding markets, can complement and substitute each other in interesting ways as the campaign evolves to produce a positive information cascade even without identifiable experts. We show that (i) early backers (those providing the first 10% of the funds) respond to information and signaling and (ii) that late backers seem to disregard this information but notice and follow the behavior of early backers. In particular, when early backers provide funds fast, the speed of funding of late backers increases irrespectively of the quality information of the project. This herding becomes more pronounced as more information from earlier backers accumulates and when there is more agreement between the group of early backers. All these results seem robust to multiple robustness tests and indicate that this market is compatible with the dynamics of information cascades rather than irrational herding.

7.1 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, generating information cascades, 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, even if late backers cannot identify early backers and cannot ascertain their expertise, they will rely on their information because 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 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. This result can be expected to reinforce information cascades in all crowdfunding markets, but is more important for reward-based crowdfunding because it works without identifiable experts.

7.2 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. Creators should not ignore the presence of well-informed backers in this market, even if they are not easily identifiable. These backers may have more information because they have tried and/or previously backed similar products or services in the same category and they are more likely to be actively looking for projects and to base their 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 (Bradford, 2012; Griffin, 2013; Hazen, 2012) make a case for a 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 (Gutierrez & 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 show that, even without identifiable experts, backers seem to self-select in a way that generates a positive information cascade that protects late backers. Thus, even though a high 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". Moreover, we argue that herding behavior is crucial in setting in motion the informational cascade. 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.

Based on these findings, 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 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.

7.3 Limitations and future research

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

A first 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 & Mollick, 2017; Hornuf et al., 2018; Roma et al., 2017, 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.

Another research avenue would explore the strategic actions of the competitors, who may 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 negative consequences, not only for the success of a project but also for future projects launched by a given creator.

A third line of research would investigate further the possibility that manipulation and love money are influencing herding in reward-based crowdfunding. Meoli and Vismara (2021) present evidence of manipulation in equity crowdfunding. Our results are not consistent with manipulation. Notice manipulation would imply that low quality projects can get funded either because creators themselves (or love money attached to them or even platforms) make important contributions at the beginning of the campaign and withdraw them later, having set in motion herding by late backers. What we observe is that early backers' behavior is more (not less) influenced by information about quality. To investigate this point further we have checked (i) whether the mean contribution per backer increases or decreases after the first days of the campaign and (ii) whether a significant number of projects see their total funding decrease on any given day. The results do not show evidence of manipulation or love money. In particular, the average number of days where total pledges decrease is very small (1,62% of total days of campaign duration). Clearly, more detailed investigation on the extent or potential for manipulation and love money as well as distance between the entrepreneur and investor (Agrawal et al., 2011), seems necessary, but it will require additional information on the characteristics of backers that are more likely to withdraw their pledges.

Finally, several authors have studied whether the gender and other characteristics of the creator influence the behavior of backers and the probabilities of success in crowdfunding (Rossi et al., 2021; Cumming et al., 2021; Letwin et al., 2023 and Wang et al., 2023). Herding behavior may be related to risk aversion and to self-confidence, and these are characteristics that may be linked to gender (Eckel & Grossman, 2008; Sarin & Wieland, 2012). Thus, it can be interesting to study whether women contributors to crowdfunding campaigns are more prone to herd. This could be especially useful in reward-based crowdfunding to promote products or services that are specially targeted towards female or male customers.

8 Conclusions

Understanding crowdfunding is important for facilitating the funding of start-up projects. In this paper we have addressed the question of the connection between signaling and herding behavior in the financing of projects using reward-based crowdfunding.

Our theoretical contribution is to show that informational cascades can originate in a market without identifiable experts, generating a complex interplay between signaling and herding behavior that improves funding efficiency. Herding can work efficiently in these markets with minimum regulation because the most informed potential contributors are the ones with the highest incentives to be early backers and pledge depending on quality information and signals, initiating an information cascade, so that less informed participants, understanding this behavior, are better off herding.

Our contribution to the practice of crowdfunding is to highlight the relevance for entrepreneurs of setting up products and services and funding campaigns in a way that attracts the most informed potential contributors, rather than a median contributor, as a way to originate a positive information cascade. Importantly, using this approach not only facilitates the funding of start-up projects but also its escalation, growth and transformation into established firms. We know that a successful crowdfunding campaign sends a signal of project quality that may be instrumental to get new funding in later stages from institutional financiers like business angels, venture capitalists or banks. Hence, this sequential approach to financing, which may be initiated in reward-based crowdfunding before moving to equity crowdfunding and then to traditional sources of funding, can be efficient from a social point of view because it requires minimum intervention in terms of regulators' monitoring.

To conclude, a key message from our paper is the importance of promoting crowdfunding platforms, in its different forms, as a key element to ensure an efficient financing of new entrepreneurial ventures and a source of economic growth.

Acknowledgements For their helpful comments, we are grateful to the editor, two anonymous reviewers and to Fernando Zapatero, Jesper Rudiger, Anna Toldra, Eduardo Melero and Monica Lopez. We also thank participants at the Boston University Questrom School of Business Brownbag seminar series, Academy of Management (AOM) 80th Annual Meeting, 28th Finance Forum (AEFIN) PhD Mentoring Day, 8th PhD Workshop on Industrial and Public Economics (WIPE). 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) and the EU's European Social Fund as well as the Programa Excelencia para el Profesorado Universitario, convenio con Universidad Carlos III de Madrid, V Plan Regional de Investigación Científica e Innovación Tecnológica, EPUC3M12, are gratefully acknowledged. Financial support provided by the program of Financial Innovation financed by The Koch Foundation. Funding for APC: Universidad Carlos III de Madrid (Agreement CRUE-Madroño 2024). Usual disclaimer applies.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Data availability All data used in this research is publicly available. Replication files can be obtained from the authors.

References

Abrams, E. (2017). Securities crowdfunding: More than family, friends, and fools? *SSRN Working paper, 2902217*. https://ssrn.com/abstract=2902217

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.

Agrawal, A. K., Catalini, C., & Goldfarb, A. (2011). The geography of crowdfunding. *NBER Working paper, 16820.* http://www.nber.org/papers/w16820. Accessed 20 Mar 2024

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

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.

Arellano, M., & Carrasco, R. (2003). Binary choice panel data models with predetermined variables. *Journal of Econometrics*, *115*(1), 125–157.

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

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

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.

Bogost, I. (2012). Kickstarter: Crowdfunding platform or reality show? *Fast Company*, July 18, 2012. https://www.fastcompany.com/1843007/kickstarter-crowdfunding-platform-or-reality-show. Accessed 30 Apr 2024

Bradford, C. S. (2012). Crowdfunding and the Federal Securities Laws. *SSRN Working paper 1916184*. https://ssrn.com/abstract=1916184

Cambridge University Centre for Alternative Finance. (2020). The 2nd Global alternative finance market benchmarking report. Retrieved from: https://www.jbs.cam.ac.uk/faculty-research/centres/alternative-finance/publications/the-global-alternative-finance-market-benchmarking-report/

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), 105934. https://www.sciencedirect.com/science/article/abs/pii/S0883902617303567

Chen, X. P., Yao, X., & Kotha, S. (2009). Entrepreneur passion and preparedness in business plan presentations: A persuasion analysis of venture capitalist'' 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.

Cumming, D., Meoli, M., & Vismara, S. (2019). Investors' choices between cash and voting rights: Evidence from dual-class equity crowdfunding. *Research Policy, 48*(8), 103740.

Cumming, D., Meoli, M., & Vismara, S. (2021). Does equity crowdfunding democratize entrepreneurial finance? *Small Business Economics, 56*(2), 533–552.

Davis, B. C., Hmieleski, K. M., Webb, J. W., & Coombs, J. E. (2017). Funder'' positive affective reactions to entrepreneur'' 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.

Eckel, C. C. & Grossman, P. J. (2008). Men, women and risk aversion: Experimental evidence. In C. Plott, & V. Smith (Eds.), *Handbook of Experimental Economics Results*, Vol. 1, Ch.113, 1061–1073. Elsevier. https://ssrn.com/abstract=1883693

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.

Goethner, M., Luettig, S., & Regner, T. (2021a). Crowdinvesting in entrepreneurial projects: Disentangling patterns of investor behavior. *Small Business Economics, 57*, 905–926.

Goethner, M., Hornuf, L., & Regner, T. (2021b). Protecting investors in equity crowdfunding: An empirical analysis of the small investor protection act. *Technological Forecasting and Social Change, 162*, 120352. https://doi.org/10.1016/j.techfore.2020.120352

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. *NBER Working Paper 26275*. https://www.nber.org/papers/w26275

Hornuf, L., & Schwienbacher, A. (2016). Crowdinvesting—Angel Investing for the Masses? In Handbook of Research on Venture Capital: Volume 3. In H. Landström & C. Mason (Eds.), *Business Angels* (pp. 381–397). Edward Elgar.

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

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), 331–335.

Huang, S., Pickernell, D., Battisti, M., & 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.

Katz, M., & Shapiro, C. (1986). Technology Adoption in the Presence of Network Externalities. *Journal of Political Economy, 94*(4), 822–841.

Katz, M. L., & Shapiro, C. (1992). Product Introduction with Network Externalities. *The Journal of Industrial Economics, 40*(1), 55–83.

Kleinert, S., Volkmann, C., & Grünhagen, M. (2020). Third-party signals in equity crowdfunding: The role of prior financing. *Small Business Economics, 54*, 341–365.

Kleinert, S., Bafera, J., Urbig, D., & Volkmann, C. K. (2022). Access denied: How equity crowdfunding platforms use quality signals to select new ventures. *Entrepreneurship Theory and Practice, 46*(6), 1626–1657.

Kleinert, S., & Vismara, S. (2024). Quality signals in equity crowdfunding. In D. Cumming & B. Hammer (Eds.), *The Palgrave Encyclopedia of Private Equity*. Palgrave Macmillan. https://doi.org/10.1007/978-3-030-38738-9_184-1

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

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

Kuppuswamy, V., & Mollick, E. (2014). After the campaign: Outcomes of crowdfunding. *SSRN Working paper 2376997*. http://ssrn.com/abstract=2376997

Letwin, C., Ciuchta, M.P., & Johnson, M. (2023). Passion and attractiveness on display: An examination of gender bias in crowdfunding. *Small Business Economics*. https://doi.org/10.1007/s11187-023-00808-9

Lin, Y., & Boh, W. F. (2020). How Different are Crowdfunders? Examining Archetypes of Crowdfunders. *Journal of the Association for Information Science and Technology, 71*, 1357–1370.

Meoli, M., & Vismara, S. (2021). Information manipulation in equity crowdfunding markets. *Journal of Corporate Finance*, *67*, 101866. https://www.sciencedirect.com/science/article/abs/pii/S0929119920303102

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

Mollick, E. (2015). Delivery Rates on Kickstarter. *SSRN Working paper, 2699251*. https://ssrn.com/abstract=2699251

Nofsinger, J. R., & Sias, R. W. (1999). Herding and Feedback Trading by Institutional and Individual Investors. *The Journal of Finance, 54*, 2263–2295.

Petit, A., & Wirtz, P. (2022). Experts in the crowd and their influence on herding in reward-based crowdfunding of cultural projects. *Small Business Economics, 58*, 419–449.

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),104267. https://doi.org/10.1016/j.respol.2017.07.012.

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).

Rossi, A., Vanacker, T., & Vismara, S. (2021). Equity Crowdfunding: New Evidence from US and UK Markets. *Review of Corporate Finance, 1*(3–4), 407–453.

Salganik, M., Dodds, P., & Watts, D. (2006). Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science, 311*, 854–856.

Sarin, R. K. and Wieland, A. M. (2012). Gender differences in risk aversion: A theory of when and why. *SSRN Working paper 2123567*. https://ssrn.com/abstract=2123567

Schelling, T. (1978). *Micromotives and Macrobehavior*. W. W. Norton & Company, Inc.

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

Signori, A., & Vismara, S. 2016. Returns on investments in equity crowdfunding. *SSRN Working paper 2765488*. https://ssrn.com/abstract=2765488

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

Stevenson, R. M., Ciuchta, M. P., Letwin, C., Dinger, J. M., & Vancouver, J. B. (2019). Out of control or right on the money? Funder self-efficacy and crowd bias in equity crowdfunding. *Journal of Business Venturing, 34*(2), 348–367.

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

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

Vismara, S. (2018b). Signaling to overcome inefficiencies in crowdfunding markets. In D. Cumming, & L. Hornuf (Eds.), *The Economics of Crowdfunding, 2*, 29–56. Palgrave Macmillan. https://doi.org/10.1007/978-3-319-66119-3_3

Vulkan, N., Åstebro, A., & Fernandez-Sierra, M. (2016). Equity crowdfunding: A new phenomena. *Journal of Business Venturing Insights, 5*, 37–49.

Wang, Y., Li, Y., & Wu, J. (2023). Does digitalization sufficiently empower female entrepreneurs? Evidence from their online gender identities and crowdfunding performance. *Small Business Economics, 61*, 325–348.

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), 103727. https://doi.org/10.1016/j.respol.2019.01.003

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.

Zhang, J., & Liu, P. (2012). Rational Herding in Microloan Markets. *Management Science, 58*(5), 892–912.

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.