The Role of Information Signals in Determining Crowdfunding Outcomes

Jin-Hyuk Kim* Peter Newberry† Calvin Qiu‡

August 14, 2018

Abstract

Crowdfunding platforms prominently display the current level of funding and the number of backers, which can have a nontrivial impact on the efficiency of the platform-level outcomes. To assess the impacts of these signals, we estimate a discrete choice model of backer behavior on a large crowdfunding platform and show that the funding status positively affects the backer's utility before the funding goal is met while the number of backers-per-day has a positive effect both before and after the goal is met. We then use the model estimates to simulate project outcomes under different information structures. The results suggest that, while both information signals have nontrivial impacts, the funding status has a disproportionately larger impact on the efficiency of platform-level outcomes.

Keywords: crowdfunding platform; value of information; social learning

JEL Codes: L15; L81, M13

*University of Colorado at Boulder, Department of Economics, Economics Building, Boulder, CO 80309. Tel: 303 492 3827. Email: jinhyuk.kim@colorado.edu.

†Pennsylvania State University, Department of Economics, 403 Kern Building, University Park, PA 16801. Tel: 814 865 4921. Email: pwnewberry@psu.edu.

‡Berkeley Research Group, One International Finance Centre, 1 Harbour View Street, Hong Kong, China. Tel: 852 2297 2270. Email: calvin.qiu@cantab.net.

1 Introduction

In the past decade, online crowdfunding has emerged as a popular means by which entrepreneurs can finance early rounds of their projects. In the so-called 'reward-based' crowdfunding model, which we focus on in this paper, entrepreneurs essentially 'pre-sell' their products to be made as a reward for donating to their projects. Hence, the mechanism provides an opportunity for entrepreneurs to learn about the demand before they invest in a project (e.g., Strausz, 2017; da Cruz, 2018).

When visitors to a crowdfunding platform review fundraising projects, it is well known in the literature that the backer's decision is affected by the multiple information signals displayed on the project pages, namely, the percent raised of the funding goal (or 'funding status') and the number of backers (or 'backer volume'). For instance, Kuppswamay and Bayus (2018) show that the daily pledges to a project increase as a project gets closer to the funding goal and also the number of backers increases.

This paper builds upon the previous findings by quantifying the impacts of the signals on the projects' outcomes, which facilitates a platform-level comparison of the importance of each signal. That is, by simulating the dynamic project outcomes and separately considering the value of these signals, we determine the relative impact of each signal on overall outcomes. That is, while the prior literature demonstrates that both signals matter for individual backer decisions, we focus on the magnitude of the impacts of each signal on project outcomes. This can inform us as to the importance of these signals in determining the efficiency of crowdfunding market.

It is important to point out that we do not intend our exercises to predict the 'counterfactual' outcomes, but rather, our goal is to correctly quantify and compare the relative effects of these signals on aggregate outcomes based on a microfounded model. We aim to investigate the dynamic effects of the funding status and the backer volume information that may result in high-quality projects reaching their funding goals. Specifically, a backer's decision reflects the decisions of previous backers as well as other important factors such as project deadlines and potential competition between projects (Belleflamme et al., 2015). Hence, our results can complement those found in the literature.[1]

To achieve this goal, we specificy and estimate a discrete choice model of backer behavior, wherein the utility of donating to a project depends on the project's funding status and the backer volume. We use data from one of the the largest crowdfunding platforms, Kickstarter, where we were able to collect daily project statuses over the course of a year. We also collect data on the number of clicks on each project's 'short url(s)', which serves as a proxy for the project creator's unobserved effort. Short url(s) are often used when sharing content on social media, such as Twitter, for brevity. Short urls keep track of the number of visitors who use them to arrive at a target webpage, and are therefore closely associated with the number of visitors who arrive to a project page from social media postings.

Our estimation of the model suggests that backers' utility from pledging to a project increases as the funding status increases towards the goal and then it drops off significantly once the campaign reaches the goal (a 'threshold effect'). For instance, backers may place value on donating to a project that has yet to reach its goal. On the other hand, backer volume information increases the utility from donating both before and after the goal is reached as backers may see a high backer volume as a signal of quality (an 'observational learning effect'). The positive effects of these two signals have been documented in the previous literature. However, the difference is that to our knowledge the literature has not clearly attributed the source of the threshold effect to either information signal. For instance, Kuppswamay and Bayus (2018) demonstrate a threshold effect by an indicator variable for

[1] Kawai et al. (2014) simulates an online credit market to quantify the extent to which the market suffers from adverse selection and the extent to which signaling can affect market outcomes. Li and Duan (2016) simulates crowdfunding project outcomes to investigate how project creators can use different 'seeding' strategies to increase the likelihood of success.

whether or not the funding goal has been reached while we interact it with the signals.

We then use the estimates of our choice model to examine the goodness-of-fit and simulate project outcomes under different information regimes—when funding status and backer volume are perfectly observable, and when one (or both) of the signals is unknown. We simulate the backer's dynamic pledge decisions while drawing the individual pledge amount from a fitted distribution of the backers' choice given the menu of project rewards and allowing the backers to form an expectation of the unknown signal(s) based on other project-level observables. Thus, in our simulation, backers will make a pledge decision based on this 'expected value' of the signal rather than simply ignore it.

First, we find that suppressing both signals results in fewer projects reaching their funding goal. This may be preferred if the projects that switch to being unsuccessful are the low quality projects. However, we also find that the average quality of the successful projects decreases by nearly 50%, where quality is measured by an estimated project fixed effect. Therefore, the combination of the observational learning effect of both signals and the threshold effect of funding status results in more projects being funded and a higher average quality of funded projects. That is, providing both signals leads to more efficient platform-level outcomes.

Allowing backers to observe the number of backers, but not funding status, results in a further reduction in the number of successful campaigns by 4% and decreases the average quality by another 8%. This suggests that providing only the number of backers actually leads to *worse* outcomes compared to the scenario in which no information is revealed. The reason for this is that the backer volume signal leads to pledges being more concentrated on a smaller number of projects. That is, the projects that receive pledges on day one will receive more pledges on day two, and so on. This might be preferred if it is only the high quality projects which receive the 'snowball' pledges; however, the decrease in the average quality suggests that there are projects that are of high quality but fail to get funded.

Adding funding status, but not backer volume, results in more projects being funded (a 15% increase), because pledges are focused on projects that need more donations to reach their funding goal. The average quality of successful projects also increases by nearly 66%. Therefore, the interaction of the threshold and the observational learning effects leads to pledges being focused on high quality projects and projects that need donations in order to reach their goals at the same time. Thus, our results imply that the efficiency of the crowdfunding market is driven primarily by the funding status, rather than the backer volume. In other words, backer volume improves efficiency only when it is accompanied by funding status because the two signals reinforce positively for the projects that have yet to reach their funding goals.

Our paper is related to a few strands of literatures. First, there is a growing literature on the economics of crowdfunding (see, e.g., Cumming and Hornuf, 2018). For instance, Ellman and Hurkens (2015), Chemla and Tinn (2017), and Strausz (2017) theoretically examine the efficiency of the crowdfunding mechanisms with a focus on its ability to be informative about market demand. In particular, Strausz suggests that the perfect information on the number of backers may yield an inefficient outcome, which is consistent with our finding. Cumming et al. (2015) and Marwell (2015) compare 'all-or-nothing' or 'keep-it-all' fundraising mechanisms. While this cross-platform choice is potentially important, we are unable to address this issue with our data.

Second, there is a burgeoning empirical literature on crowdfunding. For instance, Mollick (2014) and Vulkan et al. (2016) examine cross-sectional determinants of successful campaigns; and Li and Duan (2016) and Kuppuswamy and Bayus (2018) study project-day-level determinants of the number of backers. Agrawal et al. (2015) and Lin and Viswanathan (2016) highlight the role of geography and home bias; Inbar and Barzilay (2014) and Freedman and Jin (2017) examine social network/community effects; and Mollick and Nanda (2015) and Iyer et al. (2016) compare the ability of crowds to screen projects to that of experts. We contribute to this literature by focusing on the role of information signals in determining market outcomes.

Third, our paper is in spirit related to the literature on observational learning which was pioneered by Banerjee (1992) and Bikhchandani et al. (1992), which is empirically supported (Salganik et al., 2006; Cai et al., 2009; Zhang and Liu, 2012; and Newberry, 2016). While we do not explicitly model the belief formation process in this paper, our results also demonstrate the effect of peer generated signals on the efficiency of the platform. Unlike the aforementioned works, we quantify the observational learning effects in a discrete choice setting where competition for resources matters. Because of this, the signal on backer volume can actually reduce efficiency since the observational learning effect pulls pledges away from projects that need them.

The remainder of this paper is organized as follows. Section 2 describes the dataset. Section 3 lays out the demand model and presents the estimation results. Section 4 explains the simulation procedure and quantifies the market outcomes. Section 5 concludes.

2 Data

Our data comprise nearly 46,000 projects that were open on Kickstarter.com in 2013. The platform allows individuals to propose a project and seek funding in a variety of categories.[2] The project creator chooses a length of the funding period (typically between one and two months) as well as a funding goal. Potential backers can choose to pledge any amount of money to any open project. After the funding period ends, each backer is charged their contribution and the creator receives the amount pledged (minus a 5% commission on the total amount of the funds raised, plus a further payment processing fee) only if the project

[2] They are Art, Comics, Crafts, Dance, Design, Fashion, Film & Video, Food, Games, Journalism, Music, Photography, Publishing, Technology, and Theater. Given the large number of projects for our simulation procedure, we focus on a single 'inside' category; however, our model can be applied more generally if the sample size is small.

reaches or surpasses its funding goal; otherwise, the pledges are not collected (this is the 'all-or-nothing' mechanism).

Kickstarter is a 'rewards-based' crowdfunding platform, meaning that the project creator commits to providing the backer a reward commensurate with the level of his pledge if the project reaches its goal. Most often, these rewards are the outcomes of the creative project (e.g., a music CD or a movie DVD) although backers may instead receive a nominal reward (e.g., a thank-you note) for smaller contributions, or a unique service (e.g., a meeting with professionals) for larger contributions. The project creator thus sets the 'menu of rewards'—dollar thresholds for each level of the rewards their backers will receive—such that the total funds raised would cover the production cost (i.e., fixed and variable cost) of the rewards and leave enough capital for the next step.[3]

We focus on the Film category in our estimation, which is the largest category with nearly 7,000 projects in our sample. We chose this category because Kickstarter films are not usually seen as 'popular' productions that Hollywood blockbusters are and the budget for a film project tends to be large and less malleable than that for projects in other categories. Since crowdfunding is a tool to improve outcomes for 'long tail' products (i.e., niche projects that would not see the light of day under conventional funding mechanisms), Kickstarter would provide a way to gauge the level of interest in a film before significant resources are committed to its production. Though technically feasible, expanding our estimation to multiple categories need not yield additional insights given the unobserved heterogeneity across product categories.

For each of the Kickstarter projects in our data, we retrospectively scraped the daily number of backers and daily amount of pledges from Kickspy.com, a website that tracked all Kickstarter projects until it was shut down in March 2015.[4] The project characteristics

[3]Backers are often offered 'discounts' off the planned retail price. For instance, project creators sometimes state what the retail price of the outcome would be when the project is eventually completed. In such cases, the retail price is often higher than the necessary pledge level for the reward.

[4]Kickspy does not track individual backers but only the total pledges and the number of backers in *each*

as well as the outcomes come directly from the archive of Kickstarter.com. Specifically, we scraped the funding goal, the funding period, the launch date, the project category, the location and past activities of the project creator. We also scraped the dollar thresholds for each reward level and the number of backers who chose each reward level as of the end of the funding period.

We could not, however, obtain some data retrospectively. For instance, Kickstarter features a select few projects in each category as a “Staff Pick” which could be important, but we do not have access to this data for our sample period.[5] While it is intuitive that such projects would receive more eyeballs, these are small in numbers and there are also many counterexamples, so we think that its omission would not bias our estimation results. Another data we do not have access to is the daily number of Facebook shares or “likes” for our sample projects. Social media data are proprietary and difficult to gather for a large sample size. Further, some projects do not have a Facebook page and Facebook is not the only advertising medium available for the project creator to get the words out to attract potential backers.

However, we complement our dataset with the daily number of clicks on each project’s shortened url(s) as a proxy for the project creator’s promotional efforts, which captures the number of people arriving to the project website through social media on a given day. The data source is Bitly.com; and url shortening refers to the technique in which a website address is shortened for sharing on social media while still directing the readers to the original website upon clicking. Thus, when an Internet user shares a Kickstarter project through social media, the project page is often linked via a short url (rather than the full online address). Bitly is one of the largest url shortening services used in social media such

day of the project’s campaign period. Hence, we cannot infer the identify of individual backers or individual pledge amounts (unless there was only a single backer on a given day) from the Kickspy daily data.

[5] One possibility was to collect the data from a digital archive such as Wayback Machine, but the Wayback Machine started capturing Kickstarter category pages from mid-November of 2013, and as a result there are only 14 snapshots of the Film category prior to November.

as Twitter, Facebook and Google+ for our sample period, and it provides daily statistics for the number of each project's short url clicks.[6]

The first panel of Table 1 shows some descriptive statistics of the full sample. During 2013, nearly 16 million backers pledged \$470 million to approximately 46,000 projects, which resulted in \$414 million transferred to about 20,000 successfully funded projects. Each day, an average of 3,606 projects were open on Kickstarter, receiving an average of \$1.1 million in total pledges from 37,000 backers. The total number of clicks on the short urls from social media is about 20 million per day; and the ratio of the number of backers to Bitly clicks is 0.79 for all projects, and 0.89 for successful projects.

The second panel of Table 1 displays the same statistics for film projects only, which comprise about 15% of the total projects. This subset of projects received about \$74 million in pledges from 872,000 backers, resulting in \$62 million transferred to project creators. Interestingly, the backer-to-click ratio is much smaller for film projects (around 0.17) than for other categories, which could be helpful for our modelling exercise. For instance, this might be a reflection of low conversion ratio (from viewers to backers), so the platform's ability to achieve efficient outcomes plays an important role.

Table 2 describes the project-level outcomes. An average project across all product categories received a total of about \$10,000 in pledges from 348 backers, whereas the median is much lower. About 44 percent of the projects reached their funding goals, and the mean and the median of pledge-to-goal ratio are 3.02 and 0.34, respectively, due to the presence of highly popular projects. For film projects, the amount pledged and the success rate are similar to the overall figures, but the mean and the median of pledge-to-goal ratio are 0.82 and 0.51, suggesting more even distribution.

[6]In a Twitter Tweet Analysis conducted around the time period of our data, Bitly had a 65% share while the second runner (ow.ly) had a 19% share. It is claimed that Bitly is the worldwide best-known service without question. See, e.g., https://www.quintly.com/blog/2014/09/twitter-link-shortening-analysis-bit-ly-clear-market-leader/.

The second panel of Table 2 also shows that the average film project has a 41 percent larger funding goal than the average across categories, while the average number of backers is 37 percent lower than the average across categories. This may be because film projects often require a relatively large budget, which leads to smaller likelihoods of success. The bottom panel suggests that the mean and median of the goal was lower for successfully funded projects than for unfunded film projects. Successful projects also experienced significantly more short url clicks from social media to their project pages.

We provide a descriptive relationship between the information signals and backer pledges in Figure 1. In Figure 1(a), we calculate the percentage of the total pledges which occur at each level of funding status (rounded to the nearest percentile) for each project and then take the average of this across projects. We observe that i) a large percentage of donations occur when funding status is zero; ii) the share of donations is increasing in funding status up until the project reaches its goal, and then there is a large drop; iii) the relationship between funding status and donations becomes fuzzy after a project reaches its goal.

Similarly, Figure 1(b) displays the average percentage of pledges against the within-project deciles of backers-per-day, where we take the average across projects and plot the mid point of each decile bin (e.g., the percentage of donations occurring when backers-per-day is between the 1st and 2nd decile is plotted at percentile 15). Here, we see a large percentage of donations occurring at the lowest decile, which is likely because of the early push in donations, when the number of backers is very low. The donations then drop and slowly increase as the backers-per-day increases. This pattern is true regardless of whether or not a project has reached its funding goal.

3 Model

We begin by discussing our modelling assumptions. The first issue is how important the forward-looking dynamics of backers' decision-making is. We abtract from this dynamics because we think that as a first-order approximation backers are not overly-sophisticated vis-à-vis other backers who arrive later than they do. Friends and family of the project creator may have an incentive to influence other backers, but Agrawal et al. (2015) find that such local backers and more distance backers do not behave differently beyond their first investment. Given a relatively large number of potential backers on a crowdfunding platform, it seems to be unclear whether even local backers would necessarily intend to influence the decisions of other backers through the signals.[7]

The second issue is whether the backer's decision is a binary choice for each project, or a multinomial choice within a category. In the former, a backer visits a project page and make a donation decision for that project, while in the latter potential backers are able to navigate the Kickstarter website which also lists other active film projects and then decide which one to donate to (if any). While both types of decision-making processes are likely to be present, we chose to adopt the latter because it seems to be a more natural choice given the competition among multiple projects wherein project creators are vying for the backer's attention as well as limited budget for donations. Further, this allows us to test whether donations to a project are in fact affected by the projects' market shares in the data.

The third issue is that backers have to decide the amount of pledge when they make a pledge. To our knowledge, these two decisions have not been simultaneously modelled in the literature, but the existing studies often assume that the 'price' of the pledge is given. This is perhaps due to the data limitation (i.e., researchers do not observe individual pledge

[7]In a dynamic model in which there is no strategic delay, Alaei et al. (2016) model the backer's belief on the campaign success as an anticipating random walk, which is a function of the project quality and a probability of eventual success. While they do not consider the information signals, our utility specification can be viewed as a reduced form of this stochastic process.

amounts) as well as for tractability. Here, we propose that the individual backers form an expectation of how much he or she is going to pledge based on the menu of rewards. For instance, some projects will generally have more expensive menus than others, and potential backers can readily observe this menu pricing before making a pledge decision. If so, the expected menu price can become a static project attribute, which is absorbed by the project fixed effects.

Under these assumptions, we specify a nested logit model where one nest consists of all open film projects and the other consists of all open projects in other categories (i.e., the outside option). Formally, individual i chooses whether to pledge to one of the open film projects j or take the outside option of pledging to a project in another category. Each project is associated with a menu of prices from which a potential backer forms an expected pledge amount, $\bar{p}_j$, which does not vary across potential backers. We further assume that there is no heterogeneity in how the backers react to the observable dynamic information, so that they have common coefficients in the utility function.

Specifically, individual i arrives to the platform on day t and observes information for all open projects. The expected utility from donating to a film project j is

$$u_{ijt} = \gamma I_{jt}^{'} + \alpha X_{jt}^{'} + \delta_j + \nu_{jt} + \zeta_{if} + (1 - \sigma)\epsilon_{ijt}, \tag{1}$$

where ϵ_{ijt} is an i.i.d. demand shock that follows the type I extreme value distribution and σ captures the correlation in the error term across projects within the inside category.

I_{jt} is a vector of information signals for project j observable at the beginning of day t. X_{jt} is a vector of other contemporaneous variables such as the time remaining in the funding period and Bitly clicks.[8] δ_j is the project-level fixed effect which includes the expected donation amount, $\bar{p}_j$. Below, we refer δ_j as the project's 'quality', as it is related to the

[8]Backers need not observe the number of short url clicks; we include this variable as a proxy for the promotion or other media exposure.

utility of donating to a project that is unconditional of the dynamic variables. That is, high quality projects are those that receive a lot of donations regardless of their funding status, the number of backers, etc, and note that the project fixed effects account for the effects of any static variable such as the quality of the pitch video.[9] ν_{jt} is an unobserved demand shock that varies over time but is common to all backers. ζ_{if} is backer i's value of donating to a film project and is distributed i.i.d. extreme value.

The utility from donating to a non-film project ($j = 0$) is

$$u_{i0t} = \lambda C_t' + \epsilon_{i0t}, \tag{2}$$

where C_t is a vector of open projects in other categories which can shift the outside option value on day t.

Then the probability that an individual i donates to a project j conditional on choosing an inside project is given by

$$d_{ijt} = \frac{exp(\gamma I_{jt}' + \alpha X_{jt}' + \delta_j + \nu_{jt})}{exp(\lambda C_t') + \sum_{j \in \Omega_t} exp(\gamma I_{jt}' + \alpha X_{jt}' + \delta_j + \nu_{jt})}, \tag{3}$$

where Ω_t is the set of all open film projects on day t.

Because individuals do not vary in the deterministic preferences, the market share d_{jt} for project j is the same as equation (3). Berry (1994) shows that a linear equation can be derived under these assumptions and is given by

$$log(d_{jt}) - log(d_{0t}) = \gamma I_{jt}' + \alpha X_{jt}' + \delta_j + \nu_{jt} - \lambda C_t' + \sigma log(d_{jt/f}), \tag{4}$$

[9] When we regress the estimated $\hat{\delta}_j$ on the project characteristics, the results point towards this being an appropriate interpretation. For instance, $\hat{\delta}_j$ is positively significantly correlated with the prime locations for film professionals such as LA or NYC. If the project creator had previously backed others' projects, then $\hat{\delta}_j$ is also higher.

where $d_{jt/f}$ denotes the 'inside share' (i.e., the market share of project j within the film category).

In terms of covariates in the utility function, the vector of information comprises two types of signals, $I_{jt} = [I^1_{jt}; I^2_{jt}]$. I^1_{jt} contains the two primitive information signals: the current funding status of project j and the current number of backers for project j divided by the total number of days the project has been open. I^2_{jt} contains derivative signals based on I^1_{jt}. Specifically, they are a threshold dummy variable which equals 1 if the project has reached its funding goal and 0 otherwise, and more importantly interactions between each element in I^1_{jt} and the threshold dummy. The interaction terms account for the possibility that backers react differently to the observable signals after a project passes its funding goal. Thus, the incremental effect of funding status and the number of backers might change after reaching the funding threshold.

Funding status and backer volume affect the expected quality of donating to a project (an observational learning effect). Funding status may have an additional effect, wherein backers may value donating to projects in order to help them reach their goal (a threshold effect). Note that we are taking a reduced-form approach in terms of how the signals affect backer behavior. That is, our specification implies that both the project quality and whether or not the project reaches its funding goal enter directly into the utility function. Backers then use the signals to form their expectations of project quality and the probability the project will reach its goal. This approach is primarily done for tractability, but specifying it this way can identify the behavior that we are trying to capture in terms of the relationship between donations and signals.

The vector X_{jt} includes covariates other than the information signals that might affect backer behavior. Specifically, it includes a dummy variable for each month to control for seasonality and a function of the number of days remaining in the funding period in order to account for any variation based on time. In practice, the function is a third degree polynomial of days remaining and an interaction between the threshold dummy and the polynomial. X_{jt} also includes the number of clicks on the project's short url(s) on day t, and a dummy variable for the project-days in which the funding status is zero at the beginning of the day but positive at the end of that day.

Finally, the vector C_t contains the numbers of open projects in the Art and in the Music category, which are the two largest remaining categories outside of Film, as well as the number of open projects in all other categories. The number of projects in other categories can affect the utility of taking the outside option in either direction, due to the positive or negative externalities (Belleflamme et al., 2015). That is, if more projects in an outside category attract more backers to the platform, then there will be a positive externality, but if they are mostly vying for limited resources from the same backers, then there will be a negative externality.

We estimate the parameters γ, α, λ, and σ in equation (4) via a two-stage least squares regression with project fixed effects. We instrument for the endogenous regressor, $log(d_{jt/f})$, using the number of open projects in the inside (film) category on day t. The key identifying assumption in our estimation is that the unobserved demand shock, ν_{jt}, is conditionally independent of the variables contained in I_{jt}, X_{jt}, and C_t.

Estimation Results

The estimation results are presented in Table 3. Specification (1) does not include the interaction we proposed between the information signals and the threshold dummy, whereas specification (2) does. With a large number of inside projects in the dataset, the presence of zero market shares can be a particular concern for the identification of our demand model. We thus follow Gandhi et al. (2017) to correct for this zero market share bias by specifying them as a measurement error.[10]

Our first finding is that the effect of the threshold dummy is negative in specification (1), which is consistent with the existing literature and also supports the claim that backers may place value on donating to projects in order to help them reach the goal. While the standalone coefficient on the threshold dummy is positive in specification (2), the marginal effect of the threshold dummy is negative due to its interaction with the funding status.

The funding status has a positive correlation with pledges before the goal is reached, but that this effect drops to zero after the threshold is reached, since the coefficients on the status variable and on the interaction between status and threshold cancel out (i.e., the marginal effect is zero). This suggests the possibility that there is an observational learning effect before the threshold, as backers expect that projects that are closer to their goal are higher quality. However, the effect of the interaction between the threshold dummy and the status variable imply that backers tend to focus pledges on the projects that are close to, but have not yet reached, their funding goal.

Our next finding is that the effect of backers-per-day is positive and significant regardless of whether or not projects have reached the goal. This comes from the fact that the interaction with the threshold is insignificant and close to zero. Thus, potential backers are likely to see the number of backers-per-day as a positive signal of project quality throughout the funding period. Thus, donors will concentrate pledges towards projects with many backers regardless of whether or not the project has reached its goal, which will have important implications on how this signal will affect market outcomes.

The coefficients on the other covariates are as follows. The log number of short url clicks is positively and significantly correlated with making a pledge to a film project, representing the effect of the project creator's promotional efforts. The number of art projects decreases

[10]We explored other options such as dropping the project-days with zero market shares with little change in the qualitative results.

the value of the outside option, so the more art projects there are the more donations will go towards film projects; on the other hand, more projects in music and other categories increase the value of the outside option, pulling donations away from film projects. Finally, the log inside share is significant and positive, meaning that the projects are substitutes.

4 Simulation

We quantify the effects of information signals on market outcomes by simulating the backers' decisions under different information structures. While the parameter estimates tell us how the different signals effect the backer's decision to pledge, they do not capture the dynamic nature of backer decisions and how the information signals ultimately affect project outcomes. The simulation exercises performed below allow us to quantify and compare the magnitude of the impact of information signals on outcomes while keeping all other aspects of the market fixed.

We begin by describing the simulation procedure under the observed information structure, which will serve as both a measure of model fit and a benchmark for comparison. We set the initial values of the dynamic 'state variables' (i.e., the funding status, the threshold dummy, the backers-per-day, and days remaining) equal to the observed values at the beginning of the funding period for each project. Specifically, for all projects that start after March 1, 2013, the days remaining variable is set equal to the length of the funding period and the first three variables are set to zero.[11]

The number of backers donating to project j on the first day of our simulation period is

$$B_{j1} = \hat{d}_{j1} M_1,$$

[11]We have a two-month lead time for our simulation in order to eliminate projects which started before 2013, as we do not observe their data for the entire funding period. For any project that is already open on March 1, the start of our simulation period, we set these variables to their observed value on that date.

where M_1 is the total number of backers who donate to any project on day $t = 1$, and $\hat{d}_{j1}$ is the predicted market share of project j based on the coefficient estimates including the project fixed effect, $\hat{\delta}_j$, and the project-day level shock, $\hat{\upsilon}_{j1}$.

This predicts the number of backers to a particular project on a particular day, but in order to simulate the decisions of backers and update the state variables, we need to simulate pledge amounts to project j on day $t = 1$. We draw a set of $\tilde{B}_{j1}$ pledges from an estimated distribution of pledges for project j, where $\tilde{B}$ is B rounded to the nearest whole number. Since we observe the aggregate distribution of menu choices (associated with intervals of pledges) as of the end of the funding period, we estimate a continuous pledge distribution for each project, and then draw pledges from the estimated project-specific distribution, which is assumed to be independent and identically distributed.[12]

Specifically, we estimate a continuous distribution of pledges using the data on the menu of reward levels, R_{jy}, and the number of backers, K_{jy}, who contributed at each level, where $y \in \{0, \ldots, Y\}$ indexes the reward level. For example, if the first reward level is at \$10 and the second is at \$100, then we observe the number of backers who pledged at least \$10 but less than \$100. Given the skewness in data, we assume that for each project pledges come from a log-normal distribution with mean μ_j and a standard deviation ρ_j. The mean μ_j is observed by dividing the log of total pledges by the total number of backers, so the log-likelihood of observing the entire distribution of menu choices is

$$\begin{aligned} L(\rho_j) = &K_{j0} log(F(R_{j0}|\mu_j, \rho_j)) + K_{jY} log(1 - F(R_{jY}|\mu_j, \rho_j)) \\ &+ \sum_{y=1}^{Y-1} K_{jy} log(F(R_{jy+1}|\mu_j, \rho_j) - F(R_{jy}|\mu_j, \rho_j)), \end{aligned} \tag{5}$$

where $y = 0$ is the smallest reward level and $y = Y$ is the largest. We find the standard

[12]When we regress the average pledge in the data for a project on day t on a set of dummy variables for time remaining in the funding period and a project fixed effect, there are only a few days where the average pledge appears to be higher than normal. For the vast majority of days, they are stable across time.

deviation ρ_j which maximizes this function, resulting in the estimated pledge distribution, $\hat{F}_j(\cdot|\mu_j, \rho_j)$.

We randomly draw $\tilde{B}_{j1}$ number of pledges from this distribution, where each pledge is denoted by p_{j1b}.[13] Then the total dollar amount pledged to project j on day $t = 1$ is $P_{j1} = \sum_{b=1}^{\tilde{B}_{j1}} p_{j1b}$.

Using P_{j1} as the amount pledged and B_{j1} as the number of backers for each j on day $t = 1$, we update the dynamic state variables and move to the second day of our simulation period ($t = 2$). We iterate this process over t, with projects entering and exiting according to their funding periods, until the end of our sample period. The outcome of this procedure is the total number of backers and pledges for each project for a single simulation run. Because there is a randomness in simulating the dollar amount of pledges, we repeat this procedure 500 times and present the average outcomes across the simulations.

We simulate the project outcomes using the same procedure as above, but under three different information structures: The first structure is when both the funding status and the backers-per-day are suppressed. In the second structure, only funding status is suppressed, which is adding backers-per-day into the information set. In the final structure, only backers-per-day is suppressed, which is adding status back into the information set. Instead of naively assuming that backers would ignore the suppressed signal(s), we allow backers to form an expectation of the signal based on other observable information as it would provide a more plausible environment in which to assess the impact of the respective signal.

Specifically, we capture the backers' expectation of the suppressed signal by running a regression of that signal on the observable variables, and then use the estimates to construct the 'expected value' of the suppressed signal conditional on the covariates. For example, to find the expected value of the funding status when status is suppressed, we run a regression

[13] Hence, backers do not consider the exact amount of the pledge to project j when they make their donation decisions, but this pledge is only realized after they decide which project to donate to. The interpretation is that they first form an expectation of the pledge, and realize their personal pledge afterwards.

of funding status on project characteristics (dummies indicating the project location, number of projects created and donated to by the project creator, and dummies indicating the genre of the movie), month dummies, the function of the time remaining in the funding period, and any other information signal (i.e., number of backers). Thus, the gap between the predicted and the true value of the signal is what would drive the impact of the information signals.

Since our aim is to quantify the relative impact of the two signals, we do not purport to describe this as a counterfactual policy because for instance backers may change their behavior in a true counterfactual world in which the information environment changes. As such, our analysis does not really provide policy recommendations regarding the content or amount of information that should or should not be present in a crowdfunding platform. However, our analysis can shed light on how important the information signals are in determining the efficiency of the current market and the relative importance of the two signals.

Simulation Results

Table 4 presents the observed and simulated platform-level outcomes under the current information structure, which allows us to examine the fit of our model. In general, the model does a good job in matching the aggregate outcomes. The model slightly overpredicts the number of contributing backers, and it somewhat underpredicts the funding success rate as well as the actual donations to the successful projects.[14] Figure 2 displays a scatter plot of the project's simulated pledges and the total number of backers against the observed data. Given the concentration around the 45 degree line, the overall fit of the model appears to be good, suggesting that our procedure to model the backer behavior and pledges may be a reasonable approximation of the data generating process.

In terms of individual project outcomes, the model does a good job in matching the

[14] A project is deemed successful in our simulation if the average total pledge across the simulation runs is greater than the funding goal.

success or failure of individual projects, as 90 percent of the projects result in the same outcome in our model simulation as they do in the observed data. The table also shows the average quality of successful and unsuccessful projects, where 'quality' refers to the estimated project fixed effect. Our model tends to underestimate the quality of successful projects and overestimate the quality of unsuccessful ones. Nonetheless, we think that this is a reasonable match for dynamic simulation. Note that in what follows, we compare the simulated platform-level outcomes under the alternative information structures with the model's simulated benchmark outcomes in Table 4, not with the outcomes in the data.

Table 5 shows the simulated platform-level outcomes. The first column merely repeats some of the statistics of the simulated benchmark outcomes. The second column displays the simulated outcomes when we suppress both information signals, and the third and fourth columns show the outcomes when each piece of information is suppressed separately.

We focus on measures of market efficiency. The first is the success rate of projects or how many projects reach their funding goal, and the second is the average quality of successful projects. The reason for focusing on these two measures is that it seems reasonable that a planner would want to maximize the number of funded projects conditional on some level of quality. For completeness, we also show how much is being pledged and ultimately donated in each case.

We first examine the case where both signals are suppressed. The number of successful campaigns fall by over 8% as compared to the benchmark simulation. This change is due to over 12% of projects switching their funding outcome, with the majority of those switching from being funded to not being funded. This may be considered a market improvement if those projects that switch to not being funded are the low quality projects, but this does not seem to be the case, as the average quality of successful projects falls by almost 50%. These results come from the fact that the information signals lead to pledges being focused on both high quality projects, through the observational learning effect, and projects that needed funding to reach their goal, through the threshold effect.

We then turn to the isolated effects of each signal. In presenting these results, we compare the case of no signal to the case in which we add the signal back (this allows us to isolate the effect of the signal apart from any interaction between the signals). Looking at the simulated market outcomes in which we add the backer volume back to the information set, the number of successful campaigns decreases by about 4% compared to the no information case. Again, this may be beneficial to the market if the projects that fail to be funded are the low quality projects, but this does not seem to be the case as the average quality of successful projects falls by over 7%.

This is because donations are now more concentrated towards projects that have a large number of backers-per-day. Therefore, fewer projects will get the donations necessary to reach their funding goal. This can indicate the inefficiencies that come along with the observational learning effects: when backers base their decision on previous backers, rather than observed quality, it may be the case that projects with high observed quality will receive a bad draw of early backers and fail to gain enough traction to reach their funding goal. These results suggest that showing the backers-per-day tends to lead to inefficient market outcomes.

This, along with the comparison between the model and the no information outcomes, implies that the funding status impacts outcomes in a positive way. Indeed, the number of successful projects in the Add Status column increases by 16% over the No Information column. Further, the projects that are now successful are the high quality ones as the average quality of projects increases by 66%. The reason funding status leads to both a higher success rate and a higher average quality must be the feedback between the observational learning effect and the threshold effect. Backers donate more to the projects that have yet to reach the goal, and this pulls donations away from projects that have already reached the funding goal and projects that have a high initial volume of backers towards the projects that are both higher quality and need funding to succeed.

Finally, note that the simulated outcomes under the observed information structure are the most efficient. That is, while the backers-per-day information lowers the efficiency compared to no information outcome, it increases efficiency when we add it to the funding status information (i.e., comparing the Add Status to the Model columns). Therefore, the results of our simulation exercise suggest that while the information signals are valuable together, in terms of generating efficient platform-level outcomes, funding status information is the main contributor to the market efficiency and there is a moderate level of inefficient herding if only backer volume signal were observed in a platform. In other words, the platform-level outcome more closely resembles what occurs when one signal (the funding status) is observable rather than the other (backer volume) is.

5 Conclusion

We provided an empirical framework to disentangle the impacts of the two most commonly used information signals in crowdfunding platforms. Our analysis renders support for the argument that the funding status pulls backers' pledges towards projects that are close to their funding goal and away from projects that are either very far away from the funding goal or are already past it. Funding status acts to improve the market level outcomes, as high quality projects which have not met their funding goal are the beneficiary of the positive feedback effect of funding status. On the other hand, backer volume has a positive feedback effect both before and after the threshold, pulling pledges away from projects that have yet to reach their funding goal, but backer volume only improves the efficiency of market outcomes when both signals are observable. Therefore, the funding status information is the primary contributor to the current information structure and it also exerts a disproportionately larger impact on the efficiency of the market outcomes.

References

[1] Agrawal, A., C. Catalini, and A. Goldfarb. 2015. "Crowdfunding: Geography, Social Networks, and the Timing of Investment Decisions." *Journal of Economics & Management Strategy* 24: 253–274.

[2] Alaei, S., A. Malekian, and M. Mostagir. 2016. "A Dynamic Model of Crowdfunding." Mimeo.

[3] Banerjee, A. 1992. "A Simple Model of Herd Behavior." *Quarterly Journal of Economics* 107: 797–817.

[4] Belleflamme, P., N. Omrani, and M. Peitz. 2015. "The Economics of Crowdfunding Platforms." *Information Economics and Policy* 33: 11–28.

[5] Berry, S. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *RAND Journal of Economics* 25: 242–262.

[6] Bikhchandani, S., D. Hirshleifer, and I. Welch. 1992. "A Theory of Fads, Fashion, Custom, and Cultural Change in Informational Cascades." *Journal of Political Economy* 100: 992–1026.

[7] Cai, H., Y. Chen, and H. Fang. 2009. "Observational Learning: Evidence from a Randomized Natural Field Experiment." *American Economic Review* 99: 864–882.

[8] Chemla, G., and K. Tinn. 2017. "Learning through Crowdfunding." Mimeo.

[9] Cumming, D., G. Leboeuf, and A. Schwienbacher. 2015. "Crowdfunding Models: Keep-It-All vs. All-Or-Nothing." Mimeo.

[10] ———, and L. Hornuf. 2018. *The Economics of Crowdfunding*. Palgrave Macmillan.

[11] da Cruz, J. V. 2018. “Beyond Financing: Crowdfunding as an Informational Mechanism.” *Journal of Business Venturing* 33: 371–393.

[12] Ellman, M., and S. Hurkens. 2015. “Optimal Crowdfunding Design.” Mimeo.

[13] Freedman, S., and G. Z. Jin. 2017. “The Information Value of Online Social Networks: Lessons from Peer-to-Peer Lending.” *International Journal of Industrial Organization* 51: 185–222.

[14] Gandhi, A., Z. Lu, and X. Shi. 2017. “Estimating Demand for Differentiated Products with Zeroes in Market Share Data.” Mimeo.

[15] Inbar, Y., and O. Barzilay. 2014. “Community Impact on Crowdfunding Performance.” Mimeo.

[16] Iyer, R., A. Khwaja, E. Luttmer, and K. Shue. 2016. “Screening Peers Softly: Inferring the Quality of Small Borrowers.” *Management Science* 62: 1554–1577.

[17] Kawai, K., K. Onishi, and K. Uetake (2014) “Signaling in Online Credit Markets.” Mimeo.

[18] Kuppuswamy, V., and B. Bayus. 2018. “Crowdfunding Creative Ideas: The Dynamics of Project Backers.” In *The Economics of Crowdfunding*, edited by D. Cumming, L. Hornuf, 151–182. Palgrave Macmillan.

[19] Li, Z., and J. Duan. 2016. “Network Externalities in Collaborative Consumption: Theory, Experiment, and Empirical Investigation of Crowdfunding.” Mimeo.

[20] Lin, M., and S. Viswanathan. 2016. “Home Bias in Online Investments: An Empirical Study of an Online Crowdfunding Market.” *Management Science* 62: 1393–1414.

[21] Marwell, N. 2015. “Competing Fundraising Models in Crowdfunding Markets.” Mimeo.

[22] Mollick, E. 2014. "The Dynamics of Crowdfunding: Determinants of Success and Failure." *Journal of Business Venturing* 29: 1–16.

[23] ———, and R. Nanda. 2016. "Wisdom or Madness? Comparing Crowds with Expert Evaluation in Funding the Arts." *Management Science* 62:1533–1553.

[24] Newberry, P. W. 2016. "An Empirical Study of Observational Learning." *RAND Journal of Economics* 47: 394–432.

[25] Salganik, M., P. Dodds, and D. Watts. 2006. "Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market." *Science* 311: 854–856.

[26] Strausz, R. 2017. "A Theory of Crowdfunding—A Mechanism Design Approach with Demand Uncertainty and Moral Hazard." *American Economic Review* 107: 1430–1476.

[27] Vulkan, N., T. Estebro, and M. Fernandez Sierra. 2016. "Equity Crowdfunding: A New Phenomena." *Journal of Business Venturing Insights* 5: 37–49.

[28] Zhang, J., and P. Liu. 2012. "Rational Herding in Microloan Markets." *Management Science* 58: 892–912.

Table 1: Platform Descriptive Statistics

	Total	Total Successful	Per-Day Mean
		ALL	
Projects	45,784	20,211	3,606
Backers (k)	15,968	13,498	37.57
Pledged (mill)	\$469	\$414	\$1.10
Clicks (k)	20,172	15,236	47.47
		FILM	
Projects	6,775	3,104	685
Backers (k)	872	754	2.35
Pledged (mill)	\$73.8	\$61.8	\$0.28
Clicks (k)	5,239	4,319	17.1

Notes: Table presents aggregate descriptive statistics for the entire sample of projects and film projects only. The per-day mean is average across days in the sample.

Table 2: Project Descriptive Statistics

	Mean	Median	SD	Min	Max
ALL PROJECTS					
Backers	348	33	3,293	0	341,667
Pledged	$10,255	$1,453	$67,724	$0	$5,702,154
Goal	$28,392	$5,135	$533,834	$0	$100,000,000
Clicks	441	29	6,921	0	1,216,769
Length	32	30	10	1	84
Successful	0.44	-	-	-	-
% Goal Reached	3.02	0.34	237	0	41,535
FILM PROJECTS					
Backers	129	30	1,325	1	91,585
Pledged	$10,921	$2,275	$85,676	$1	$5,702,154
Goal	$40,035	$8,000	$236,823	$1	$100,000,000
Clicks	774	40	16,695	0	1,216,769
Length	32	31	11	2	62
Successful	0.46	-	-	-	-
% Goal Reached	0.82	0.51	10	0	835
SUCCESSFUL FILM PROJECTS					
Backers	243	75	1,869	1	91,585
Pledged	$19,933	$6,180	$120,358	$15	$5,702,154
Goal	$15,071	$5,000	$59,403	$1	$2,000,000
Clicks	1,297	86	23,530	0	1,216,769
Length	31	31	10	2	62
% Goal Reached	1.59	1.10	14	1	835

Notes: Table presents project-level descriptive statistics for the entire sample of projects, film projects, and successfully funded film projects.

Table 3: Model Estimation

Variable name	(1)	(2)
Funding Status	0.001**	0.481***
	(0.000)	(0.039)
Funding Status × Threshold Dummy		-0.481***
		(0.039)
Backers-per-Day	0.012***	0.012***
	(0.003)	(0.004)
Backers-per-Day × Threshold Dummy		-0.005
		(0.004)
Threshold Dummy	-0.143***	0.370***
	(0.011)	(0.035)
Log Clicks	0.141***	0.141***
	(0.010)	(0.010)
Status = 0 Dummy	-0.095***	-0.024***
	(0.008)	(0.005)
Art Projects	0.335***	0.332***
	(0.012)	(0.012)
Music Projects	-0.224***	-0.230***
	(0.007)	(0.007)
Other Projects	-0.016***	-0.018***
	(0.004)	(0.004)
Log Inside Share	0.582***	0.559***
	(0.030)	(0.033)
Month Dummies	Y	Y
Days Remaining Function	Y	Y
Project Fixed Effect	Y	Y
No. of Obs.	202,926	202,926
R-Sq	0.757	0.748

Notes: Presented are the results of the choice model regressions à la Berry (1994), where zero market shares are corrected for following Gandhi et al. (2017). Standard errors are in the parentheses. *** is significant at 1%, ** is significant at 5%, and * is significant at 10%.

Table 4: Model Fit

	Data	Model
Total Backers (100K)	8.72	8.99
Average Backers	129.15	133.07
Median Backers	30	36.14
Total Pledges ($M)	73.78	75.77
Average Pledges ($K)	10.92	11.22
Median Pledges ($K)	2.28	2.80
Total Donations ($M)	61.87	49.91
Average Donations ($K)	9.16	7.39
Median Donations ($K)	0	0
Success Rate	0.46	0.40
Match	-	0.90
Average Quality of Successful	0.15	0.12
Average Quality of Unsuccessful	-0.06	-0.02

Notes: Presented are the aggregate outcomes which we observe and the same statistics from simulating the model. The model statistics are the average of 500 different sets of pledge amounts over the life of a project. Quality is measured by the estimate project fixed effect and success in the model is defined as if the average pledge amount across simulations exceeds the funding goal.

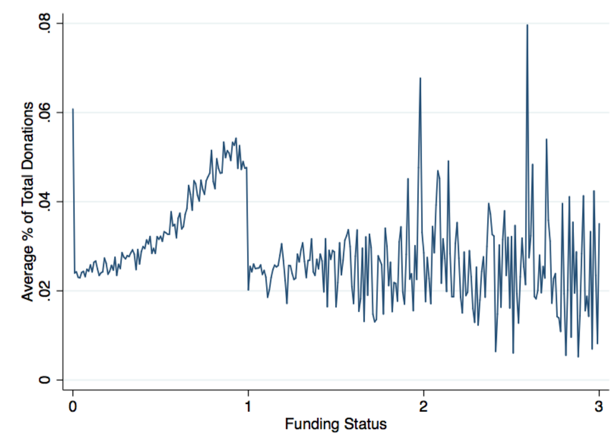
Table 5: Simulated Outcomes

	Model	No Information	Add Backers	Add Status
Total Backers (100K)	8.99	8.85	9.00	8.77
Total Pledges ($M)	75.77	75.27	75.58	75.04
Total Donations ($M)	49.91	47.07	46.82	51.22
Success Rate	0.395	0.363	0.348	0.421
Switch to Unsuccessful	-	0.078	0.086	0.000
Switch to Successful	-	0.046	0.039	0.026
Average Quality of Successful	0.123	0.065	0.060	0.108
Average Quality of Unsuccessful	-0.018	0.022	0.026	-0.013

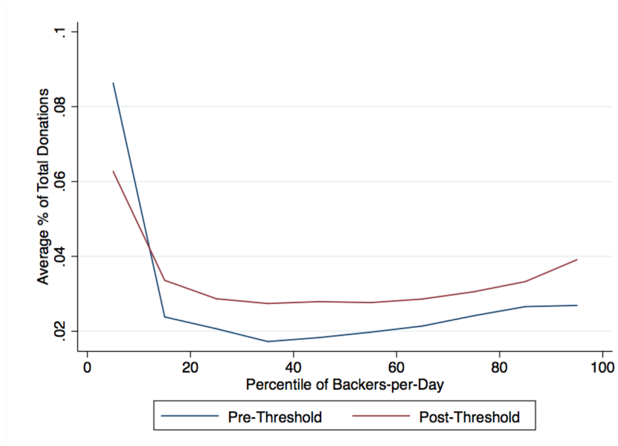
Notes: The model presents the simulated outcomes under different information structures. The simulations are calculated by suppressing the information signal and allowing the backers to form an expectation based on observables and then simulating the dynamic pledge decisions, keeping all other covariates fixed. Quality is measured by the estimate project fixed effect and success in the model is defined as if the average pledge amount across simulations exceeds the funding goal.

Figure 1: Stylized Data Pattern

(a) Funding Status

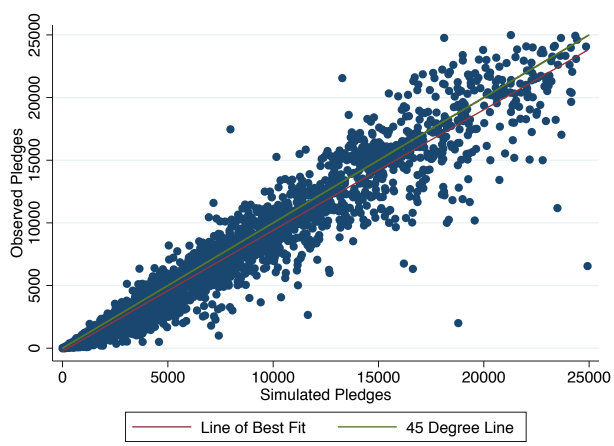


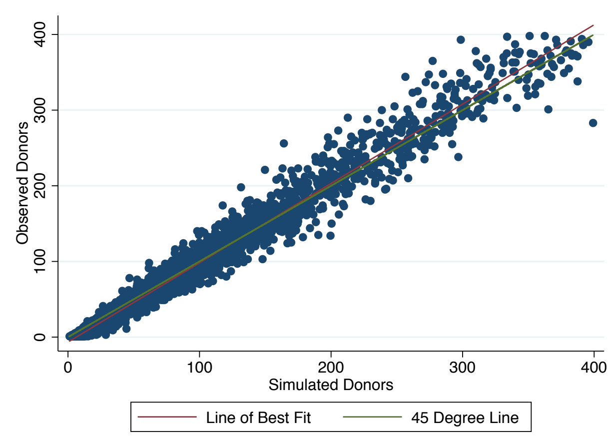
(b) Backers



Notes: Figure (a) shows the average percentage of pledges that occur at each level of funding status; Figure (b) shows the average percent of pledges at each backers-per-day decile.

Figure 2: Goodness of Fit





Notes: Displayed are simulated project-level pledges and backer volume plotted against the corresponding observed outcome, along with the 45 degree line and a line of best fit.