

# Are All Spillovers Created Equal? The Impact of Blockbusters and the Composition of Backers in Online Crowdfunding

*Completed Research Paper*

**Jongho Kim**

KAIST College of Business  
85 HoegiroDongdaemoon-gu Seoul,  
Korea  
jonghkim@kaist.ac.kr

**Minhyung Lee**

KAIST College of Business  
85 HoegiroDongdaemoon-gu Seoul,  
Korea  
winterlove@business.kaist.ac.kr

**Daegon Cho**

KAIST College of Business  
85 HoegiroDongdaemoon-gu Seoul,  
Korea  
dgcho@business.kaist.ac.kr

**Byungtae Lee**

KAIST College of Business  
85 HoegiroDongdaemoon-gu Seoul,  
Korea  
btle@business.kaist.ac.kr

## Abstract

*Crowdfunding has emerged alongside the IT development. It is believed that overwhelmingly successful projects, blockbusters, would have significant impacts on the overall crowdfunding platform. However, there are notable limitations in previous studies. First, we consider how the advent of blockbusters impact according to the projects' similarity with inside and outside clusters, rather than pre-determined category. Second, we examine the blockbusters' heterogeneity with the type of backers that bring different effects. We use project-level dataset and apply novel clustering method to analyze blockbuster effects. We find empirical evidence that blockbusters have a spillover effect on same categories, especially inside clusters experience larger effects than outside clusters. In the long run, these spillover effects decay faster in outside clusters, but last long for inside cluster. Furthermore, this result changes according to the composition of backers. Our study presents a promising avenue for the application of semantic network analysis to the crowdfunding context.*

**Keywords:** Crowdfunding, Blockbuster Effects, Semantic Network, Market Structure, Text Mining, Consumer Decision Making, Customer Experience

## Introduction

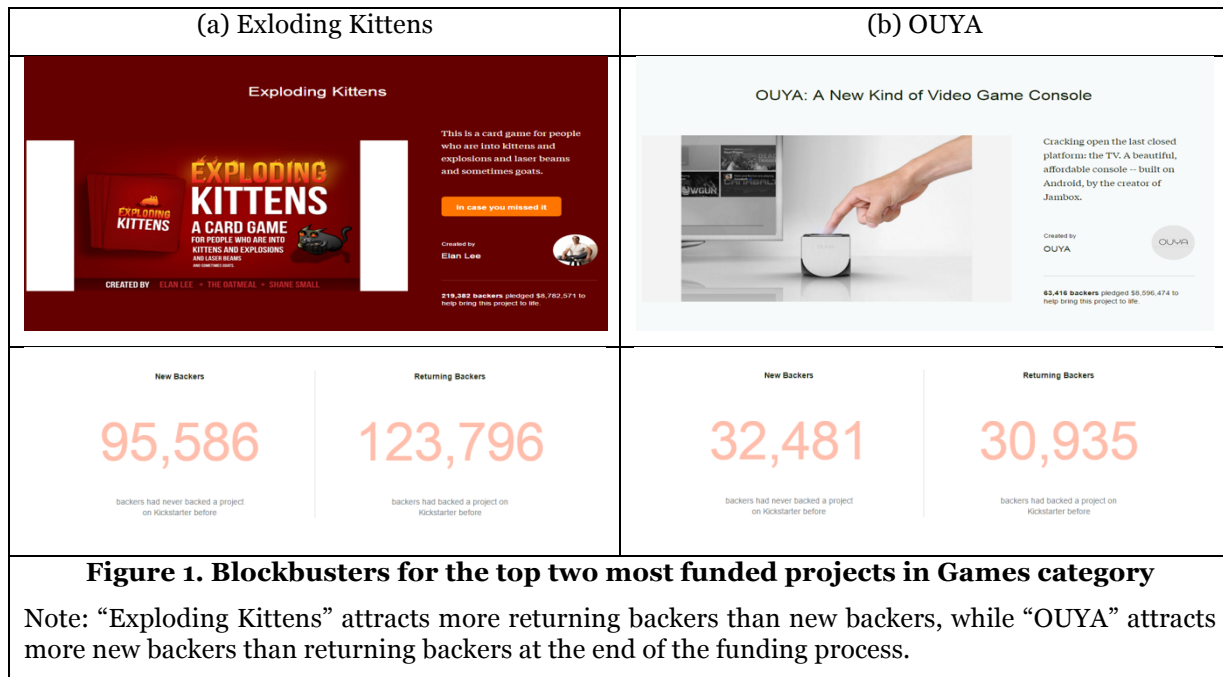
Crowdfunding has recently emerged alongside the development of information technology, helping to reduce search and matching costs for entrepreneurs. The basic idea of crowdfunding is simple: instead of using a small group of sophisticated investors to raise funds, entrepreneurs attempt to obtain funds from a large crowd through Internet channels, where individuals can provide a small amount of funding (Burtch et al. 2013). Crowdfunding allows individual founders to bypass financial intermediaries such as institutional banks and venture capital firms by lowering the entry barrier to starting a social or a for-profit project (Beaulieu and Sarker 2013). Therefore, crowdfunding fosters the entrepreneur's early stage of firm development by providing a possible alternative to the seed capital of angels or venture capitals (Tomczak and Brem 2013).

As many traditional markets have been considered as “winner-take-all,” information technology has changed the supply and demand of popular and niche products, which in turn has resulted in changes of market structure and competition (Brynjolfsson et al. 2010; Fleder and Hosanangar 2009). In this regard, it is widely believed that overwhelmingly successful projects, called “blockbusters” in this study, would have a significant impact on the overall crowdfunding platform (Schwienbacher and Larralde 2010; Tomczak and Brem 2013). However, there have only been a few studies that have further investigated how blockbusters impact competitions across projects. Liu et al. (2015) is an example that showed the impact of blockbuster projects on existing projects at the category-level. They found positive concurrent and lasting effects within the same category and cannibalization effects across categories. Also, after the arrival of an outlier, Doshi (2014) found that there is a positive spillover effect on dollars pledged for projects inside the same category and onto a competing platform within the same category with a decline of the focal platform's profit. Our study is motivated by these studies. We further aim at untangling the relationship between blockbusters and other projects by addressing the limitations of previous studies. In other words, we examine how the advent of blockbusters affect other projects and backers' behavior. To do so, unlike prior work, we use a large-scale project-level data set and apply some machine learning techniques.

Indeed, there are notable limitations in the previous studies. First, the category-level analysis is not able to account for the differentiated impact of a blockbuster on each project. Categories have imperfect class definitions, overlapping categories and random variation among observations within the same category (Anderberg 2014). In other words, it is common to see that projects in the same category have notably different attributes, so the impact of a blockbuster is very likely to be different across projects even within the same category. For example, music player, e-library, drone and 3d printer projects are presented in the same “technology” category at one of the largest crowdfunding websites. While these four projects may be related to the use of advanced technologies, it is hard to see that these are comparable projects from the perspective of potential backers. Therefore, rather than using the given category group, classifying similar projects into the associated group may be a more sensible idea. In this regard, our research conceptualizes the relationship between the projects into both the cluster and the category level. When a blockbuster project exists, a category including the blockbuster can mutually exclusively be divided into *inside cluster* (where the blockbuster is in) and *outside cluster* (where the blockbuster is not in). All other categories that are nothing to do with the blockbuster can be regarded as *outside category*. Therefore, we examine the distance effects of blockbuster, in three distinct aspects based on the project similarity. We tried to find the natural association among observations according to the semantic network analysis. To the best of our knowledge, no previous study has examined the distance effect in a crowdfunding setting despite the importance of similarity between projects.

Second, “blockbusters” are commonly defined by the size of the total pledge amount (Shakan and Bayus 2003; Collins et al. 2002), and previous studies did not take into account the difference of the backers' composition. However, to measure the effect of blockbusters more comprehensively, it may be important to consider how many, and what proportion of, new backers contributed to blockbuster projects. Our study emphasizes the division between new backers and returning backers, because we conjecture that blockbusters attracting more new backers may have a different impact on other projects compared with those attracting more returning backers. Luring more new backers is also important for the growth of a crowdfunding platform. For example, in Figure 1, we list the top two most funded projects in the Game category as of 2016. “Exploding Kittens” in Figure 1(a) is a common and frequently-seen type of card game, and the number of returning backers is greater than the number of new backers. In contrast, “OUYA” in Figure 1(b) is a new and innovative video game console that is rarely seen in the Game category. It appears

that the originality of the project may attract a greater number of new backers compared to returning backers. These two blockbusters show different compositions of new and returning backers, implying that the two types of backers perhaps exhibit different contribution behavior. For example, returning backers are familiar with the platform, so they may have the higher level of perceived risk as they tend to have the better understanding of the attributes of products due to the previous experiences (Pavlou 2003, Kim and Gupta 2009). Therefore, they are more likely to behave to reduce the (potential) risk by gaining more information from their experiences. Whereas the investment decision of new backers is more likely to be affected by their social friends, which make new backers feel safer projects. Although it is impossible to trace the entire contribution history of each backer, our study focuses on investigating how the composition of new and returning backers for blockbusters impacts the growth of other projects, which has not been explicitly examined in previous work.



We summarize our research questions as follows: 1) what impact do blockbuster projects have on projects according to their similarities and 2) what impact does a blockbuster project’s composition of new and returning backers have on other projects? For addressing the first research topic, we analyze two differentiated time effects of blockbusters: (a) the concurrent effect (i.e., short-term effect), and (b) the lasting effect (i.e., long-term effect) on projects in the inside cluster, the outside cluster and the outside category. In the second research topic, we investigate how the composition of returning and new backers of blockbuster projects gives a differential effect on other projects. To empirically analyze these effects, it is worth noting that we collect relevant data sets from four different sources using a self-developed web crawler. We also apply novel text mining and clustering techniques to unfold latent market structure.

We show our findings in three directions. First, the concurrent effects of blockbusters on projects of both inside and outside clusters are positive, but the impact of the inside cluster is greater than that of the outside cluster, as one would expect. Specifically, the spillover effect of blockbusters indeed varies across projects within the same category, suggesting that it is important to take into account the similarity of the projects for more precise evaluations. However, the blockbusters show concurrently cannibalization effect to the projects outside the category where the blockbusters belong. This finding could suggest that considering both the cluster and the category will be important when launching new crowdfunding projects. Second, and more interestingly, when it comes to the lasting effect, blockbusters provide a diminishing positive impact on projects of both inside and outside clusters. Also, the effects on the outside cluster experience faster decay of spillover effects as the timeframe is extended after blockbuster’s exit. Therefore, it appears

that the blockbuster shows stronger spillover effects on the inside cluster than outside clusters in both short and long term. Third, our findings regarding the composition of new and returning backers suggest that blockbuster projects with more new backers show 1) marginal cannibalization effects on the projects in the inside clusters, 2) positive spillover effects on the projects in the outside clusters and the projects outside the clusters. As a result, it would be important to take the backers' characteristics into consideration for more precise examination, which was not dealt with in the previous studies.

As we will discuss later, new backers reacted more sensitively to the risk description of projects, indicating that they are more likely to depend on projects with more risk-description. This observation may imply that when new backers consider the next investment, they seem to rely on projects with more description about risk in order to relieve the risk and show more herding behavior than returning backers. In other words, new backers are likely to seek projects with high quality in the outside cluster rather than contribute to a functionally similar project in the same cluster where they first invested. On the other hand, blockbuster projects with more returning backers have greater spillover effects on other projects than blockbuster projects with more new backers. Anecdotal evidence posits that blockbuster projects with a large number of new backers tend to be unique and creative, and thus these projects are likely to have a high degree of the social exposure. Most blockbuster projects with a greater number of returning backers are not as creative as blockbusters with a large number of new backers, but spillover effects of these blockbusters are greater than blockbusters with many new backers. We believe that our findings were not comprehensively examined in prior studies and provide practical implications for both entrepreneurs and platform. For instance, an entrepreneur planning to open a new project similar to an ongoing blockbuster project may take advantage of the opportunity. Also, the composition of new and returning backers is important to determine the characteristics of a blockbuster, which may in turn provide a varying degree of the spillover effect on other projects according to the similarity between projects.

We use detailed project-level data and analyze the competition of projects via semantic network analysis. Despite the importance of revealing the latent market structure, there has been no study using this technique in a crowdfunding context since it could be empirically challenging to test the latent relationship among projects. However, very recently, a growing number of studies have begun using semantic analysis in the context of the recommender system, and these studies have highlighted novel insights that were not fully measured in previous works (Xu et al. 2014; An et al. 2014; Pongetti 2011; Amara et al. 2014; and Buttice et al. 2015). We believe that our study pioneers the application of semantic network analysis to the context of crowdfunding. In short, this study can be positioned as a unique addition to the growing stream of literature due to our novel methodology. Furthermore, we argue that every blockbuster project does not have the same effect on other projects. We use the composition of new and returning backers to differentiate blockbuster projects and find varying effects by this ratio. This finding can provoke further studies and practical discussion to explore more explicit features of blockbusters.

## **Literature Review & Theoretical Perspectives**

### ***Blockbuster Effects in Crowdfunding***

“Blockbusters” refer to dominantly successful projects compared to other existing projects (Shakan and Bayus 2003), and they are mainly determined according to the size of the total pledged amount (Collins et al. 2002). As noted above, only a few studies examined the effect of blockbusters in the context of crowdfunding. Liu et al. (2015) is an example; they showed that there is a positive network effect within the category and negative effects across categories. The unit of analysis of the study is the pre-determined category, which is an aggregated sum of the pledged amount of all projects within the same categories. Doshi (2014) also investigated the impact of outlier sellers, which could be referred as blockbuster projects. A main finding of the study suggests that outliers in certain product categories encouraged new entry and transactions for the projects in the same category.

However, it is difficult to conclude that each blockbuster project gives the same degree and the same direction of an effect to all the other projects. For example, in the same category, there can be a pair of two projects that are very closely related to each other and another pair of two projects that are almost independent from one another. Therefore, it may be necessary to consider the similarity among projects to measure the impact of the blockbuster project more rigorously. To compromise the challenge, we adopted

the concept of the cluster, which is defined as a group of natural association (Anderberg 2014). In this case, it is also necessary to analyze the blockbuster effect at the project level instead of the category level to account for the heterogeneity across projects. Our study analyzes each project's semantic description shown on Kickstarter.com through using text mining techniques. This aspect is one of the major contributions of this study, which is also differentiated from the related prior work.

### ***Semantic Network***

In the online market, product brands can be described through consumer-posted text descriptions, and the text co-occurrence can reflect the associative and semantic networks between product brands (Netzer et al. 2012). Consumers' comments can also reflect the undisclosed product characteristics that used to be only realized after a consumer experienced the product. In fact, the validity of using the frequency of term occurrence on a website is proved to match the likelihood of corresponding phenomenon (Saiz and Simonsohn 2007). It is also found that direct comparisons of products in the semantic network of consumer comments is one of the main motives of content generators for seeking information (Schindler and Bickart 2005). Mostafa (2013) found that there is a positive consumer sentiment in consumer tweets towards famous cosmopolitan brands. However, there is little research that focuses on the semantic network between product description rather than consumer comments. In particular, the descriptions of crowdfunding projects include a greater deal of contents than other general product descriptions: They include the background, purpose, risk, current situation and historical development of the product itself in more detail. Even though projects are separated according to the characteristics of the main project, there can be a latent relationship between projects in terms of project descriptions. Our research attempts to investigate the latent market structure of crowdfunding projects within the same category via analyzing the text description of each project to show the different semantic networks.

### ***Network Effect***

Our study also deals with the network effect between projects. Network effect, which is also referred to as network externality, has appeared in markets and societies where the utility gain from a product is affected by its combination with other products (Katz and Shapiro 1994). In many cases, the network effect can be divided into direct network effect and indirect network effect. The classical example of the direct network effect is from the telecommunications network, where the utility of a product is directly connected to its consumption by users. In other words, more users linked to the service leads to increased utility of the phone service (Katz and Shapiro 1985). Indirect network effect can be achieved when an increase of complementary product usage impacts the usage of certain similar products types (Gandal 1995; Clements and Ohashi 2005).

In the context of crowdfunding, both direct and indirect network effects can exist in a similar manner. Burch (2011) determined that backers' participation in the same project affects the utility of an individual backer, which is referred to be the direct network effect in the context of the crowdfunding platform. Additionally, Liu et al. (2015) suggested that there is the indirect network effect when the performance of certain projects influences other projects' performance by attracting new backers and making the platform more popular. Our study investigates indirect network effect by blockbuster projects to existing projects according to their cluster separated by semantic analysis.

### ***Inside Cluster vs. Outside Cluster***

Shakan and Bayus (2003) suggested that a blockbuster project will increase the size of the network of projects as numerous backers are attracted to the blockbuster project. Strahilevitz (2003) found that additional input resources, such as new backers, could be shared with other related projects within the same category. Similarly, Hagi (2009) highlighted that there can be same-side network effects of blockbusters in a similar group. Doshi (2014) also suggested that the effect of a high-performing outlier's entry (i.e., a blockbuster project) impacted positively across similar projects. This can be connected to the spillover effect of blockbuster projects both inside and outside the cluster.

However, as Doshi (2014) noted, there could be countervailing effects of blockbusters to similar projects. The blockbuster may crowd out future contributions to certain cluster by reducing the budget of backers who have a certain taste to inside cluster. Therefore, it may cause negative spillover effects, called "cannibalization effects," led by the blockbuster project (Ghose et al. 2006). Similarly, there also could be conflicting effects to outside clusters, as the largest project absorbs a significant portion of attraction and

attention by backers, which results in an overwhelmingly concentrated cash flow to inside cluster (Shilling 2002; Noe and Parker 2005). In short, the effect of blockbusters on other projects is inconclusive, and both positive and negative spillover effects can be possible. In this regard, we examine the degree to which latent competition leads to cannibalization or spillover effects across clusters and time periods.

Furthermore, the effect can change as blockbusters meet the end of fundraising deadline. For example, assuming that a blockbuster project leads to a cannibalization effect by attracting a considerable amount of backers' interests, similar projects in the same cluster may receive a higher degree of attraction by backers. As a greater number of backers enter the particular cluster, the pledged amount of projects in the cluster can be increased. Thus, we expand our analysis by separating concurrent and lasting effects. These two effects are barely considered in prior research, with the exception of Liu et al. (2015), who suggested investigated concurrent and lasting spillover effects within projects of the same category.

### ***Difference between New Backers and Returning Backers***

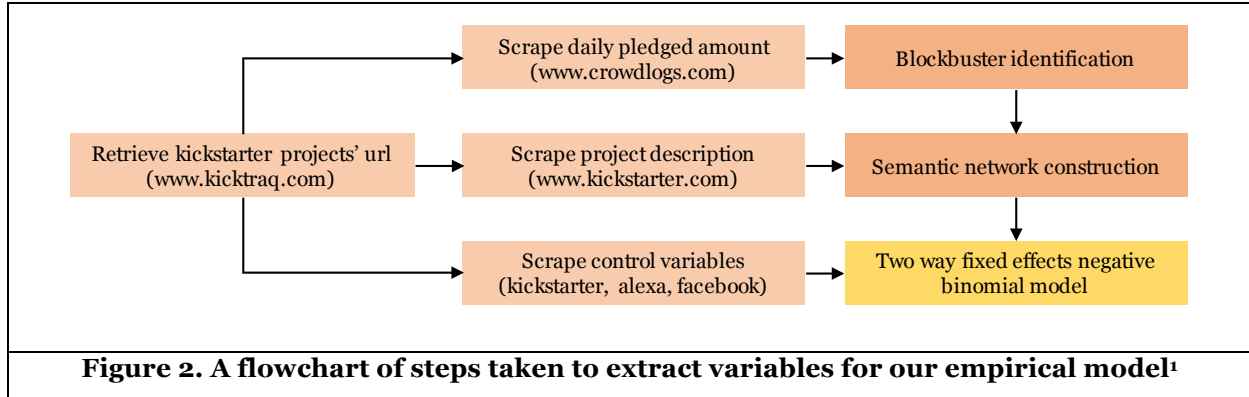
New and returning backers can be divided by past investment experience. The behavior of backers over time may differ from when they first entered the crowdsourcing platform. (Yu et al. 2005). Backers' experience may give an impact to their investment decision, as it influences perception and attitude. (Sheth and Parvatiyar 1995; Montoya-Weiss et al. 2003). Liao et al. (2006) documented that consumers feel more in control and have positive intentions to purchase as they repeat purchasing behavior. By repeating their contributions, returning backers are more likely to feel in control than new backers, which may, in turn, result in an increase in the willingness to contribute.

In another perspective, Hahn and Lee (2013) highlighted that perceived risk of backers would be different according to the past experiences in the platform. Perceived risk is considered as a function of the uncertainty about the potential outcomes of a behavior and represents the uncertainty of consumers about gain or loss in a transaction (Murray 1991). However, the amount of perceived risk will be larger for experienced consumers as they are better at evaluating and understanding the attributes of platforms due to experience with creators (Pavlou 2003, Kim and Gupta 2009). Rodgers et al. (2005) also investigated the different levels of satisfaction between experienced consumers and unexperienced consumers. Accordingly, returning backers are more likely to be affected by the perceived risk and benefits of backing rather than platform attributes and seller attributes like reputation or brand names due to their previous experiences on the crowdfunding platform (Ward and Lee 1999). On the other hand, new backers have fewer experiences of participating in backing the crowdfunding projects than returning backers, so they will have less amount of perceived risk. This will make new backers likely to be affected by the platform attributes or seller attributes and more dependent on the network effects (Hahn and Lee 2013, Cheung et al. 2003, Ward and Lee 1999). In this regard, we conjecture that the level of the perceived risk may be different between returning backers and new backers. Returning backers may rely more on their previous backing experiences and the history of project, whereas the investment decision of new backers are more likely to be affected by their social friends, which make new backers feel safer projects. This will make them exhibit herding behaviors and find projects with higher quality as less experienced consumers consider information quality more important (Rodgers et al. 2005). Thus, it can appear that new backers may search and wander other clusters rather than stay in one cluster in order to find the projects with higher quality.

## **Empirical Method**

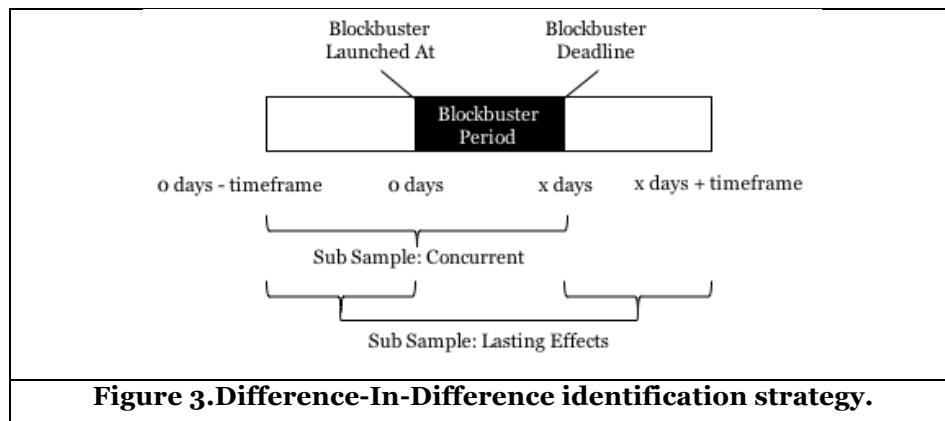
### ***Data Sources & Preprocessing***

In this section, we outline the procedure adopted for data collection and preprocessing. We compiled data into a daily project level dataset from Kickstarter.com and various tracking sites. Our dataset covers a two-year period from March 1, 2014 to February 29, 2016, and contains a total of 148,398 Kickstarter campaigns that ended within this period. These projects received US\$ 1.221 billion in total from 13 million contributions from backers. The procedures used to collect our data are described in Figure 2. Firstly, we identified the URLs of Kickstarter campaigns from www.kicktraq.com, because Kickstarter.com only shows currently-active projects and provides limited search functionality. Using this URL information, we retrieved project information from Kickstarter.com and historical daily-level pledged amounts from Crowdlogs.com. We then merged this data set with the data from Alexa.com to control for platform-level popularities.



Among 148,398 projects, we only considered completed projects, excluding 205 purged projects, 156 projects stopped by Kickstarter due to intellectual property disputes, 17,101 canceled projects, and 1,268 suspended projects. Additionally, we removed 507 projects that did not have any text descriptions which made it impossible to construct a semantic network. To the best of our knowledge, Crowdlogs.com was the only website that enabled us to collect historically pledged amounts of projects. Among 148,398 projects, 25,519 were missing, and 995 projects had improperly designed URLs. However, these missing projects may have had a very marginal impact on other projects due to having \$2,889 of pledged amounts and attracting merely 33.69 backers on average. (For the entire sample, the mean of pledge amounts is \$9,817 and the mean of the number of backers is 111.98.) In summary, our study used a dataset of 106,801 projects and 2,915,821 observations.

A natural way to identify the effect of blockbusters is to examine changes when blockbusters come in or pass away, relative to when blockbusters do not exist. Based on this basic idea, we borrow the difference-in-difference specification by constructing the sub-samples for each of concurrent effect and lasting effect. For better understanding, we describe the method in Figure 3. We first look at how blockbusters concurrently affect the projects in the inside cluster and outside clusters using 120 days of data prior to the start of the blockbusters to its end date.<sup>2</sup> Secondly, we construct the sub-samples to see lasting effect of blockbusters using the data prior to the start of the blockbusters and after its deadline to examine how the blockbuster effects are evolving as time goes by. Also for each sample, we drop the blockbuster projects because our research focuses on the effect of blockbusters to non-blockbuster projects. Our difference-in-difference identification strategy enables us to reduce concerns about the effect of missing data by accounting for the trends in the crowdfunding industry and time trends by restricting the time window only to nearby existing blockbusters.

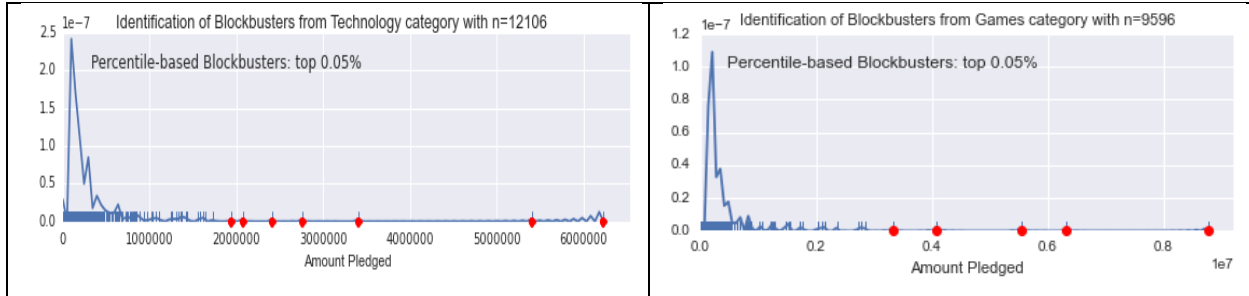


<sup>1</sup> We employ a series of filters to build our set of words and phrases based on Natural Language Toolkit (NLTK 3.0)

<sup>2</sup> We select 120 days to safely cover the life cycle of crowdfunding projects. Also, we conduct robustness checks with various timeframes. The results show the high level of consistency regardless of the choice of timeframes.

**Definition of Blockbusters**

In this study, we define blockbusters as the top 0.05% projects ordered by the total pledged amount, similar to Liu et al (2015).<sup>3</sup> As a result, we identify 54 blockbusters from 5 different categories including Design, Film & Video, Food, Games and Technology. Descriptive statistics for these 54 blockbusters are summarized in Table 1, indicating that a blockbuster received a mean of \$3,226,728.20 pledges from a mean of 25,644.85 backers. These top 0.05% of projects account for 16.47% of total pledged amounts in our study period.



**Figure 4. Identification of Categorical blockbusters in Technology and Games categories.**

Note: Red dots indicate blockbusters and blue lines show the density of projects

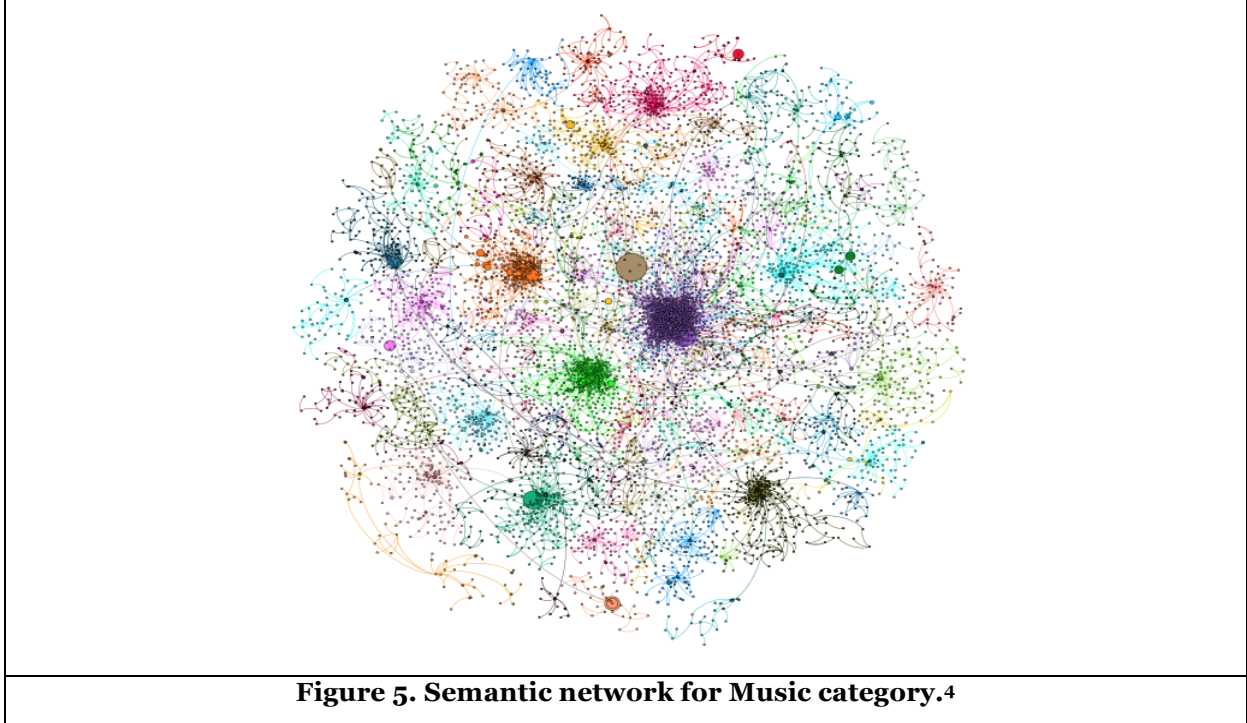
	<b>Goal</b>	<b>Total Pledged</b>	<b>Backers Count</b>	<b>New Backers</b>	<b>Returning Backers</b>
Count	54	54	54	54	54
Mean	286,479.00	3,226,728.20	25,644.85	9,512.42	16,132.42
Std	457,818.28	3,323,469.33	35,445.95	15,739.34	20,709.8
Min	10,000	1,363,381	1,069	199	461
25%	50,000	1,549,272.12	8,337.25	1,577	4,653
50%	100,000	1,863,453.29	12,330	4,145.5	8,881
75%	237,500	3,322,761.12	34,359	8,690.5	23,051.75
Max	2,000,000	20,338,986.27	219,382	95,586	123,796

**Semantic Network Construction**

To reveal the latent market structure, we compile various text mining and clustering techniques in a novel method which involves five main steps for extracting clustering information for each project within the same category. The proposed method enables us to divide projects into disjoint clusters. Also, unlike other clustering algorithms, it provides the relationship between clusters in simultaneously. Figure 5 represents our results of clustering in the Music category. The size of the nodes shows total pledged amounts, and the clustering information is shown by different colors in the figure. Based on the clustering results, we investigate the effects of blockbusters on projects located on either the inside or the outside clusters.

<sup>3</sup>We used four different outlier detection approaches: 1) median average distance based; 2) standard deviation with log-normal distribution assumption based; 3) market share based; and 4) percentile based approaches. The percentile based approach provided us the most consistent results across categories which may have different underlying distributions of pledged amounts.





**Figure 5. Semantic network for Music category.<sup>4</sup>**

The detailed procedures are as follows: we construct our network by filtering out the unimportant similarities based on the minimal spanning tree algorithm and clustering them using the Louvain method. Therefore, we expect only the projects closely related are in the same cluster.

*Step 1. Preprocessing:* At first, we remove stop words from documents and tokenize documents into words using NLTK 3.0, which are statistical Natural Language Processing (NLP) libraries that include WordNet, a lexical database for finding a conceptual relationship between words such as hypernyms, hyponyms, synonyms, and antonyms, etc.

*Step 2. Extracting important words using Tf-idf (Term frequency inverse document frequency):* If a word appears frequently in a document, then we count it as an important word. However, if a word appears in many documents, then it's not a unique identifier. For instance, 'a' or 'the' appear many times in many documents, but these words are not helpful for finding the similarity of documents. So we reduce the importance of these spurious words by giving them low scores.

*Step 3. Measuring distance:* We measure the distance between documents by converting the vector of correlation of words with tf-idf scores. To do this, we use document correlation which is defined as:

$$\rho_{ij} = \frac{\langle Y_i Y_j \rangle - \langle Y_i \rangle \langle Y_j \rangle}{\sqrt{(\langle Y_i^2 \rangle - \langle Y_i \rangle^2)(\langle Y_j^2 \rangle - \langle Y_j \rangle^2)}} \quad (1)$$

$Y_i$  and  $Y_j$  are the term vector of documents for the project  $i$  and  $j$ . The similarity of two documents corresponds to the correlation between the vectors. This cosine similarity is one of the most popular similarity measure applied to text documents as used in various literature (Baeza-Yates and Ribeiro-Neto 1999; Larsen and Aone 1999). Then correlation distance are derived by converting this document correlation coefficients using (2) which enable correlation distance to fulfill the three axioms that define a metrics in Mantegna (1999):

<sup>4</sup>Regarding with the validation of cluster, since there is no label to test the validity of the clusters, we can't measure how well the clusters are formed. Therefore, to address this issue, we will conduct a survey by asking whether the cluster is reasonable or not for the future research.

$$d(i, j) = \sqrt{2(1 - \rho_{ij})} \quad (2)$$

This satisfies  $-1 \leq \rho_{ij} \leq 1$  and  $0 \leq d(i, j) \leq 2$ . If the correlation coefficient closes to 1, then correlation distance also closes to 0. Conversely, if the correlation coefficient closes to -1 then distance closes to 2.

*Step 4. Construct semantic network:* Using the correlation distance, we can build a complete network where every node is connected with one another. However, this complete network has certain limitations to interpreting because its high dimensionality and complexity make it impossible to derive a meaningful relationship. Therefore, we filter out unimportant edges of the network using a hierarchical clustering method called a minimal spanning tree (MST). The MST is widely used in various research areas including biology, physics and mathematics, and it shows salient advantages by revealing underlying structures of the complete graph (Dusser et al. 1987; Mantegna 1999; Onnela et al. 2002; Tumminello et al. 2007). The key idea of MST is that it keeps only the important edges which satisfy the tree structure.

*Step 5. Clustering.* After filtering out the less important edges, we apply widely-used community detection algorithm, the Louvain method, developed by Blondel et al. (2008). The Louvain method is one of the representative community detection algorithms for large scale data. Moreover, as measured by modularity, the quality of the communities detected is turned out to be good. (Vincent et al. 2008)

The results of this procedure have very similar results with the network using cosine similarity with maximally filtered graph. Checking the validity of our proposed cluster-based approach is an important issue. To do this, we will conduct a survey to check the validity of our approach more rigorously in the near future.

### **Empirical Model**

Our empirical evaluations address the effects of blockbusters on the subsequent capital pledged to other projects. We estimate the concurrent effect of blockbusters:

$$Pledged_{it} = \beta_1 * InsideCluBB_{it} + \beta_2 * OutsideCluBB_{it} + \beta_3 * OutsideCatBB_{it} + \beta_4 * NumProInsideCluBB_{it} + \beta_5 * NumProOutsideCluBB_{it} + \beta_6 * NumProOutsideCatBB_{it} + X_{it} + \theta_i + \gamma_t + \varepsilon_{it} \quad (3)$$

where  $i$  indexes the project, and  $t$  indexes the day. The dependent variable,  $Pledged_{it}$ , is the daily-level pledged amount.  $InsideCluBB_{it}$  is an indicator variable that equals one if a project  $i$  is in the same cluster where the blockbuster is placed in day  $t$ , and zero otherwise. Similarly,  $OutsideCluBB_{it}$  equals one if a project  $i$  is in the cluster where the blockbuster is not placed in day  $t$ , and zero otherwise.  $OutsideCatBB_{it}$  also equals one if a project  $i$  is in the category where the blockbuster does not belong to in day  $t$ , and zero otherwise. For example, if a blockbuster exists in the Games category, then there can be a group of projects in the same cluster with the blockbuster. These projects are classified as projects in the inside cluster. At the same time, there can be the other group of projects in the same category, and these are not in the same cluster with the blockbuster. These projects are classified as projects in the outside cluster. In other words, the former group receives one for  $InsideCluBB_{it}$ , and the latter group receives one for  $OutsideCluBB_{it}$ . Also, there can be a set of projects in other categories (e.g., Fashion or Technology category) receives zeros for  $InsideCluBB_{it}$  and  $OutsideCluBB_{it}$ , and receives one for  $OutsideCatBB_{it}$ .

To control for unobserved factors, we include two-way fixed effects in the model. The term  $\theta_i$  denotes project-level fixed effects, and  $\gamma_t$  denotes time fixed effects including yearly and monthly dummies. We also include a vector of time-varying control variables, which is denoted as  $X_{it}$ . Specifically,  $X_{it}$  includes platform-level popularity measured as the time trend of website ranking, the number of page view per million, the number of page view per user and the number of reach per million. For cluster-level controls, we include the normalized number of projects inside the cluster and outside of the cluster, denoted as  $NumProInsideCluBB_{it}$ ,  $NumProOutsideCluBB_{it}$  and  $NumProOutsideCatBB_{it}$ .

We also add a project-level control variable, the achievement rate in the day  $t-1$ . The coefficients of interest are  $\beta_1$ ,  $\beta_2$  and  $\beta_3$ . The term  $\beta_1$  shows the differential impact in the inside cluster that has the blockbusters compared to the outside of the clusters. The term  $\beta_2$  represents the impact of the blockbusters in the outside cluster where the blockbusters are located in the same category. The term  $\beta_3$  explains the effect of blockbuster projects to the projects outside the category where the blockbusters belong to.

Next, we examine the lasting effect of the blockbusters with the following model:

$$Pledged_{it} = \beta_1 * LastingInsideCluBB_{it} + \beta_2 * LastingOutsideCluBB_{it} + \beta_3 * LastingOutsideCatBB_{it} + \beta_4 * NumProInsideCluBB_{it} + \beta_5 * NumProOutsideCluBB_{it} + \beta_6 * NumProOutsideCatBB_{it} + X_{it} + \theta_i + \gamma_t + \varepsilon_{it} \quad (4)$$

In this model specification, *LastingInsideCluBB<sub>it</sub>* and *LastingOutsideCluBB<sub>it</sub>* have values of one if the blockbuster projects are within the suggested time frame in the same cluster or not, respectively. Note that the time frames include 6 different periods from 30 to 180 days with the interval of 30 days. *LastingOutsideCatBB<sub>it</sub>* receives the value of one if the blockbusters are finished within the same time frame and the projects are outside the category of blockbusters.

In addition, we extend Model (4) to Model (5) by including a variable specifying the composition of new and returning backers:

$$Pledged_{it} = \beta_1 * (InsideCluBB_{it} * NewReturnRatio_{it}) + \beta_2 * (OutsideCluBB_{it} * NewReturnRatio_{it}) + \beta_3 * (OutsideCatBB_{it} * NewReturnRatio_{it}) + \beta_4 * InsideCluBB_{it} + \beta_5 * OutsideCluBB_{it} + \beta_6 * OutsideCatBB_{it} + \beta_7 * NumProInsideCluBB_{it} + \beta_8 * NumProOutsideCluBB_{it} + \beta_9 * NumProOutsideCatBB_{it} + X_{it} + \theta_i + \gamma_t + \varepsilon_{it} \quad (5)$$

Specifically, we include interaction terms regarding *NewReturnRatio<sub>it</sub>* in Model (5) to see the different impact of blockbusters according to the type of backers inside the blockbusters.

## Results

### Concurrent Effect of Blockbuster Projects

To estimate the model, we employ *negative binomial model* because our daily pledged amount that a project receives follows the highly skewed distributions and has a probability mass at zero because many projects do not achieve any funding on a given day (Doshi 2014; Kim et al. 2014).<sup>5</sup> Furthermore, our dependent variable, the pledged amount, presents non-negative integer values by discarding the negligible decimal point of dollar values.

We think that a conditional fixed-effects negative binomial model is more appropriate than a poisson estimator, because it is more suitable for two reasons: 1) the distribution is quite over-dispersed (mean: 304.8124, variance: 4771.041); and 2) the considerable number of projects (7,867) received no funding at all, resulting in a significant number of zeros in our dependent variable.

We first look at how blockbusters concurrently affect the projects in inside blockbuster clusters, outside blockbuster clusters and outside categories of blockbusters using the data from 120 days before the start of the blockbusters to its end date by analyzing the results with equation (3) which adds more control variables to the base model. Table 2 shows the results. In Column (1), we include control variables only for project level fixed effects. In Column (2), we include time-varying controls. We then include time-fixed effects in Column (3).

DV : Pledged			
VARIABLES	(1)	(2)	(3)
InsideCluBB	<b>0.222***</b> (17.10)	<b>0.226***</b> (17.37)	<b>0.241***</b> (18.45)
OutsideCluBB	<b>0.117***</b> (57.69)	<b>0.122***</b> (59.02)	<b>0.134***</b> (59.06)
OutsideCatBB	<b>-0.0204***</b> (-22.22)	<b>-0.0191***</b> (-19.57)	<b>-0.0097***</b> (-7.22)
NumProInsideClu	<b>0.0097***</b> (7.37)	<b>0.0063***</b> (4.79)	<b>0.0044**</b> (3.29)

<sup>5</sup>We also estimate the model of main results using fixed effect panel regression. All noncategorical variables in our model are log transformed as in Burtch et al (2013). The results are consistent from negative binomial estimation by having much greater and positive impact on inside cluster.

NumProOutsideClu	<b>-0.0428***</b> (-31.35)	<b>-0.0449***</b> (-32.50)	<b>-0.0498***</b> (-35.54)
NumProOutsideCat	<b>-0.0404***</b> (-29.13)	<b>-0.0507***</b> (-30.98)	<b>-0.0643***</b> (-33.31)
Year FE	No	No	Yes
Month FE	No	No	Yes
Time-varying Controls	No	Yes	Yes
Project level FE	Yes	Yes	Yes
Observations	1,893,102	1,873,042	1,873,042
Wald Chi2	7,345.9	11,136.8	15,355.79
Note: t statistics in parentheses Significance Level: * p<0.05, ** p<0.01, *** p<0.001			

The result in Table 2 shows that there is a positive spillover effect of blockbuster projects inside the cluster. Similarly, blockbuster projects outside the cluster also show a positive spillover effect. Meanwhile, there is a cannibalization effect of blockbuster projects outside the category with negative coefficients in the base model. The result of blockbuster effects among those three was significant. As more control variables are considered, the spillover effect of blockbusters both inside and outside the cluster increased with the increase of coefficients. The cannibalization effect of blockbusters outside the category also increased when all control variables are considered. Our category level results are consistent with prior studies by having spillover effects on the same category and cannibalization effects to the other categories. However, blockbusters show larger spillover effects to the projects in the same cluster than outside clusters. The numbers of projects on inside cluster and in other clusters oppositely influence the pledged amount of projects. From model (1) to model (3), our findings are consistent, with the slight change of coefficients.

### Lasting Effect of Blockbuster Projects

We now estimate the lasting effect of blockbusters with sub-samples of 6 different time periods, which include time periods from 30 days to 180 days before and after the lifetime of blockbusters with the interval of 30 days respectively. The variables of interest are *LastingInsideCluBB*, *LastingOutsideCluBB*, and *LastingOutsideCatBB*. Table 3 reports our findings, and we show the results of 6 different time frames in a similar manner with the Column(3) in Table 2.

DV : Pledged	Sub Samples					
VARIABLES	(1) 30 day	(2) 60 days	(3) 90 days	(4) 120 days	(5) 150 days	(6) 180 days
LastingInsideCluBB	<b>0.331***</b> (19.40)	<b>0.221***</b> (19.69)	<b>0.170***</b> (19.65)	<b>0.138***</b> (18.65)	<b>0.128***</b> (19.01)	<b>0.129***</b> (20.83)
LastingOutsideCluBB	<b>0.135***</b> (47.50)	<b>0.0866***</b> (51.86)	<b>0.0559***</b> (41.45)	<b>0.0289***</b> (23.11)	<b>0.0143***</b> (12.24)	<b>0.0134***</b> (11.49)
LastingOutsideCatBB	<b>-0.0253***</b> (-15.37)	<b>-0.0281***</b> (-28.81)	<b>-0.0315***</b> (-34.60)	<b>-0.0421***</b> (-43.95)	<b>-</b> (-45.26)	<b>-0.0359***</b> (-36.15)
NumProInsideClu	<b>0.0077***</b> (4.75)	<b>0.0077***</b>	<b>0.0064***</b>	<b>0.0056***</b>	<b>0.0046***</b> (3.46)	<b>0.0041**</b> (3.12)
NumProOutsideClu	<b>-0.0219***</b> (-12.74)	<b>-0.0381***</b> (-25.70)	<b>0.0449***</b> (-31.86)	<b>0.0452***</b> (-31.82)	<b>0.0435***</b> (-30.53)	<b>0.0448***</b> (-31.49)
NumProOutsideCat	<b>0.0226***</b> (8.14)	<b>-</b> <b>0.0205***</b> (-10.10)	<b>-</b> <b>0.0265***</b> (-14.97)	<b>-0.0141***</b> (-7.76)	<b>-0.0091***</b> (-4.91)	<b>-0.0176***</b> (-9.71)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Time-varying Controls	Yes	Yes	Yes	Yes	Yes	Yes
Project level FE	Yes	Yes	Yes	Yes	Yes	Yes

Observations	1,299,997	1,743,204	1,937,837	1,937,959	1,937,959	1,937,959
Wald Chi2	10,937.0	13,262.4	14,207.6	15,217.0	14,982.3	14,193.1
Note: t statistics in parentheses Significance Level: * p<0.05, ** p<0.01, *** p<0.001						

Table 3 shows that there are positive spillover effects of blockbusters on both inside and outside the clusters, but the cannibalization effect on projects outside the blockbuster category. All results are statistically significant. The lasting effect of blockbuster projects both inside and outside the cluster tend to decline as the timeframe increases. Also, the lasting effect of blockbusters shows the much faster decrease of spillover effects on the outside cluster than inside cluster. This finding may be interpreted that the category containing blockbusters may receive greater attention from backers, so the projects in the same category are more likely to achieve the greater pledged amount within a certain length of the time period. In contrast, the projects outside the category are less likely to receive attentions when the blockbusters exist in a particular category.

Our findings thus far can be summarized as follows: 1) blockbuster projects show positive spillover effects to the projects inside and outside the cluster regardless of concurrent and lasting models, but they give a cannibalization effect to the projects outside the category; and 2) The number of projects inside the cluster increases, whereas the number of projects outside the cluster decreases. In other words, blockbusters may give a positive signal to both project creators and backers in the same category. However, blockbusters may give a negative impact on projects in other category in terms of gathering backers. This finding may imply that the interest and attention of backers can be concentrated when the blockbuster is in place in a particular category, and the spillover effect may not be extended to the projects that are far from the blockbusters.

**Contribution Behavior of New Backers Relative to Returning Backers**

Table 4 reports results about how the composition of backers is related to the pledged amount. We find that spillover effects of blockbuster projects on inside clusters are larger when there are more returning backers than new backers. However, it turns out to be opposite when it comes to projects outside the clusters and outside category. They receive greater spillover effects by blockbusters when there are more new backers in the blockbuster projects. To interpret the results, it can be conjectured that returning backers might have previous experience and knowledge of backing; therefore, they are more likely to be aware of the presence of blockbuster projects compared to new backers, especially inside the clusters. Thus, returning backers inside the clusters might be more affected by blockbuster projects as they already judge that it would be safe to be in the clusters with those blockbusters according to their previous experience. This may result in the positive spillover effect of blockbuster projects inside the cluster. Unlike returning backers, new backers may place importance on the quality and uniqueness of the project itself, and they are less likely to be stick to the same cluster. This finding can also be related to the fact new backers are not as good as returning backers at understanding of the crowdfunding platforms as they have fewer experiences than returning backers (Kim and Gupta 2009). As a result, new backers will not be highly dependent on the blockbuster projects inside the cluster and they may easily move to other projects outside the cluster or the category. However, returning backers show quite different contribution behaviors from new backers, so the effect of blockbusters to the projects outside the cluster and category tends to be more significant when there are more new backers than returning backers.

DV : Pledged	(1)	(2)	(3)
InsideCluBB * NewReturnRatio	<b>-0.0499***</b> (-7.12)	<b>-0.0568***</b> (-8.09)	<b>-0.0540***</b> (-7.67)
OutsideCluBB * NewReturnRatio	<b>0.0322***</b> (25.35)	<b>0.0327***</b> (25.73)	<b>0.0229***</b> (17.54)
OutsideCatBB * NewReturnRatio	<b>0.0014***</b> (3.82)	<b>0.0015***</b> (3.93)	<b>0.0010**</b> (2.68)
InsideCluBB	<b>0.142***</b> (19.24)	<b>0.144***</b> (19.47)	<b>0.141***</b> (18.95)

OutsideCluBB	<b>0.0312***</b> (33.30)	<b>0.0319***</b> (31.18)	<b>0.0275***</b> (21.96)
OutsideCatBB	<b>-0.0392***</b> (-78.67)	<b>-0.0382***</b> (-59.70)	<b>-0.0423***</b> (-44.17)
NumProInsideClu	<b>0.0075***</b> (5.66)	<b>0.0074***</b> (5.57)	<b>0.0072***</b> (5.42)
NumProOutsideClu	<b>-0.0436***</b> (-31.27)	<b>-0.0469***</b> (-33.46)	<b>-0.0454***</b> (-31.93)
NumProOutsideCat	<b>-0.0053***</b> (-3.88)	<b>-0.0236***</b> (-15.40)	<b>-0.0151***</b> (-8.33)
Year FE	No	No	Yes
Month FE	No	No	Yes
Time-varying Controls	No	Yes	Yes
Project level FE	Yes	Yes	Yes
Observations	1,893,102	1,873,042	1,873,042
Wald Chi2	7,345.9	11,136.8	13,237.2
Note: t statistics in parentheses Significance Level: * p<0.05, ** p<0.01, *** p<0.001			

## Discussion & Conclusion

### *Factors Attracting More New Backers*

We further examine which factors are associated with attracting more backers, and the results are presented in Table 5. In Columns (1) and (2), most characteristics (the length of project description funding period, photos and videos, project social exposure, staff pick, the number of creator's social friends ranking) are positively related to the number of new or returning backers attracted. Both of them have more pledged amount when the rank of projects are higher, which has the lower number literally. Also, in line with our findings in the previous models, blockbusters attract more new and returning backers. In particular, the coefficients of blockbusters are larger for new backers, indicating that new backers are more highly affected by blockbuster projects than returning backers. New backers prefer the successful project, because they are more likely to be reluctant to taking risks due to the lack of experiences. Next, Column (3) uses the ratio of new and returning backers as a dependent variable in order to see the impact on the composition of backers. Our results show that new backers seem to have more interests in projects which have a longer description related to the risk of projects. But returning backers are attracted by a longer description of the project itself. Also, new backers are attracted by a larger number of the creator's friends on social media. This finding might support an argument that the certain number of new backers can be directly invited through the social media channel, and this path may alleviate the perceived risk imposed to the new backers. Unlike the new backers, returning backers have the better understanding of the attributes and specific parts of crowdfunding projects, so they do not have to rely on the description about the risk. Also, returning backers are more attracted by backing and creating history of creators. In other words, returning backers may put a bigger emphasis on reciprocity than new backers, as the past backing experience can be related to their next backing decision. Thus, participating in backing may be an essential part for creators in order to attracting more returning backers in the next period.

Anecdotally, our findings are consistent with the evidence from social psychology. The new entry may bring with the fear of risk. In this regard, in order to avoid potential risk, new backers are more likely to follow what the majority of people do. In our case, a blockbuster project can be regarded as the output of the investment that the majority of people made. To avoid risk and examine how a project proceeds, newcomers may participate in backing one of blockbuster projects. If there are detailed descriptions presenting the risk of a project, new backers may feel more comfortable with the blockbuster project than other projects. There is also a possibility that new backers enter the crowdfunding platform through a link provided by one of their Facebook friends. This intimacy may also be related to the entry incentive of new backers, which results in a high ratio of new backers.

**Table 5. Different Factors that Attract Different Types of Backer**

VARIABLES	(1) DV: <i>new_backers</i>	(2) DV: <i>returning_backers</i>	(3) DV: <i>new_return_ratio</i>
BB_Indicator	<b>4.684***</b> (7.53)	<b>2.740***</b> (9.19)	<b>39.55*</b> <b>(2.08)</b>
Goal	<b>3.996*</b> (2.12)	1.574 (1.18)	<b>20.70*</b> (2.24)
Desc_Length	<b>0.0975***</b> (6.36)	<b>0.135***</b> (7.49)	<b>-6.292***</b> (-3.79)
Risk_Desc_Length	<b>0.0396***</b> (3.50)	0.00675 (0.52)	<b>6.584***</b> (4.07)
Funding_Duration	<b>0.0570***</b> (3.62)	<b>0.0540**</b> (3.01)	<b>4.327*</b> (1.99)
Project_Social_Exposure	<b>0.0698***</b> (6.59)	<b>0.0502***</b> (3.95)	<b>-4.189**</b> (-2.76)
Num_Video	<b>0.123***</b> (8.02)	<b>0.102***</b> (5.34)	-2.214 (-1.39)
Num_Photo	<b>0.243***</b> (14.82)	<b>0.544***</b> (34.21)	<b>-31.07***</b> (-20.87)
Backing_Histroy	<b>0.0078</b> (0.87)	<b>0.245***</b> (14.53)	<b>-13.69***</b> (-11.90)
Creating_History	<b>-0.147***</b> (-13.85)	<b>0.224***</b> (14.99)	<b>-37.01***</b> (-26.92)
Num_Creator_Friends	<b>0.219***</b> (19.49)	<b>0.0994***</b> (9.19)	<b>5.495**</b> (3.24)
Staff_Pick	<b>0.610***</b> (18.68)	<b>0.979***</b> (33.09)	<b>-91.03***</b> (-33.78)
Ranking	-0.0020 (-0.23)	-0.0106 (-0.99)	<b>-3.709**</b> (-2.68)
Constants	<b>3.604***</b> (59.79)	<b>3.575***</b> (79.26)	<b>228.7***</b> (100.61)
Inalpha_cons	<b>0.117***</b> (8.45)	<b>0.267***</b> (20.62)	
Number of Observations	32,456	32,456	32,260
R2			0.080
Pseudo R2	0.047	0.072	
Note: t-statistics are in parentheses Significant Levels : ***, ** and * denote significance a 0.01, 0.05 and 0.1 respectively			

## Conclusion

Our findings highlight that it is necessary to consider the similarity between blockbusters and other projects in order to analyze the impact of blockbusters more precisely and rigorously. That is, our main findings suggest blockbusters give a positive spillover effect to projects in the same category, but the magnitudes are largely different according to the distance from the blockbuster. For this reason, it is important to take the proximity into consideration when we measure the impact of blockbusters, but previous studies did not account for the difference. More specifically, projects in the outside cluster of the same category are less likely to be affected by the blockbuster project in the short run, but the effect becomes greater as time progresses. In addition, our findings suggest that it is important to consider the characteristics of

blockbusters by showing the differential effect of the composition of new and returning backers. As emphasized earlier, this may be the first study to untangle the impact of blockbuster projects in the context of crowdfunding by the use of semantic network analysis.

In the perspective of managerial implications, our result can help managers who are going to open a new project. An entrepreneur planning to initiate a new project similar to an ongoing blockbuster project may take advantage of the opportunity. They can see whether there is a similar project with theirs and decide to keep going on creating the project or to stop. They can strategically decide where to create their projects according to the projects in the inside-cluster group and the outside-cluster group. Besides, managers can think about the composition of new and returning backers of blockbuster projects to determine the characteristics of a blockbuster, which may, in turn, provide a varying degree of the spillover effect on other projects.

From the perspective of the platform operators, they have several goals to achieve – e.g., attracting more new participants to the platform, maintaining the new participants' activities, and retaining existing participants. We believe that a small fraction of blockbuster projects has a clearly noticeable effect on other projects and the platform, so it is essential to explore the impact of blockbusters to a greater detail. Indeed, our empirical findings show that it is necessary to analyze at the project-level rather than the category-level to find useful implications.

We suggest several avenues of attention for future work to address. First, our novel methodology enables us to extend the analysis by maintaining the relationship between clusters with the tree structure. Based on this network structure, as Granovetter (1973) investigated, we will further consider more detailed topological characteristics. For instance, we may define the creativity or innovativeness of projects based on the centrality measures which also can affect the consequence population and contribution behavior of backers. Second, our project-level analysis can be generalized with category-specific attributes, and idiosyncratic features may be related to the impact of blockbusters in the category. Third, in this study, we focus on the demand side analysis, but the blockbusters also affect the supply side behavior, so we will extend our analysis to cover the changes of supply side responses which might be different from backers behavior. Lastly, to strengthen our findings, we will suggest a series of robustness checks.



## References

- Abdellatif, T. 2015. "A Semantic Study of the Recommendation Process for a Prestation of Touristic Services," SSRN Electronic Journal (April) (doi: 10.2139/ssrn.2703520).
- Academy, T., and Journal, M. 2002. "Technology Success and Failure in Winner-Take-All Markets : The Impact of Learning Orientation , Timing , and Network Externalities Author ( s ): Melissa A . Schilling Source : The Academy of Management Journal , Vol . 45 , No . 2 ( Apr . , 2002 ), pp . 38," The Academy of Management Journal (45:2), pp. 387–398.
- An, J., Quercia, D., and Crowcroft, J. 2014. "Recommending Investors for Crowdfunding Projects," Arxiv - Computers & Society, pp. 261–269 (doi: 10.1145/2566486.2568005).
- Anderberg, M. R. 2014. Cluster analysis for applications: probability and mathematical statistics: a series of monographs and textbooks (19). Academic press.
- Archak, N., Ghose, A., and Ipeirotis, P. G. 2011. "Deriving the Pricing Power of Product Features by Mining Consumer Reviews," Management Science (57:8), pp. 1485–1509 (doi: 10.1287/mnsc.1110.1370).
- Bash, E. 2015. "No Title No Title," PhD Proposal (1) (doi: 10.1017/CBO9781107415324.004).
- Baeza-Yates, R., and Ribeiro-Neto, B. 1999. "Modern information retrieval," New York (9), p. 513 (doi: 10.1080/14735789709366603).
- Beaulieu, T. Y., and Sarker Suprateek. 2013. "Discursive Meaning Creation in Crowdfunding : A Socio-material Perspective," Thirty Fourth International Conference on Information Systems, pp. 1–10.
- Belleflamme, P., Lambert, T., and Schwienbacher, A. 2014. "Crowdfunding: Tapping the right crowd," Journal of Business Venturing (29:5), pp. 585–609 (doi: 10.1016/j.jbusvent.2013.07.003).
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. 2008. "Fast unfolding of communities in large networks," Journal of Statistical Mechanics: Theory and Experiment (10008:10), p. 6 (doi: 10.1088/1742-5468/2008/10/P10008).
- Bretschneider, U., Knaub, K., and Wieck, E. 2014. "Motivations for Crowdfunding : What Drives the Crowd To Invest in Start-Ups?," Twenty Second European Conference on Information Systems, pp. 1–11.
- Buttice, V., Milano, P., Rossi-lamastra, C., and Milano, P. 2015. "Paper to be presented at Segmented receivers ? preferences and the response to signals . An empirical test in film-crowdfunding Massimo Colombo," (2015).
- Brynjolfsson, E., Hu, Y. J., and Smith, M. D. 2010. "Long Tails vs. Superstars: The effect of information technology on product variety and sales concentration patterns," Information Systems Research (21:4), pp. 736–747 (doi: 10.1287/isre.1100.0325).
- Brynjolfsson, E., and Kemerer, C. F. 1996. "Network externalities in microcomputer software: An econometric analysis of the spreadsheet market," Management Science (42:12), pp. 1627–1647 (doi: 10.1287/mnsc.42.12.1627).
- Burtch, G., Ghose, A., and Wattal, S. 2013. "An Empirical Examination of the Antecedents and Consequences of Contribution Patterns in Crowd-Funded Markets An Empirical Examination of the Antecedents and Consequences of Contribution Patterns in Crowd-Funded Markets," Information Systems Research (24:3), pp. 499–519 (doi: http://dx.doi.org/10.1287/isre.1120.0468).
- Burtch, G., Ghose, A., and Wattal, S. 2015. "The Hidden Cost of Accommodating Crowdfunder Privacy Preferences: A Randomized Field Experiment," Management Science (61:5), pp. 949–962 (doi: 10.1287/mnsc.2014.2069).
- Cameron, A. C., and Trivedi, P. K. 2005. Microeconometrics - Methods and Applications (doi: 10.1007/s13398-014-0173-7.2).
- Cheung, C.M.K., Zhu, L., Kwong, T., Chan, G.W.W. and Limayem, M. 2003. "Online consumer behavior: a review and agenda for future research," Proceedings of the 16th Bled eCommerce Conference, Bled, Slovenia.

- Cropley, A. 2000. "Defining and measuring creativity: Are creativity tests worth using?," *Roeper Review* (23:2), pp. 72–79 (doi: 10.1080/02783190009554069).
- Collins, A., Hand, C., and Snell, M. C. 2002. "What makes a blockbuster? Economic analysis of film success in the United Kingdom," *Managerial and Decision Economics* (23:6), pp. 343–354 (doi: 10.1002/mde.1069).
- Crowdsourcing.org. 2012. "Crowdfunding industry report - market trends, composition and crowdfunding platforms," Research report (May), pp. 1–30.
- Cumming, D. J., and Johan, S. A. 2014. "Venture Capital and Private Equity Contracting," *Venture Capital and Private Equity Contracting*, pp. 725–727 (doi: 10.1016/B978-0-12-409537-3.00023-2).
- Doshi, A. 2014. "The Impact of High Performance Outliers on Two-Sided Platforms: Evidence from Crowdfunding," Available at SSRN 2422111.
- Dusser, C., Rasigni, M., Palmari, J., Rasigni, G., Llebaria, A., and Marty, F. 1987. "Minimal Spanning Tree Analysis of Biological Structures," *Journal of Theoretical Biology* (125:3), pp. 317–323.
- Fleder, D., and Hosanagar, K. 2009. "Blockbuster Culture's Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity," *Management Science* (55:5), pp. 697–712 (doi: 10.1287/mnsc.1080.0974).
- Fader, P. S., Lodish, L. M. 1990. "A cross-category analysis of category structure and promotional activity for grocery products," *The Journal of Marketing*, pp. 52–65.
- Ghose, A., Smith, M. D., and Telang, R. 2006. "Internet exchanges for used books: An empirical analysis of product cannibalization and welfare impact," *Information Systems Research* (17:1), pp. 3–19 (doi: 10.1287/isre.1050.0072).
- Granovetter, M. 1973. "The Strength of Weak Ties," *The Strength of Weak Ties*, pp. 1360–1380 (doi: 10.1086/225469).
- Hahn, J., and Lee, G. 2013. "Archetypes of Crowdfunders' Backing Behaviors and the Outcome of Crowdfunding Efforts: An Exploratory Analysis of Kickstarter," *Conference on Information Systems and Technology 2013*, pp. 1–18.
- Hirschman, E. C. 1980. "Innovativeness, novelty seeking, and consumer creativity," *Journal of Consumer Research* (7:December), pp. 283–295 (doi: 10.1086/208816).
- Mayer, K., and Bercovitz, J. 2008. "The influence of inertia on contract design: Contingency planning in information technology service contracts," *Managerial and Decision Economics* (163:August), pp. 149–163 (doi: 10.1002/mde).
- Montoya-Weiss M, Voss GB, Grewal D. 2003. "Online channel use and satisfaction in a multichannel service context," *MSI Report* (2:3), pp. 19–35.
- Journal, T., and Economics, I. 2005. "Indirect Network Effects and the Product Cycle : Video Games in the U . S . , 1994-2002 Author ( s ): Matthew T . Clements and Hiroshi Ohashi Published by : Wiley Stable URL : <http://www.jstor.org/stable/3569886> Accessed : 13-04-2016 03 : 12 UTC," *The Journal of Industrial Economics* (53:4), pp. 515–542.
- Katz, B. M. L., and Shapiro, C. 1985. "American Economic Association Network Externalities , Competition , and Compatibility Author ( s ): Michael L . Katz and Carl Shapiro Source : The American Economic Review , Vol . 75 , No . 3 ( Jun . , 1985 ) , pp . 424-440 Published by : American Economic," *The American Economic Review* (75:3), pp. 424–440.
- Katz, Michael L., and C. S. 1994. "Systems competition and network effects," *The Journal of Economic Perspectives*, pp. 93–115.
- Kim, H.W., Gupta, S. 2009. "A comparison of purchase decision calculus between potential and repeat customers of an online store," *Decision Support Systems* (47), pp.477–487.
- Kim, K., and Hann, I.-H. 2013. "Does Crowdfunding Democratize Access to Capital? A Geographical Analysis," *SSRN Electronic Journal*, pp. 1–35 (doi: 10.2139/ssrn.2334590).Kristensson, P.,

- Gustafsson, A., and Archer, T. 2004. "Harnessing the creative potential among users.," *Journal of Product Innovation Management* (21:1), pp. 4–14 (doi: 10.1111/j.0737-6782.2004.00050.x).
- Kim, K., and Viswanathan, S. 2014. "The Experts in the Crowd: The Role of Reputable Investors in a Crowdfunding Market," *TPRC 41: The 41st Research Conference on Communication, Information and Internet Policy* (June), p. 33 .
- Larsen, B., and Aone, C. 1999. "Fast and effective text mining using linear-time document clustering," *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining KDD 99* (5), pp. 16–22 (doi: 10.1145/312129.312186).
- Liao C, Palvia P, Lin H-N. 2006. "The roles of habit and web site quality in e-commerce," *International Journal of Information Management* (26:6), pp. 469–483
- Liu, J., Yang, L., Wang, Z., Hahn, J.. 2015. "Winner Takes All? The 'Blockbuster Effect' in Crowdfunding Platforms," *Thirty Sixth International Conference on Informantion Systems*, pp. 1–11.
- Maccrimmon, K. R., Wagner, C., Maccrimmon, K. R., and Wagner, C. 2016. "Linked references are available on JSTOR for this article : Stimulating Ideas through Creativity Software," (40:11), pp. 1514–1532.
- Mantegna, R. N. 1999. "Hierarchical structure in financial markets," *The European Physical Journal B - Condensed Matter and Complex Systems* (11:1), pp. 193–197 (doi: 10.1007/s100510050929).
- Mollick, E. 2014. "The dynamics of crowdfunding: An exploratory study," *Journal of Business Venturing* (29:1), The Author, pp. 1–16 (doi: 10.1016/j.jbusvent.2013.06.005).
- Mostafa, M. M. 2013. "More than words: Social networks' text mining for consumer brand sentiments," *Expert Systems with Applications* (40:10), pp. 4241–4251 (doi: 10.1016/j.eswa.2013.01.019).
- Murray, KB. 1991. "A test of services marketing theory: consumer information acquisition activities," *The Journal of Marketing* (55:1), pp. 10f25
- Netzer, O., Feldman, R., Goldenberg, J., and Fresko, M. 2012. "Mine Your Own Business: Market-Structure Surveillance Through Text Mining," *Marketing Science* (31:3), pp. 521–543 (doi: 10.1287/mksc.1120.0713).
- Noe, T., and Parker, G. 2005. "Winner take all: Competition, strategy, and the structure of returns in the internet economy," *Journal of Economics and Management Strategy* (14:1), pp. 141–164 (doi: 10.1111/j.1430-9134.2005.00037.x).
- Onnela, J.-P., Chakraborti, A., Kaski, K., and Kertiész, J. 2002. "Dynamic Asset Trees and Portfolio Analysis," *The European Physical Journal B - Condensed Matter* (30:3), pp. 285–288.
- Pavlou, P. A. 2003. "Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model," *International Journal of Electronic Commerce* (7:3), pp. 101-134.
- Pongetti, F. 2011. A crowdsourcing methodology for a semantic sentiment analysis engine (available at <https://www.politesi.polimi.it/handle/10589/25061>).
- Press, T. M. I. T., and Review, T. 1995. "Competing Compatibility Standards and Network Externalities in the PC Software Market Author ( s ): Neil Gandal Source : The Review of Economics and Statistics , Vol . 77 , No . 4 ( Nov ., 1995 ), pp . 599-608 Published by : The MIT Press Stable URL : htt," *The Review of Economics and Statistics* (77:4), pp. 599–608.
- Rakesh, V., and Lee, W. (n.d.). "Probabilistic Group Recommendation Model for Crowdfunding Domains Categories and Subject Descriptors," pp. 257–266 (doi: 10.1145/2835776.2835793).
- Ratneshwar, S., & Shocker, A. D. 1991. "Substitution in use and the role of usage context in product category structures," *Journal of Marketing Research*, pp. 281-295.
- Rodgers W, Negash S, Suk K. 2005. "The moderating effect of online experience on the antecedents and consequences of on-line satisfaction," *Psychological Marketing* (22:4), pp. 313–31.
- Runco, M. a., and Jaeger, G. J. 2012. "The Standard Definition of Creativity," *Creativity Research Journal* (24:1), pp. 92–96 (doi: 10.1080/10400419.2012.650092).

- Ribes-iñesta, E. 2011. "HUMAN BEHAVIOR AS LANGUAGE: SOME Language as Language games," Thirty Second International Conference on Information Systems (121), pp. 109–121.
- Saiz, A., and Simonsohn, U. 2008. "Downloading Wisdom from Online Crowds," SSRN Electronic Journal (3809), pp. 1–33 (doi: 10.2139/ssrn.990021).
- Schindler, R. M., and Bickart, B. 2005. "Published word of mouth: Referable, consumer-generated information on the Internet," *Online consumer psychology: understanding and influencing consumer behavior in the virtual world* (November 2015), pp. 35–40 (doi: 10.4324/9781410612694).
- Shankar, V., and Bayus, B. L. 2003. "Network effects and competition: An empirical analysis of the home video game industry," *Strategic Management Journal* (24:4), pp. 375–384 (doi: 10.1002/smj.296).
- Srikanth AN, "Indiegogo vs Kickstarter: Crowdfunding Platform Wars (2015)." Beebom. N.p., 24 Apr. 2015. Web. 22 Dec. 2015.
- Strahilevitz, L. J. 2003. "Charismatic Code, Social Norms, and the Emergence of Cooperation on the File-Swapping Networks," *Virginia Law Review* (89:3), pp. 505–595 (doi: 10.2139/ssrn.329700).
- Schulz, M., Haas, P., Schulthess, K., Blohm, I., & Leimeister, J. M. 2015. "How Idea Creativity and Hedonic Value Influence Project Success in Crowdfunding," In *Wirtschaftsinformatik*, pp. 948-962.
- Sheth JN, Parvatiyar A. 1995. "Relationship marketing in consumer markets: antecedents and Consequences," *Journal of the Academic Marketing Science* (23:4), pp. 255–271.
- Schulz, M., Haas, P., Schulthess, K., Blohm, I., and Leimeister, J. M. 2015. "How Idea Creativity and Hedonic Value Influence Project Success in Crowdfunding," 12th International Conference on *Wirtschaftsinformatik*, March 4-6 2015, Osnabrück, Germany, pp. 1–15.
- Taylor S, Todd PA. 1995. "Assessing it usage: the role of prior experience," *MIS Quarterly* (19:4), pp. 561–70.
- Tomczak, A., and Brem, A. 2013. "A conceptualized investment model of crowdfunding," *Venture Capital* (15:January 2015), pp. 1–25 (doi: 10.1080/13691066.2013.847614).
- Tumminello, M., Di Matteo, T., Aste, T., and Mantegna, R. N. 2007. "Correlation Based Networks of Equity Returns Sampled at Different Time Horizons," *The European Physical Journal B* (55:2), pp. 209-217.
- Ward, M.R., and Lee, M.J. 1999. *MInternet shopping, consumer search, and product branding*, Working paper, University of Pennsylvania
- Xiao, S., Tan, X., Dong, M., and Qi, J. 2014. "How to Design Your Project in the Online Crowdfunding Market? Evidence from Kickstarter," Thirty Fifth International Conference on Information Systems (Szaky 2011), pp. 1–8.
- Yu, J., Ha, I., Choi, M., Rho, J. 2005. "Extending the Tam for a t-Commerce," *Information & Management* (42:7), pp. 965–976.
- Zhang, J., and Liu, P. 2012. "Rational Herding in Microloan Markets," *Management Science* (58:5), pp. 892–912 (doi: 10.1287/mnsc.1110.1459).
- Zvilichovsky, D., Aviv, T., Aviv, T., and Aviv, T. 2013. "Playing BothSides of The Market: Success and Reciprocity on Crowdfunding" Thirty Fourth International Conference on Information Systems, pp. 1–18.