



9-2013

# An Empirical Examination of the Antecedents and Consequences of Investment Patterns in Crowd-Funded Markets.

Gordon Burtch

Anindya Ghose

Sunil Wattal

Follow this and additional works at: [http://repository.upenn.edu/fnce\\_papers](http://repository.upenn.edu/fnce_papers)

 Part of the [Economics Commons](#), and the [Finance and Financial Management Commons](#)

---

## Recommended Citation

Burtch, G., Ghose, A., & Wattal, S. (2013). An Empirical Examination of the Antecedents and Consequences of Investment Patterns in Crowd-Funded Markets.. *Information Systems Research*, 24 (3), 499-519. <http://dx.doi.org/10.1287/isre.1120.0468>

This paper is posted at ScholarlyCommons. [http://repository.upenn.edu/fnce\\_papers/135](http://repository.upenn.edu/fnce_papers/135)  
For more information, please contact [repository@pobox.upenn.edu](mailto:repository@pobox.upenn.edu).

---

# An Empirical Examination of the Antecedents and Consequences of Investment Patterns in Crowd-Funded Markets.

## **Abstract**

Crowd-funded markets have recently emerged as a novel source of capital for entrepreneurs. As the economic potential of these markets is now being realized, they are beginning to go mainstream, a trend reflected by the explicit attention crowdfunding has received in the American Jobs Act as a potential avenue for economic growth, as well as the recent focus that regulators such as the U.S. Securities and Exchange Commission have placed upon it. Although the formulation of regulation and policy surrounding crowd-funded markets is becoming increasingly important, the behavior of crowdfunders, an important aspect that must be considered in this formulation effort, is not yet well understood. A key factor that can influence the behavior of crowd funders is information on prior contribution behavior, including the amount and timing of others' contributions, which is published for general consumption. With that in mind, in this study, we empirically examine social influence in a crowd-funded marketplace for online journalism projects, employing a unique data set that incorporates contribution events and Web traffic statistics for approximately 100 story pitches. This data set allows us to examine both the antecedents and consequences of the contribution process. First, noting that digital journalism is a form of public good, we evaluate the applicability of two competing classes of economic models that explain private contribution toward public goods in the presence of social information: substitution models and reinforcement models. We also propose a new measure that captures both the amount and the timing of others' contribution behavior: contribution frequency (dollars per unit time). We find evidence in support of a substitution model, which suggests a partial crowding-out effect, where contributors may experience a decrease in their marginal utility from making a contribution as it becomes less important to the recipient. Further, we find that the duration of funding and, more importantly, the degree of exposure that a pitch receives over the course of the funding process, are positively associated with readership upon the story's publication. This appears to validate the widely held belief that a key benefit of the crowdfunding model is the potential it offers for awareness and attention-building around causes and ventures. This last aspect is a major contribution of the study, as it demonstrates a clear linkage between marketing effort and the success of crowd-funded projects.

## **Keywords**

economics of IS, electronic commerce, crowdfunding

## **Disciplines**

Economics | Finance and Financial Management

# **An Empirical Examination of the Antecedents and Consequences of Investment Patterns in Crowd-Funded Markets**

Gordon Burtch\*

Anindya Ghose\*\*

Sunil Wattal\*\*\*

**February 15<sup>th</sup>, 2012**

---

\* Fox School of Business. [gburtch@temple.edu](mailto:gburtch@temple.edu)

\*\* NYU Stern School of Business and Wharton School of Business. [aghose@stern.nyu.edu](mailto:aghose@stern.nyu.edu)

\*\*\* Fox School of Business. [swattal@temple.edu](mailto:swattal@temple.edu)

<sup>1</sup> The authors are grateful to seminar participants at the 2011 Workshop on Statistical Challenges in e-Commerce Research (SCECR), the 2011 Marketing Science Conference, the 2011 Academy of Management, as well as seminar participants at the Fox School of Business at Temple University for useful suggestions. They also thank Ashish Agarwal, Avi Goldfarb, Yili Hong, Pei-yu Chen and Evan Polman for providing helpful comments. Anindya Ghose acknowledges funding from the WCAI (Wharton Customer Analytics Institute) and NSF CAREER Award IIS-0643847.

## Abstract

Crowd-funded markets have recently emerged as a novel source of capital for entrepreneurs. As the economic potential of these markets is now being realized, they are beginning to go mainstream, a trend reflected by the explicit attention crowd-funding has received in the American Jobs Act as a potential avenue for economic growth, as well as the recent focus that regulators such as the SEC have placed upon it. While the formulation of regulation and policy surrounding crowd-funded markets is becoming increasingly important, the behavior of crowd-funders, a key aspect that must be considered in this formulation effort, is not yet well understood. A key factor that can influence the behavior of crowd funders is information on prior contribution behavior, including the amount and timing of others' contribution, which is published for general consumption. With that in mind, in this study, we empirically examine social influence in a crowd-funded marketplace for online journalism projects, employing a unique dataset that incorporates contribution events and web traffic statistics for about 100 story pitches. This dataset allows us to examine both the antecedents and consequences of the market's contribution behavior. First, noting that digital journalism is a form of public good, we evaluate the applicability of two competing classes of economic models that explain private contribution toward public goods in the presence of social information: substitution models and reinforcement models. We also propose a new measure that captures both the amount and the timing of others' contribution behavior: contribution frequency (dollars per unit time). We find evidence in support of a substitution model; individuals shift their contributions away from a project when the frequency of others' contribution to that project is high. This suggests a partial crowding-out effect, where contributors may experience a decrease in their marginal utility from making a contribution as it becomes less important to the recipient. Further, we find that frequency is negatively associated with readership upon the story's publication, which suggests that measures of contribution behavior may be useful in predicting the quality of a project's output following its implementation.

**Keywords:** *electronic commerce; crowd funding; public goods; altruism; social influence*

## 1. Introduction

Crowd-funded markets have recently emerged as a viable alternative for sourcing capital to support innovative, entrepreneurial ideas and ventures (Schwienbacher and Larralde 2010). As the economic potential of these markets has become more apparent, they have boomed, gaining attention both from policymakers and regulators, as evidenced in the American Jobs Act (2011) and the Entrepreneur Access to Capital Act (i.e., H.R. 2930). Marketplaces like ProFounder Financial, PeerBackers, KickStarter and IndieGoGo are beginning to facilitate greater transactions, both in terms of volumes and amounts. Kickstarter.com, for example, has raised more than \$120 million for projects over the last two years<sup>2</sup>. While the formulation of policy and regulation surrounding these marketplaces is beginning to garner greater attention, the behavior of participants, a key factor that must be considered in this formulation effort, is not very well understood.

Crowd-funded markets present a unique combination of collective evaluation and crowd-based fundraising. Contributors, by choosing to pledge funds in support of a project, implicitly undertake the evaluation and selection of that project. This interesting aspect of the crowd-funding process is compounded with the significant potential for social influence in these marketplaces, which stems from the fact that both the timing and amount of other participants' prior contributions toward a particular project are published for all to see. While a number of recent studies have considered the influence of popularity indicators in consumer decision-making within electronic markets (Duan et al. 2009; Tucker and Zhang 2010, 2011), these studies have generally dealt with binary adoption (adopted versus did not adopt) measures (and, in the case of Duan et al, rank data based on these binary measures), presented to visitors in aggregate (e.g., total downloads). In contrast, the temporal and continuous nature of the popularity indicators that are typically made available for consideration in crowd-funded markets makes them deeper and much more complex, both in terms of content and potential influence.

This study examines a crowd-funded marketplace that supports journalists, enabling authors to pitch article ideas to the crowd and to then raise the money necessary to research and publish the piece.

---

<sup>2</sup> <http://www.kickstarter.com/blog/2011-the-stats>

We begin by noting that this crowd-funding platform provides a guaranteed outlet for authors' work and that all published work is subsequently offered freely for public consumption. As such, our empirical analysis is based on the premise that these published stories meet the criteria of a public good (Hamilton 2003; Kaye and Quinn 2010), a characterization that is further supported by prior research, which has found clear evidence that this marketplace's output is treated as such by participants (Aitamurto 2011). Bearing this in mind, we explore the applicability of two competing classes of economic models that attempt to explain private contribution toward public goods: i) reinforcement models, which predict cooperative or reciprocal behavior and ii) substitution models, which predict altruistic behavior and crowding out. The main research questions that we address are as follows: *What effects do observable indicators of others' prior contribution decisions in a crowdfunded market have on later participants' contribution decisions? Is the manifestation of these effects in the funding phase associated with project performance, and, if so, what is the nature of that association?*

In most crowd funding markets, the information on prior contribution behavior comprises a project's 'funding status', which offers an observer a richer set of information than more common popularity indicators that are available in other marketplace contexts (e.g., software download statistics on download.com). An observer can see who has contributed to a project, when they have contributed and, in some cases, even the exact amount they have contributed. The availability of this rich information has important implications for potential contributors' decision-making. For example, the availability of information on the timing of others' contributions allows a potential contributor to gauge the emergence of spikes or lulls in market contribution over time. In general, given these rich signals, it is relatively easy for an observer to compare the contributions of prior others to one another, as well as to his or her own potential contribution. To capture information about the timing and amount of prior contributions that is available to a potential contributor in the marketplace, we define a new measure as follows: *contribution frequency*, the average amount (in dollars) contributed per period at a given point in time.

This work adds to multiple literary bases in the fields of information systems, marketing, and economics. First, we consider a novel product type that has seen limited consideration by information

systems researchers. Whereas much of the online public goods research in IS has looked at P2P file sharing, we consider the production of digital journalism, a cultural, experience good. Second, whereas prior work in marketing and IS has looked at binary indicators, we consider deeper measures that are both temporal and continuous. Third, we explore an entirely new context and process that combines crowd-based fundraising and collective evaluation. While economists have conducted a variety of online public goods experiments in the past, none, to our knowledge, have explored this combination of private contribution and collective evaluation. Further, no prior work to our knowledge has explored a context in which subjects are faced with contribution toward a variety of public goods (e.g., multiple simultaneous projects undergoing funding); typically, lab experiments have looked at subjects' contributions toward a single public good. Finally, we leverage proprietary data in the form of web traffic statistics, allowing us to empirically examine the consequences of different contribution patterns on later outcomes. This last aspect is key, as this allows us to model both the antecedents and consequences of different patterns of private contribution toward a public good, simultaneously exploring the process and efficacy of collective evaluation.

The main results of our analysis are as follows: we find evidence that supports the idea that altruism is a key incentive to contribute in this marketplace (i.e., a substitution effect). More specifically, as individuals observe others contributing more frequently, the amount they are inclined to contribute falls (an increase of 1% in the prior frequency of contribution is associated with a 0.31% decrease in subsequent contribution). This substitution effect suggests that contributions are subject to crowding out. That is, the marginal utility contributors gain from giving in the marketplace is diminished by others' contributions. In turn, this suggests that individuals in this marketplace are primarily driven by altruism. Importantly, since our findings are based on contribution frequency, this also suggests that the duration over which prior contributions have arrived is important.

Further, we find that stories tend to garner more readers upon publication when they undergo a lengthier funding process (lower contribution frequencies). This seems to imply that pitches that receive a relative lack of scrutiny (i.e., quicker funding) are associated with lower quality project outcomes. While

this relationship may not be causal, at the least, this finding demonstrates that measures of contribution behavior can be leveraged in a predictive capacity to anticipate project performance in crowd-funding markets. Our findings are robust to a variety of different models, estimators and data splits.

In the following sections, we begin with a review of the relevant literature pertaining to crowd-funding markets and private contribution toward public goods. We then discuss our expectations, based on theory, of the influence of contribution behavior on subsequent potential contributors, as well as the relationship that such signals will have with project quality. Finally, we present details of our empirical evaluation and a discussion of the implications of our findings for both practitioners and scholars dealing with crowd-funded markets.

## **2. Literature Review**

### **2.1 *Crowd-funding***

There is an emerging stream of research that has examined the concept of crowd-funded markets. Schwienbacher and Larralde (2010) have defined crowd-funding as “the financing of a project or a venture by a group of individuals instead of professional parties.” These authors provide an overview of the concept, based on an earlier review (Belleflamme et al. 2010), noting the still embryonic stage of the phenomenon’s emergence. Schwienbacher and Larralde report that a number of platforms have arisen in support of online crowd-funding over the last 4-6 years. The earliest successful example is Sellaband.com, a Dutch-based marketplace in which musical artists raise funds to produce and sell an album. Agarwal et al. (2010) examined this marketplace to evaluate the influence of proximity, amongst other factors, on individual investment behavior, re-evaluating the ‘flat-world hypothesis’. Another prominent example of a crowd-funding marketplace is the peer-to-peer lending site Prosper.com. In this marketplace, Lin et al. (2009) sought to identify novel forms of information that individuals consider in crowd-funded marketplaces when making investment or contribution decisions. These authors have concluded that the likelihood of credit being issued is greater when the borrower exhibits greater social capital (e.g., a larger social network), as lenders appear to take this as a sign of credibility or



trustworthiness<sup>3</sup>.

While the literature on the topic of crowd-funded markets is quite limited, there is a wide body of related work. Particularly relevant is the stream of research that has examined the effects of popularity indicators on adoption and consumption decisions. One example of this work include that by Tucker and Zhang (2011), who consider wedding vendor service adoption in the presence of popularity statistics, based on prior adoption decisions captured in the form of hyperlink click-through rates. These authors find that publication of these popularity statistics causes the formation of a “steep tail,” attracting new outsiders to what was previously a niche product choice. Interestingly, they also find that this occurs without cannibalizing consumers from competitors’ market share. Similarly, Duan et al. (2009), considering users’ downloads of free software from download.com, observe severe changes in software download rates when published product rankings shift, suggesting that consumers take prior downloads as a signal of product approval and quality. Lastly, Salganik et al. (2006), manipulating download statistics for songs in an online experiment using an artificial marketplace, find that variation in the presence and ordering of prior download volumes directly influences the inequality and predictability of song success. Thus the decisions of prior users have the potential to influence subsequent users, as is evidenced by the above noted work.

## **2.2 Public Goods**

The influence of information about others’ prior decisions on those of later deciders is not limited to studies of private goods. The public goods literature is also replete with comparable studies that pertain to individuals’ contributions toward public goods in the presence of “social information” (e.g., information on others’ contributions). This is particularly relevant in our study context because we are dealing with private contribution toward the production of online journalism, which in many cases has been argued to constitute a form of public good (Hamilton 2003; Kaye and Quinn 2010). In our case, the crowd-funded platform provides a guaranteed venue for publication, and offers all material freely on the

---

<sup>3</sup> It is important here to note that Prosper.com, while related to crowd-funding as defined above, is slightly different. In the Prosper marketplace, individuals seek interest-bearing loans, rather than contributions or investments. That is, each loan is to be repaid in full by the borrower, plus interest, regardless of how the money is used and regardless of whether any associated project succeeds. Further, the purpose of the loans is highly varied and dependent and need not pertain to a specific project. In addition, the funding process is quite different from a typical crowd-funded market as borrowers obtain loan offers on Prosper.com through a second price auction process, rather than an open, unconstrained contribution process that takes place continuously over the funding period.

Internet, without need for subscription. As such, while it may be true that online journalism does not always meet the criteria of a public good, this is certainly the case in our context (Aitamurto 2011)<sup>4</sup>.

In the public goods literature, there are generally two contrasting classes of models that have been proposed that predict different directions of influence from prior others' contributions, on the amounts contributed by later deciders: reinforcement models and substitution models (Shang and Croson 2009). As one might expect, reinforcement models are those that predict that greater initial contribution will have a positive effect on later contributions, whereas substitution models are those that predict that initial contributions will have a negative effect on later contributions.

Broadly speaking, reinforcement models are based on the mechanisms of reciprocity, fairness and social norms – a need to contribute fairly, rather than leave others to bear the burden of supporting the public good on their own. Key examples of such models include that proposed by Sugden (1984), which is rooted in the idea that everyone will seek contribute at least as much as the least-generous contributor, out of a sense of moral purpose, as well as that proposed by Bernheim (1994), which is based on the idea that individuals will rarely deviate from social norms because they are worried about how others will perceive them. Thus, when one observes others contributing in greater amounts, these models predict that contributors will be inclined to donate in kind.

On the other hand, substitution models are generally explained by the concept of *altruism* and the *warm glow* effect (Andreoni 1989, 1990)<sup>5</sup>. In these models, individuals allocate their resources between contributions to private goods and public goods. The reason individuals contribute to public goods, it is argued, is that they derive some utility from their own consumption of the good, as well as others' consumption (i.e., from aiding others). However, as the level of others' contributions rises, the volume of the public good that is available increases such that the marginal benefit to an individual from contributing falls, because others are meeting the "need" of the public good. The end result of this is that

---

<sup>4</sup> Aitamurto, 2011 examined the motivations of contributors in the exact marketplace from which our data is drawn for this study and found that contributors perceive themselves as contributing toward a greater social good.

<sup>5</sup> We would like to thank the Associate Editor for suggesting this reference.

the individual will tend to reallocate their funds more toward private consumption activities as this takes place (Roberts 1984), an effect generally referred to as crowding out (Andreoni 1990).

In terms of empirical work on private contribution toward public goods in an online setting, research examining online journalism has been quite limited. One of the few examples of this research is the study by Borck et al. (2006), who studied individuals' voluntary contribution of funds in support of an online German newsletter. These authors found that individuals would contribute in accordance with reinforcement models, attempting to match their contributions to the amounts donated by others. Though online journalism has received only limited attention, there is fortunately a wide body of related work that has studied private contributions to P2P file-sharing networks.

To date, the P2P file-sharing context has probably received the greatest attention from scholars examining private contribution toward public goods, online. Interestingly, however, this stream of work has presented mixed evidence that variably supports reinforcement and substitution models. Xia et al. (2011), Gu et al. (2009) and Nandi and Rochelandet (2008) have all found evidence of reinforcement models, reporting that P2P contributors behave in a reciprocal manner. Xia et al. (2011) found that individuals are more likely to continue contributing if they receive greater benefits from the network, reflecting a reciprocal relationship. Gu et al. (2009) similarly found that P2P file sharers contribute their content in a reciprocal manner, though in their study the behavior was found to be enforced through explicit sanctions on deviants (bandwidth throttling for downloads by free riders). Lastly, Nandi and Rochelandet (2008) again found evidence of reciprocity, though the primary focus of reciprocal contribution in their case was on ensuring diversity of content, rather than volume or availability.

In contrast to the above, other work has found clear evidence in support of substitution models. Research that examined one of the first P2P file sharing networks, Gnutella, identified rampant free-riding behavior (Adar and Huberman 2000), a direct indication of a substitution effect. Subsequent work, conducted a few years later in the same setting, found that this behavior had only grown more prevalent over the intervening period (Hughes et al. 2005). Similarly, Asvanund et al. (2004), studying participant

behavior across multiple file sharing networks, found further evidence of free riding. In this case, the authors found that the prevalence of free riding grew stronger as the network grew in size.

Given the conflicting findings reported in the above work, the directionality of social influence stemming from information on prior contributions in online networks is not easily inferred for a given marketplace. This is likely because the direction of influence will depend on a variety of factors, such as the contributors' motivations or incentives and the degree to which contributors' decisions are observable to others or convey status and social identity. A number of these factors are potentially at play in a crowd-funding market. Contributors may be driven by a sense of altruism, to help the author and support others who might benefit from the story. Alternatively, contributors may have a personal interest in the proposed story topic. Crowd-funders' decisions may be influenced by concern over social image, if they perceive that their actions are highly observable to others in the market. In contrast, if crowd-funders feel that they have a sufficient degree of anonymity, they may be more inclined to withhold their funds and simply consume content as it is produced.

Thus, while it is difficult to determine the directionality of social influence in advance, it is quite likely that individuals' contribution decisions in crowd-funded markets are highly dependent on the behavior of other contributors, just as they are in P2P networks and other, offline examples of private contribution to public goods. This is because, by drawing on information about prior contribution behavior, crowd-funders are easily able to assess the impact of their own potential contribution relative to those of others. We make this potential clear in the following section, where we provide details of the marketplace's operation and the available information signals.

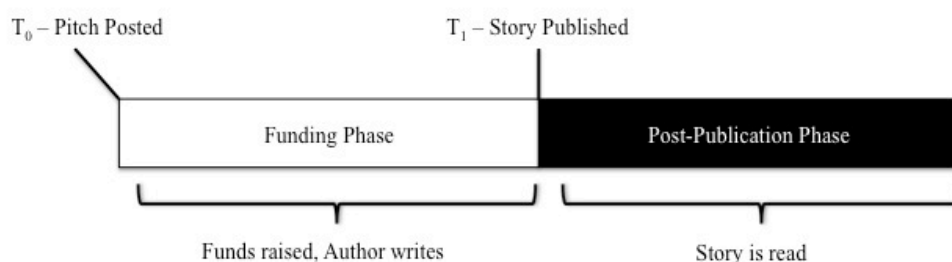
### **3. Study Context**

In this section we detail our study context in order to make clear the relevance of the above discussion and to provide a basis of understanding for the subsequent empirical analysis. The market we consider is a crowd-funded platform (Agarwal et al. 2010; Kappel 2008; Schwienbacher and Larralde 2010) based in San Francisco, California, which was established in November of 2008. The site enables

new and established journalists to crowd-source capital in support of their work by pitching their ideas to the community, just as an entrepreneur might pitch venture capitalists. The site functions as follows: individuals first register to become members of the community. Members can then establish a profile page where they can provide details about themselves along with a photograph and a link to a personal website.

Any member in the community can choose to pitch a story to other members of the community<sup>6</sup>. A pitch is a proposal to publish a story on the website, and includes a description of the story topic, the proposer's qualifications, as well a list of planned deliverables and a budget. Other members of the community can then choose to contribute funds toward the pitch in any increment they wish. The contribution phase continues in this manner until the story is completed and published on the website. Figure 1 provides a graphical depiction of the timeline of article funding and publication.

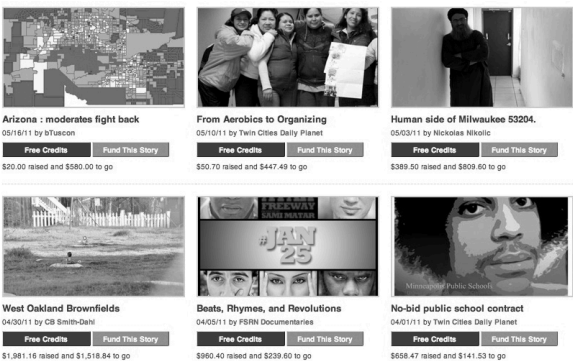
**FIGURE 1**  
**Funding and Publication Timeline**



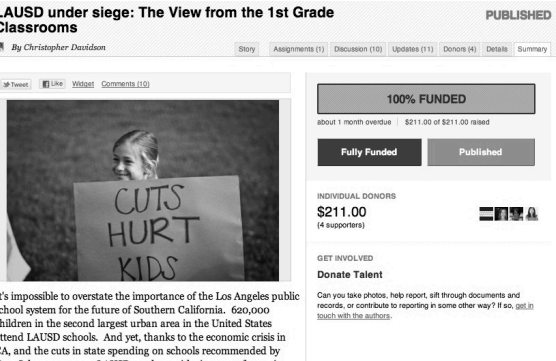
Community members' preferences are indicated to others by their observable contribution history. Once a community member contributes toward a pitch, this becomes public record. When reviewing a pitch, a list of all prior contributors is provided to the observer, along with their specific contribution amounts, the date and time of their contribution, as well as the aggregate contribution that has been made up to that point by all community members. Screenshots of the marketplace are provided in Figure 2, depicting the format in which these pieces of information are presented.

<sup>6</sup> It is important to note that very few members of the marketplace act as both contributors and authors, even though this is possible. Of the 4,515 unique contributors present in our dataset, only 72 are authors contributing to the pitches of others. A notable implication of this is that contribution behavior cannot be attributed to a tit-for-tat strategy.

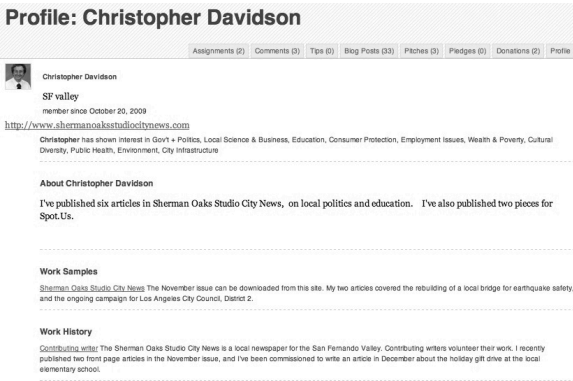
FIGURE 2  
Marketplace Screenshots



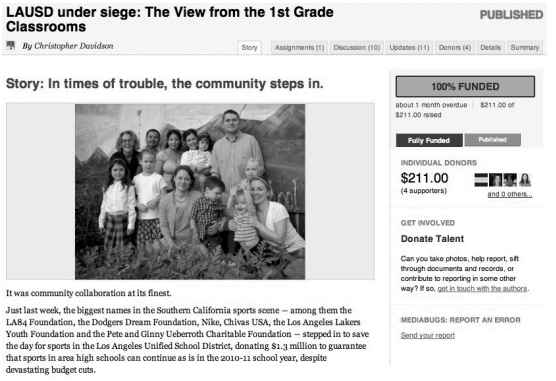
**(a)**



**(b)**



**(c)**



**(d)**

(a – main landing page, b – pitch, c – member profile, d – story)

While prior work has examined the role of popularity indicators as signals of quality, driving subsequent adoption behavior (Duan et al. 2009), in our case, the potential impact of observable prior decisions is more complex. This is because the two classes of public good contribution models that we have outlined above provide contradictory predictions. On the one hand, reinforcement models predict that participants, upon observing prior contribution actions and amounts, will seek to contribute in kind out of a desire to be fair or to fit in. In contrast, substitution models predict that participants will react quite differently, lessening their contributions as they observe the arrival of funding from elsewhere. We give more detailed consideration to this in the following section, where we outline the relevant variables we consider in our empirical analysis and the directionality of influence that each class of public good contribution model would predict.

## 4. Methods

### 4.1 *Model Formulations*

Empirical evaluations of the role of ‘social information’ in determining subsequent private contributions toward a public good have generally focused on prior others’ contribution amounts. For example, a large body of work has examined the crowding out effect, drawing on archival data to understand how increases in government spending influence individuals’ private contribution levels (Abrams and Schitz 1978; Kingma 1989; Reece 1979). More relevant to the present paper, however, are the various laboratory and field experiments that have explored the effects of prior others’ private contributions on a later contributor (Andreoni 1995; Bardsley and Sausgruber 2005; Chauduri and Paichayontvijit 2006; Croson and Shang 2008; Shang and Croson 2009). Considering the relevance of past work to the present crowd-funding context, it seems reasonable to explore the effects of earlier contributors’ contributions on later contributors by looking at prior contribution amounts. The coefficient on prior contribution would allow us to explore the fit of both reinforcement and substitution models, simultaneously, based on the directionality of the resultant coefficient estimate. In a reinforcement model, one would expect subsequent contributors to try and match the donations of earlier contributors, out of a desire to be fair or to conform, resulting in a positive coefficient on prior contribution. In contrast, a substitution model would predict that later contributors would withhold their funds as they observe others contributing more in their stead, lowering the marginal value of their own contribution, and thus causing a crowding out effect. This would result in a negative coefficient estimate.

The above being said, it is also important to note that most prior research has explored the effects of prior others’ decisions in what is, or is assumed to be, a relatively static environment. The efforts of Andreoni (1998), Vesterlund (2003) and Varian (1994) are rare exceptions to this as these authors consider contribution sequencing. This is quite likely because the vast majority of study contexts considered have not offered observers temporal information that would indicate how long an adoption or contribution process has been going on. Yet, in a crowd-funded market, this timing information is readily available and is quite likely to play an important role in contributor decision-making. Given that a potential contributor can quickly see how long a project has been raising funds (based on its posting date) as well as the funds that have been raised to date, it is likely that they will assume the rate funding arrival.

To illustrate this point, consider the following example. Consider a hypothetical subject who is a potential contributor in a crowd-funded marketplace. This subject has the option of donating to one of two projects: X and Y. Both projects have raised \$300 thus far, however, project X has been posted in the marketplace for 30 days, while project Y has only been posted for 10 days. Given this information, our subject should reasonably infer a contribution arrival rate of \$10/day for project X, and an arrival rate of \$30/day for project Y. Our subject is likely to infer that contributions are arriving more quickly for project Y than for project X.

In an effort to incorporate this temporal information into our empirical model, we propose a new measure that captures the frequency of contribution. We define *contribution frequency* as the amount in dollars contributed to a given pitch, divided by the duration of a pitch's funding in the marketplace, as illustrated above. Our logic for evaluating the fit of either a substitution or reinforcement model to the data is as follows: a greater prior frequency of contribution will incite cooperation and reciprocation in the case of a reinforcement effect (a positive coefficient) or it will cause later contributors to withhold their funds in the case of a substitution effect (a negative coefficient).

In addition to the above, we need to consider a number of other factors that reflect a potential contributor's prior behaviors, as these are likely to influence his or her ultimate decision about whether to contribute and how much to contribute. In particular, the number of individuals viewing a pitch in a given day is likely to reflect general interest in the topic at that point in time. Thus, web traffic statistics capturing the number of page views are also included in our model as a contribution predictor. We also include the lag of contribution as a predictor, to capture word of mouth effects (Duan et al. 2009) or general inertia in contribution behavior.

Further, though the focus of this study is the influence of preference indicators and the predictive power of prior user behavior, we must also control for some important pitch characteristics associated with contribution behavior in our analysis. First, we include the amount of the pitch's required budget that remains outstanding as of a given time period, in dollars. The impact of potential contributors' awareness of the outstanding budget for a given pitch is difficult to predict. A greater volume of funds required may result in a desire to assist the author. Alternatively, this may be perceived as somewhat of an insurmountable goal, and thus pitches may have a harder time attracting initial funding. We also control for static features of a pitch, such as pitch genre or length, in addition to unobservable heterogeneity,



through the use of panel fixed effects.

Finally, in order to ensure our model is identified, we also account for exogenous sources of influence on contribution behavior, such as the pitch topic's popularity in the mainstream media and on the broader Internet. We address these factors in two ways: first, we consider the number of page referrals to the crowd-funding marketplace, at the time of a given observation, from external sources such as search engines, Facebook links, e-mail links and the like. Such referrals are in contrast to those page visits that stemmed from individuals manually accessing the marketplace by directly typing the URL into their web browser. A greater number of external referral visits would reflect greater popularity of the article content posted on the marketplace website, in mainstream society. Second, we include results from Google Search trends for each observation, employing 2-5 keywords that were manually extracted from each pitch. These keywords included any proper nouns in the pitch title, as well as story topic tags assigned to the pitch by the website purveyor. Again, the prevalence of individuals searching for these topics on Google on a given day is expected to capture the broader popularity of the article content in society<sup>7</sup>.

It should also be noted that these popularity variables serve an additional purpose. The focal variable in our analysis, contribution frequency, has the potential to not only drive later contribution decisions, but to be driven, in turn, by prior contribution decisions (e.g., patterns of contribution may be driven by a key early investment made by one user). Hence, the two may be endogenous. The referrals data and Google Search Trends aid us in addressing this possibility, as they can be leveraged as instruments in our robustness checks. We provide more details subsequently. A table of all variables included in our model of the antecedents is provided in Table 1.

-----INSERT TABLE 1 ABOUT HERE-----

#### ***4.1.1 Antecedents Model***

All non-categorical variables in our antecedent analyses are log transformed, thereby allowing us to identify percentage changes in effect. This was deemed appropriate primarily because the required budget and daily contribution for a given pitch, as well as the popularity of different stories, vary widely, so understanding contribution effects in percentage terms is significantly more useful (Keene 1995). For a

---

<sup>7</sup> Please see Appendix A for a more detailed description of our referrals and search trend variables.

given pitch,  $i$ , on a given day,  $t$ , the final, complete model is reflected below, by equation (1). Here,  $\phi_t$  is a vector of date fixed effects for each day in the sample, which controls for changes in the marketplace over time<sup>8</sup>, while  $\alpha_i$  is a vector of pitch-level fixed effects<sup>9</sup>.

$$\begin{aligned} \log(\text{contribution}_{it}) = & \beta_0 + \beta_1 * \log(\text{contribution\_frequency}_{i(t-1)}) + \beta_2 * \log(\text{contribution}_{i(t-1)}) + \\ & \beta_3 * \log(\text{page\_views}_{it}) + \beta_4 * \log(\text{remaining\_budget}_{it}) + \\ & \beta_5 * \log(\text{referrals}_t) + \alpha_i + \phi_t + \varepsilon_{it} \end{aligned} \quad (1)$$

Our evaluation of the antecedent model is conducted via time series regression employing the Arellano-Bover/Blundell-Bond system estimator (Arellano and Bover 1995; Blundell and Bond 1998), an extension of the original Arellano-Bond estimator (Arellano and Bond 1991), allowing us to instrument for the lag of the dependent variable using the Generalized Method of Moments (GMM). The original GMM estimator developed by Arellano and Bond (1991) is generally known as the “difference GMM” (DGMM). This estimator treats the model as a system of equations, one for each time period. The equations differ only in their moment condition sets. However, Arellano and Bover (1995), as well as Blundell and Bond (1998), later noted that lagged levels could be weak instruments for first differences if the autoregressive process is too persistent over time, as is possible in our case. In order to address this issue, these authors proposed the “system GMM” (SGMM) estimator in which the original, uninstrumented equations are added to the system. The idea is to estimate instrument differences with lagged levels and instrument levels with lagged differences. It is this latter estimator that we employ in our analyses.

Admittedly, higher order lagged variables might not be ideal instruments since it is possible to have common demand shocks correlated over time, in which case lagged variables would be correlated with the current period demand shock. However, common demand shocks correlated over time are conceptually similar to trends. Hence, a suitable control for correlated demand shocks or trends can alleviate this problem in the GMM estimation (Archak et al. 2011). Our web site referrals control variable thus alleviates any potential concerns in this regard, as does our re-estimation using Google Search Trends. To alleviate any concerns about a high instrument count in the GMM estimations, we also use a

---

<sup>8</sup> To clarify, changes in the marketplace over time might include things such as seasonality effects (e.g., contributors having lower volumes of disposable income in the tax season).

<sup>9</sup> We do not incorporate author fixed effects into our model as the vast majority of authors in our sample have pitched only a single story. As such, author fixed effects are highly collinear with pitch fixed effects and lead to qualitatively similar results.

collapsed instrument set comprised only of low order lags, as this has been shown to produce more reliable results in scenarios where the instrument count is on the higher side (Mehrhoff 2009). Thus, these additional estimations add credibility to our findings.

We then assess the reliability of our findings by re-estimating all formulations using two stage least squares with date and pitch-level fixed effects (2SLS). The 2SLS estimation allows us to report additional tests statistics to evaluate model validity (exogeneity of regressors) and instrument strength, in support of our initial GMM-based dynamic panel estimates. We conduct numerous other robustness checks, including standard OLS regressions excluding the lagged dependent variable, as well as re-estimations using all estimators employing alternate dependent variables (cumulative prior contribution and average prior contribution in place of frequency), different data splits (e.g., excluding null observations and outliers), to assess the stability of our coefficient estimates<sup>10</sup>.

#### ***4.1.2 Consequences Model***

Next, we develop a model linking the pattern of contributions in the funding phase to the story's performance after it is published. Unlike in the funding phase, where the unit of analysis is a pitch, in the publication phase, the unit of analysis is a story. In order to determine whether the contribution frequency is also associated with the performance of the story upon its eventual publication, we employ a static, summary measure of this dynamic value in our evaluation of the consequences model. Specifically, we employ the final contribution frequency value for an associated pitch as a predictor of story performance (i.e., the readership – number of unique story URL visits). In addition, we include a single period lag of the story readership as a predictor in the model, to again capture word of mouth or inertial effects. Lastly, we include a measure of external page referrals, for reasons identical to those outlined for the model of antecedents, above. A table of all variables included in our model of the consequences is provided in Table 2.

-----INSERT TABLE 2 ABOUT HERE-----

---

<sup>10</sup>A number of these robustness checks are presented in Appendix B, including an analysis of cumulative and average prior contribution (in place of frequency), as well as an analysis of story genre. In general, these analyses suggest that: i) contribution frequency is a better explanatory variable than cumulative or average contributions, which ignore temporal aspects, and ii) story genre is not a major factor in determining contribution behavior.

For a given pitch,  $i$ , on a given day,  $t$ , the final, complete model for the consequences portion is reflected by equation (2). This model is similar to the antecedent model, though the dependent variable constitutes count rather than continuous data (i.e., readers rather than monetary contribution). In accordance with this, we employ a count estimator (negative binomial regression with a conditional fixed effects dispersion parameter). All independent variables are again log transformed in these analyses. We also add a vector of time dummies ( $\phi_t$ ) once again to control for changes in the marketplace over time. To control for unobserved heterogeneity at the story level, we introduce story-level fixed effects via the dispersion parameter (Hausman et al. 1984).

$$\log(\text{readers}_{it}) = \beta_0 + \beta_1 * \log(\text{final\_contribution\_frequency}_i) + \beta_2 * \log(\text{readers}_{i(t-1)}) + \beta_3 * \log(\text{referrals}_t) + \alpha_i + \phi_t + \varepsilon_{it} \quad (2)$$

We once again follow the initial model with some robustness checks. In particular, we re-estimate the model as a standard OLS regression with fixed effects. We also execute a Quantile Count regression, which allows us to examine the effect of our final contribution frequency variable at various points in the distribution of readership (Cameron and Trivedi 2009). Noting that count models have generally been found to perform much better when the dependent variable is left untransformed (O'Hara and Kotze 2010), and considering that there is wide variation in readership across stories in our sample, which we might typically address via a transformation, quantile count regressions were deemed appropriate.

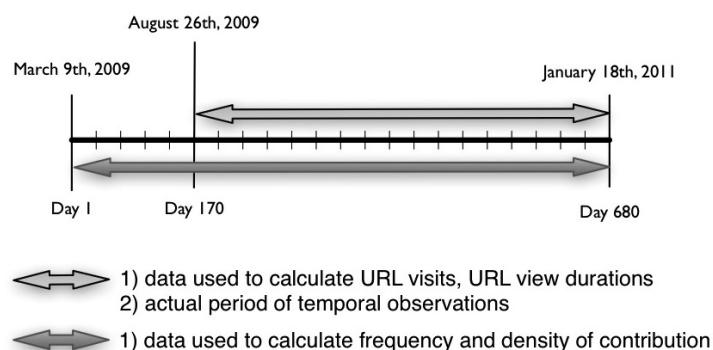
## 4.2 Dataset

We collected our data from a variety of sources. First, the purveyor of the marketplace provided us with access to proprietary data in the form of web traffic statistics for the marketplace via a Google Analytics account. Using this account, one of the authors then developed a piece of software to programmatically retrieve time series data of web traffic statistics for each URL, by day, leveraging the Google Analytics Data Export API. A similar approach was used to obtain time series data on external page referrals (i.e., unique website visits initiated from external sources, such as web pages or e-mail links).

Second, we developed a software application to retrieve all available public information from the website about pitch and story characteristics, such as the date on which a pitch was proposed, the amount of funding received on any given day, who contributed to a pitch and the amount of each contribution, the required budget outstanding for a pitch as of a given day, whether a pitch was located on the default landing page of the marketplace as of a given point in time, the date an associated story was published by the author, etc. Finally, one of the authors manually retrieved time series data on Google search trends for 3-5 keywords associated with each pitch and story.

Our data set includes every contribution made in the marketplace from March of 2009 until January of 2011, as well as web traffic statistics, by URL, by date, spanning August of 2009 to January of 2011. Because the web traffic statistics obtained from Google Analytics only extend back as far as August 26, 2009, and in order to ensure consistency between our two analyses, we limit our consideration to those stories published on or after this date, as well as the contribution observations for associated pitches that take place on or after this date<sup>11</sup>. The timeline diagram provided in Figure 3 clarifies the nature of our data set.

**FIGURE 3**  
**Marketplace Screenshots**



As per the diagram, for our primary antecedent model, we consider only those pitch contribution observations (aggregate contribution by all users, in a particular pitch, on a particular day; pitch-day

<sup>11</sup> Because it might be argued that such an approach introduces the potential for positivity or selection bias, in that our analysis is limited only to those pitches that completed the funding process, for the sake of robustness, a 2SLS estimation was performed on the entire set of pitch observations, from August 26, 2009, onward, and similar findings were obtained in terms of both sign and magnitude of the regressors. Further, a simple model without the Google Analytics data was analyzed, evaluating all contributions from March 9<sup>th</sup>, 2009 onward, and the results were again qualitatively similar to those reported in our Results section.

pairs) that took place on or after August 26, 2009, in order to ensure the availability of web traffic details associated with URL views. Of the 154 pitches that received some contribution in the marketplace over this period, 100 completed the funding process during the period of observation and are thus considered in our analysis. Of the \$126,654.97 contributed in the marketplace over this period, these 100 pitches accounted for \$66,016.21, supplied by a total of 4,353 contributors. Based on this, contribution appears to be distributed across pitches in a reasonably uniform manner, with 65% of the pitches receiving 52% of the funding.

The remaining variables we include in our analysis of the antecedents are the required budget outstanding for a pitch as of a given observation, Google Search trends for associated pitch keywords and the number of external URL referrals to the marketplace. A pitch enters our sample on August 26, 2009, or when it is initially posted for consideration, whichever is most recent. The pitch then exits on the date an associated story is published.

Our dependent variable in the antecedent model, contribution toward a given pitch on a given day, is the sum of all contributions toward that pitch on that day. Our antecedent sample is comprised of 5,874 pitch-day pairs. The average pitch takes 54.5 days to complete the funding process, though there is considerable variation, with the fastest pitch having completed funding in 1 day, while the slowest completed funding in 294 days. The average completed pitch receives funds from 40 unique contributors. However, once again, there is considerable variation around the mean, with the number of contributors ranging from just 1 to as many as 175. This is understandable, however, as at the time of data collection, the pitches observed have raised from as little as \$20 to as much as \$7,685. Thus, the number of contributors required to supply these funds will vary quite a bit. The average individual contribution is \$6.01 and contributions range from \$0.28 to \$4,749. We present the descriptive statistics of all the variables in our antecedent model in Table 3.

-----INSERT TABLE 3 ABOUT HERE-----

The data set we employ in our analysis of the consequences model is very similar to that used in the antecedent model. In contrast to the antecedent analysis, however, we aggregate the consequence data to weekly values, for the sake of manageability. As stories remain available for viewing continuously following publication, the considered period of readership for a given story tends to extend for a much

longer duration than the associated pitch's funding phase. As noted earlier, our sample includes all those stories published from August of 2009, onward, as well as each story's associated pitch contribution data. The associated pitch contribution data is used to calculate a summary measure of contribution frequency for each pitch. Beyond this summary statistic, we again leverage a continuous indicator of web site referrals. The descriptive statistics for our second sample are provided in Table 4<sup>12</sup>.

-----INSERT TABLE 4 ABOUT HERE-----

## 5. Results

### 5.1 Antecedents

We begin by exploring the effect of various lags of the dependent variable, in an effort to identify the extent of autocorrelation in the data. The results of these estimations are presented below in Table 5. Based on these results, we determined that only the first and second lags of contribution have a significant effect on contribution in the present period. Based on this, we chose to employ 3<sup>rd</sup>+ order lags as instruments in our Arellano-Bover Blundell-Bond estimations, as well as in our 2SLS estimations. Importantly, this analysis of lags has little bearing on our determination of a substitution or reinforcement effect. This is because, as outlined above, the information available to potential contributors about others' contributions is generally presented in aggregate (i.e., it consists of aggregate funding to date, and the total duration over which it has arrived). Though an observer could drill down into the contributor list to determine the total funds contributed the day prior, unlike the aggregate details, this information would take some effort to extract.

-----INSERT TABLE 5 ABOUT HERE-----

Next, we consider the results of our antecedent model, which are, presented in Table 6. First of all, looking at the tests of autocorrelation, we see that AR(3) is insignificant in all models. This is in line with our above analysis, as it indicates that autocorrelation is insignificant at the 3<sup>rd</sup> order. Looking at the

---

<sup>12</sup> Note: we also tried incorporating a variable in the consequences model that captured the total funds raised by the project, as we considered that more expensive projects should be of higher quality and thus draw more readers. However, we obtained qualitatively similar results, thus we have excluded it for the sake of simplicity.

coefficient estimates in column 4, we find that contribution frequency has a significant, negative effect ( $\beta_1=-0.31$ ;  $p < 0.01$ ). This estimate is stable, regardless of whether control variables, such as remaining budget, search trends or referrals are included, as can be seen in columns 2 and 3. Based on these estimated coefficients, a 1% increase in prior contribution frequency is associated with a 0.31% decrease in contribution. This suggests that higher contribution frequencies are associated with lower subsequent contribution amounts. As such, the direction of these effects is consistent with a substitution model, rather than a reinforcement model. Further, the size of the coefficient here seems to suggest that partial, rather than complete crowding out, is observed (as the coefficient is less than 1, dollar-for-dollar crowding out is not implied).

-----INSERT TABLE 6 ABOUT HERE-----

Moving next to our pitch URL traffic variables, we again find significant results. In particular, the effect of prior period contribution is significant and positive ( $\beta_2=0.29$ ;  $p < 0.01$ ). More specifically, a 1% increase in the volume of prior period contribution is associated with a 0.29% increase in subsequent contribution. The positive coefficient of this variable indicates the presence of word of mouth effects or general inertia in contribution behavior. We also find that greater numbers of users viewing a pitch on a given day are associated with more money being contributed ( $\beta_3=1.27$ ;  $p < 0.001$ ). That is, a 1% increase in the number of pitch views is associated with a 1.27% increase in contributions, a rather intuitive result. With regard to our other control variables (i.e., remaining budget, referrals and search trends), we find no significant effects.

Our estimates are qualitatively similar across all hierarchical regressions, thus our coefficients appear to be quite stable estimates. Further, the Hansen J test of overidentifying restrictions cannot be rejected in any of our regressions, and the difference in Hansen test does not allow us to reject the exogeneity of our instrument variable set.

## 5.2 Consequences

Shifting next to a consideration of our consequences model, we find a number of interesting results (Table 7). Firstly, we find that the final level of contribution frequency, for the pitch associated with the story, is a negative and statistically significant predictor of story readership ( $e^{\beta_1}=-0.14$ ;  $p <$



0.001). With respect to our control variables, we find that the lag of the dependent variable is highly significant ( $e^{\beta^2}=0.49$ ;  $p < 0.001$ ). Lastly, we find no significant effect of referrals in our model.

-----INSERT TABLE 7 ABOUT HERE-----

The significance and directionality of the *final\_contribution\_frequency* coefficient suggests that stories with pitches that are funded more quickly tend to perform less well, perhaps reflecting a lack of due diligence by the evaluating community of crowd-funders. That is, perhaps those stories that are funded more quickly are given less scrutiny before a contribution decision is made. However, as the contribution frequency will also increase with the total funds raised (perhaps suggesting that bigger budget projects tend to perform less well), some additional analysis is warranted to confirm our supposition. With this in mind, we can re-estimate the consequences model using funding duration in place of contribution frequency. Doing so produces the expected results, as estimation of this model produces a significant, positive coefficient for funding duration ( $e^{\beta}=0.22$ ;  $p < 0.001$ ), indicating that stories that take longer to go through the funding process attract a higher readership. In other words, these stories simply take longer to draw readership, because they are not in the funding process long enough to gain attention from the marketplace before publication.

In addition, we explored the degree to which inclusion of the *final\_contribution\_frequency* in the consequences model improves model fit, over and above a naïve model. We did this by estimating a simple model incorporating only the lag of readers and the log of website referrals. We then estimated our final model, as reported in Table 7, column two. Drawing on the log-likelihood of each of these nested models, we performed a likelihood ratio test. This test indicates that the inclusion of *final\_contribution\_frequency* vastly improves the model fit ( $\chi^2(1)=1,031.37$ ,  $p < 0.0001$ ).

### 5.3 Robustness Checks

We begin our robustness checks by ruling out issues of multicollinearity. To assess multicollinearity, we estimated both the antecedent and consequence models using cross-sectional (pooled) OLS regression and obtained the variance inflation factors (VIFs) associated with each variable. All values were found to be below 4, and hence multicollinearity is not of concern in our sample.

We next evaluated the robustness of our coefficient estimates, in the antecedent model, to outliers. To do this, we re-estimated the model using the GMM-based Dynamic Panel estimator,

excluding outlier observations that were identified based on a frequency value that lay within the top 5% of the distribution (this resulted in 395 excluded observations, or roughly 7% of the antecedent sample). This re-estimation produced coefficients that were very similar in terms of sign, statistical significance, and magnitude.

We also explored the sensitivity of the results to the choice of estimator. We re-estimated the model using the two-stage least squares estimator with fixed effects (2SLS). In this re-estimation, we instrumented for the lag of the dependent variable using time-series based instruments such as higher order lags of the same variable. This model (Table 8) again supports our core findings, producing an insignificant Hansen J statistic, which indicates that the overidentifying restrictions are not rejected. Further, the Cragg-Donald statistics in these estimations are ~45.0, which exceeds all Stock-Yogo cutoff values, indicating that our instrumentation is quite strong (Stock and Yogo 2002)<sup>13</sup>.

-----INSERT TABLE 8 ABOUT HERE-----

We also estimated a standard OLS regression with fixed effects, excluding the lagged dependent variable, in order to address any concerns about this variable being confounded with contribution. Upon doing so, we obtained coefficient estimates consistent with those reported above. These results are reported in Table 9

-----INSERT TABLE 9 ABOUT HERE-----

Next, in order to mitigate any concerns over the frequency variable potentially being endogenous, we re-estimated the model with the 2SLS estimator once more, instrumenting for both the lagged DV and the contribution frequency using higher order lags. This estimation again produced qualitatively similar results, which are reported in Table 10<sup>14</sup>.

-----INSERT TABLE 10 ABOUT HERE-----

---

<sup>13</sup> Notably, our Kleibergen-Paap F and LM statistics both exceed the Stock-Yogo cutoffs as well.

<sup>14</sup> It is important to note as well that the 2SLS regression reported in Table 10 also included an orthogonality condition, to evaluate the exogeneity of contribution frequency. The resultant C statistic suggested that we could not reject this variable's exogeneity (i.e., an insignificant C statistic;  $\text{Chi}^2(1) = 0.354$ ), further supporting our belief that frequency is exogenous.

Lastly, we re-executed our initial GMM-based dynamic panel data model including a variable that captures the number of days a pitch has been posted for. This model produced qualitatively similar results, and the duration variable was found to be insignificant (see Table 11, below). This estimation helps to rule out the alternative explanation that the effects we observe are simply a result of available funds running dry (i.e., the project reaching its market contribution capacity). It is interesting to note here that, in addition to further addressing the market capacity argument, inclusion of this posting duration variable also allows us to rule out an effect of waning interest, over time. To clarify, this check is necessary because one might reasonably suggest that pitches that have been posted for a longer duration of time will contain older ideas. As news is only valuable if it is recent, such pitches could be expected to draw less interest from readers.

-----INSERT TABLE 11 ABOUT HERE-----

In terms of robustness checks for our consequences model, beyond the simple checks for multicollinearity and outliers, we also considered that the variance of contribution might be different for different types of pitches. That is, pitches that received consistent levels of funding from contributors might perform quite differently than pitches that received a variety of contribution amounts. In order to control for this, we incorporated the variance of contributions into the negative binomial estimation (we refer to this term as *final\_variance*). When the model was re-estimated, *final\_variance* was found to be significant and positive, perhaps suggesting that pitches garnering support from a wider audience might also perform better in attracting readers upon publication. Regardless, inclusion of this variable did not influence the sign, significance or magnitude of our other main variables, thus we've excluded it from our main regression table for the sake of brevity. Next, we re-estimated the consequences model using a standard OLS fixed effects estimator. Importantly, the final frequency retained the same sign and significance as in the other estimations.

Finally, we also executed a Quantile Count regression. As noted previously, this method allows us to examine the stability of the identified effects across various points in the distribution of readership (i.e., for stories of varying quality in our sample). Thus, we can ensure that the high variability in readership across our sample is addressed without employing a log or square root transform, which have

been noted to severely impede the performance of count estimators (O'Hara and Kotze 2010). We estimated the marginal effects of each variable at the 0.45, 0.50 and 0.65 deciles of the outcome variable, readers. We report the marginal effects at the mean of the independent variables. Notably, the effects identified in our negative binomial regression are consistent in these median regressions across each selected decile. These results are reported in Table 12.

-----INSERT TABLE 12 ABOUT HERE-----

It should also be noted that we performed a number of additional robustness checks across both the antecedent and consequence models, to explore the effects of story topic genre, as well as different operationalizations for prior contribution (cumulative prior contribution, average prior contribution). These additional robustness checks are presented in Section 10, Appendix B.

## **6. Discussion**

We also considered some possible alternative contribution incentives, beyond those we explicitly factored into our empirical model. We considered the possible role of reputational gains, social identity, sanctions on deviants and money. Beginning with reputational gains, it is important to note that many of the contributors in this marketplace employ “user names” or pseudonyms. For this reason, contributors are, in general, fairly anonymous. It therefore seems unlikely that a contributor could obtain much in the way of reputational gains from a contribution. For similar reasons, it seems unlikely that individuals would view contribution as a viable opportunity to establish social identity or to associate themselves with a cause. In terms of sanctions on deviants, there do not appear to be any mechanisms by which sanctions could be imposed. We say this because a) users are generally anonymous, as noted above, and b) there does not appear to be any avenue by which an individual could be penalized by the community. Finally, with regard to monetary incentives, to our knowledge, there are no means by which a contributor could earn funds as a result of their contribution. We therefore feel confident that our empirical model strikes a good balance between parsimony and exhaustiveness.

The usefulness of popularity or preference indicators in predicting subsequent product popularity has been demonstrated previously within certain online contexts, such as YouTube (Cha et al. 2007; Duan et al. 2009; Tucker and Zhang 2011). However, no prior work has attempted to examine how users' reliance upon these types of indicators might play out in the context of private contribution toward a

public good, nor how this reliance might be associated with the quality of the public good once implemented. The empirical study conducted herein presents a unique effort to analyze these effects. While other scholars have recently examined private contribution toward a public good in the context of communities of practice (Wasko et al. 2009) and peer to peer file sharing networks (Asvanund et al. 2004; Krishnan et al. 2002; Nandi and Rochelandet 2008; Xia et al. 2011), this work advances the literature, examining contribution behavior around a new product type (online journalism), subject to the influence of observable contribution decisions made by prior others. Further still, our web traffic and product popularity variables allow us to address the fact that certain stories will have a broader appeal than others. That being said, the possibility remains that certain types of stories will attract different types of people, who contribute in different ways. As such, while we have identified the predominant influence in this marketplace (altruism and warm glow), it is possible that this marketplace is in fact made up of a combination of altruists and conditional cooperators.

The availability of data spanning 1.5 years allows us to gain an understanding of the longitudinal process underlying consumer decision-making in the presence of the aforementioned signals. While prior work has considered process characteristics of this form, it has done so with comparatively limited scope, examining, for instance, how the behavior of one reviewer will influence that of subsequent reviewers (Forman et al. 2008; Schlosser 2005). Our examination has thus allowed us to gain a more in-depth understanding of the influence that multiple signals have on contributor decisions, in tandem. These effects are rather complex, and suggest that users consideration of such signals (e.g., relying on them to different degrees and in different ways in different scenarios) may be highly dependent upon context. Our identification of a significant association between final contribution frequencies/densities and story readership also implies that such information either (i) reflects the inherent, objective quality of an author or pitch, or (ii) the broader population's perception of such that is carried through in the form of general interest in reading the content.

## **7. Implications for Practice**

Our paper solidifies the important role of online social influence in collective evaluation and crowd-based funding. Further, our paper empirically supports the notion of the substitutive effects in contribution toward public goods, providing some practical evidence, in a real-world setting, of a

phenomenon that has been examined at length via simulation and laboratory experiments (e.g., Carpenter 2007; Eckel 2008). Our findings provide evidence that users do consider the behavior of others in this context, and they suggest that this influence is contingent on the nature of the good in question. Whether this consideration takes place consciously or inadvertently is left as a subject for future research.

In terms of the implications of our findings for policy and regulation, the identification of social influence in this market raises the potential that some markets may enable fundraisers to ‘game’ the system. This would be true, for example, if the market exhibited a mechanism of reciprocity, or if herding behavior were prevalent (Burtch 2011). However, importantly, our findings demonstrate that reinforcing mechanisms are not guaranteed to emerge in all markets. As such, the need for involvement from regulatory bodies should likely be weighed on a case-by-case basis. Further, there is ongoing research in the area of manipulation resistant feedback mechanisms that could quite likely offer an avenue to mitigating the threat of gaming, even where viable, further diminishing the need for regulatory oversight (Resnick and Sami 2008).

The existence of a substitution effect also has interesting implications for individuals seeking to raise capital within electronic markets, as well as individuals seeking to consume, invest or contribute in said markets. Our findings indicate that the patterns of collective evaluation in the funding stage reflect the ultimate sentiment of the market once the project is implemented. This seems to suggest that our measures may constitute a useful information source from a managerial perspective. If contribution patterns can be leveraged to predict project quality (and thus sales performance), then such information could be leveraged for inventory management and sales forecasting purposes. Of course, the degree of improvement delivered from the use of our measures should also be given consideration in future work.

If the crowd-funding platform wished to mitigate the identified substitutive effect, it is important to consider the findings of associated work that has examined different approaches to boosting contribution toward public goods in the presence of crowding out. List and Lucking-Reiley (2002) consider two such treatments: subsidization and the offering of refunds. These authors find that both approaches successfully increase contribution levels, as both approaches effectively lower the marginal cost of contribution. In the case of subsidization (i.e., matching contributions), a given contributor’s donation to a project is effectively increased by some multiple, by the purveyor’s contribution. In the case of refunds, a potential contributor is given the peace of mind that they can withdraw their funds from a project at any

time during the course of the funding process, if they should change their mind. Thus, a reasonable suggestion would then seem to be for the crowd-funding platform to match contributions. However, the effects of subsidization can likely be improved if a more nuanced approach is taken.

Prior work has found that different types of contributors are more or less susceptible to information on contributions made by prior others when making their own decisions. In particular, Shang and Croson (2009) found that first-time contributors are more likely to positively reinforce the contributions of others than are repeat contributors, because first-time contributors are faced with a greater degree of uncertainty or ambiguity. With this in mind, providing subsidies to mitigate the substitutive effect might work best if targeted at one type of contributor versus another. First time contributors could be targeted with an initial, one time account credit upon registering with the marketplace, whereas matching contributors' donations in general could be a strategy to target repeat-contributors.

It is also important to consider the underlying psychological effect that drives the positive influence of contribution matching on individuals' willingness to contribute. In work related to the idea of subsidizing contributions toward a public good, Falk (2004) finds that supplying potential contributors with a small gift amplifies the amount they are willing to contribute toward a charity, as it incites reciprocal behavior. Returning to the notion of impure altruism and the warm glow effect, where contributors receive some utility from aiding others, the influence of subsidization that is identified in the above studies may in fact result from individuals' perception that the subsidy is a type of gift to them, as it amplifies their contribution. With this in mind, it is therefore possible that the initial gift need not be monetary in nature to achieve the same outcome. So, perhaps if the crowd-funding platform were to provide other types of gifts to potential contributors, this might serve to boost contribution levels. For example, the platform might consider offering free merchandise associated with the platform brand, or some form of recognition, when contributors surpass some threshold of contributions (e.g., a "premium contributor" designation on the website).

## **8. Conclusion**

This work presents a holistic analysis of the process of consumer decision-making in a crowd-funding context. *Our findings seem to suggest contributions in this marketplace are subject to a crowding out effect. This, in turn, suggests that individuals in this market may be primarily*

*motivated by altruism*. Further, we find that those signals that appear to drive contribution are also significantly associated with project performance. In particular, higher contribution frequency is found to be associated with poorer performance in terms of project output. This finding suggests that it may be possible to leverage data on the patterns of contribution in the earlier stages of a project to predict later performance. We have presented one of the first empirical attempts to understand the influence and implications of behavioral signals in crowd-funded marketplaces. With the emergence of “crowd-funding” as a viable business model, such marketplaces are providing users with the opportunity to express themselves in new ways, and to examine others’ behavior in new ways. The nascent body of literature on crowd-funding will likely benefit greatly from further consideration of these types of signals.

The results of our empirical analysis need to be understood in light of some limitations. Our data set does not allow us to draw direct association between pitch contributors and story readers. We observe these behaviors in aggregate, but are unable to determine whether the exact same individuals contributing funds are also reading the story that results. Similarly, while we have identified a negative effect on allocations toward a particular project by the market, it is likely that all crowd-funders do not behave in the same manner; some may be altruists and others conditional cooperators. Our descriptive regression simply indicates the predominant behavior in the market. Further, it is possible that, while contributions fall off for a particular project when prior contributions arrive more frequently, it may be that the market simply reallocates funds to different projects that are most in need.

Generalizability may be an issue, as the subject of the contribution is somewhat subjective in terms of its quality, being a form of ‘art’, experience good. Future research can build on our analysis by using data on funding and project outcomes from other crowd-funded markets. Future research can also explore the relationship between popularity indicators, behavioral signals and more traditional, explicit forms of investor or purchaser preference (e.g., online product reviews) in the prediction of subsequent sales performance (Archak et al. 2011). Finally, future work might also seek to elaborate upon the initial relationship we have identified here between the contribution frequency of the pitch and the performance of the published story, in terms of readership.

In conclusion, given crowd-funding’s significant economic potential, the design of these platforms warrants a good deal of caution. While numerous avenues for future research do remain, this work presents an initial step toward understanding individuals’ behavior in these markets. As such, it is



our hope that this work will provide insights to scholars and practitioners, informing design, as well as policy and regulation going forward.

## References

2011. The American Jobs Act. Democratic National Committee. Retrieved on November 5, 2011. <http://www.americanjobsact.com/helping-small-businesses.html>.
- Abrams, B.A., M.D. Schitz. 1978. The 'Crowding-out'; Effect of Governmental Transfers on Private Charitable Contributions. *Public Choice* **33**(1) 29-39.
- Adar, E., B.A. Huberman. 2000. Free Riding on Gnutella. *First Monday* **5**(10).
- Agarwal, A., C. Catalini, A. Goldfarb. 2010. Entrepreneurial Finance and the Flat-World Hypothesis: Evidence from Crowd-Funding Entrepreneurs in the Arts. *NBER Working Paper*.
- Aitamurto, T. 2011. The Impact of Crowdfunding on Journalism. *Journalism Practice* **5**(4) 429-445.
- Andreoni, J. 1989. Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence. *Journal of Political Economy* **97**(6) 1447-1458.
- Andreoni, J. 1990. Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *The Economic Journal* **100**(401) 464-477.
- Andreoni, J. 1995. Cooperation in Public-Goods Experiments: Kindness or Confusion? *The American Economic Review* **85**(4) 891-904.
- Andreoni, J. 1998. Toward a Theory of Charitable Fund-Raising. *The Journal of Political Economy* **106**(6) 1186-1213.
- Archak, N., A. Ghose, P. Ipeirotis. 2011. Deriving the Pricing Power of Product Features by Mining Consumer Reviews. *Management Science* Forthcoming.
- Arellano, M., S. Bond. 1991. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies* **58**(2) 277-297.
- Arellano, M., O. Bover. 1995. Another Look at the Instrumental Variable Estimation of Error-Components Models. *Journal of Econometrics* **68**(1) 29-51.
- Asvanund, A., K. Clay, R. Krishnan, M.D. Smith. 2004. An Empirical Analysis of Network Externalities in Peer-to-Peer Music-Sharing Networks. *Information Systems Research* **15**(2) 155-174.
- Bardsley, N., R. Sausgruber. 2005. Conformity and Reciprocity in Public Good Provision. *Journal of Economic Psychology* **26**(5) 664-681.
- Belleflamme, P., T. Lambert, A. Schwienbacher. 2010. Crowdfunding: Tapping the Right Crowd. *SSRN Working Paper*.
- Bernheim, B.D. 1994. A Theory of Conformity. *Journal of Political Economy* **104**(5) 841-877.
- Blundell, R., S. Bond. 1998. Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics* **87**(1) 115-143.

- Borck, R., B. Frank, J. Robledo. 2006. An Empirical Analysis of Voluntary Payments for Information Goods on the Internet. *Information Economics and Policy* **18**(2) 229-239.
- Burch, G. 2011. Herding Behavior as a Network Externality. *Proceedings of the International Conference on Information Systems*.
- Cameron, A.C., P.K. Trivedi. 2009. *Microeconometrics Using Stata*. Stata Press, College Station, TX.
- Carpenter, J. 2007. Punishing Free-riders: How Group Size Affects Mutual Monitoring and the Provision of Public Goods. *Games and Economic Behavior* **60**(1) 31-51.
- Cha, M., H. Kwak, P. Rodriguez, Y. Ahn, S. Moon. 2007. I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system. *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement*.
- Chauduri, A., T. Paichayontvijit. 2006. Conditional Cooperation and Voluntary Contributions to a Public Good. *Economics Bulletin* **3**(8) 1-14.
- Croson, R., J. Shang. 2008. The Impact of Downward Social Information on Contribution Decisions. *Experimental Economics* **11**(3) 221-233.
- Duan, W., B. Gu, A. Whinston. 2009. Informational cascades and software adoption on the Internet: An empirical investigation. *MIS Quarterly* **33**(1) 23-48.
- Eckel, C. 2008. Differences in the Economic Decisions of Men and Women: Experimental Evidence *Handbook of Experimental Economics Results*. Elsevier, 509-519.
- Falk, A. 2004. Charitable Giving as a Gift Exchange: Evidence From a Field Experiment. *Working Paper*.
- Forman, C., A. Ghose, B. Wiesenfeld. 2008. Examining the Relationship Between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets. *Information Systems Research* **19**(3) 291-313.
- Gu, B., Y. Huang, W. Duan, A. Whinston. 2009. Indirect Reciprocity in Online Social Networks - A Longitudinal Analysis of Individual Contributions and Peer Enforcement in a Peer-to-Peer Music Sharing Network. *McCombs Research Paper Series*.
- Hamilton, J.T. 2003. *All the News That's Fit To Sell: How the Market Transforms Information into News*. Princeton University Press, Princeton, NJ.
- Hausman, J., B. Hall, Z. Griliches. 1984. Econometric Models for Count Data with An Application to the Patents-R&D Relationship. *Econometrica* **52**(4) 909-938.
- Hughes, D., G. Coulson, J. Walkerdine. 2005. Free Riding on Gnutella Revisited: The Bell Tolls? . *IEEE Distributed Systems Online* **6**(6) 1-18.
- Kappel, T. 2008. Ex Ante Crowdfunding and the Recording Industry: A Model for the U.S.? *Loyola of Los Angeles Entertainment Law Review* **29** 375-385.
- Kaye, J., S. Quinn. 2010. *Funding Journalism in the Digital Age*. Peter Lang Publishing, Inc., New York.
- Keene, O. 1995. The Log Transformation is Special. *Statistics in Medicine* **14** 811-819.

- Kingma, B.R. 1989. An Accurate Measurement of the Crowd-out Effect, Income Effect, and Price Effect for Charitable Contributions. *Journal of Political Economy* **97**(5) 1197-1207.
- Krishnan, R., M. Smith, Z. Tang, R. Telang. 2002. The Virtual Commons: Why Free-Riding Can Be Tolerated in File Sharing Networks. *Proceedings of the International Conference on Information Systems*.
- Lin, M., N. Prabhala, S. Viswanathan. 2009. Judging Borrowers by the Company they Keep: Social Networks and Adverse Selection in Online Peer-to-Peer Lending. *SSRN Working Paper*.
- List, J., D. Lucking-Reiley. 2002. The Effects of Seed Money and Refunds on Charitable Giving: Experimental Evidence from a University Capital Campaign. *Journal of Political Economy* **110**(1) 215-233.
- Mehrhoff, J. 2009. A Solution to the Problem of Too Many Instruments in Dynamic Panel Data GMM. *Deutsche Bank Discussion Series 1: Economic Studies* **31**.
- Nandi, T., F. Rochelandet. 2008. The Incentives for Contributing Digital Contents Over P2P Networks: An Empirical Investigation. *Review of Economic Research on Copyright Issues* **5**(2) 19-35.
- O'Hara, R., J. Kotze. 2010. Do not log-transform count data. *Methods in Ecology and Evolution* **1**(2) 118-122.
- Reece, W.S. 1979. Charitable Contributions: New Evidence on Household Behavior. *The American Economic Review* **69**(1) 142-151.
- Resnick, P., R. Sami. 2008. Manipulation-Resistant Recommender Systems Through Influence Limits. *ACM SIGecom Exchanges* **7**(3) 1-4.
- Roberts, R. 1984. A Positive Model of Private Charity and Public Transfers. *The Journal of Political Economy* **92**(1) 136-148.
- Salganik, M., P.S. Dodds, D.J. Watts. 2006. Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science* **311**(854) 854-856.
- Schlosser, A. 2005. Posting versus lurking: Communicating in a multiple audience context. *Journal of Consumer Research* **32** 260-265.
- Schwiebacher, A., B. Larralde. 2010. Crowdfunding of Small Entrepreneurial Ventures. *SSRN Working Paper*.
- Shang, J., R. Croson. 2009. A Field Experiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods. *The Economic Journal* **119**(540) 1422-1439.
- Stock, J.H., M. Yogo. 2002. Testing for Weak Instruments in Linear IV Regression. *NBER Working Paper*.
- Sugden, R. 1984. Reciprocity: The Supply of Public Goods Through Voluntary Contributions. *The Economic Journal* **94**(376) 772-787.
- Tucker, C., J. Zhang. 2010. Growing Two-Sided Networks by Advertising the User Base: A Field Experiment. *Marketing Science* **29**(5) 805-814.

- Tucker, C., J. Zhang. 2011. How Does Popularity Information Affect Choices? A Field Experiment. *Management Science* **57**(5) 828-842.
- Varian, H. 1994. Sequential Contributions to Public Goods. *Journal of Public Economics* **53**(2) 165-186.
- Vesterlund, L. 2003. The Informational Value of Sequential Fundraising. *Journal of Public Economics* **87**(3-4) 627-657.
- Wasko, M., R. Teigland, S. Faraj. 2009. The provision of online public goods: Examining social structure in an electronic network of practice. *Decision Support Systems* **47**(3) 254-265.
- Xia, M., Y. Huang, W. Duan, A. Whinston. 2011. To Continue Sharing or Not to Continue Sharing? An Empirical Analysis of User Decision in Peer-to-Peer Sharing Networks. *Information Systems Research* Articles in Advance.

**TABLE 1**  
**Antecedents Model Variable Definitions**

<b>Variable</b>	<b>Definition</b>
1. contribution	Total dollars contributed to a pitch on a particular day.
2. remaining_budget	A given pitch's total budget outstanding as of a given day.
3. contribution_frequency	Total amount raised divided by days of funding (\$/day).
4. page_views	Number of pitch views on a given day.
5. search_trends	Google search trends associated with story keywords.
6. referrals	Number of page visits referred from external Internet sources.

**TABLE 2**  
**Consequences Model Variable Definitions**

<b>Variable</b>	<b>Definition</b>
1. readers	Total story page views over a given week.
2. referrals	Number of page visits referred from external Internet sources.
3. final_contribution_frequency	The final contribution frequency of the pitch associated with a given story.

**TABLE 3**  
**Descriptive Statistics of Antecedent Variables**

<b>Variable</b>	<b>min</b>	<b>max</b>	<b>mean</b>	<b>standard dev.</b>
1. contribution	0.00	4,749.00	1,123.87	85.52
2. remaining_budget	0.00	38,468.22	2,059.44	6400.59
3. contribution_frequency	0.00	500.00	14.58	31.33
5. page_views	0.00	692.00	9.13	20.85
6. search_trends	0.00	100.00	45.19	28.19
7. referrals	57.00	6,967.00	763.37	654.23

*Note: N = 5,874; All continuous variables are log-transformed in the analyses.*

**TABLE 4**  
**Descriptive Statistics of Consequence Variables**

<b>Variable</b>	<b>min</b>	<b>max</b>	<b>mean</b>	<b>standard dev.</b>
1. readers	0.00	960.00	11.21	37.29
2. final_contribution_frequency	0.00	177.00	18.81	28.30
3. referrals	126.00	17,405.00	5,615.80	3,298.59

*Note: n = 3,199; continuous variables are log-transformed in analyses, except for readers and lag of readers.*

**TABLE 5**  
**Results of Antecedents Model: Pitch-Day**

OLS-FE with Lags			
Variable	(1)	(2)	(3)
Log(contribution <sub>t-1</sub> )	<b>0.22*** (0.02)</b>	<b>0.20*** (0.02)</b>	<b>0.20*** (0.02)</b>
Log(contribution <sub>t-2</sub> )	--	<b>0.05** (0.02)</b>	<b>0.05** (0.02)</b>
Log(contribution <sub>t-3</sub> )	--	--	0.01 (0.02)
Observations	5,774	5,774	5,774
F	<b>12.38 (18, 99)</b>	<b>13.13 F(19,97)</b>	<b>12.54 (20, 96)</b>
AIC	28,312.39	27,754.29	27,243.15
BIC	28,432.29	27,880.52	27,375.67
Time Dummies	Yes	Yes	Yes

*Notes: \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ ; +  $p < 0.10$ ; Robust Standard Errors reported in brackets for coefficients, degrees of freedom for test statistics.*

**TABLE 6**  
**Results of Antecedents Model: Pitch-Day**

GMM-based Dynamic Panel Estimator (DV = contribution)				
Explanatory Variable	(1)	(2)	(3)	(4)
Log(contribution_frequency <sub>(t-1)</sub> )	<b>-0.32* (0.11)</b>	<b>-0.32** (0.11)</b>	<b>-0.32** (0.11)</b>	<b>-0.31** (0.11)</b>
Log(contribution <sub>(t-1)</sub> )	<b>0.29** (0.09)</b>	<b>0.29** (0.09)</b>	<b>0.29** (0.09)</b>	<b>0.29** (0.09)</b>
Log(page_views)	<b>1.29*** (0.09)</b>	<b>1.29*** (0.09)</b>	<b>1.29*** (0.09)</b>	<b>1.27*** (0.09)</b>
Log(remaining_budget)	--	0.01 (0.03)	0.01 (0.02)	0.01 (0.03)
Log(search_trends)	--	--	-0.02 (0.04)	--
Log(referrals)	--	--	--	0.11 (0.10)
Observations	5,774	5,774	5,774	5,774
Wald Chi <sup>2</sup>	<b>691.20 (20)</b>	<b>808.15 (21)</b>	<b>808.09 (22)</b>	<b>811.97 (22)</b>
AR(3)	-0.09	-0.10	-0.10	-0.11
Hansen J	Insignificant	Insignificant	Insignificant	Insignificant
Time Dummies	Yes	Yes	Yes	Yes

*Notes:*

1. \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ ; +  $p < 0.10$ .
2. Robust standard errors in brackets for coefficients, degrees of freedom in brackets for test statistics.
3. Lag of contribution instrumented using higher order lags.
4. Difference-in-Hansen tests are insignificant for all models (i.e., instrument sets are exogenous).

**TABLE 7**  
**Results of Consequences: Story-Week**

<b>Explanatory Variable</b>	<b>Negative Binomial</b> <b>(Dependent Variable = readers)</b>	
	<b>(1)</b>	<b>(2)</b>
Log(final_contribution_frequency)	<b>-0.17*** (0.03)</b>	<b>-0.14*** (0.03)</b>
Log(readers <sub>(t-1)</sub> )	--	<b>0.49*** (0.01)</b>
Log(referrals)	--	-0.03 (0.05)
Observations	3,073	3,073
-2LL	-8,565.28	-7,675.52
Wald Chi <sup>2</sup>	<b>536.39 (74)</b>	<b>2,153.85 (75)</b>
Time Dummies	Yes	Yes

*Notes:*

1. \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ ; +  $p < 0.10$ .
2. Standard errors reported in brackets for coefficients, degrees of freedom for test statistics.
3. Conditional fixed effects via dispersion parameter.



## ROBUSTNESS CHECKS

**TABLE 8**  
**Results of Antecedents Model: Pitch-Day**

Explanatory Variable	2SLS (Dependent Variable = contribution)	
	(1)	(2)
Log(contribution_frequency <sub>(t-1)</sub> )	-0.59*** (0.12)	-0.56*** (0.15)
Log(contribution <sub>(t-1)</sub> )	0.36*** (0.10)	0.36** (0.10)
Log(page_views)	1.24*** (0.08)	1.23*** (0.08)
Log(remaining_budget)	0.02 (0.03)	0.02 (0.03)
Log(search_trends)	-0.07 (0.05)	--
Log(referrals)	--	0.01 (0.08)
Observations	5,477	5,477
F	51.67 (22, 5362)	52.02 (22, 5362)
Hansen J	Insignificant	Insignificant
Cragg-Donald	45.34	45.77
Kleibergen-Paap F	28.28	28.65
Kleibergen-Paap LM	83.37	84.49
Time Dummies	Yes	Yes

Notes:

1. \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ ; +  $p < 0.10$ .
2. Robust standard errors in brackets for coefficients, degrees of freedom in brackets for test statistics.
3. Lag of contribution instrumented using higher order lags.

**TABLE 9**  
**Results of Antecedents Model: Pitch-Day**

Explanatory Variable	OLS with Pitch Fixed Effects (Dependent Variable = contribution)	
	(1)	(2)
Log(contribution_frequency <sub>(t-1)</sub> )	-0.19** (0.07)	-0.19** (0.07)
Log(page_views)	1.50*** (0.09)	1.50*** (0.10)
Log(remaining_budget)	-0.01 (0.03)	-0.01 (0.04)
Log(search_trends)	-0.03 (0.08)	--
Log(referrals)	--	-0.03 (0.09)
Observations	5,874	5,874
F	23.06 (21, 99)	22.88 (21, 99)
Overall R <sup>2</sup>	0.29	0.29
Time Dummies	Yes	Yes

Notes:

1. \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ ; +  $p < 0.10$ .
2. Robust standard error in brackets for coefficients, degrees of freedom in brackets for test statistics.

**TABLE 10**  
**Results of Antecedents Model: Pitch-Day**

<b>Explanatory Variable</b>	<b>2SLS /w Instruments for Lagged DV &amp; Contribution Frequency</b> <b>(Dependent Variable = contribution)</b>	
	<b>(1)</b>	<b>(2)</b>
Log(contribution_frequency <sub>t-1</sub> )	<b>-0.28*</b> (0.12)	<b>-0.27*</b> (0.12)
Log(contribution <sub>t-1</sub> )	<b>0.31***</b> (0.08)	<b>0.31***</b> (0.08)
Log(page_views)	<b>1.23***</b> (0.08)	<b>1.23***</b> (0.08)
Log(remaining_budget)	0.02 (0.03)	0.02 (0.03)
Log(search_trends)	-0.03 (0.05)	--
Log(referrals)	--	0.02 (0.08)
Observations	5,477	5,477
F	<b>52.53 (22, 5362)</b>	<b>52.76 (22, 5362)</b>
Hansen J	Insignificant	Insignificant
Cragg-Donald F	<b>51.91</b>	<b>50.58</b>
Kleibergen-Paap LM	<b>124.50</b>	<b>124.04</b>
Kleibergen-Paap F	<b>32.86</b>	<b>32.63</b>
Time Dummies	Yes	Yes

*Notes: \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ ; +  $p < 0.10$ ;*

*Robust standard errors in brackets for coefficients, degrees of freedom in brackets for test statistics; Lagged DV and contribution frequency instrumented with higher order lags; All Stock-Yogo cutoffs of instrument strength are met.*

**TABLE 11**  
**Results of Antecedents Model: Pitch-Day**

<b>GMM-based Dynamic Panel Estimator</b>		
<b>(DV = contribution)</b>		
<b>Explanatory Variable</b>	<b>(1)</b>	<b>(2)</b>
Log(contribution_frequency <sub>t-1</sub> )	<b>-0.32** (0.11)</b>	<b>-0.31** (0.11)</b>
Log(contribution <sub>t-1</sub> )	<b>0.28** (0.10)</b>	<b>0.29** (0.10)</b>
Log(page_views)	<b>1.28*** (0.09)</b>	<b>1.26*** (0.09)</b>
Log(remaining_budget)	0.01 (0.03)	0.01 (0.03)
Log(search_trends)	-0.03 (0.04)	--
Log(referrals)	--	0.11 (0.10)
Log(posting_duration)	-0.06 (0.05)	-0.06 (0.05)
Observations	5,774	5,774
Wald Chi <sup>2</sup>	<b>893.01 (23)</b>	<b>893.23 (23)</b>
AR(3)	-0.10	-0.11
Hansen J	Insignificant	Insignificant
Time Dummies	Yes	Yes

*Notes: \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ ; +  $p < 0.10$ ;*

*Robust standard errors in brackets for coefficients, degrees of freedom in brackets for test statistics; Lagged DV and contribution frequency instrumented with higher order lags; All Stock-Yogo cutoffs of instrument strength are met.*

**TABLE 12**  
**Results of Consequences Model: Story-Week**

<b>Quantile Count</b>			
<b>(Dependent Variable = readers)</b>			
<b>Explanatory Variable</b>	<b>Deciles</b>		
	<b>(0.25)</b>	<b>(0.50)</b>	<b>(0.75)</b>
Log(final_contribution_frequency <sub>(t-1)</sub> )	<b>-0.17*** (0.04)</b>	<b>-0.12*** (0.03)</b>	<b>-0.12* (0.05)</b>
Log(referrals)	0.10 (0.09)	0.05 (0.07)	0.02 (0.11)
Observations	3,175	3,175	3,175
Time Dummies	Yes	Yes	Yes

*Notes:*

1. \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ ; +  $p < 0.10$ .
2. Standard errors reported in brackets.

## **9. APPENDIX A: Variable Construction**

### **9.1 *Google Search Trends***

We retrieved data on Google Search Trends employing 2-5 story keywords that were associated with each pitch/story. These keywords included a) any proper nouns present in the title of the pitch, and b) the main story category, as specified by the website purveyor (e.g., education, politics, sustainability). If no proper nouns were present, we simply used the subject of the title. This combination of words was used as input to Google Insights for Search, in order to retrieve search trends over the period of observation in our dataset.

### **9.2 *Website Referrals***

This value was not constructed; rather, it was retrieved directly from the website's Google Analytics account. Referrals include the total daily number of unique visitors to the website that were referred from an external source (e.g., search engines, Facebook links, e-mail links, etc.). The number of referrals is distinct from the number of direct visitors, who accessed the website by typing the URL directly into a web browser.

## 10. APPENDIX B: Supplementary Robustness Checks

In addition to the primary Antecedent and Consequence models we report above, we also explored secondary models with alternative operationalizations of others' contribution, in place of prior contribution frequency. We felt this was necessary because we acknowledge the potential that crowd-funders might not concern themselves with funding durations (i.e., that they might only be concerned with aggregate prior contributions to the public good). In the below analyses, we propose and consider alternative measures, including cumulative contribution and average contribution (per contributor)<sup>15</sup>. Further, we perform additional robustness checks by controlling for the genre of the pitch/story.

### *Cumulative and Average Contributions*

To begin, Table 13, below, provides our re-estimation of the main model using cumulative contribution in place of contribution frequency. We would expect the effect of cumulative contributions to be in line with that of frequency, however, we would also expect cumulative contribution to be slightly less explanatory, given that it ignores time. Estimating this different model using the AB-BB produces results consistent with this expectation. The overall model Chi-square is much lower here than that of our main model, reported in Table 6 (~630 versus ~810) and, while both coefficient estimates are highly significant, the magnitude of the coefficient for frequency is three times that the coefficient for cumulative contribution.

**TABLE 13**  
**Results of Antecedents Model: Pitch-Day**

<b>AB-BB System GMM</b>		
<b>(Dependent Variable = contribution)</b>		
<b>Variable</b>	<b>(1)</b>	<b>(2)</b>
Log(cumulative_contribution <sub>t-1</sub> )	<b>-0.11*** (0.03)</b>	<b>-0.10*** (0.03)</b>
Log(contribution <sub>t-1</sub> )	<b>0.28** (0.09)</b>	<b>0.28** (0.09)</b>
Log(page_views)	<b>1.22*** (0.10)</b>	<b>1.19*** (0.10)</b>
Log(remaining_budget)	0.01 (0.02)	0.01 (0.02)
Log(search_trends)	-0.03 (0.04)	--
Log(referrals)	--	0.15 (0.11)
Observations	5,774	5,774
Wald Chi <sup>2</sup>	<b>633.69 (22)</b>	<b>632.38 (22)</b>
AR(3)	-0.10	-0.11
Hansen J	Insignificant	Insignificant
Difference in Hansen	Insignificant	Insignificant
Time Dummies	Yes	Yes

*Notes: \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ ; +  $p < 0.10$ ;*

*Robust standard errors in brackets for coefficients, degrees of freedom in brackets for test statistics; Lagged DV instrumented with higher order lags.*

Further, re-estimating this model using 2SLS produces qualitatively similar results, as reported in Table 14.

**TABLE 14**  
**Results of Antecedents Model: Pitch-Day**

<sup>15</sup> We would like to thank an anonymous reviewer for suggesting these measures.

<b>2SLS – Instruments for Lagged DV</b>		
<b>(Dependent Variable = contribution)</b>		
<b>Variable</b>	<b>(1)</b>	<b>(2)</b>
Log(cumulative_contribution <sub>t-1</sub> )	<b>-0.044<sup>+</sup> (0.026)</b>	<b>-0.043<sup>+</sup> (0.026)</b>
Log(contribution <sub>t-1</sub> )	<b>0.18** (0.08)</b>	<b>0.28** (0.09)</b>
Log(page_views)	<b>1.23*** (0.08)</b>	<b>1.22*** (0.08)</b>
Log(remaining_budget)	0.02 (0.03)	0.02 (0.03)
Log(search_trends)	-0.02 (0.05)	--
Log(referrals)	--	0.03 (0.08)
Observations	5,477	5,477
F	<b>52.99 (22, 5362)</b>	<b>53.17 (22, 5362)</b>
Hansen J	Insignificant	Insignificant
Cragg-Donald F	<b>59.05</b>	<b>58.28</b>
Kleibergen-Paap F	<b>37.02</b>	<b>36.68</b>
Kleibergen-Paap LM	<b>107.11</b>	<b>106.50</b>
Time Dummies	Yes	Yes

Notes: \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ ; +  $p < 0.10$ ;

Robust standard errors in brackets for coefficients, degrees of freedom in brackets for test statistics; Lagged DV instrumented with higher order lags; All Stock-Yogo cutoffs of instrument strength are met.

We also re-estimated our primary model, using average contribution in place of contribution frequency, where average contribution is defined as the average contribution amount, across all prior contributors. The results of this re-estimation with the AB-BB estimator are presented below in Table 15. Again, we see that average contribution is significant and in the expected direction, however, this model also produces a poorer fit statistic than the main model. The Wald statistic in Table 15 is roughly 690 while the Wald statistic of roughly 810 in Table 6. This suggests that the average of prior contribution has less explanatory or predictive power than the frequency of prior contribution.

**TABLE 15**  
**Results of Antecedents Model: Pitch-Day**

<b>AB-BB System GMM</b>		
<b>(Dependent Variable = contribution)</b>		
<b>Variable</b>	<b>(1)</b>	<b>(2)</b>
Log(average_contribution <sub>t-1</sub> )	<b>-0.34*** (0.07)</b>	<b>-0.34*** (0.07)</b>
Log(contribution <sub>t-1</sub> )	<b>0.29** (0.09)</b>	<b>0.29** (0.10)</b>
Log(page_views)	<b>1.21*** (0.10)</b>	<b>1.19*** (0.11)</b>
Log(remaining_budget)	0.02 (0.02)	0.02 (0.02)
Log(search_trends)	-0.01 (0.04)	--
Log(referrals)	--	0.15 (0.11)
Observations	5,774	5,774
Wald Chi <sup>2</sup>	<b>691.20 (22)</b>	<b>690.63 (22)</b>
AR(3)	-0.10	-0.11
Hansen J	Insignificant	Insignificant
Difference in Hansen	Insignificant	Insignificant
Time Dummies	Yes	Yes

Notes: \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ ; +  $p < 0.10$ ;

Robust standard errors in brackets for coefficients, degrees of freedom in brackets for test statistics; Lagged DV instrumented with

---

*higher order lags.*

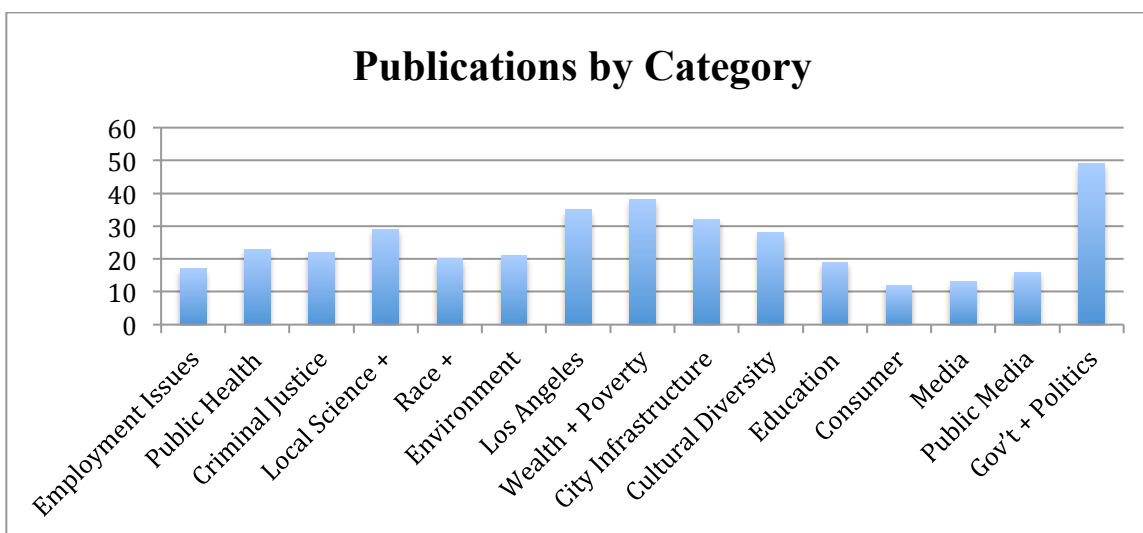
### ***Story Topic Genre***

We also explored the possibility that contribution behaviors might vary by story topic or genre. We collected additional data from the marketplace to incorporate this aspect into our models. As we show in the detailed analysis below, doing this does not qualitatively change the results we have reported above.

The marketplace purveyor maintains a set of story topic tags that are assigned to each article upon publication. We manually retrieved the tags for each of the stories in our dataset and incorporated them into our regression models using dummy variables. We provide a plot on the following page (Figure 4), indicating the counts of stories that have been assigned each tag. It is important to note that these tags are not mutually exclusive; a given story can be assigned multiple tags. The marketplace maintains 17 unique topic tags, of which 15 appear in our data set. The mean and median tag averages for all stories in our data set are 3.82 and 4.00, respectively, while the story having the broadest subject matter was assigned a total of 13 tags. The most prevalent topic tag is “Government + Politics” with 49 stories, while the least prevalent is “Consumer Protection.”

We next explored the possibility of differences in funding and readership across story genres by incorporating dummy variables into our regression models for different topic tags. We first estimated our antecedent model incorporating dummy variables for each story category, to determine if this would influence our key coefficient estimates (i.e., to confirm the presence or lack of any significant differences between story categories, beyond what would be captured by a fixed effect term). We find that including topics does not appear to significantly influence the estimated effect of contribution frequency. That is, our key variable (contribution frequency) remains stable in the presence of these dummies. The results of this regression are reported below in Table 16. Further, we obtain qualitatively similar results using a 2SLS estimator, with fixed effects (Table 17), or an OLS estimator without the lagged dependent variable.

**FIGURE 4**  
**Story Frequency by Topic**



**TABLE 16**  
**Results of Antecedents Model: Pitch-Day**

<b>AB-BB System GMM</b>	
<b>(Dependent Variable = contribution)</b>	
<b>Variable</b>	<b>(1)</b>
Log(contribution_frequency <sub>t-1</sub> )	<b>-0.41*** (0.11)</b>
Log(contribution <sub>t-1</sub> )	<b>0.29** (0.09)</b>
Log(page_views)	<b>1.28*** (0.10)</b>
Log(remaining_budget)	0.01 (0.02)
Log(referrals)	0.11 (0.10)
Observations	5,774
Wald Chi <sup>2</sup>	<b>1,071.56 (37)</b>
AR(3)	-0.10
Hansen J	Insignificant
Difference in Hansen	Insignificant
Time Dummies	Yes
Topic Dummies	All

*Notes: \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ ; +  $p < 0.10$ ; Robust standard errors in brackets for coefficients, degrees of freedom in brackets for test statistics; Lagged DV instrumented with higher order lags.*

**TABLE 17**  
**Results of Antecedents Model: Pitch-Day**

<b>2SLS – Instruments for Lagged DV</b>	
<b>(Dependent Variable = contribution)</b>	
<b>Variable</b>	<b>(1)</b>
Log(contribution_frequency <sub>t-1</sub> )	<b>-0.56*** (0.15)</b>
Log(contribution <sub>t-1</sub> )	<b>0.36*** (0.10)</b>
Log(page_views)	<b>1.23*** (0.08)</b>
Log(remaining_budget)	0.02 (0.03)
Log(referrals)	0.01 (0.08)
Observations	5,477
F	<b>52.02 (22, 5362)</b>
Hansen J	Insignificant
Cragg-Donald F	<b>45.77</b>
Kleibergen-Paap F	<b>28.65</b>
Kleibergen-Paap LM	<b>84.49</b>
Time Dummies	Yes
Topic Dummies	All

*Notes: \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ ; +  $p < 0.10$ ; Robust standard errors in brackets for coefficients, degrees of freedom in brackets for test statistics; Lagged DV instrumented with higher order lags; All Stock-Yogo cutoffs of instrument strength are met.*



We next re-estimated the consequences model to see if the story topic might influence readership. Table 18, below, provides a re-estimation of our negative binomial model. Incorporating the topic dummy variables into the model does not qualitatively change our key coefficient estimates. Further, excluding the lagged dependent variable from the estimation had no qualitative impact on the coefficient estimates (Table 19).

**TABLE 18**  
**Results of Consequences Model: Story-Week**

<b>NBREG-FE (Dependent Variable = readers)</b>	
<b>Variable</b>	<b>(1)</b>
Log(final_contribution_frequency <sub>t-1</sub> )	<b>-0.16*** (0.03)</b>
Log(readers <sub>t-1</sub> )	<b>0.50*** (0.01)</b>
Log(referrals)	-0.04 (0.05)
Observations	3,073
-2LL	-7,628.14
Wald Chi <sup>2</sup>	<b>2,315.34 (90)</b>
Time Dummies	Yes
Topic Dummies	All

*Notes: \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ ; +  $p < 0.10$ ; Bootstrap standard errors reported in brackets for coefficients, degrees of freedom for test statistics; Fixed effects introduced via negative binomial dispersion parameter.*

**TABLE 19**  
**Results of Consequences Model: Story-Week**

<b>NBREG-FE (Dependent Variable = readers)</b>	
<b>Variable</b>	<b>(1)</b>
Log(final_contribution_frequency <sub>t-1</sub> )	<b>-0.15*** (0.03)</b>
Log(referrals)	0.04 (0.06)
Observations	3,175
-2LL	-8,592.21
Wald Chi <sup>2</sup>	<b>606.81 (90)</b>
Time Dummies	Yes
Topic Dummies	All

*Notes: \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ ; +  $p < 0.10$ ; Bootstrap standard errors reported in brackets for coefficients, degrees of freedom for test statistics; Fixed effects introduced via negative binomial dispersion parameter.*

Lastly, we explored whether there is a significant interaction between the effect of contribution frequency and story topic, focusing on the top two most common story topics in our sample (“Government + Politics” and “Wealth + Poverty”). The results indicated that there are no main effects for these two story topic categories, nor are there interaction effects.