The Effect of Crowd Voting on Participation in Crowdsourcing Contests

Liang Chen
West Texas A&M University
WT Box 60268, Canyon, TX 79016.
lchen@wtamu.edu
Phone: (806) 651-2469

Pei Xu
Raymond J. Harbert College of Business, Auburn University
405 W. Magnolia Ave., Auburn, AL 36849
pzx0002@auburn.edu
Phone: (334) 844-6513

De Liu
Carlson School of Management, University of Minnesota Twin Cities
321 19th Avenue South Minneapolis, MN 55455
deliu@umn.edu
Phone: (612) 626-4480

Accepted at JMIS

Feb, 2020

The Effect of Crowd Voting on Participation in Crowdsourcing Contests

Abstract

While expert rating is still a dominant approach for selecting winners in contests for creative works, a few crowdsourcing platforms have recently used “crowd voting” for winner selection – i.e., let users of the crowdsourcing community publically vote for contest winners. We investigate how a contest’s reliance on crowd voting for winner selection, defined as the percentage of crowd-voted prizes to the total prize sum (in dollar amounts), affects contest participation. Drawing upon expectancy theory and tournament theory, we develop a theoretical understanding of this relationship. Using a novel dataset of contests employing both crowd voting and expert rating, we find that a contest’s reliance on crowd voting is positively associated with participation. Specifically, every 10% increase in the crowd-voting reliance can boost users’ odds of participation by about 7%. Moreover, crowd voting is more appealing to users whose expertise is not high and whose status in the crowdsourcing community is high.

Keywords: crowdsourcing contests, winner-selection mechanisms, crowd voting, expert rating, expectancy theory, tournament theory

The Effect of Crowd Voting on Participation in Crowdsourcing Contests

INTRODUCTION

Historically, the main mechanism for selecting the best creative ideas has been "*expert rating*," where a small group of experts chooses winners from a pool of candidates [32]. With crowdsourcing platforms becoming a widely popular way to acquire innovative ideas [1], a relatively new form of idea selection mechanism has emerged: instead of relying on a small panel of experts, one can tap into the vast crowdsourcing user community and let community users publically vote for best ideas. This approach, which we call "crowd voting", seems popular among crowdsourcing contests for creative works [46].[1] For instance, at Threadless, a crowdsourcing platform for apparel design, users can vote for their favorite T-shirt designs submitted by peers in the community, and the most-voted designs each week are printed and sold worldwide through the platform's online and retail stores. Besides Threadless, crowd voting is used in a few other crowdsourcing platforms, including Jovoto (brand innovation), PimTim (graphic design), Zooppa (marketing creative design), and Lego Ideas (Lego design). Zooppa, in particular, allows sponsors to offer expert-rated and crowd-voted prizes in the same contest.

Crowd voting has a few advantages over expert rating: it uses a large number of unpaid volunteers as selectors, and thus could be cheaper to operate and scale better than expert rating [15, 37, 53]. It may also facilitate collaboration and co-evolvement of knowledge and get more users involved in the crowdsourcing community [26, 47, 53]. Given these advantages, it is natural to ask whether contest sponsors should choose a winner-selection design that relies more on crowd voting, especially when they can offer expert-rated and crowd-voted prizes in the same

[1] According to Magallanes, Sánchez, Cervantes and Wan [52], a creative work is a manifestation of creative effort as in the formation of concepts, artwork, literature, music, paintings, software, and architectural designs. Creative works have in common a degree of arbitrariness, such that it is improbable that two people would independently create the same work.

contest.[2] However, there is little research on the implications of crowd voting for contest sponsors and crowdsourcing platforms. A few empirical studies comparing crowd voting and expert rating focus on the quality of judgment [32] and agreement between crowds' and experts' winner selections [59]. These studies implicitly assume that expert-rated contests would attract the same set of submissions as crowd-voted ones. Anecdotal evidence suggests, however, that users' motivation for participating in a contest may change when they learn that winners will be selected by crowd voting as opposed to expert rating. The goal of this research is to examine *how a contest's reliance on crowd voting for winner selection affects participation*.

We define an objective measure of "*crowd-voting reliance*" as the percentage of the prize sum judged by crowd voting (as opposed to expert rating) in a contest. For example, at Zooppa, a crowdsourcing platform for marketing creatives, a contest called "I love Italian Shoes" provides $1,720 worth of prizes judged by expert rating and $2,100 worth of prizes judged by crowd voting. By our definition, this contest's crowd-voting reliance is 55 percent. Similarly, a contest that offers only expert-rated (crowd-voted) prizes has zero percent (100 percent) crowd-voting reliance.

We focus on the effect of crowd-voting reliance on *contest participation* for two reasons. First, participation is an important goal on its own for contests of creative works [31, 71]. By attracting more participants with different backgrounds, contest sponsors can increase their chances of finding exceptional or unique solutions [71]. In addition, a large number of participants can help contest sponsors increase their brand awareness. Second, because participation is a precursor to other contest outcomes, it is important to account for participation, even when the goal is to, say, compare winner performance under different winner-selection

[2] One may combine crowd voting and expert rating in other ways (e.g., using crowd voting for initial screening and expert rating for selecting final winners). We focus on the side-by-side use of crowd voting and expert rating.

mechanisms.

We examine the effect of crowd-voting reliance on *the number of contest participants* as well as *the type of participants*. For the latter, we consider two dimensions: a user's *expertise* and *status* in the crowdsourcing community. Expertise is the underlying reason for superior task performance and a form of "human capital." Status, on the other hand, represents a user's position in a crowdsourcing user community that results from accumulated acts of deference by other users [64] and is a form of "social capital." A user's status in the community may not be aligned with his/her expertise – e.g., a user with high expertise may not have a high status if the user does not interact with community peers or act negatively in the community. In sum, we focus on two research questions in this study: (1) *Does a crowdsourcing contest's crowd-voting reliance affect a user's probability of participating in the contest?* (2) *Does the effect of crowd-voting reliance on participation differ by users' expertise and status in the crowdsourcing community?*

To address these research questions, we first develop a theoretical framework that combines expectancy theory and tournament theory. The former is used to establish the relationship between the motivation for participation and the winning expectancy, and the latter helps us theorize how crowd voting differs from expert rating in winning expectancies. Building on such a theoretical framework, we formulate hypotheses on the effects of crowd-voting reliance on participation.

We test our hypotheses using a novel dataset from Zooppa. This platform organizes crowdsourcing contests for user-generated marketing creatives (video or print advertisements) that offer (a) prizes judged by an expert panel chosen by the client (i.e., contest sponsor), (b) prizes judged by crowd voting, or (c) both types of prizes. Moreover, when a contest offers both types of prizes, expert raters and crowd voters operate independently from each other, providing us a unique opportunity for separating their effects. We test our hypotheses by leveraging the variations in the crowd-voting reliance across 102 contests while controlling for many individual- and contest-specific factors that may also affect participation.

To our knowledge, the crowdsourcing contest literature has not examined the effect of winner-selection mechanisms on participation. Prior research suggests that participation in crowdsourcing contests is a function of design characteristics (e.g., reward size, the number of rewards, contest duration, and task difficulty) and contestant characteristics (e.g., skill, experience, tenure, and social activities) [5, 20, 49, 51, 82-84]. The literature also suggests that different contestants may prefer different contests: e.g., open-source developers sort by the degree of openness of the project's license [9, 68]. Nearly all of the existing studies, however, focus on expert-rated contests or contests with objective winner-selection criteria. They have not examined how the design of winner-selection mechanisms affects participation.

A form of crowd voting is also used in open innovation, where an organization uses an online platform to facilitate knowledge creation and innovation by external entities such as customers. Open innovation platforms ask crowds to vote on submitted ideas, though there may not be a contest. A few studies in this literature examine user participation [20, 30], but they have not studied the role of winner-selection mechanisms.

Our research belongs to the broader tournament-theory literature that examines equilibrium behaviors in contests/tournaments (see Connelly, Tihanyi, Crook and Gangloff [22] for a review). A few papers in this literature have studied how contest participation is affected by design features such as prize structure [50, 79], reward size [23], handicapping rules [50], and entry fees [79]. To our knowledge, this literature has not modeled subjective winner-selection mechanisms or their relationship with participation.

As mentioned earlier, only a handful of papers have compared crowd voting and expert rating in field settings. Using a Eurovision Song Contest dataset that includes both expert-rated and televoter-judged contests, Haan, Dijkstra and Dijkstra [32] show that experts are better judges of quality, in the sense that they are less sensitive to contestants' order of appearance than televoters. Using rating data on motion pictures, Holbrook and Addis [36] find that there is a weak relationship between expert and audience ratings of movies. Recently, Mollick and Nanda [59] compare outcomes of crowdfunding art projects and expert opinions on the same projects and find a significant agreement between the two. Unlike ours, these studies are not concerned with contest participation, but with quality of crowds' judgment or agreement between crowds and experts. Moreover, they do not have two winner-selection mechanisms operating independently in the same contest.[3]

The "wisdom of crowds" literature has demonstrated that a large crowd can sometimes beat a panel of experts in predicting outcomes such as presidential elections, sports, and new office openings [55, 70, 81]. While the same "diversity-trump-expertise" argument forwarded by this literature may also be at play in our setting, an important difference is that we focus on the use of crowds in *subjective* winner selection rather than predicting a fact. The wisdom-of-crowds literature holds that one of the key conditions for "wisdom of crowds" to work is that crowd participants must make their predictions independently. Clearly, crowds in our context are connected by social networks and not independent.[4]

THEORETICAL BACKGROUND AND HYPOTHESES

[3] Haan, Dijkstra and Dijkstra [32] rely on comparing contests judged by experts and a different set of contests judged by televoters. Mollick and Nanda [59] obtain expert opinions through separate surveys outside of the crowdfunding platform. Holbrook and Addis [36] have both expert and audience ratings of the same movies but these two types of ratings typically influence each other.

[4] The wisdom of crowds has also been applied in organizational designs, in the form of "organizational democracy," where employees are empowered to collectively make decisions on workplace issues through direct or representative joint consultation, dialogue, voting, co-determination, or other democratic processes [42, 78]. This literature considers a broader set of decisions (e.g. a buyout deal) and processes (e.g. dialogue) whereas we focus on the application of crowd voting as a winner-selection mechanism in contests for creative works.

In this section, we argue that a contest's crowd-voting reliance is positively associated with contest participation. Furthermore, the relationship is moderated by a user's expertise and status, as shown in Figure 1. We base our arguments on expectancy theory and tournament theory, along with characteristics of crowd voting and expert rating.

<Insert Figure 1 about here>

Overall Effect of Crowd-Voting Reliance on Participation

To establish the relationship between winner-selection mechanisms and participation, we draw on expectancy theory, a high-level theoretical framework for understanding the motivation of choosing between alternatives with uncertain outcomes. Expectancy theory states that one's motivation to select each alternative is determined by an expectation or perceived probability that his action can result in rewards, called *expectancy*, and the subjective value of rewards associated with the alternative, called *valence* [48, 75]. Expectancy theory has proven to be useful in several organizational contexts such as employee motivation and behavior. Applied to our context, expectancy theory suggests that motivation to participate in a contest is determined by the expectancy of winning the contest and rewards associated with winning.[5] Indeed, prior research has used expectancy theory to establish antecedents of participation in crowdsourcing, including extrinsic rewards, task characteristics, and competition intensity [49].

While expectancy theory is useful for establishing the connection between participation and winning expectancy, it does not offer insights on how winner-selection mechanisms affect winning expectancy. To this end, we leverage tournament theory, which models contests from a game-theoretic perspective. According to tournament theory, expertise, effort, and uncertainty

[5] Crowdsourcing research shows that participants of crowdsourcing contests may also be motivated by intrinsic motivations of self enhancement, enjoyment, and autonomy [16, 84]. We focus on the tangible rewards because, as we will argue, the choice of winner-selection mechanism has a direct impact on the expectancy of tangible rewards but the same cannot be said about intrinsic rewards.

can all contribute to a participant's performance in a contest, which in turn determines her chance of winning [72]. The weight of each component depends on the characteristics of the task and the evaluation criteria. For example, performance on "trial-and-error" tasks depends more heavily on uncertainty and effort than on expertise [72]. We argue that the choice of winner-selection mechanism affects the composition of expertise, effort, and uncertainty in performance evaluation, thus can affect expectancy.

We argue that crowd voting increases the element of uncertainty in performance evaluation for two main reasons. First, we note that crowd voting and expert rating differ in *selectors*. Crowd voting relies on a large number of users in a crowdsourcing community, who tend to have diverse backgrounds and varying levels of expertise. A typical crowdsourcing community may consist of novices, amateurs, and even professionals, but an average user in the community is not an expert. In contrast, expert rating relies on a small group of expert judges who have extensive and authoritative knowledge and are familiar with the standard of excellence in the relevant field [35]. The literature has suggested that the judgments of non-experts have greater variance, lower agreement, and larger errors than those of experts [65, 66]. Thus, when a larger portion of the prize sum is judged by crowd voters, the element of uncertainty will play a bigger role in a user's winning expectancy.

Second, at the level of the selection system, increased reliance on crowd voting can strengthen a contest's exposure to the "herding effect" among crowd voters, which adds more uncertainty to outcomes. Crowd voting is typically organized as an open, dynamic voting process with a tally of ongoing total votes for each candidate. Because crowd voters can observe the number of existing votes when casting their votes, subsequent voters may discard their own judgment in favor of following the opinions of others, causing a herding effect [6]. Theoretical models of herding predict that it can cause high uncertainty in final rankings because a small disturbance in the decisions made by early evaluators can cause a cascading effect in the subsequent evaluations [10]. Recent experimental results confirm such a prediction: a community's collective ranking of songs becomes more uncertain when members are exposed to prior ratings by other members of the community [19, 63].

According to tournament theory, as a contest's performance evaluation becomes more random (as a result of non-expert selectors or herding), the winning chances will become more dispersed and more users have a chance of winning [13, 28, 43, 50]. By expectancy theory, this means that more users will choose to participate in the contest. This argument is supported by an experimental finding that added uncertainty in contest outcomes leads to excess entry into the contests [43]. In sum, when a contest relies more on crowd voting for winner selection, we expect increased participation:

H1*: A user's probability of participating in a contest* increases with the *contest's* reliance on *crowd voting for winner selection.*

We next explore how the effect of winner-selection mechanisms differs by users' human capital (expertise) and social capital (status).

Moderating Effect of Expertise

Expertise reflects a user's competence and knowledge in a specific domain. People with high expertise are expected to demonstrate the general superiority of their performance in a repeatable and reproducible fashion [24]. Expertise is a form of "human capital" obtained through professional training and practical experience [25] and indicated by the achievement of superior performance [82].

Per our earlier arguments, increasing a contest's reliance on crowd voting can increase the element of uncertainty in performance evaluation. This effectively dilutes the role of expertise in performance. For most users, this means an increase in their winning chance; but for users with high expertise, their winning chance will suffer when expertise matters less. Thus, users with higher expertise are less motivated to participate in a contest but others are more motivated when there is a stronger element of uncertainty in performance evaluation. This asymmetric effect of added uncertainty has been noted in both theoretical research [72] and experimental study [44] in the tournament literature. Extending this insight to our setting and noting the impact of expectancy on motivation, we hypothesize:

H2: *As a user's expertise increases, a contest's reliance on crowd voting has a smaller impact on the user's probability of participation*.

Moderating Effect of Status

Status represents the position in a social community that results from accumulated acts of deference [64]. Status captures differences in social rank that generate privilege or discrimination [77]. As a collective assessment of overall deference as seen or judged by peers [54, 67], status is a form of "social capital." We note that status fundamentally differs from expertise. Expertise captures a person's domain-specific knowledge obtained through experience and training, while status represents a user's community-specific position built through repeated social interactions with peers. Status may not be aligned with expertise. For example, a user with high expertise may not obtain a high status among peers if the peers are not aware of his expertise or do not value their interactions with the user [8]. In addition, one typically judges a user's expertise by the user's performance, but a person's status attainment can be based on factors other than his performance [32].

Though the tournament theory literature does not consider a contestant's status as a determinant of performance, we argue that, in *subjectively* evaluated contests, it may play a role. Specifically, crowd voters are more likely to consider a contestant's status and vote for a high-status contestant than expert raters for two reasons. Our first argument is based on social capital. Social capital research has shown that individuals of high status have developed social connections, which in turn help them get attention, credits, and positive perceptions [61]. For example, Hutter, Füller, Hautz, Bilgram and Matzler [38] demonstrate that high-status users can prompt social connections to rate their submission positively. In our context, crowd voters are in the same community as contestants; a contestant with high status is more likely to accrue votes because of their social connections with voters [56]. In contrast, expert raters are recruited to be independent judges and are typically not part of the crowdsourcing community. Therefore, a contestant's status capital does not matter as much under expert rating.

Our second argument is based on judgment style. Prior research on individual judgment suggests that expert evaluators are known to focus more on intrinsic quality and less affected by peripheral cues [32]. In our setting, they are hired to do so. In contrast, crowd voters, who are voluntary evaluators and on average non-experts, are more likely to use peripheral cues to inform their judgments. Hence, crowd voters, because of their stronger reliance on status cues, are more likely to vote for contestants of high status than experts. In sum, increased reliance on crowd voting benefits high-status users who are more likely to get crowd votes and win crowd-voted prizes, which increases their participation motivation. We therefore hypothesize:

H3**:** *As a user's status in the crowdsourcing community increases, a contest's reliance on crowd voting has a greater impact on the user's probability of participation.*

EMPIRICAL CONTEXT

Our research context is Zooppa.com, a global crowdsourcing platform that helps companies such as Google and General Mills produce user-generated marketing campaigns in the form of video, print, and concept creatives. Since its opening in 2007, Zooppa has become one of the major platforms for user-generated campaigns. It holds about two dozen contests per year. As of mid-2016, Zooppa had provided over $6 million cash prizes.

A unique feature of Zooppa is that a contest sponsor may offer both expert and crowd prizes, judged separately by an expert panel and peers on the platform. Expert prizes (called "client-selected awards" by Zooppa) are chosen after the contest deadline by experts appointed by the contest sponsor, whereas crowd prizes (called "voter awards" by Zooppa) are voted on by members of the Zooppa community during the contest. Occasionally, the contest sponsor may choose to offer special prizes such as early entry and honorary prizes; these are also judged by experts. We next discuss the contest procedure in detail.

After a contest sponsor decides to hold a contest on Zooppa, it determines and uploads contest terms in three documents (Figure 2-b): a brief (e.g., introduction, awards, selection process, judging criteria, and deadline), rules (e.g., technical requirements for video files), and downloadable materials (e.g., logos). All contest terms, including types and the dollar amount of prizes, are announced at the launch of the contests and remain unchanged. Though experts' judging criteria for these contests vary from contest to contest, some major components, including engaging storytelling, capturing brand value, originality, and production quality, are common among contests.[6]

<Insert Figure 2 about here>

[6] For example, contests "*Smile You're at Red Robin*" and "*The UK's Fastest Network*" have four equally weighted criteria: engaging storytelling, positive representation of the brand, originality, and production quality. "Squeeze More Out!" has these four: the functional benefits of the new bottle are communicated effectively (40%), the delicious-looking sandwiches make us drool (20%), videos are creative and unique (20%), videos are of high production quality (20%).

Once a contest is launched, members of the Zooppa community will receive a notification about the new contest and can start submitting until the contest deadline. Multiple entries are allowed but rare. As entries come in, members can use Zooppa's website to browse contests (Figure 2-a), view submitted entries, and comment on them (Figure 2-c). They can also browse a member's profile page, which contains information such as user name, photo, short bio, past submissions, and awards. Members cannot contact the contest sponsor directly, because Zooppa takes overall administrative matters, including answering questions (via an online forum) and screening submissions for adherence to contest rules. If a contest uses crowd voting, community members can also vote on entries. Voting starts as soon as entries are posted. All registered members are eligible to vote until the contest deadline. Zooppa classifies its members into junior and senior members, based on prior submission and winning experiences. Each vote has an associated point value. A junior member can give no more than 5 points per vote, whereas a senior member can give no more than 20 points per vote. Zooppa automatically calculates total points, i.e., *voting scores*, and displays them in real-time. The final voting scores are used to determine crowd ranking and prizes.

After a contest is closed, Zooppa forwards all the submitted files to the contest sponsor. If the contest offers expert prizes, the contest sponsor will have their panel of expert judges select winners. The expert panel includes a wide variety of professionals such as chief marketing officer, film director, creative director of a media production company, social media and marketing expert, professor, etc. A sample of expert judges from one contest is shown in Table 1. Zooppa staff cannot be on the expert panel. Zooppa asks each panelist to enter his/her ratings and comments through a web portal. Final winners are determined based on all panelists' ratings and comments. According to our interviews with Zooppa staff, expert panelists are not part of the

Zooppa user community and are not provided with crowd-voted scores. Hence, expert rating and crowd voting operate independently of each other. Within two weeks of the contest deadline, all winners are announced, and prizes are given (Figure 2-d).[7]

RESEARCH METHODOLOGY

Data and Variables

We focus on video creatives, as these are the dominant form of creatives at Zooppa. We obtain data on all of the 132 contests held between Dec 2007 and Aug 2013,[8] but exclude 30 from further analysis, including ten that accepted only print creatives, nine that offered no prize, and 11 that were sponsored by Zooppa (e.g., Best of Zooppa contests). Among the 102 remaining contests, 44 offered both expert and crowd prizes, 52 offered only expert prizes, and 6 offered only crowd prizes. On average, the total prize in dollar amount per contest was $13,301, the number of prizes per contest was 10, and the number of expert prizes per contest was 6. On average, each contest received 71 valid entries and lasted 61 days.

Dependent Variable. We construct a user-contest panel to study the relationship between the reliance on crowd voting and participation. Specifically, we include a user-contest pair in the panel if the user meets the following two criteria: first, the user registered before the contest's deadline so that a participation decision is necessary. Second, the user had at least one community activity (commenting, voting, or submission) in the two years previous to the current contest (additional time windows are tested in our robustness checks). If users were no longer active when a contest was launched, their participation decisions would have been negative

[7] Contest sponsors own the copyright of winning entries and can use them on their media channel without paying use fees. Contest sponsors can also use the work of any other member in addition to the work of the award-winning member, on its media channels, at any time after an award is granted. Members whose work is selected for use by the sponsor will receive a use fee of $500.

[8] After Aug 2013, Zooppa changed its website and rarely used crowd voting.

regardless of the factors we suggested in the current study [57]. With this user-contest panel construction, we generate 158,202 observations of 3,833 distinct users.[9] For each user-contest pair, we may observe zero, one, or a few video entries. This panel data is unbalanced, as some users entered the community earlier than others and thus had more observations. Because very few people submit multiple entries to a contest on Zooppa, we dichotomize the participation. That is, we observe one or more video entries from user i in contest j, we code the participation variable $EnterContest_{ij}$ as 1. Otherwise, we code $EnterContest_{ij}$ as 0.[10] Such a binary specification is simpler and provides a fixed-effects estimator.[11] As a check of robustness, we also estimate a count model that accounts for multiple submissions.

Reliance on Crowd Voting. We measure a contest's crowd-voting reliance as 100 times the ratio of the sum of crowd-voted prizes to the total prize sum (in dollar amounts). The variable, labeled as *CrowdReliance*, may take values of 0 (all expert-rated prizes), 100 (all crowd-voted prizes), or any number in between (indicating a mixture of crowd-voted and expert-rated prizes). For example, Best Western hosted a contest that provided two expert-rated prizes with a total amount of $1,000, and five crowd-voted prizes with a total amount of $2,460. The *CrowdReliance*, in this case, takes a value of 71.1.

Expertise. A user's expertise is latent and manifests itself by one's task performance. Prior studies have used individual performance over a set of questions/events to evaluate the person's expertise [18, 76]. In our context, since the more accepted form of performance evaluation is done by experts hired by the contest sponsors, we measure *expertise* by a contestant's lifetime hit

---

[9] The user fixed-effect estimation further removes users who had no variation in participation, retaining a set of 2,635 distinct users in the final estimation.

[10] Each registered user receives an email notification when a new contest is announced. Thus, it is reasonable to interpret a lack of video entries from this user as the user choosing not to participate in this contest.

[11] Although a fixed-effects estimator has been proposed for the negative binomial model for count data [33], Allison and Waterman [4] pointed out that it is not a true fixed-effects estimator.

rate in expert-rated prizes for video submissions. The variable *expertise* also has values between 0 and 100, with 100 indicating every submission has resulted in an expert-rated prize. This measure provides the best estimate of a user's success rate by leveraging data points over a long horizon and reflects the notion that a user's expertise is a relatively stable attribute.[12] As a robustness test, we use an alternative measure of expertise that calculates a user's lifetime hit rate in any type of prize (expert-rated or crowd-voted).

Status. Status represents a user's community-specific position built through repeated social interactions with members in a particular group [27]. As such, an appropriate indicator of status should be based on how well an individual is recognized by the community, e.g., in the forms of acknowledgments, praise, and compliments. On the Zooppa platform, comments on entries are the primary channel for users to express acknowledgments and compliments. The vast majority of peer comments are compliments or expressions of gratitude (see Online Appendix for an example).[13] We, therefore, capture a user's status[14] using the number of unique users who have commented on the focal user's video submissions before the start of the current contest. This measure is essentially a centrality measure in the social network literature, which has been used quite frequently to measure status. For instance, Pollock, Lee, Jin and Lashley [61] use a network centrality measure to capture the status of venture capital firms.

The moderating effects of *Expertise* and *Status* are captured by the interaction terms between *CrowdReliance* and the two user attributes, namely, *Expertise* and *Status*. Details of the variables' definition and their descriptive statistics are presented in Table 2 (see Online

[12] Because the lifetime hit rate is calculated on the per-entry basis, it is orthogonal to the decision of whether to participate, our dependent variable of interests.

[13] We also evaluate a random sample of 5% of all comments using Amazon Mechanical Turk and find that only 2.4% of comments are negative, and 88.7% are positive.

[14] The *Status* variable is log-transformed in data analysis as it is highly skewed and has a large value range.

Appendix for a correlation table).

<Insert Table 2 about here>

Model-Free Evidence

Before testing the hypotheses, we first conduct a model-free analysis of the relationship between the reliance on crowd voting and participation. Specifically, we plot prize-normalized participation, defined as the number of entries divided by the prize sum (in thousand dollars), against *CrowdReliance* for each contest. As seen in Figure 3, there appears to be a positive relationship (β = 0.190, p<0.001) between *CrowdReliance* and the prize-normalized participation, which provides model-free support for our hypothesis.

<Insert Figure 3 about here>

Econometric Model for Participation

To understand how the reliance on crowd voting affects participation, we adopt a conditional logit (also called fixed-effects logit) specification. Logistic regression is suitable when the dependent variable is dichotomous, which is the case in our setting. The user fixed-effects specification further controls for time-invariant user heterogeneities. This is important because the probability of participation likely varies greatly across individuals. The fixed-effects specification allows us to control for such variability even if the relevant user characteristic is not observable. Specifically, we model a user i's expected utility from participating in contest j, denoted as U_{ij}, as:

$$U_{ij} = V_{ij} + \varepsilon_{ij} = \beta_1 CrowdReliance_j + \beta_2 Status_{ij} + \beta_3 CrowdReliance_j \cdot Expertise_i$$
$$+ \beta_4 CrowdReliance_j \cdot Status_{ij} + \beta_5 Controls_{ij} + \alpha_i + \varepsilon_{ij} \quad (1)$$

where ε_{ij} is an idiosyncratic error term and V_{ij} is the systematic component of a user's utility that consists of the following parts: the user's status $Status_{ij}$, the interaction terms $CrowdReliance_j * Expertise_i$ and $CrowdReliance_j * Status_{ij}$, a set of controls $Controls_{ij}$, and user fixed-effects α_i.

Controls$_{ij}$ consists of a host of factors that could affect participation decisions. First, since α_i captures the time-invariant user characteristics, we only control for time-varying user characteristics. Zooppa users who have been in the community for a long time might be systematically different from newcomers in participation patterns. Therefore, we control for how long the user has been a member of the Zooppa community (*Tenure*$_{ij}$), and the number of the user's prior video submissions (*VideoExperience*$_{ij}$). Next, because video creation is time-consuming, whether the user participated in the last contest (*lagEnter*$_{ij}$) could also have an impact on current participation. Our measure of status as the number of unique users who have commented on the focal user's video submissions may be confounded by the user's outgoing comments. To remove such a confound, we control the number of unique users on whom the focal user has commented (*logCmtGiven_user*) [17].

Second, we also control for an extensive list of contest attributes, including the total prize sum (*logTotalAwardAmount*), the number of prizes (*NumAwards*), and duration (*ContestDuration*), which are frequently used in prior research [12, 49]. Participation decisions may also be affected by the amount of work required and the clarity of the judging criteria, so we include whether the contest had additional requirements (*AddtlRequirement*), its maximum video file size (*logMaxSize*), and whether the contest specified the evaluation criteria in its brief (*CriteriaSpecified*). Users may relate better with products that they are familiar with and may prefer some industries over others, so we include whether the contest was about consumer goods or services (*ConsumerGood*) and industry dummies (*service*, *manufacture*, or other industries). Finally, to control for seasonality, we include quarter dummies.

Main Findings

We apply the following estimation strategies in data analysis. First, to aid interpretation, we center the moderator variables, *Expertise* and *Status*, so that we can directly interpret the coefficient of *CrowdReliance* as the average effect evaluated at the means of *Expertise* and *Status* [17]. Second, it is possible that a user's participation decisions for different contests are related, that is, users can differ in their tendency to participate in crowdsourcing contests. Therefore, we estimate the model with robust standard errors clustered by users.

Following the recommendation of Aiken, West and Reno [2] for estimating both direct and moderating effects, we conduct our analyses in a hierarchical fashion (See Table 3). We first estimate a model with only the control variables and then add the independent variable of interest, *CrowdReliance*. We further include the direct effect of the moderator variable, *Status* (noting that *Expertise* as a time-invariant variable is absorbed by the fixed effects). Finally, we introduce interaction terms. We find no significant concern of multicollinearity: the mean and largest VIF values for the explanatory variables are 1.98 and 6.10 respectively, which are well below the recommended threshold of 10. Stepwise regressions (Tables 4 and 6) yield similar coefficients and standard errors across models, further suggesting that collinearity may not be a major concern.

The findings are consistent with our expectations. The results reported in Table 3 suggest a positive correlation between reliance on crowd voting and participation: on average, the odds of participation increase by about 0.7% for every percentage point increase in the reliance on crowd voting (odds ratio = 1.007; $p<0.001$). **This supports H1**. Take the "I love Italian Shoes" contest as an example. With $1,720 worth of expert-rated prizes and $2,100 worth of crowd-voted prizes, the contest's crowd-voting reliance was 55 percent. If we reallocate $380 from expert-rated prizes to crowd-voted prizes, the crowd-voting reliance would increase by 10 percentage points and by our estimation, the odds of participation would increase by 7%.

Next, *CrowdReliance* negatively interacts with expertise (odds ratio=0.999, $p<0.001$). This suggests that a one percentage point increase in the lifetime hit rate in expert prizes is associated with a 0.1% decrease in the impact of *CrowdReliance* on the odds of participation. We, therefore, find evidence **in support of H2** that *CrowdReliance* has a greater impact on the participation of users with lower expertise. In our dataset, the average number of video submissions by a typical contestant is about five. That means, by winning one additional expert-rated prize, a typical user can increase her lifetime hit rate by twenty percentage points (one prize out of five submissions). Accordingly, her odds of participation would increase by 2%.

Finally, the interaction between *Status* and *CrowdReliance* is significant and positive (odds ratio=1.002, $p<0.05$), which **supports H3,** that the reliance on crowd voting has a greater impact on the participation of users with high status. Since we calculate *Status* as the log of the number of unique peers who have commented on the focal user's video entries before the current contest, an odds ratio of 1.002 indicates that gaining 10% additional commenters before a contest is associated with a 2% increase in the odds of participation to that contest. Considering a large number of potential commenters, this impact is quite substantial compared to similar effects reported in the literature [21].

We plot the moderation effects of expertise and status in Figure 4. Because conditional logit models does not estimate the value of the fixed effect term, α_i, we cannot calculate the marginal effect on predicted probability. As a workaround, we calculate the marginal effect on log odds (which does not require estimates of α_i) as suggested by the Stata Journal[15] and other researchers[16]. Following [80], we visualize the interactions between crowd reliance and two

[15] See here: https://www.stata.com/manuals13/xtxtlogitpostestimation.pdf, accessed on January 22, 2020.
[16] e.g., see here: https://xiangao.netlify.com/2019/01/25/marginal-effects-in-models-with-fixed-effects/, accessed

moderators (expertise and status) by plugging in selected values of the moderators and then plotting crowd reliance against the predicted log odds. As shown in Figure 4(a), as expertise increases, the marginal effect of crowd reliance on the log odds of contest participation decreases, indicating a negative interaction between crowd reliance and expertise. Similarly, in Figure 4(b), as status increases, the marginal effect of crowd reliance on the log odds of contest participation increases, indicating a positive interaction between crowd reliance and status.

<Insert Figure 4 about here>

Besides our main findings, a few findings on antecedents of participation are also noteworthy. First, contests that explicitly mention the evaluation criteria attract a higher level of participation (odds ratio=1.180, $p<0.001$), suggesting that users avoid ambiguous contests. Second, users who have a higher status within the community are less likely to participate (odds ratio=0.662, $p<0.001$), regardless of winner-selection mechanisms. This result is intriguing because one can also argue that users with high status may find the community more attractive and thus become more engaged. One reason could be that such users, having achieved high status, also have a high opportunity cost (e.g. they are more likely hired as contractors). As a result, such users become less interested in crowdsourcing contests and the associated award money. Finally, as users become more engaged with the community, as indicated by more comments given to others, they are also more likely to participate (odds ratio=1.437, $p<0.001$). To our knowledge, the above findings have not been noted in the crowdsourcing literature.

<Insert Table 3 about here>

Robustness Checks

In this part, we address several potential concerns in the primary analysis. Our findings

---

on January 22, 2020.

reported in this section are largely consistent with the above findings in terms of sign and significance of the coefficients. We start with alternative variable definitions and model specifications, then introduce the pre-estimation matchings.

First, when creating the user-contest panel for data analysis, we only include users with at least one community activity in the past two years before the current contest. To make sure this panel construction is robust, we also construct two other panels with 365 days (one year) and 182 days (6 months) as the cut-off points respectively. Table 4 reports the findings after applying the same analysis to the two panels. The results remain the same as our primary findings.

Second, to accommodate multiple submissions to one contest, we estimate a negative binomial model, using the number of video entries (*numEntries*) submitted to a contest as an alternative dependent variable [4, 33]. As shown in the first column of Table 5, the coefficients remain consistent with our main findings.

Third, we apply alternative measures to the variables of expertise and status. In the main analysis, we measure expertise using the lifetime hit rate in expert-rated prizes. It is possible that winning a prize selected by crowd voting also indicates expertise. Therefore, we also run a model with expertise measured as the lifetime hit rate in any prizes (expert-rated and crowd-voted prizes). Similarly, we introduce an alternative measure of status using the number of unique peers who have voted on the focal user's video entries before the current contest. Like commenting behavior, voting is an important part of repeated social interaction in contests with crowd voting. As seen in column 2 and column 3 of Table 5, the alternative measures do not change our result.

Fourth, one of the control variables, *logCmtGiven_user*, is highly correlated with *Status* (rho=0.71). To rule out the potential collinearity concern, we exclude *logCmtGiven_user* and re-estimate the model. The result, as shown in Table 5, column 4, stays consistent with the main result.

<Insert Table 4, Table 5 about here>

Finally, we reconstruct contest subsets with matching approaches. Unlike experimental studies, the treatment assignment mechanism in observational data is often unknown or ambiguous and is not random. This exposes us to potential biases – e.g. when contests using crowd voting are systematically different from those not using it. Though controlling for confounding co-variates helps alleviate the problem, such biases can be greatly reduced by combining careful parametric models with non-parametric data pre-processing such that the treatment variable is closer to being independent of the confounding covariates. This motivates us to apply various matching methods to our data before estimating our models.

The key goal of matching is to prune observations from the data so that the remaining data have a better balance between the treated and the untreated groups, meaning that the empirical distributions of the confounding covariates in two groups are as similar as possible. Besides reducing biases arising from non-random assignments, an approach based on matching also greatly reduces the impact of model specification errors [34, 41]. Several matching approaches have been suggested by empirical scholars. In this current study, we apply the Propensity Score Matching (PSM) and Coarsened Exact Matching (CEM) for their attractive statistical properties and their popularity among social science researchers [7, 39, 40]. The details about the implementation of the matching approaches can be found in the Online Appendix.

After performing matching, we repeat the main analyses for the participation model. As shown in Table 6, matching results in much smaller sample sizes due to large differences between treatment and control groups. The results based on matching approaches are very consistent with the primary analyses, suggesting that our results are, to a large extent, robust to biases arising from non-random assignment and potential model specification errors.

<Insert Table 6 about here>

DISCUSSION AND CONCLUSION

Crowd voting allows users in a crowdsourcing community to vote for the winners of crowdsourcing contests. Motivated by the adoption of crowd voting at a few crowdsourcing platforms for creative works, we investigate how a contest's reliance on crowd voting – defined as the proportion of prize sum judged by crowd voting – can affect participation. We find that overall participation increases with a contest's reliance on crowd voting. Moreover, crowd voting is less appealing to users with higher expertise (a form of human capital) but more appealing to those with high status in the crowdsourcing community (a form of social capital). Our findings are robust across several alternative specifications, including matching-based estimations that mitigate potential biases caused by the endogenous choice of winner-selection mechanisms.

Contribution to Academic Literature

We firstly contribute to tournament-theory and crowdsourcing-contest literature by establishing a link between a contest's reliance on crowd voting and participation in crowdsourcing contests. Tournament-theory and crowdsourcing-contest literature find that contest design features such as prize structure and reward size affect participation [23, 50, 79], but they have not examined how winner-selection mechanisms, a key contest design element, affects contest participation. We firstly develop a theoretical understanding of this relationship. We argue that when a contest increases its reliance on crowd voting, it adds uncertainties in performance evaluation. Such added uncertainties, according to the tournament theory, disperse winning chances among contestants [13, 44] and lead to greater overall participation. We then confirm our theoretical prediction using our empirical findings based on a unique dataset: every

1% increase in crowd-voting reliance can boost the odds of participation by 0.7%.

Our study also adds to tournament-theory and crowdsourcing-contest literature by showing how participants may *sort into* contests by winner-selection mechanisms they offer – in the sense that there could be a matching between participant types and winner-selection mechanisms. The literature has demonstrated sorting preferences along other dimensions such as project license type and project size [9, 14, 68, 69], but not winner-selection mechanisms. We develop the theoretical reasons for such sorting: As a contest relies more on crowd voting, added uncertainties dilute the value of expertise, thus can better motivate users with lower expertise. In the meantime, users with high social status have advantages in gaining votes from peers, thus are more motivated to participate. Hence, users with low expertise and high status can sort into contests with higher reliance on crowd voting. We also find empirical support for such a sorting prediction. Our finding on the role of contestant social status makes a unique contribution to the tournament-theory literature, which rarely considers a contestant's status. We show that, in a crowd-voted contest, a contestant's social status is a salient factor in determining his/her participation.

Our findings hold important implications for research on comparing different winner-selection mechanisms. There have been a few studies comparing crowd voting and expert rating in terms of the agreement in selection outcomes and quality of selection outcomes [32, 59]. They typically ask "what if the same submissions were judged by a different winner-selection mechanism?" Our findings suggest that, if we switch to a winner-selection mechanism, the number and types (in terms of expertise and status) of participants would not have stayed the same, implying that the quantity and characteristics of submissions may also change.[17]

[17] Our results include direct evidence on the number of submissions being affected by winner-selection mechanisms, but we can only indirectly infer characters (e.g. quality) of submissions may also change based on the

Therefore, future “what-if” comparisons of winner-selection mechanisms should adjust for the effect of endogenous participation.

Though we study crowd voting in a specific crowdsourcing contest platform, we believe that our insights may have broader implications. Crowd voting, in a broad sense, is also used in various other situations such as idea selection in open-innovation communities, the ranking of user-generated content in social media, online reviews, and online forums, sorting of feature suggestions in open-source software, and reality-show contests [45, 60, 74]. Our findings suggest that a crowd-voting-based selection approach may promote contributions to these platforms but may appeal more strongly to contributors whose expertise is not high but status is high. Such issues could be promise directions for future research.

Practical Implications

Our findings provide several practical guidelines on the choice of winner-selection mechanisms for contests in the creative domains. First, we show that contest designers can use crowd voting to attract broader participation in their contests. Specifically, we found that the odds of participation increase by about 0.7% for every 1% increase in the reliance on crowd voting. In other words, if a contest increases its reliance on crowd voting by 30%, the odds of participation will increase by 21%, which is remarkable.

Our finding indicates that crowd voting is more useful for attracting users with lower expertise. Therefore, crowd voting is especially appropriate for encouraging new users and users who have alternative backgrounds to participate. Such users may help sustain the crowdsourcing community and increase the diversity of ideas. In light of this, crowd voting is most useful for ideation and “rugged” problems [72] where wide participation and diversity are most important.

intuition that high- and low-expertise contestants, as well as high and low-status contestants, may submit systematically different solutions. The latter implication is worthy of future investigations.

Contest designers can also use crowd voting to reward high-status users in the crowdsourcing community. We show that high-status users are at high risk for leaving the platform, but crowd voting may help mitigate this trend because high-status users have advantages in getting crowd votes. However, we caution that high status might not translate to high expertise or high performance. Hence, crowd voting may not be suitable for contests where expertise is more important than diversity for sponsors.

We note that crowd voting and expert rating are not mutually exclusive. Having a mix of expert-rated and crowd-voted prizes may allow the sponsor to strike a balance between attracting elite contestants and having broader participation. In fact, they can complement each other in interesting ways. For example, some researchers suggest using crowd voting to narrow the field of candidates before using expert judges to determine the final winners [46].

Our findings on auxiliary variables also hold a number of practical implications. First, regardless of winner-selection mechanisms, contest sponsors should explicitly state the evaluation criteria to attract more participants. Second, interestingly, we found that users who actively leave comments on their peers are also more likely to participate in a contest. This points to the importance of maintaining an active user community on crowdsourcing platforms.

Limitations

This research has a few limitations. First, our findings were based on a single crowdsourcing platform with a specific implementation of expert rating and crowd voting, and thus may not generalize to other settings. It would be interesting to investigate the issue in other crowdsourcing contests and non-contest-based platforms (e.g., Dell’s open innovation platform). Recent research in open innovation suggests that user communities can also effectively perform the task of selecting the best ideas on behalf of the firms [11, 29, 31, 71, 73]. Second, one could also expand the current research into winner-selection mechanism designs that mix crowd voting and expert rating in alternative ways (e.g., using crowd voting as an initial screening) and the impact of winner-selection mechanisms on other outcomes (e.g., community building, learning, and knowledge collaboration) [3, 26, 53, 62]. Third, though we have used matching methods to alleviate the concern of endogenous choice of winner-selection mechanisms, more research (e.g., using experiments) is required to eradicate such a concern. Finally, our theoretical model can be further tested by directly measuring the effect of crowd voting on winning expectancy.

Conclusions

In conclusion, this study provides an understanding of the relationship between crowd voting and participation in crowdsourcing contests. We demonstrate that the number and characteristics of submissions attracted by a contest are a function of its reliance on crowd voting. We hope our findings invite further research on crowd voting in crowdsourcing contests and beyond.

REFERENCES

1. Adamczyk, S., Bullinger, A.C., and Möslein, K.M. Innovation contests: a review, classification and outlook. *Creativity and Innovation Management*, 21, 4 (2012), 335-360.
2. Aiken, L.S., West, S.G., and Reno, R.R. *Multiple regression: testing and interpreting interactions*. Sage, 1991.
3. Alexy, O., Criscuolo, P., and Salter, A. Managing unsolicited ideas for R&D. *California Management Review*, 54, 3 (2012), 116-139.
4. Allison, P.D., and Waterman, R.P. 7. Fixed-effects negative binomial regression models. *Sociological methodology*, 32, 1 (2002), 247-265.
5. Archak, N., and Ghose, A. Learning-By-Doing and Project Choice: A Dynamic Structural Model of Crowdsourcing. *International Conference on Information Systems 2010 Proceedings*, St. Louis, MO, USA, 2010.
6. Banerjee, A.V. A simple model of herd behavior. *The quarterly journal of economics*, 107, 3 (1992), 797-817.
7. Bapna, R., Gupta, A., Ray, G., and Singh, S. IT outsourcing and the impact of advisors on clients and vendors. *Information Systems Research*, 27, 3 (2016), 636-647.
8. Barnett, M.L., Jermier, J.M., and Lafferty, B.A. Corporate reputation: The definitional landscape. *Corporate reputation review*, 9, 1 (2006), 26-38.
9. Belenzon, S., and Schankerman, M. Motivation and sorting of human capital in open innovation. *Strategic Management Journal*, 36, 6 (2015), 795-820.
10. Bikhchandani, S., Hirshleifer, D., and Welch, I. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, 100, 5 (1992), 992-1026.
11. Blohm, I., Riedl, C., Füller, J., and Leimeister, J.M. Rate or trade? identifying winning ideas in open idea sourcing. *Information Systems Research*, 27, 1 (2016), 27-48.

12. Bockstedt, J., Mishra, A., and Druehl, C. Do Participation Strategy and Experience Impact the Likelihood of Winning in Unblind Innovation Contests? *Available at SSRN 1961244* (2011).
13. Boudreau, K.J., Lacetera, N., and Lakhani, K.R. Incentives and problem uncertainty in innovation contests: An empirical analysis. *Management Science*, 57, 5 (2011), 843-863.
14. Boudreau, K.J., and Lakhani, K.R. Field Experimental Evidence on Sorting, Incentives and Creative Worker Performance. *Harvard Business School* (2011), 1-48.
15. Brabham, D.C. Crowdsourcing the public participation process for planning projects. *Planning Theory*, 8, 3 (2009), 242-262.
16. Brabham, D.C. Moving the crowd at Threadless: motivations for participation in a crowdsourcing application. *Information, Communication & Society*, 13, 8 (2010), 1122-1145.
17. Brown, J., Broderick, A.J., and Lee, N. Word of mouth communication within online communities: Conceptualizing the online social network. *Journal of interactive marketing*, 21, 3 (2007), 2-20.
18. Budescu, D.V., and Chen, E. Identifying expertise to extract the wisdom of crowds. *Management Science*, 61, 2 (2014), 267-280.
19. Chen, L., Huang, Z., and Liu, D. Pure and hybrid crowds in crowdfunding markets. *Financial Innovation*, 2, 1 (2016), 19.
20. Chen, L., Marsden, J.R., and Zhang, Z. Theory and analysis of company-sponsored value co-creation. *Journal of Management Information Systems*, 29, 2 (2012), 141-172.
21. Chen, P.-Y., Pavlou, P.A., and Yang, Y. Determinants of Open Contest Participation in Online Labor Markets. *Fox School of Business Research Paper No. 15-074.* (2014).
22. Connelly, B.L., Tihanyi, L., Crook, T.R., and Gangloff, K.A. Tournament theory: thirty years of contests and competitions. *Journal of Management*, 40, 1 (2014), 16-47.
23. DiPalantino, D., and Vojnovic, M. Crowdsourcing and all-pay auctions. *Proceedings of the 10th ACM conference on Electronic Commerce*, Stanford, CA, USA: ACM, 2009, pp. 119-128.
24. Ericsson, K.A. Expertise in interpreting: an expert-performance perspective. *Interpreting*, 5, 2 (2000), 187-220.
25. Ericsson, K.A., and Smith, J. *Toward a general theory of expertise: prospects and limits*. Cambridge University Press, 1991.
26. Faraj, S., Jarvenpaa, S.L., and Majchrzak, A. Knowledge collaboration in online communities. *Organization Science*, 22, 5 (2011), 1224-1239.
27. Fershtman, C., Murphy, K.M., and Weiss, Y. Social status, education, and growth. *Journal of political Economy*, 104, 1 (1996), 108-132.
28. Franke, N., Lettl, C., Roiser, S., and Tuertscher, P. "Does God Play Dice?"-Randomness vs. Deterministic Explanations of Crowdsourcing Success. *Academy of Management Proceedings*, 2014.
29. Fuchs, C., Prandelli, E., and Schreier, M. The psychological effects of empowerment strategies on consumers' product demand. *Journal of Marketing*, 74, 1 (2010), 65-79.
30. Füller, J., Matzler, K., and Hoppe, M. Brand community members as a source of innovation. *Journal of Product Innovation Management*, 25, 6 (2008), 608-619.
31. Girotra, K., Terwiesch, C., and Ulrich, K.T. Idea generation and the quality of the best idea. *Management Science*, 56, 4 (2010), 591-605.
32. Haan, M.A., Dijkstra, S.G., and Dijkstra, P.T. Expert judgment versus public opinion–evidence from the Eurovision Song Contest. *Journal of Cultural Economics*, 29, 1 (2005), 59-78.
33. Hausman, J.A., Hall, B.H., and Griliches, Z. Econometric models for count data with an application to the patents-R&D relationship. national bureau of economic research Cambridge, Mass., USA, 1984.
34. Ho, D.E., Imai, K., King, G., and Stuart, E.A. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political analysis*, 15, 3 (2007), 199-236.
35. Holbrook, M.B. Popular appeal versus expert judgments of motion pictures. *Journal of Consumer Research*, 26, 2 (1999), 144-155.

36. Holbrook, M.B., and Addis, M. Taste versus the market: an extension of research on the consumption of popular culture. *Journal of Consumer Research*, 34, 3 (2007), 415-424.
37. Howe, J. *Crowdsourcing: how the power of the crowd is driving the future of business*. Random House, 2008.
38. Hutter, K., Füller, J., Hautz, J., Bilgram, V., and Matzler, K. Machiavellianism or morality: which behavior pays off in online innovation contests? *Journal of Management Information Systems*, 32, 3 (2015), 197-228.
39. Iacus, S.M., King, G., and Porro, G. Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association*, 106, 493 (2011), 345-361.
40. Iacus, S.M., King, G., and Porro, G. Causal inference without balance checking: Coarsened exact matching. *Political Analysis*, 20, 1 (2012), 1-24.
41. Iacus, S.M., and Porro, G. Invariant and metric free proximities for data matching: An R package. *Journal of Statistical Software*, 25, 11 (2008), 1-22.
42. Johnson, P. Whence democracy? a review and critique of the conceptual dimensions and implications of the business case for organizational democracy. *Organization*, 13, 2 (2006), 245-274.
43. Kalra, A., and Shi, M. Designing optimal sales contests: a theoretical perspective. *Marketing Science*, 20, 2 (2001), 170-193.
44. Karelaia, N., and Hogarth, R.M. The attraction of uncertainty: Interactions between skill and levels of uncertainty in market-entry games. *Journal of Risk and Uncertainty*, 41, 2 (2010), 141-166.
45. Kim, A., Moravec, P.L., and Dennis, A.R. Combating fake news on social media with source ratings: the effects of user and expert reputation ratings. *Journal of Management Information Systems*, 36, 3 (2019), 931-968.
46. King, A., and Lakhani, K.R. Using open innovation to identify the best ideas. *MIT Sloan management review*, 55, 1 (2013), 41-48.
47. Lakhani, K.R., and Panetta, J.A. The principles of distributed innovation. *Innovations: technology, governance, globalization*, 2, 3 (2007), 97-112.
48. Lawler, E.E., and Porter, L.W. The effect of performance on job satisfaction. *Industrial relations: A journal of Economy and Society*, 7, 1 (1967), 20-28.
49. Li, D., and Hu, L. Exploring the effects of reward and competition intensity on participation in crowdsourcing contests. *Electronic Markets*, 29, 3 (2017), 199-210.
50. Liu, D., Geng, X., and Whinston, A.B. Optimal design of consumer contests. *Journal of Marketing*, 71, 4 (2007), 140-155.
51. Liu, T.X., Yang, J., Adamic, L.A., and Chen, Y. Crowdsourcing with all-pay auctions: a field experiment on taskcn. *Management Science*, 60, 8 (2014), 2020–2037.
52. Magallanes, Y., Sánchez, J.A., Cervantes, O., and Wan, W. A pattern-based approach for developing creativity applications supported by surface computing. *International Journal of Human-Computer Studies*, 110 (2018), 134-154.
53. Majchrzak, A., and Malhotra, A. Towards an information systems perspective and research agenda on crowdsourcing for innovation. *Journal of Strategic Information Systems*, 22, 4 (2013), 257-268.
54. Malaga, R.A. Web-based reputation management systems: problems and suggested solutions. *Electronic Commerce Research*, 1, 4 (2001), 403-417.
55. Malone, T.W., Laubacher, R., and Dellarocas, C. Harnessing crowds: mapping the genome of collective intelligence. *MIT Sloan Research Paper. Available at SSRN: https://ssrn.com/abstract=1381502* (2009).
56. Martínez-Torres, M.R. Analysis of open innovation communities from the perspective of social network analysis. *Technology Analysis & Strategic Management*, 26, 4 (2014), 435-451.
57. Matyas, L., and Sevestre, P. The econometrics of panel data. *A Handbook of the Theory with* (1996).
58. McFadden, D. Conditional logit analysis of qualitative choice behavior. In, Zarembka, P., (ed.), *Frontiers in Econometrics*, New York: Academic Press 1974, pp. 105-142.

59. Mollick, E., and Nanda, R. Wisdom or madness? comparing crowds with expert evaluation in funding the arts. *Management Science*, 62, 6 (2015), 1533-1553.
60. Piezunka, H., and Dahlander, L. Distant search, narrow attention: How crowding alters organizations' filtering of suggestions in crowdsourcing. *Academy of Management Journal*, 58, 3 (2015), 856-880.
61. Pollock, T.G., Lee, P.M., Jin, K., and Lashley, K. (Un) tangled: exploring the asymmetric coevolution of new venture capital firms' reputation and status. *Administrative Science Quarterly*, 60, 3 (2015), 482-517.
62. Riedl, C., and Seidel, V.P. Learning from mixed signals in online innovation communities. *Organization Science*, 29, 6 (2018), 1010-1032.
63. Salganik, M.J., Dodds, P.S., and Watts, D.J. Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311, 5762 (2006), 854-856.
64. Sauder, M., Lynn, F., and Podolny, J.M. Status: Insights from organizational sociology. *Annual Review of Sociology*, 38 (2012), 267-283.
65. Shanteau, J., Weiss, D.J., Thomas, R.P., and Pounds, J.C. Performance-based assessment of expertise: How to decide if someone is an expert or not. *European Journal of Operational Research*, 136, 2 (2002), 253-263.
66. Spence, M.T., and Brucks, M. The moderating effects of problem characteristics on experts' and novices' judgments. *Journal of marketing Research*, 34, 2 (1997), 233-247.
67. Standifird, S.S. Reputation and e-commerce: eBay auctions and the asymmetrical impact of positive and negative ratings. *Journal of Management*, 27, 3 (2001), 279-295.
68. Stewart, K.J., Ammeter, A.P., and Maruping, L.M. Impacts of license choice and organizational sponsorship on user interest and development activity in open source software projects. *Information Systems Research*, 17, 2 (2006), 126-144.
69. Stewart, K.J., and Gosain, S. The impact of ideology on effectiveness in open source software development teams. *Mis Quarterly* (2006), 291-314.
70. Surowiecki, J. *The wisdom of crowds: why the many are smarter than the few and how collective wisdom shapes business, economies, societies and nations*. New York, NY: Little Brown Press, 2004.
71. Terwiesch, C., and Ulrich, K.T. *Innovation tournaments: creating and selecting exceptional opportunities*. Harvard Business Press, 2009.
72. Terwiesch, C., and Yi, X. Innovation Contests, Open Innovation, and Multiagent Problem Solving. *Management Science*, 54, 9 (2008), 1529-1543.
73. Toubia, O., and Florès, L. Adaptive idea screening using consumers. *Marketing Science*, 26, 3 (2007), 342-360.
74. Velichety, S., Ram, S., and Bockstedt, J. Quality assessment of peer-produced content in knowledge repositories using development and coordination activities. *Journal of Management Information Systems*, 36, 2 (2019), 478-512.
75. Vroom, V.H. *Work and motivation*. New York: John Wiley & Sons, 1964.
76. Wang, J., and Cole, C.A. The effects of age and expertise on product evaluations: does the type of information matter? *Management Science*, 62, 7 (2015), 2039-2053.
77. Washington, M., and Zajac, E.J. Status evolution and competition: theory and evidence. *Academy of Management Journal*, 48, 2 (2005), 282-296.
78. Weber, W.G., Unterrainer, C., and Schmid, B.E. The influence of organizational democracy on employees' socio-moral climate and prosocial behavioral orientations. *Journal of organizational behavior*, 30, 8 (2009), 1127-1149.
79. Wen, Z., and Lin, L. Optimal fee structures of crowdsourcing platforms. *Decision Sciences*, 47, 5 (2016), 820-850.
80. Williams, R. Interaction effects between continuous variables. *University of Notre Dame* (2015).
81. Wolfers, J., and Zitzewitz, E. Prediction markets. *Journal of economic perspectives*, 18, 2 (2004), 107-126.

82. Yang, Y., Chen, P.-y., and Banker, R. Winner Determination of Open Innovation Contests in Online Markets. *International Conference on Information Systems 2011 Proceedings (ICIS2011)*, Shanghai, China, 2011.
83. Ye, H.J., and Kankanhalli, A. Solvers' participation in crowdsourcing platforms: Examining the impacts of trust, and benefit and cost factors. *Journal of Strategic Information Systems*, 26, 2 (2017), 101-117.
84. Zheng, H., Li, D., and Hou, W. Task design, motivation, and participation in crowdsourcing contests. *International Journal of Electronic Commerce*, 15, 4 (2011), 57-88.

Figures and Tables

Figure 1 Model of Crowd Voting on Participation

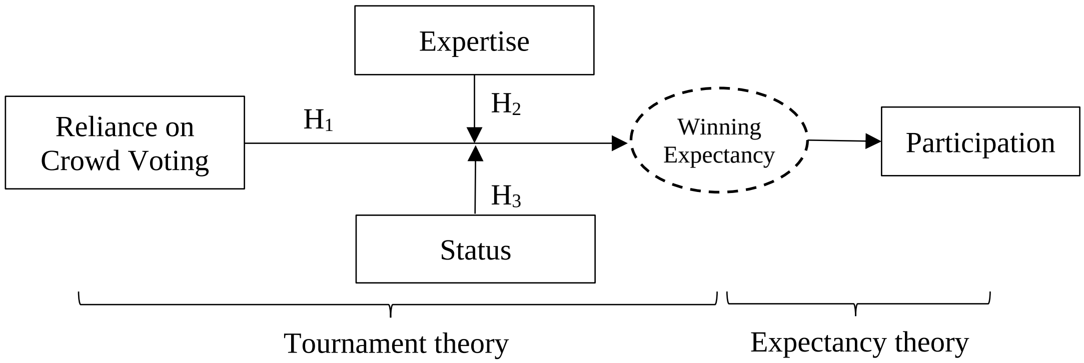
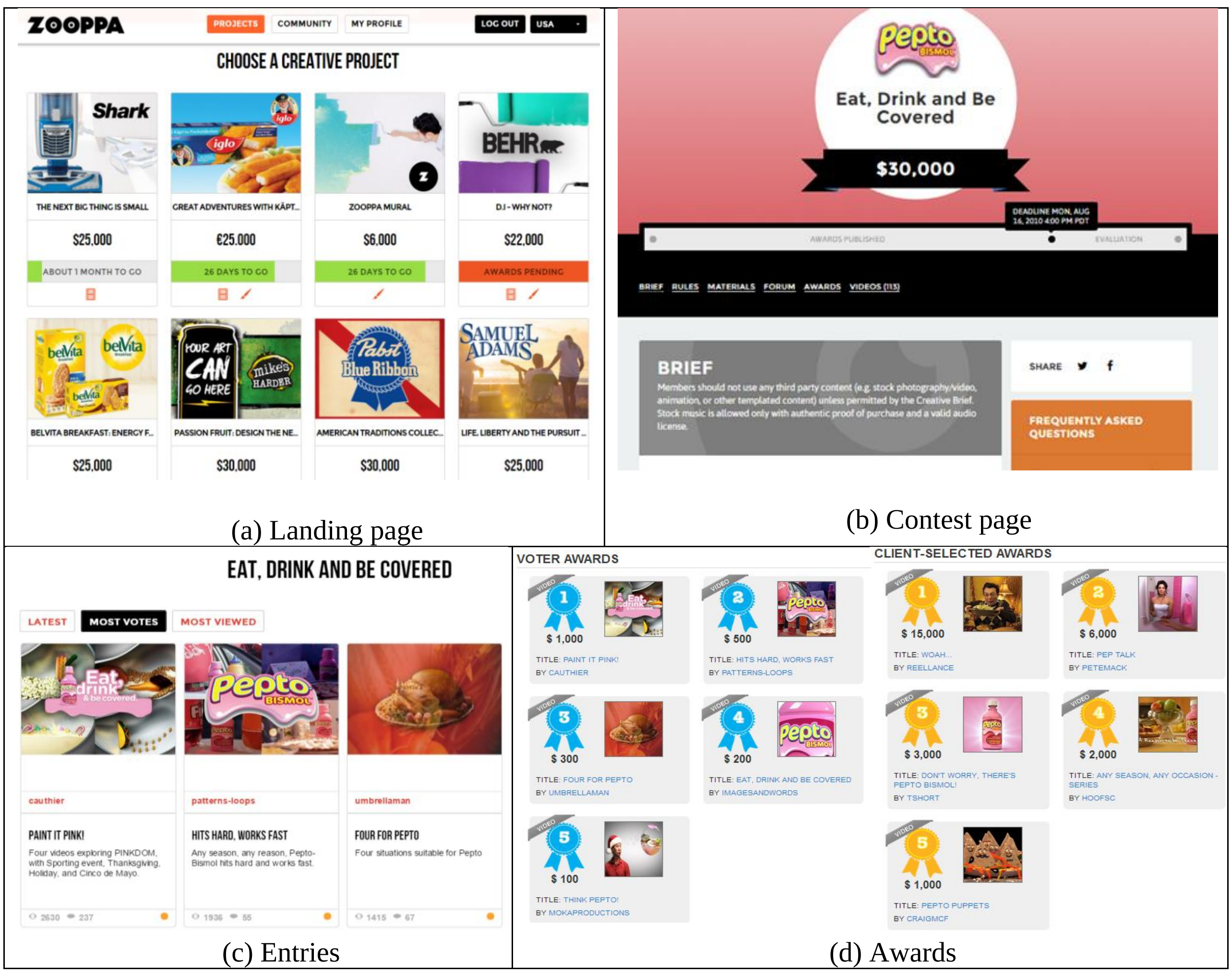


Figure 2 Zooppa Website Screenshots



(a) Landing page (b) Contest page (c) Entries (d) Awards

Note. (a) a landing page lists current and past contests, with amount of prizes and current status; (b) a contest page, including tabs for brief, rules, materials, forum, award, and videos; (c) a contest's entries, listing all creatives with information cards that include user name, title, description, and number of views and comments; (d) an awards page, with winners for two types of prizes.

Figure 3 Scatter Plot of EntryPer1000USD by Crowd Ratio for 102 Contests

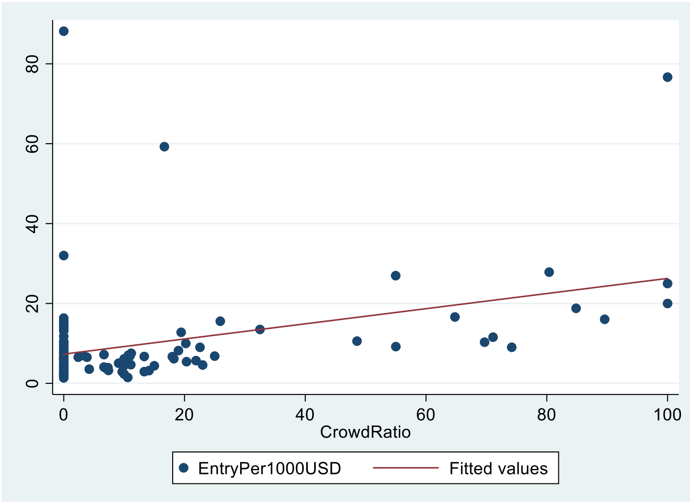
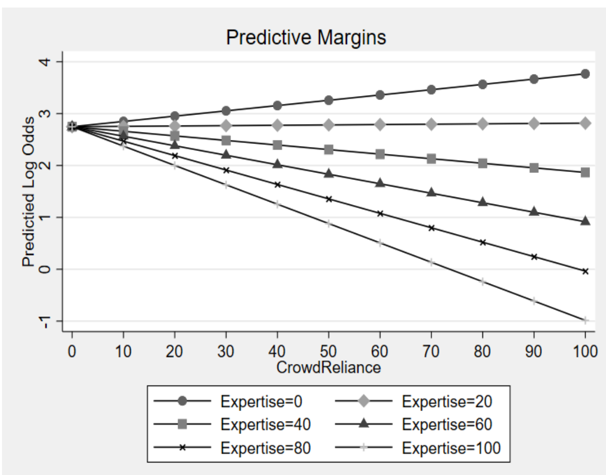
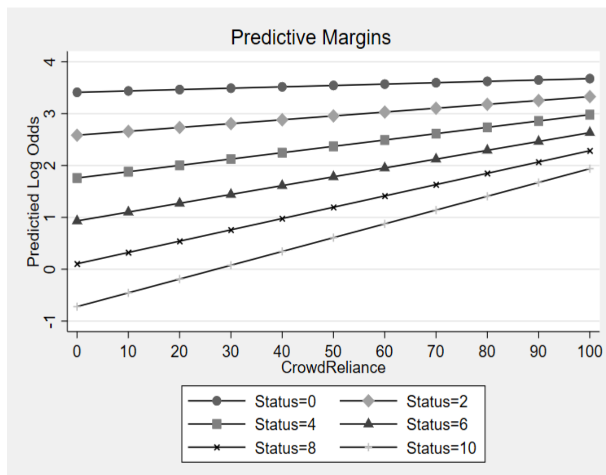


Figure 4 Interaction Plots for the Moderating Effects of Expertise and Status



(a)



(b)

Table 1 List of Expert Judges for a Zooppa Contest

Judges	Description
Rob M.	The founder and COO/Partner of the Wonderland Creative Group. An award-winning Producer/Director, Rob has created memorable commercials for industry leaders.
Jim R.	Served as President of New Line Television and New Line Media. Jim managed all aspects of the company's television production as well as the distribution of New Line feature films and television series.
Alan D.	President of National Lampoon and movie producer; he has produced many movies.
Tom C.	The Senior Vice President of Music at 20th Century Fox. Tom is also the Senior Vice President of Business & Legal at Fox Music, Inc. Tom works on all Fox mega-hits like Glee, American Idol, and Avatar.
Daniel E.	An Executive Producer, Director & Founding Partner Hybrid Films
Brad H.	An award-winning movie producer and vice president of Cross Cut Productions Inc. He is currently in development for three feature films and an original movie.
Brad F.	An award-winning lead technical director and visual artist for some of the most famous animated films for Disney/Pixar, Lucas Film and other film studios.

Table 2 Variables for Participation Models

Variable	Description	mean	sd	min	max
Dependent variable					
EnterContest	The user submitted one or more entries to the contest (Yes=1; No=0)	0.028	0.165	0	1
Main explanatory variables					
CrowdReliance	The percentage of crowd prizes over total prizes, in terms of dollar amounts	6.488	15.317	0	100
Expertise	Total expert prizes the user has won between Dec 2007 and Aug 2013 divided by the number of entries submitted (in percentage)	7.595	22.730	0	100
Status	The natural logarithm of the number of unique peers who have commented on the focal user's video entries before the current contest	1.539	1.543	0	6.098
Contest-related control variables					
AwardAmount	Total dollar value of the prizes offered by the contest	16,708.461	14,366.172	270	100,000
NumAwards	The number of crowd and expert prizes offered by the contest	9.261	6.292	1	43
ContestDuration	The duration of the contest in days	62.290	25.710	7	160
AddtlRequirement	Whether this contest has additional requirement besides standard ones (e.g., format, length, size) (Yes=1; No=0)	0.876	0.330	0	1
MaxSize	Recommended file size (in Megabytes)	65.746	40.585	0	300
CriteriaSpecified	Whether the contest specifies the evaluation criteria in its brief (Yes=1; No=0)	0.344	0.475	0	1
ConsumerGood	Whether the contest involves goods or service intended for consumers, rather than for manufacturers (Yes=1; No=0)	0.314	0.464	0	1
User-related control variables					
Tenure	Tenure of the user on the platform in years by the contest start date	1.258	0.931	0	6.425
VideoExperience	Count of video entries the user has submitted in the past	1.813	2.888	0	46
lagEnter	Whether the user participated in the last contest (Yes=1; No=0)	0.036	0.186	0	1
CmtGiven_user	The number of unique peers whose entries have been commented by the focal user before the current contest	8.547	35.522	0	785

Note: Log-transformation on some variables is applied during the estimation process wherever is deemed necessary.

Table 3 User-fixed effect Logit Regression on Participation

	DV = *EnterContest* Odds ratio (se) (1)	(2)	(3)	(4)
logAwardAmount	1.845***	1.923***	1.920***	1.940***
	(0.057)	(0.062)	(0.062)	(0.063)
NumAwards	0.981***	0.977***	0.979***	0.980***
	(0.004)	(0.004)	(0.004)	(0.004)
ContestDuration	0.998*	0.998+	0.999	0.999
	(0.001)	(0.001)	(0.001)	(0.001)
AddtlRequirement	0.603***	0.617***	0.635***	0.646***
	(0.036)	(0.038)	(0.040)	(0.041)
logMaxSize	0.909***	0.904***	0.899***	0.901***
	(0.019)	(0.019)	(0.018)	(0.018)
CriteriaSpecified	1.172***	1.173***	1.182***	1.180***
	(0.053)	(0.053)	(0.053)	(0.054)
ConsumerGood	1.416***	1.390***	1.419***	1.439***
	(0.060)	(0.060)	(0.061)	(0.062)
Tenure	0.542***	0.575***	0.580***	0.600***
	(0.034)	(0.036)	(0.036)	(0.037)
logVideoExperience	0.121***	0.119***	0.174***	0.177***
	(0.014)	(0.013)	(0.022)	(0.022)
lagEnter	0.648***	0.648***	0.658***	0.654***
	(0.059)	(0.059)	(0.060)	(0.059)
logCmtGiven_user	1.182**	1.224***	1.450***	1.437***
	(0.066)	(0.068)	(0.090)	(0.089)
CrowdReliance		**1.008*** **	**1.006*** **	**1.007*** **
		(0.001)	(0.001)	(0.001)
Status			**0.706*** **	**0.662*** **
			(0.038)	(0.037)
*Expertise * CrowdReliance*				**0.999*** **
				(0.000)
*Status * CrowdReliance*				**1.002***
				(0.001)
Log-likelihood	-10,736.16	-10,711.55	-10,659.77	-10,625.03
Pseudo R-squared	0.243	0.245	0.248	0.251
N	113,955	113,955	113,955	113,955

* p<0.05, ** p<0.01, *** p<0.001. Cluster-robust standard errors are in parentheses.

Notes. 1. Quarter and industry dummies are included as controls. Odds ratios are reported instead of the raw coefficients. When the odds ratio is greater than 1, the variable has a positive impact on the dependent variable, vice versa.

2. During lag variable generation, 3,833 missing values are created and the corresponding rows are dropped from estimation.

3. In the fixed-effect models, 1,198 user groups (40,414 observations) dropped because of all positive or all negative outcomes.

4. There are various Pseudo R-squared measures for binary response models. We report the measure defined by McFadden [58], which has been chosen by Stata as the official Pseudo R-squared measure. It reports the log-likelihood improvement of the full model over the intercept only model.

Table 4 Results of Robustness Check (1)

	DV = *EnterContest* Odds ratio (se)		
	(1) Drop inactive > 365 days	(2) Drop inactive > 182 days	(3) Linear probability model with 2-way clustering
logAwardAmount	1.906***	1.872***	1.016***
	(0.063)	(0.061)	(0.002)
NumAwards	0.980***	0.979***	0.999**
	(0.004)	(0.004)	(0.000)
ContestDuration	0.998*	0.997**	1.000
	(0.001)	(0.001)	(0.000)
AddtlRequirement	0.662***	0.657***	0.983***
	(0.043)	(0.043)	(0.003)
logMaxSize	0.904***	0.905***	0.997*
	(0.019)	(0.020)	(0.001)
CriteriaSpecified	1.173**	1.139*	1.005*
	(0.057)	(0.060)	(0.002)
ConsumerGood	1.474***	1.470***	1.009*
	(0.065)	(0.068)	(0.004)
Tenure	0.770**	1.036	1.007***
	(0.068)	(0.100)	(0.002)
logVideoExperience	0.159***	0.161***	0.888***
	(0.020)	(0.019)	(0.010)
lagEnter	0.602***	0.539***	0.972**
	(0.055)	(0.047)	(0.009)
logCmtGiven_user	1.379***	1.274***	0.996
	(0.084)	(0.076)	(0.004)
CrowdReliance	**1.007*** **	**1.007*** **	**1.001+**
	(0.001)	(0.002)	(0.000)
Status	**0.695*** **	**0.740*** **	**0.979*** **
	(0.039)	(0.040)	(0.006)
*Expertise * CrowdReliance*	**0.999*** **	**0.999*** **	**0.999*** **
	(0.000)	(0.000)	(0.000)
*Status * CrowdReliance*	**1.002*** **	**1.003*** **	**1.001***
	(0.001)	(0.001)	(0.000)
Log-likelihood	-9,352.03	-7,829.48	-
(Pseudo) R-squared	0.237	0.217	0.114
N	76,792	48,786	154,336

+ p<0.1, * p<0.05, ** p<0.01, *** p<0.001. Cluster-robust standard errors are in parentheses. Quarter and industry dummies are included as controls.

Table 5 Results of Robustness Check (2)

	DV = *NumEntries* incidence-rate ratio (se)	DV = *EnterContest* Odds ratio (se)		
	(1) Multiple submissions	(2) Alternative Expertise measure	(3) Alternative Status measure	(4) Drop Comments given
logAwardAmount	1.693***	1.935***	1.944***	1.945***
	(0.042)	(0.063)	(0.064)	(0.063)
NumAwards	0.982***	0.979***	0.978***	0.981***
	(0.003)	(0.004)	(0.004)	(0.004)
ContestDuration	0.999	0.999	0.998*	0.999
	(0.001)	(0.001)	(0.001)	(0.001)
AddtlRequirement	0.576***	0.642***	0.637***	0.657***
	(0.029)	(0.040)	(0.040)	(0.042)
logMaxSize	0.894***	0.899***	0.899***	0.899***
	(0.016)	(0.018)	(0.019)	(0.018)
CriteriaSpecified	1.170***	1.185***	1.159**	1.179***
	(0.046)	(0.054)	(0.053)	(0.053)
ConsumerGood	1.377***	1.433***	1.410***	1.444***
	(0.053)	(0.061)	(0.060)	(0.062)
Tenure	0.540***	0.598***	0.599***	0.611***
	(0.022)	(0.037)	(0.037)	(0.038)
logVideoExperience	0.255***	0.180***	0.170***	0.186***
	(0.014)	(0.023)	(0.021)	(0.024)
lagEnter	0.858**	0.652***	0.655***	0.651***
	(0.043)	(0.059)	(0.059)	(0.060)
logCmtGiven_user	1.672***	1.434***	1.428***	
	(0.051)	(0.090)	(0.086)	
CrowdReliance	**1.009*** **	**1.008*** **	**1.007*** **	**1.006*** **
	(0.001)	(0.001)	(0.001)	(0.001)
Status	**0.775*** **	**0.660*** **	**0.717*** **	**0.764*** **
	(0.025)	(0.037)	(0.027)	(0.039)
*Expertise * CrowdReliance*	**0.999*** **	**0.999*** **	**0.999*** **	**0.999*** **
	(0.000)	(0.000)	(0.000)	(0.000)
*Status * CrowdReliance*	**1.002*** **	**1.002*** **	**1.002** **	**1.002*** **
	(0.000)	(0.001)	(0.001)	(0.001)
Log-likelihood	-13,167.19	-10,647.70	-10,605.28	-10,671.54
Pseudo R-squared	-	0.249	0.252	0.247
N	113,955	113,955	113,955	113,955

* p<0.05, ** p<0.01, *** p<0.001. Cluster-robust standard errors are in parentheses.
Quarter and industry dummies are included as controls.

Table 6 Results of Robustness Check (3)

	DV = *EnterContest* Odds ratio (se)	
	(1) Propensity Score Matching	(2) Coarsen Exact Matching
logAwardAmount	1.985***	1.890***
	(0.084)	(0.146)
NumAwards	0.979***	0.978*
	(0.004)	(0.010)
ContestDuration	0.999	0.991**
	(0.001)	(0.003)
AddtlRequirement	0.754***	0.411***
	(0.055)	(0.074)
logMaxSize	0.929**	0.691***
	(0.024)	(0.047)
CriteriaSpecified	0.957	1.401***
	(0.054)	(0.126)
ConsumerGood	1.373***	1.047
	(0.069)	(0.081)
Tenure	0.572***	0.561***
	(0.034)	(0.043)
logVideoExperience	0.172***	0.145***
	(0.026)	(0.028)
lagEnter	0.663***	0.662**
	(0.067)	(0.090)
logCmtGiven_user	1.464***	1.756***
	(0.102)	(0.157)
CrowdReliance	**1.007*** **	**1.004***
	(0.002)	(0.002)
Status	**0.655*** **	**0.615*** **
	(0.042)	(0.058)
*Expertise * CrowdReliance*	**0.999*** **	**0.999***
	(0.000)	(0.000)
*Status * CrowdReliance*	**1.004*** **	**1.002***
	(0.001)	(0.001)
Log-likelihood	-7,040.49	-3,274.46
Pseudo R-squared	0.281	0.314
N	74,026	30,523

* p<0.05, ** p<0.01, *** p<0.001. Cluster-robust standard errors are in parentheses. Quarter and industry dummies are included as controls.

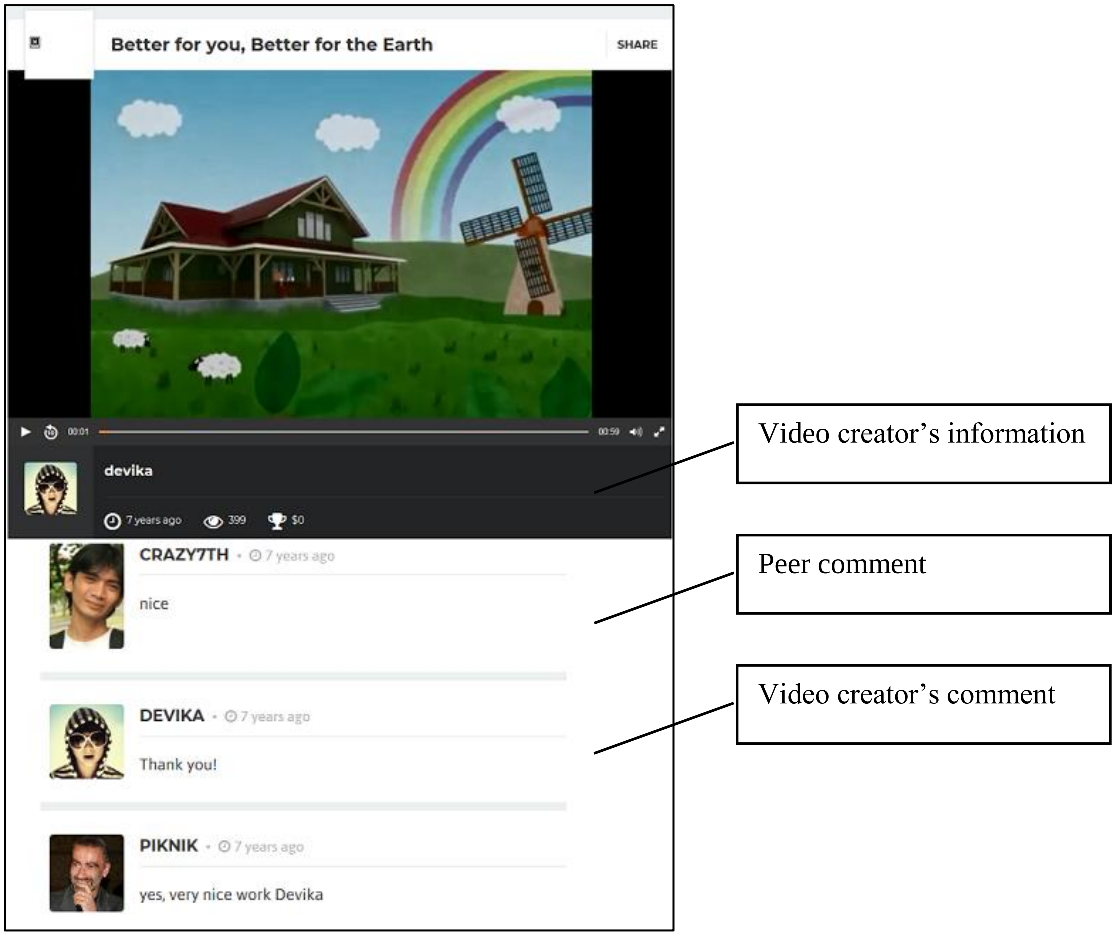
Online Appendix

Table A1. Correlations between Explanatory Variables

	Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	*logAwardAmount*	1.00														
2	*NumAwards*	0.31***	1.00													
3	*ContestDuration*	-0.20***	-0.06***	1.00												
4	*AddtlRequirement*	0.30***	-0.01***	-0.18***	1.00											
5	*logMaxSize*	0.02***	-0.03***	-0.12***	0.03***	1.00										
6	*CriteriaSpecified*	0.17***	0.08***	0.03***	0.23***	-0.05***	1.00									
7	*ConsumerGood*	-0.07***	0.05***	-0.08***	-0.07***	-0.06***	-0.22***	1.00								
8	*Tenure*	0.09***	-0.02***	-0.00	0.11***	0.03***	0.03***	-0.06***	1.00							
9	*logVideoExperience*	0.05***	-0.00	-0.06***	0.06***	0.03***	0.01*	-0.02***	0.47***	1.00						
10	*lagEnter*	-0.07***	-0.01***	0.01***	-0.04***	-0.02***	-0.02***	0.03***	-0.15***	-0.11***	1.00					
11	*logCmtGiven_user*	-0.11***	0.01***	-0.01***	-0.11***	-0.01***	-0.07***	0.05***	0.44***	0.50***	-0.02***	1.00				
12	*CrowdReliance*	-0.37***	0.07***	0.01**	-0.35***	-0.01***	-0.17***	0.10***	-0.13***	-0.07***	0.06***	0.12***	1.00			
13	*Expertise*	0.02***	-0.00	0.00	0.02***	0.00	0.01***	-0.01***	0.00	0.06***	0.02***	0.03***	-0.02***	1.00		
14	*Status*	-0.14***	0.02***	-0.02***	-0.13***	-0.02***	-0.09***	0.07***	0.48***	0.64***	-0.07***	0.71***	0.13***	0.06***	1.00	
15	*Expertise * CrowdReliance*	0.03***	0.01*	-0.00	0.03***	0.00	0.01***	-0.00	0.01***	0.03***	0.01*	0.04***	-0.07***	0.38***	0.03***	1.00
16	*Status * CrowdReliance*	-0.14***	-0.02***	-0.00	-0.11***	-0.00	-0.07***	-0.00	0.11***	0.26***	-0.01***	0.32***	0.24***	0.02***	0.44***	0.08***

* p<0.05, ** p<0.01, *** p<0.001

Figure A1. A sample video webpage with user comments



Source: https://community.zooppa.com/en-us/contests/better-for-you-better-for-the-earth/submissions/better-for-you-better-for-the-earth

Implementation of the Propensity Score Matching and the Coarsened Exact Matching

In our context, the treatment is the adoption of prizes selected by crowd voting in a contest.[1] We followed the standard procedure introduced by Iacus et al. [1, 2] in conducting the CEM process. The first step of matching is the selection of pre-treatment covariates. We selected a list of covariates that may impact the choice of selection mechanisms, including *ContestDuration*, *logTotalAwardAmount*, *logMaxSize*,

[1] The treatment in our primary model specification is a continuous variable (i.e., the crowd reliance). Though matching methods for continuous and categorical treatments have been developed recently, they typically require many data points for each strata to achieve good matching outcomes. Unfortunately, there are only 102 contests in our dataset, and it is unlikely for us to get satisfactory matching outcomes with continuous matching. We follow some of the prior research to perform a traditional matching based on a dichotomized treatment variable.

AddtlRequirement, ConsumerGood, IndustryType and *CriteriaSpecified*. For each covariate, we calculated a univariate imbalance score, which is measured by the *L1* statistic [1, 2]. Among the seven covariates, we selected five covariates with univariate imbalance scores higher than 0.15, and carried out the matching process. Using the CEM procedure, we obtained a matched sample of 43 contests from the original dataset, with 22 using crowd voting and the other 21 not using. The CEM procedure greatly reduced the global multivariate imbalance score, the L_1 statistic between the treatment and control groups, from 0.884 to 0.311, indicating a successful matching [3]. Table A1 presents evidence that the CEM procedure has resulted in a greater balance between treatment and control groups in terms of means and standard deviations.

Table A2 Mean and Std. Dev. comparison before and after CEM

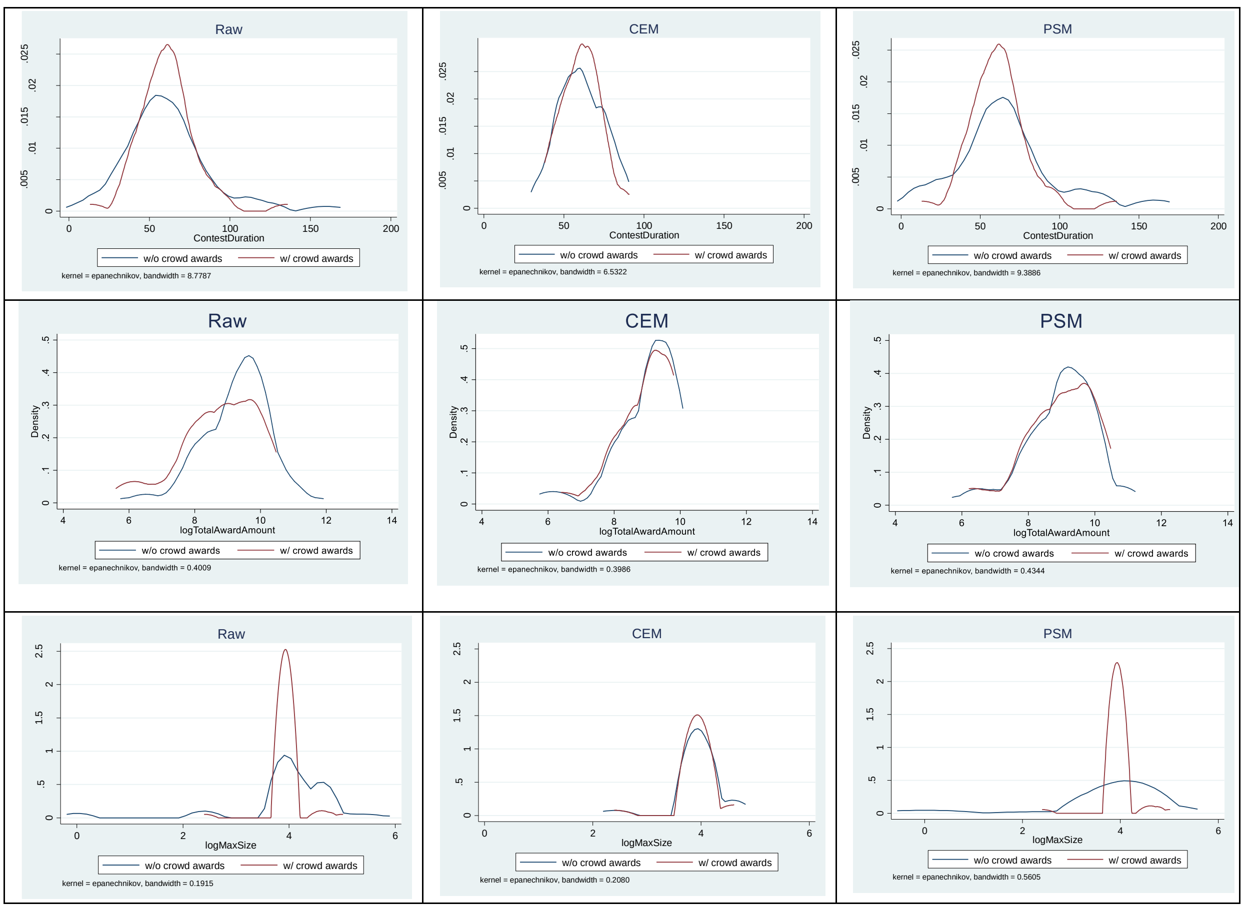
	Mean				Standard Deviation			
	Before		After		Before		After	
Variables	Control	Treated	Control	Treated	Control	Treated	Control	Treated
ContestDuration	60.92	61.08	60.81	60.73	26.64	18.59	13.48	12.35
logTotalAwardAmount	9.26	8.76	8.98	8.94	0.99	1.19	0.88	0.86
logMaxSize	3.98	3.95	3.95	3.92	1.00	0.30	0.43	0.40
AddtlRequirement	0.85	0.59	0.81	0.82	0.36	0.50	0.40	0.40
ConsumerGood	0.29	0.46	0.43	0.55	0.46	0.50	0.51	0.51

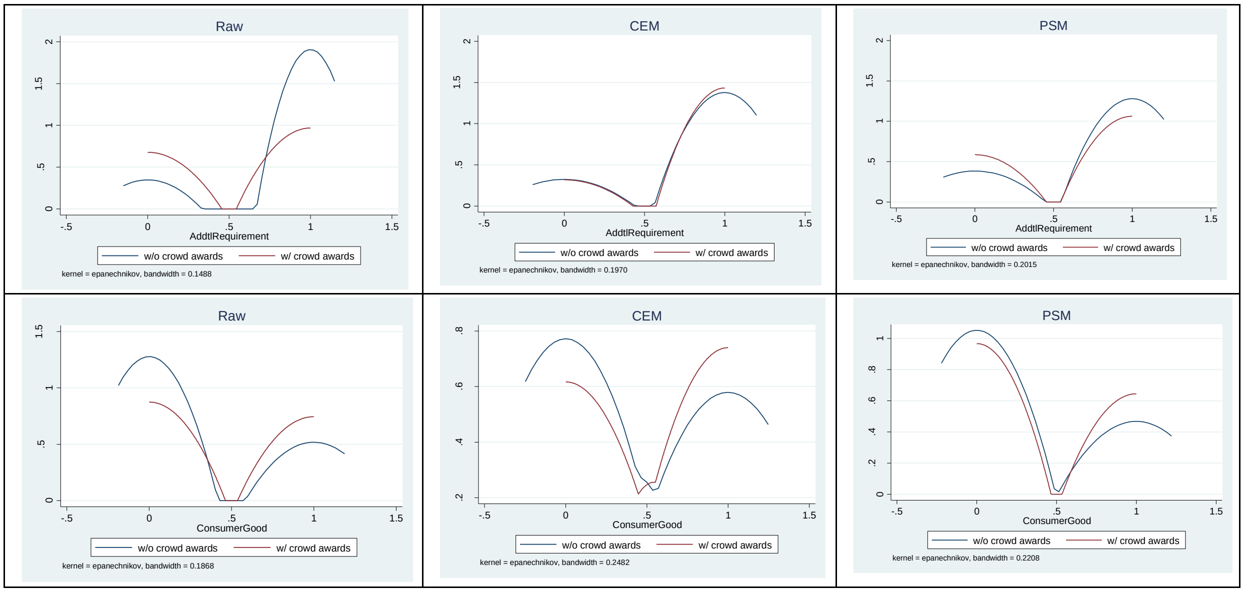
King et al. [4] suggested the use of both CEM and PSM to achieve more robust results after matching. Therefore, in addition to CEM, we conducted a propensity score matching to balance the treatment and control groups. Because our contest dataset is small, we used a single nearest neighbor with maximum distance = 0.05 in PSM. Also, if there were more than one observations in the treatment group can be matched with the control group, we kept all of them. PSM selected 71 contests, in which 26 were in the control group and 45 were in the treatment group. In Table A2, we present the summary statistic for the matched dataset compare to the original full dataset, which suggests after a better balance after matching. In Figure A2, the density plots for the matched samples via CEM and PSM are nearly indistinguishable, implying that both matchings balanced the covariates.

Table A3. Mean and Std. Dev. Comparisons before and after PSM

	Mean				Standard Deviation			
	Before		After		Before		After	
Variables	Control	Treated	Control	Treated	Control	Treated	Control	Treated
ContestDuration	60.92	61.08	66.00	61.38	26.64	18.59	32.76	19.14
logTotalAwardAmount	9.26	8.76	8.98	8.98	0.99	1.19	1.03	1.05
logMaxSize	3.98	3.95	3.74	3.95	1.00	0.30	1.20	0.32
AddtlRequirement	0.85	0.59	0.77	0.64	0.36	0.50	0.43	0.48
ConsumerGood	0.29	0.46	0.31	0.40	0.46	0.50	0.47	0.50

Figure A2: Balance Plots of Matching Variables





1. Iacus, S.M., King, G., and Porro, G. Causal inference without balance checking: Coarsened exact matching. *Political Analysis*, 20, 1 (2012), 1-24.
2. Iacus, S.M., King, G., and Porro, G. Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association*, 106, 493 (2011), 345-361.
3. Iacus, S.M., King, G., and Porro, G. Matching for causal inference without balance checking. Available at https://ssrn.com/abstract=1152391, 2008.
4. King, G., Nielsen, R., Coberley, C., Pope, J.E., and Wells, A. Comparative effectiveness of matching methods for causal inference. *Unpublished manuscript*: Harvard University, 2011.