Competition and Collaboration in Crowdsourcing Communities

What Happens When Peers Evaluate Each Other?

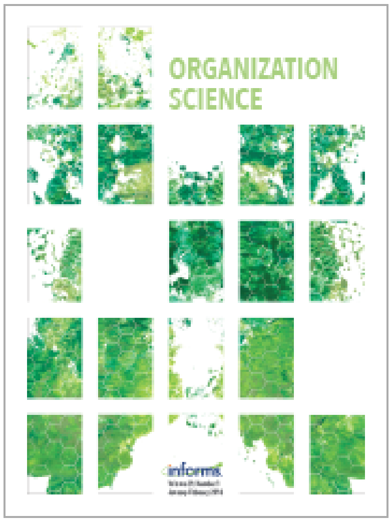
Riedl, Christoph ; Grad, Tom; Lettl, Christopher Ulrich

Document Version
Final published version

Published in:
Organization Science

DOI:
10.1287/orsc.2021.15163

Publication date:
2024



Organization Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Competition and Collaboration in Crowdsourcing Communities: What Happens When Peers Evaluate Each Other?

Christoph Riedl, Tom Grad, Christopher Lettl

https://pubsonline.informs.org/journal/orsc

ORGANIZATION SCIENCE
Articles in Advance, pp. 1–20
ISSN 1047-7039 (print), ISSN 1526-5455 (online)

Competition and Collaboration in Crowdsourcing Communities: What Happens When Peers Evaluate Each Other?

Christoph Riedl,[a,]* Tom Grad,[b] Christopher Lettl[c]

[a] D'Amore-McKim School of Business, Northeastern University, Boston, Massachusetts 02115; [b] Department of Strategy and Innovation, Copenhagen Business School, 2000 Frederiksberg, Denmark; [c] Department of Strategy and Innovation, Vienna University of Economics and Business, 1020 Vienna, Austria

*Corresponding author

Contact: c.riedl@northeastern.edu, https://orcid.org/0000-0002-3807-6364 (CR); tg.si@cbs.dk, https://orcid.org/0000-0003-0528-173X (TG); christopher.lettl@wu.ac.at, https://orcid.org/0000-0002-3267-0872 (CL)

Received: April 22, 2021
Revised: April 14, 2022; January 30, 2023; September 9, 2023; January 23, 2024
Accepted: March 14, 2024
Published Online in Articles in Advance: April 30, 2024

https://doi.org/10.1287/orsc.2021.15163

Abstract. Crowdsourcing has evolved as an organizational approach to distributed problem solving and innovation. As contests are embedded in online communities and evaluation rights are assigned to the crowd, community members face a tension: They find themselves exposed to both competitive motives to win the contest prize and collaborative participation motives in the community. The competitive motive suggests they may evaluate rivals strategically according to their self-interest, the collaborative motive suggests they may evaluate their peers truthfully according to mutual interest. Using field data from Threadless on 38 million peer evaluations of more than 150,000 submissions across 75,000 individuals over 10 years and two natural experiments to rule out alternative explanations, we answer the question of how community members resolve this tension. We show that as their skill level increases, they become increasingly competitive and shift from using self-promotion to sabotaging their closest competitors. However, we also find signs of collaborative behavior when high-skilled members show leniency toward those community members who do not directly threaten their chance of winning. We explain how the individual-level use of strategic evaluations translates into important organizational-level outcomes by affecting the community structure through individuals' long-term participation. Although low-skill targets of sabotage are less likely to participate in future contests, high-skill targets are more likely. This suggests a feedback loop between competitive evaluation behavior and future participation. These findings have important implications for the literature on crowdsourcing design, and the evolution and sustainability of crowdsourcing communities.

Funding: This work was supported by the National Science Foundation [Grant IIS-1514283] and the U.S. Office of Naval Research [Grant N00014-17-1-2542].
Supplemental Material: The online appendix is available at https://doi.org/10.1287/orsc.2021.15163.

Keywords: crowdsourcing • online communities • collaboration and competition • self-promotion • sabotage

1. Introduction

Crowdsourcing as an organizational approach to distributed problem solving and innovation (Afuah and Tucci 2023) has come a long way. Originally conceived as an approach through which firms solicit contributions from independent individuals, the past decade has seen a significant evolution in crowdsourcing design. Two aspects stand out. First, crowdsourcing contests are now often embedded in online communities (Jeppesen and Lakhani 2010, Riedl and Seidel 2018, Grad et al. 2023, Zaggl et al. 2023, Riedl et al. 2024). Second, many of these crowdsourcing communities have adopted open participation architectures that assign evaluation rights to the crowd (Majchrzak and Malhotra 2013, Blohm et al. 2016, Dahlander et al. 2019).

When community members are asked to evaluate their rivals in contest-based communities, they face a tension as they find themselves exposed to both competitive motives to win the contest prize and collaborative participation motives in the community (Deutsch 1949, Fiske 1992). Community members may evaluate their rivals fairly according to the collaborative ideal of the community (Fiske (1992) calls this *equality matching*), or they may act in self-interest and evaluate them strategically to maximize their own chances of winning the contest (*market pricing*). That is, the incentives designed to encourage effort during idea generation (i.e., the contest prize) introduce a competitive element in the otherwise collaborative idea-evaluation process. This may thus tempt idea generators to violate collaborative community norms and evaluate ideas in strategic self-interest. Embedding contests in online communities appears to have created a mismatch between organizational context and incentives (Gallus et al. 2022): the organizational context is collaborative yet the incentives are competitive. Such incongruent incentive schemes can backfire with unintended consequences and lower participation. It is not clear how individuals evaluate their peers as they face this tension between competitive and collaborative motives; and in case they act in self-interest and evaluate strategically: who do they target and with what consequences?

Collaborative aspects of mutual interest, such as helping behavior, reciprocity, and a desire for fairness (Wasko and Faraj 2000, Franke and Shah 2003, Gebauer et al. 2013, Bauer et al. 2016) are key characteristics of communities which bring individuals with shared interests in member or social welfare together (Faraj and Johnson 2011, Majchrzak and Malhotra 2013). Shared values and norms play a crucial role in sustaining good community citizenship and mutually interested collaborative behavior (Chiu et al. 2015, Ivaturi and Chua 2019). Individuals may consequently choose to evaluate fairly and truthfully along the meritocratic ideal. On the other hand, when contests are embedded in such communities, the self-interested motivation to win the contest prize is in conflict with the mutually interested participation motivation of the community. The economics literature on strategic behavior in contests (see Konrad (2009) for a summary) provides precise predictions for the self-interested evaluation behavior we may expect to see in crowdsourcing contests. It predicts that competitors engage in sabotage—defined as the effort of one individual that reduces the performance of another—and self-promotion—a form of unproductive effort that makes one's own contributions appear better without increasing the quality of the contributions itself (Lazear and Rosen 1981, Magee and Galinsky 2008, Konrad 2009).[1] However, it is not at all clear if these predictions of self-interested behavior from contest theory bear out given the community setting. Indeed, research has shown that community members sometimes act collaboratively even when they are direct competitors (Franke and Shah 2003, Harhoff et al. 2003) and that they are sometimes overly positive in their peer evaluations (Aadland et al. 2019, Klapper et al. 2024). Investigating sabotage and self-promotion together is crucial. Although sabotage decreases the quality signal, (self-)promotion increases it and the two may cancel each other out. Sabotage also comes with an externality: Beyond the saboteur, it also increases the chances of winning for all other competitors (Konrad 2009). As a result, the presence of one form of strategic behavior affects the effectiveness of the other. Community members will therefore consider the two forms of strategic behavior jointly as they attempt to resolve the tension between competition and cooperation.

Past work has repeatedly identified such conflicting competitive and collaborative motivations underlying behavior both in online communities and social situations more broadly (Deutsch 1949, Lewis 2000). Self-interested behavior refers to actions of an individual or entity being conducted for the sole purpose of achieving personal benefits (Cropanzano et al. 2005, Smith 2010). In contrast, collaborative behavior refers to actions of individuals or entities being conducted in pursuit of a common goal and thus for the purpose of achieving mutual benefits (Deutsch 1949, Fjeldstad et al. 2012). Examples of collaborative behavior include knowledge sharing and contributing to joint problem solving, providing fair peer evaluations, encouraging new members, or the maintenance of shared resources (Li et al. 2007, Chiu et al. 2015, Chambers and Baker 2020). These self-interested (competitive) and mutual-interested (collaborative) motivations are often simultaneously present and they interact with each other in complex ways (Franke and Shah 2003, Roberts et al. 2006, Bullinger et al. 2010, Nambisan and Baron 2010, Adler and Chen 2011, Hutter et al. 2011, Majchrzak and Malhotra 2013, Bauer et al. 2016, Chambers and Baker 2020).

However, existing studies do not offer precise predictions of how individuals resolve tensions between competitive and collaborative participation motives when asked to evaluate peers in crowdsourcing contests in which they are competing. Recent research on crowdsourcing communities has started to acknowledge the existence of strategic behavior (Archak and Sundararajan 2009, Hutter et al. 2011, Liu et al. 2014, Hofstetter et al. 2018a, Chen et al. 2020, Deodhar et al. 2022). Notably, Klapper et al. (2024) highlight that peer evaluations, when transparent, offer the possibility for strategic behavior as they can provide individuals with a platform to shape their own reputation. However, their work is focused on noncompetitive environments with transparent evaluations which is only the case in some peer evaluation settings.

Therefore, we do not fully understand how and why strategic behaviors arise, its dynamics, nor how it affects the structure of the crowdsourcing community via long-term participation (Balietti and Riedl 2021) when evaluations are anonymous and competitive. Thus, we complement existing research by offering insights for these manifestations of peer evaluations. The covert nature combined with competition is likely to spur an additional strategic behavior beyond the self-promotion found in prior studies (Edelman and Larkin 2015, Klapper et al. 2024), namely sabotage of competitors. Thus, we provide a more comprehensive view on strategic behaviors in crowdsourcing communities that differentiates between self-promotion and sabotage and investigates their complementary effect.

Our paper addresses two research questions. (1) *How do participants in crowdsourcing communities resolve the tension between the competitive and collaborative participation motive when asked to evaluate their peers?* (2) *How does the way participants resolve the tension change the composition of the community when some individuals are more motivated by the competitive aspects of the contest?* A core moderator for self-interested behavior suggested by contest theory is skill[2] because it affects the chance to win the contest and thus affects the gain an individual can expect from making strategic peer evaluations (Boudreau et al. 2016). This suggests that the strength of competitive and collaborative motives may depend on skill and that the tension is greatest for highly skilled participants who are most likely to win the contest, while it is less acute for those of lower skill who know they are unlikely to win.

We investigate these questions using longitudinal panel data from a leading crowdsourcing community (Threadless), analyzing more than 38 million peer evaluations of 150,000 ideas by 75,000 individuals over 10 years. Using self-interested behavior predicted by contest theory as a framework, we look at self-promotion, sabotage, and skill as a moderator to identify both the culprits as well as the targets of strategic behavior. We then draw on the rich online community literature to explain collaborative behavior that is currently not well explained by contest theory. We also explain how the individual-level use of strategic evaluations translates into important organizational-level outcomes by affecting the community structure through individuals' long-term participation.

We have two main findings. First, we find behavior that is both consistent with self-interest and mutual interest in a rich and nuanced way. Consistent with self-interested behavior predicted by contest theory, we find lower-skilled individuals do not sabotage but self-promote. As their skill level increases, they increasingly adopt more competitive strategies and shift from using self-promotion to sabotaging their closest competitors (other high-skill individuals). Conversely, we also find signs of collaborative behavior when they show leniency toward those community members who do not threaten their chance of winning. Using insights from two natural experiments—an evaluation rule change and a change in incentives—we rule out alternative explanations and establish that the observed behavior is indeed strategically motivated. Second, we find that the future participation of the targets of sabotage is potentially affected, depending on their skill level. Low-skill targets of sabotage are *less* likely to participate in future contests, while high-skill targets are *more* likely. This suggests a feedback loop between competitive evaluation behavior and future participation. Individuals increasingly act strategically as their skill level increases, and they find the resulting fierce competition so engaging that they increase their future participation even though it makes them the targets of sabotage.

Our paper makes three key contributions. First, we extend prior research on the competitive and collaborative nature in crowdsourcing (Franke and Shah 2003, Nambisan and Baron 2010, Hutter et al. 2011, Bauer et al. 2016). Our theorizing explains why community members sometimes wear the competitive hat rather than the collaborative one when they evaluate their peers. Specifically, we theorize that leniency is a crucial collaborative element that allows individuals to better justify self-interested strategic evaluations through a form of moral licensing (Blanken et al. 2015). Second, we challenge the assumption in much community research that the tension between competitive and collaborative participation motives is a tension across individuals due to stable attributes (some people are competitive, whereas others are collaborative; Lakhani and Wolf 2003, Erat and Gneezy 2012, Belenzon and Schankerman 2015, Reuben et al. 2015) by showing that the tension is instead context specific and can be described coherently based on the competitiveness of the situation (which depends on the skill of the individuals involved). Third, our work contributes to our understanding of important organizational-level outcomes by showing how ostensible negative strategic behavior can have positive long-term effects. We theorize how the self-reinforcing dynamic of strategic behavior affects the social structure of communities (Faraj and Johnson 2011, Huang et al. 2014, Hofstetter et al. 2018b, Kim et al. 2018, Piezunka and Dahlander 2019).

2. Contest-Theoretic Intuition

How would evaluation behavior look like in a world without community in which evaluators act only according to self-interest? In this section, we develop a simple theoretical framework to formalize predictions of self-interested evaluation behavior for heterogeneous participants in crowdsourcing contests with an open participation architecture. Like much formal modeling work, our model is a simplified version of what we might expect going on in real contests but that can help build credible behavioral foundations about mechanisms that may explain observed behavior (Knudsen et al. 2019).

2.1. Background

Sabotage and self-promotion are conceptually similar in that they affect the relative ranking within contests: Because the contest winner is determined based on the *relative* rank-order of contestants, any action that increases the likelihood to win for one contestant by necessity implies that the likelihood of other contestants is reduced. The economics literature points to one important difference between self-promotion and sabotage with regard to their effect on the relative ranking (Konrad 2009).

Self-promotion affects the likelihood of the culprit to win relative to everyone else (the relative likelihood to win among all others remains the same) while sabotage has an important externality. The action of the culprit affects the relative likelihood of the target and *all other contestants* to win: Everyone's likelihood to win is affected, not only the likelihood of the culprit committing the sabotage. As we will show in our economic model, this distinction has important implications on the cost-benefit calculation to determine which form of strategic behavior to employ. In the Online Appendix (Section A), we provide a brief overview of the economics literature on strategic behavior. There we show that, although sabotage is well understood theoretically, empirical insights are scarce (especially with regard to the effect of skill), and that so far prior work has not modeled sabotage and self-promotion together. It is thus unclear how the two affect each other.

2.2. Basic Model Setup

To keep the modeling tractable and the exposition simple, we build on previous models of sabotage in a single contest (so call one-shot contest), actors that are differentiated by skill, and a single winner prize (Konrad 2000, Harbring and Irlenbusch 2011). We term the participants in the contest "agents." The contest consists of three types of agents: high type agents, low type agents, and neutral outsiders. The contest proceeds as follows. First, agents will make contest submissions (every agent makes exactly one submission). Second, the value of each submission is determined through peer evaluation. Third, the submission with the highest value (determined through peer evaluation) wins the single contest prize (M).

First, we look at contest submissions. Low and high types produce a contest submission (i.e., an idea) of low (b_l) and high quality (b_h), respectively. Without loss of generality, we normalize the evaluation scale to range from zero and one. Furthermore, by construction the quality of a low submission is lower than that of a high submission ($0 < b_l < b_h < 1$). Neutral outsiders are agents who participate only in the evaluation phase of the contest but do not enter their own submission into the contest. As they are not competing for the winner prize, they have no incentive to act strategically and evaluate sincerely; thus, they help us establish the "true quality" of a submission.

Second, to determine the value of a submission, we make the simplifying assumption that every agent (each outsider, low, and high type) evaluates every submission. Because every submission thus receives the same amount of evaluations, we can determine the value of a submission by summing up all the evaluations each submission receives. Although evaluations can be sincere (i.e., b_l or b_h, respectively), an agent can also sabotage any other agent, evaluating their submission with zero or promote them by evaluating their submission with one. As an example, in the case of sincere evaluation by everyone (no promotion and no sabotage), the value of a high type would be b_h times the number of high types (h), low types (l), and outsiders (n; i.e., $v_h = b_h(n + l + h)$). The only other assumption required for our model is that there are fewer high types than low types, and even more outsiders than low types (i.e., $n > l > h$), which should be realistic in most cases.

Third, to determine how to act during peer evaluation (i.e., evaluate sincerely, sabotage, or promote), agents consider the cost and benefits of their actions. The benefit for an agent derives from that agents likelihood of winning the contest multiplied by the prize of the contest (M)—that is, the incentive of the idea generation phase of the contest. To calculate an agent's likelihood of winning, we rely on established contest literature and for simplicity model this as a Tullock contest (Tullock 1980). In a Tullock contest, the probability of winning is proportional to the value of an agent's submission (i.e., the sum of the evaluations the agent receives on her submission) in relation to the total contest output (i.e., aggregate evaluations of all contest submissions). As far as the costs are concerned, we assume that while evaluating sincerely is free of costs, sabotaging and promoting incurs costs (c_s and c_p, respectively). Costs associated with promotion and sabotage might arise from the costs of identifying suitable targets (Harbring et al. 2007, Münster 2007), the moral costs associated with lying (Gneezy et al. 2018, Abeler et al. 2019), or violating social norms (Elster 1989). In our empirical setting, evaluating is anonymous but this is not an assumption of our model. If evaluation is anonymous, this simply means that the costs associated with promotion and sabotage may be relatively small as there are no reputation costs associated with it (but moral costs may still exist).

From this setup, we can already derive four important insights.

Insight 1. The damage done by sabotage to a high type (evaluating zero instead of b_h) is larger than the damage done to a low type (evaluating zero instead of b_l because $b_l < b_h$ by construction). The cost of a single act of sabotage is the same whether a high type or low type is targeted. This suggests that targeting high types is more attractive.

Insight 2. Following a parallel argument, the benefit gained by promoting a low type is larger than the benefit gained by a high-type (i.e., $1 - b_l > 1 - b_h$).

Insight 3. Promoting anyone other than oneself is not rational: One would incur the cost of promoting but has no benefit (in fact, one reduces one's own chance of winning by promoting others). With this setup, it is immediately clear that sabotaging oneself and promoting any agent other than oneself are not helpful in winning the contest prize and no rational agent would engage in such behavior. In the following we thus speak of only "self-promotion" and simplify the notation to consider at most one promotion act per contest entry. Together with the previous insight, this suggests that self-promotion will be more attractive to low types.

Insight 4. All neutral outsiders evaluate others sincerely as they have nothing to gain from strategic behavior (because they did not submit to the contest their chance of winning the prize is zero) yet would incur cost of evaluating strategically.

Regarding the last point, note that neutral outsiders are not necessarily agents who never compete—they simply do not compete in the current contest. We will use this feature to identify strategic behavior in our empirical analysis where agents do not enter every contest: They evaluate as neutral outsiders in some weeks and evaluate as competitors with stakes from idea generation—and incentives to evaluate strategically—in the contest in other weeks.

2.3. Self-Promotion

Self-promotion increases an agent's chance of winning. Because the expected gain from self-promotion is higher for a low type than a high type, there is a cost boundary at which all low type agents decide to self-promote, but none of the high type agents self-promote. If the cost of self-promotion are low enough, all high types will of course self-promote in addition to all low types who will continue to self-promote. Even low types will not self-promote if the cost exceeds their expected gain (see Online Appendix B for a formal proof).

2.4. Sabotage

The decision to sabotage—and who to sabotage—is more complicated. If an agent sabotages another agent, they decreases that agent's output and thus decrease total contest output, which increases that agents probability of winning the contest (in the utility function the numerator of the agent's output remains the same while the total contest output in the denominator decreases). However, this decrease in the denominator also benefits all other agents and thus sabotage has—contrary to self-promotion—an important negative externality (Konrad 2000). In the Online Appendix, we first show a proof that the marginal gain from sabotaging one more agent of a given type is increasing in the number of agents of that type that are already being sabotaged. This means that an agent will either sabotage *all* agents of a given type, or *none* of that type. Second, we show that high types have more to gain from sabotaging other agents of a given type (all low types or all high types) than low types have to gain from sabotaging those the same agents (see proof in Online Appendix). This result follows from the externality associated with sabotage: the negative externality of sabotage that improves everyone else's chance of winning has a relatively larger impact for low types than high types. Hence, the relative gain from sabotaging other agents is larger for high types than low types. Third, we show that both high and low types have more to gain from sabotaging high types than low types, which follows from the fact that the damage done by sabotage to a high type is higher than the damage done to a low type.

A key question now is whether low types or high types find it more attractive to sabotage low types, or low types sabotage high types. That is, at what cost of sabotage will one group start to sabotage the other group (while also considering everyone else's behavior). In the Online Appendix, we go through the exercise of calculating all bounds to establish the precise order according to which groups of agents will decide to sabotage other groups of agents. The bounds reveal that low types have more to gain from sabotaging high types than high types have to gain from sabotaging low types. The intuition behind this result rests on the fact that damage done by sabotage to high types is higher than the damage done to low types.

2.5. Predictions for Empirical Analysis

Our model allows us to make several predictions of the self-interested behavior of crowdsourcing participants. The first is that self-promotion is prevalent: Most agents, including low type agents, self-promote. The second is that high types are the most likely to sabotage. The third is that high types target each other with their sabotage. In addition to guiding our empirical investigation—in particular with regard to the crucial role of skill—these predictions also serve as a baseline expectation for self-interested strategic behavior that we can use to explore contrasting collaborative behavior.

3. Empirical Context

Our empirical setting is Threadless, a prototypical crowdsourcing community (Majchrzak and Malhotra 2016). Threadless is a crowdsourcing and e-commerce platform that hosts weekly T-shirt design contests (Nickell 2010). Since 2001, the site has developed into a leading crowdsourcing platform, pioneering a community-based business model. The company involves its community of 1.5 million members in nearly all aspects of the innovation process and does not employ any in-house designers (Lakhani and Kanji 2008). The platform draws on the creative talent of designers from across the world and has a distinct focus on building and nurturing an online community of creative individuals (Riedl and Seidel 2018).

The tension between competition and collaboration is apparent in our study context: the Threadless T-shirt design community. Threadless describes itself as being an "inspiring design community" (Nickell 2010, title page). Like in many other crowdsourcing communities (Majchrzak and Malhotra 2013, 2016), Threadless users freely reveal their work by posting draft designs, ask for and provide feedback on designs in the forum, inspire each other with their work, read interviews, and learn from each other (Nickell 2010, Riedl and Seidel 2018). According to the founder, Threadless is about "real community, friendships, and working on fun, cool projects together" (Nickell 2010, inside cover flap). Community members value participation as an experience in itself with constructive and developmental spirit (Brabham 2010), making the community ethos front and center. The community has also developed strong social norms supportive of collaborative behavior (Bauer et al. 2016). Threadless as a company is seen as a member in the community in which it participates rather than being its owner (Nickell 2010, p. 45). The designs printed on T-shirts and sold on the community's e-commerce page are chosen by the community and most of the proceeds from sales are distributed to the community (Nickell 2010). Yet, winning a design contest is coveted. The name of the designer of a winning design is printed on each T-shirt, thus giving credit to the individual. Winners are inducted into an elite alumni club, are invited to events, appear in interviews on the community's blog, and are profiled in books and teaching cases (Lakhani and Kanji 2008, Nickell 2010). Other communities even post public leader boards (Grad et al. 2023). Cash prizes for winning a contest have increased over the years (Nickell 2010). Designers have also been honored in additional *Designer of the Year* and *Most Printed Design* awards (Nickell 2010). Ultimately, being a successful contest winner can be a launch pad for design careers beyond Threadless (Nickell 2010).

Contests are divided into two distinct and temporally separated stages: (1) entering a contest by submitting a design and (2) evaluating the submitted designs. To enter a contest, Threadless provides a designer kit and template, and submissions from any standard software program or digitized drawings are accepted, thus creating low barriers to entry. The submission section of the site is static and identical every week. Through a simple form, Threadless enables designers to upload their design, provide a thumbnail, title, and short description. The submission site provides no information as to how many designs have been submitted, or by who, or what their quality might be. At this stage, submissions undergo an editorial review by Threadless staff before they go up for a seven-day rating period.[3]

Threadless posts designs publicly on their website for peer rating, using a scale from zero to five stars. The "score submissions" page shows a grid of all submissions in the running with thumbnails, title (possibly trimmed to fit the grid), and username. To navigate designs, participants can toggle between "currently in the running" and "archive," the design they submitted themselves, and a single "filter by keyword" search. Clicking on the thumbnail or title leads to the scoring page of a specific design which shows the full title and description, a thread of comments, and the days remaining of the scoring period. The page also shows a counter of how many others have already scored the design, but no average rating is shown to avoid social influence and herding. The average rating is shown once the scoring period is over and the submission has entered the archive, but individual ratings are never revealed. That is, ratings are completely anonymous to the community (but not to us as the researchers).

Among the submissions that received the highest average rating, Threadless typically selects three to six designs per week as contest winners. We show the rating percentile of contest winners in the Online Appendix (Figure A.III). On average, contest winners scored in the 95th percentile of all designs submitted that week. The median rank of printed designs is in the 98th percentile. The designs of contest winners are then printed and subsequently sold through the Threadless e-commerce site. Designers receive a cash prize and additional store credit if their design was selected for printing. Threadless does not publish a global ranking with respect to designers' past design evaluations as other crowdsourcing platforms do (e.g., Topcoder). Threadless generated $30 million in revenue in 2012 and provided more than $775,000 in prizes over the observation period. We focus on the regular weekly contests from 2001 to 2011, excluding special and themed contests that Threadless also hosts occasionally from our analysis. The data were provided to us by Threadless directly (i.e., it was not scraped from the website) under a non-disclosure agreement change. The authors declare that they have no relevant financial or nonfinancial competing interests to report.

3.1. Identification Strategy

To empirically identify strategic contest behavior, we leverage the fact that the competition-collaboration tension is most acute for designers who participate in the evaluation when also being active submitters, and less acute when they only participate in the evaluation process and are not themselves competing: A designer may evaluate the submissions of others without having her own design in the running in one week and then enter her own design in the contest *and* evaluate in the next week.[4] We model the effects of participating in a contest on evaluation behavior using a difference-in-difference approach. That is, we model within individual differences between evaluations cast on competitors versus noncompetitors (thus controlling for unobserved individual differences such as being a harsh critic in general) and within submission differences between evaluations cast by those who submitted to the contest and those that did not (thus controlling for submission quality). Evaluations cast by an individual in a contest that the individual did not compete in serve as control for those contests in which they did. At the same time, evaluations of the same submission from individuals who are not themselves competing in the contest serve as controls for individuals who are competing in the contest. In terms of our model, a contestant who did not submit a contest entry is a neutral outsider, and a (high/low type) contestant otherwise.

"Neutral outsiders" are other individuals who submit to contests in general, just not to the one of the current week. Therefore, outsiders are still designers who are in a good position to judge the quality of submissions and not consumers who never submit. Why are outsiders participating in the rating? These outsiders appear to be predominantly motivated to participate in the rating process to engage in the community and learn vicariously from the work of others (Riedl and Seidel 2018). We make the assumption that these outsiders are neutral in the sense that they have no direct incentive to evaluate strategically because they do not have anything to gain from doing so, whereas facing moral costs of lying if they evaluated strategically rather than truthfully. This assumption seems particularly plausible given the community focus of Threadless (Lakhani and Kanji 2008). This assumption not withstanding, it is possible that rivalries among a small set of competitors may cause them to evalute strategically even when not competing (Kilduff et al. 2010). We explore this in Online Appendix D.2.[5]

We model sabotage as the change in probability that individual i submits a zero star rating—the lowest possible rating—on submission j when having submitted to the same contest relative to not having submitted to the same contest. Conversely, we model self-promotion as the change in probability of rating one's own submission with five stars—the highest possible rating—when rating one's own submission versus rating someone else five stars.[6] The two key explanatory variables are (a) a dummy variable indicating whether an individual is a participant in the contest (i.e., whether i has submitted to the same contest c and is in the running for the prize) and (b) a dummy variable indicating if the individual is rating his or her own submission (this is only possible if the other dummy variable is also one). That is, we estimate the following equations:

$$\begin{aligned}\text{0-Star Rating}_{ij} = {} & \beta_{11}\text{Submitted to same contest}_{ij} \\ & + \beta_{12}\text{Rate own submission}_{ij} + \alpha_i \\ & + \alpha_s + \epsilon_{ij}, \end{aligned} \tag{1}$$

$$\begin{aligned}\text{5-Star Rating}_{ij} = {} & \beta_{21}\text{Submitted to same contest}_{ij} \\ & + \beta_{22}\text{Rate own submission}_{ij} + \alpha_i \\ & + \alpha_s + \epsilon_{ij}, \end{aligned} \tag{2}$$

where 0-Star Rating$_{ij}$ is an indicator that is one if the rating submitted by individual i on the submission j is a zero star rating, 5-Star Rating$_{ij}$ is an indicator that is one if the rating submitted by individual i on the submission j is a five star rating, α_i are individual-level and α_s submission-level fixed effects, and ϵ_{ij} are error terms.

We use this approach to separately identify the two types of strategic behavior outlined in the previous theoretical model:

- *Self-promotion:* If a five star vote is assigned to one's own submission, this indicates self-promotion (β_{22} in Equation (1)). We provide robustness tests using a natural experiment to address concerns that this would not reflect strategic behavior but rather the result of overconfidence or an increased preference fit in Section 5.2.
- *Sabotage:* Sabotage is indicated if zero star votes are assigned more liberally when competing in the same contest versus being an outsider who has no stakes in the contest (controlling for how likely the submission is to receive zero stars overall; β_{11} in Equation (1)). We provide results from a second natural experiment to provide evidence that low ratings are strategic rather than a result of endogenous entry into contests in Section 5.1.

We estimate our models as linear probability models for ease of interpretation (Angrist 2001, Greene 2012). A discussion of the usefulness of this approach can be found in Angrist (2001) and Moffitt (2001). Furthermore, notwithstanding the limitations of this approach, efficient estimation approaches for ordinary least squares (OLS) exist for large data sets that support demeaning of multiple fixed effects. We estimate linear probability models using the lfe package for R (Gaure 2013), which supports demeaning of multiple fixed effects. This algorithmic approach is mandatory as our data set is extremely large with more than 38 million observations, 74,525 individual fixed effects, and 154,086 submission fixed effects (or 511 contest fixed effects).

To investigate how sabotage and self-promotion vary across skill levels we compute skill as the time lagged average submission quality.[7] We compute skill for both the individual casting the rating (we refer to this as the *source*) and the individual who made the contest entry being evaluated (we refer to this as the *target*). We then estimate versions of the models in Equations (2) and (1) in which we include source and target status and their interactions with the *Submitted to same contest* and *Rate own submission* dummy variables. Because target skill is time invariant at the submission level, the main coefficient drops out.

3.2. Data

Our data consist of an unbalanced cross-panel of 74,525 individuals participating in 511 contests, with more than 38 million ratings on 154,086 contest entries (Table 1). The skill distribution is right skewed (Figure A.IIa in the Online Appendix), and thus, our empirical setting matches our model, which assumes fewer high types

Table 1. Summary of Rating Behavior

Submitted to same contest	Rate own submission: No	Rate own submission: Yes
No	30,246,070 (13.58%)	
Yes	7,772,772 (10.73%)	114,914 (74.58%)

Note. Parentheses show probability of rating conditional on submitting.

than low types. The average contest size is 407 submissions (Figure A.IIb in the Online Appendix shows the entire distribution). We show descriptive statistics and correlations of main variables in Table 2.

4. Results

We start with baseline estimates of a difference-in-difference model estimated (Table 3).[8] With a continuous measure of ratings as dependent variable (Model 1), we find no signs of sabotage but rather slight leniency toward ones' competitors ($\beta = 0.024$; $p < 0.001$) and strong signs for self-promotion ($\beta = 2.820$; $p < 0.001$). Moving to a linear probability model (Model 2) we find also find overall leniency: a slightly lower likelihood to rate zero stars when an individual submitted to the same contest ($\beta = -0.008$; $p < 0.001$). That is, community members do not appear to act strategically against their competitors but instead show leniency. Our model with individual and submission fixed effects controls for the possibility that some contests include more low-quality submissions that deserve zero stars. Next, we estimate a linear probability model on casting a five star rating. We find a significant level of self-promotion (Model 3: $\beta = 0.855$; $p < 0.001$). The effect of self-promotion is extremely strong. Seventy-five percent of individuals rate their own submissions, and the overwhelming majority of them does so with the highest possible rating: 97% of self-votes are five stars. Given that the average rating for non–self-ratings is only 1.80 stars, this indicates that contestants are self-promoting with dramatically inflated ratings.

Next, we investigate whether there is heterogeneity in the use of strategic behavior among those of different skill. We estimate a variation of Equation (2) and include measures of idea generation skill for both the source and the target of the rating, an interaction between source and target skill, and interaction terms between the *Submitted to same contest* dummy and source and target skill, respectively (Table 4).[9] We find significant heterogeneity in sabotaging behavior. We find that higher-skill individuals are more likely to be the source of sabotage (Model 1: $\beta = 0.012$; $p < 0.001$) and are more likely to be the target of sabotage ($\beta = 0.019$; $p < 0.001$). This pattern mirrors the predictions of our theoretical model. Turning to self-promotion, we estimate a similar model using the *Rate own submission* dummy variable but include only the interaction with source skill, since source and target are the same in the case where a contestant is voting on themselves. We find that higher-skill individuals are significantly less likely to self-promote (Model 2: $\beta = -0.055$; $p < 0.001$) compared with lower-skill contestants. This matches our model prediction that low-skill agents are more likely to engage in self-promotion.

To better interpret the heterogeneity across skill levels, we compute predicted values for the expected *change* in the likelihood to rate zero stars when competing versus not competing (Figure 1). The heat map in (a) shows that high-skill contestants target other high-skill contestants with an up to 6% increased likelihood of assigning a zero star rating when competing compared with their baseline when not competing. The null-line marking the start of sabotaging behavior is not perfectly symmetric: High-skill contestants are sabotaged by contestants of all other skill levels. This matches our theoretical predictions that low types sabotaging high types is the second most lucrative form of sabotage (after high types sabotaging high types). We similarly compute the relative *change* in probability of self-promotion (because source and target skill are the same in the case of self-rating, this analysis is effectively the diagonal of the heatmap in (a)). Figure 1(b) shows that the likelihood that contestants engage in self-promotion decreases with their own skill. This finding is in line with the predictions from our theoretical model that self-promotion is less effective among high types.[10]

As further evidence for the use leniency as mutual-interested compensation for self-interested strategic evaluation, the Online Appendix (Table A.V) shows regressions on the contest level using standard deviation of the ratings that individuals cast as dependent variable. Those regressions show that when individuals have submitted to the

Table 2. Descriptive Statistics and Correlations of Main Variables

Variables	Mean	Standard deviation	Minimum	Maximum	(1)	(2)	(3)	(4)	(5)	(6)
Rating (1)	1.93	1.53	0.00	5.00						
Submitted to same contest (2)	0.21	0.41	0.00	1.00	0.03					
Rate own submission (3)	0.00	0.05	0.00	1.00	0.11	0.11				
Contest size (4)	406.73	131.88	2.00	848.00	0.07	0.11	0.00			
Average score in contest (5)	1.93	0.23	1.26	2.76	0.15	0.04	0.00	0.50		
Source skill (6)	2.26	0.61	0.00	5.00	0.02	0.05	0.00	0.17	0.29	
Target skill (7)	2.28	0.57	0.00	5.00	0.27	−0.01	−0.01	0.13	0.29	0.08

Note. All correlations are significant at $p < 0.001$.

Table 3. Estimates for Strategic Behavior in Contest

Variables	OLS Rating (1)	Linear probability: Sabotage Zero star rating (2)	Linear probability: Self-promotion Five star rating (3)
Submitted to same contest: Yes	0.024***	−0.008***	−0.002***
	(0.001)	(0.000)	(0.000)
Rate own submission: Yes	2.820***	−0.195***	0.855***
	(0.003)	(0.001)	(0.001)
Individual	Fixed	Fixed	Fixed
Submission	Fixed	Fixed	Fixed
Adjusted R^2	0.382	0.372	0.210
No. of observations		38,102,880	

Note. Standard errors are in parentheses, clustered at the submission level.
$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

same contest, they make better use of the full rating spectrum and submit ratings with a higher standard deviation. The effect is amplified by skill so that higher-skilled individuals submit ratings with an even higher standard deviation. This suggests that leniency is a phenomenon within a single contest: individuals accumulate credits on an invisible scorecard by promoting some community members (giving them higher ratings than they deserve) and then spend these credits on self-interested strategic evaluations (giving them lower ratings than they deserve).

Table 4. Estimates for Strategic Behavior with Skill Heterogeneity

Variables	Linear probability: Sabotage Zero star rating (1)	Linear probability: Self-promotion Five star rating (2)
Submitted to same contest: Yes	−0.071***	0.001***
	(0.001)	(0.000)
Rate own submission: Yes	−0.194***	0.957***
	(0.001)	(0.004)
Source skill	−0.028***	−0.007***
	(0.001)	(0.001)
Source skill × *Target skill*	−0.005***	0.008***
	(0.000)	(0.000)
Submitted to same contest: Yes		
× *Source skill*	0.012***	
	(0.000)	
× *Target skill*	0.019***	
	(0.000)	
Rate own submission: Yes		
× *Source skill*		−0.055***
		(0.002)
Individual	Fixed	Fixed
Submission	Fixed	Fixed
Adjusted R^2	0.391	0.243
No. of observations	18,787,584	

Note. Standard errors are in parentheses, clustered at the submission level.
$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

5. Mechanisms of Strategic Behavior

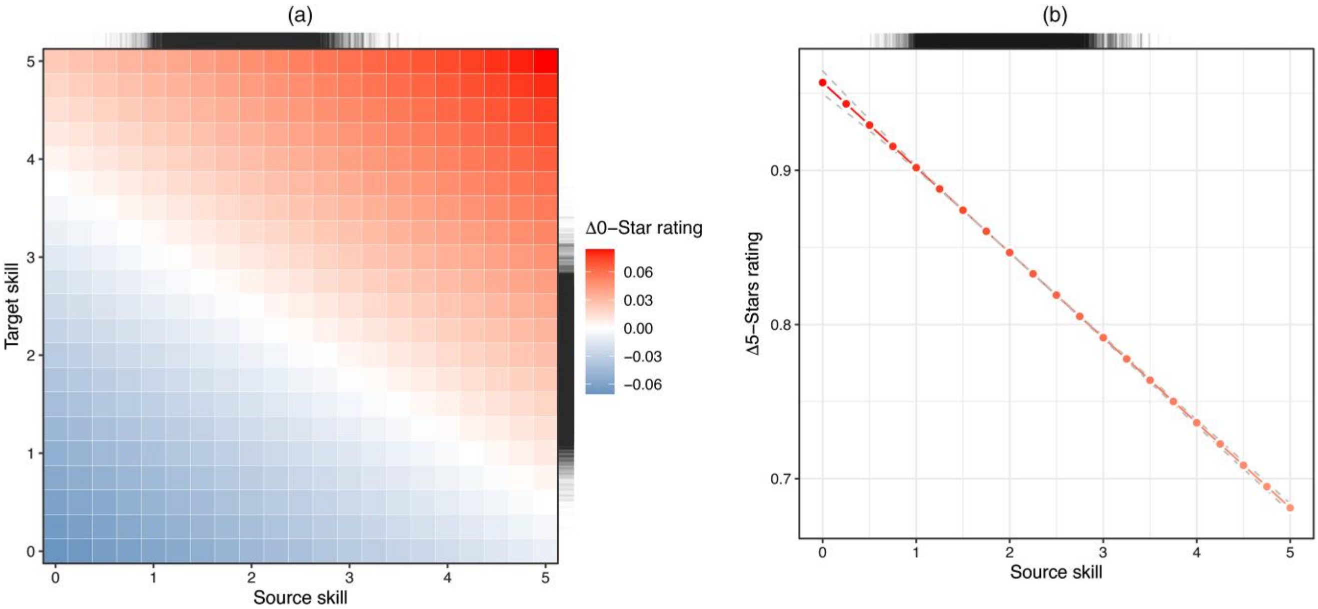
Why do some community members behave strategically? Is it really driven by incentives of the idea generation phase spilling over to idea evaluation or are there other alternative explanations? Exploring our data further, we show that (a) sabotage is strategically motivated as a result of idea generation incentives spilling over to the idea evaluation, (b) self-promotion is strategically motivated and not just a result of overconfidence, (c) those who act strategically mostly use sabotage and self-promotion together rather than substituting one with the other, and (d) we briefly explore the effect that strategic behavior has on the selection of contest winners.

5.1. Sabotage Is a Result of Idea Generation Incentives

If peer evaluation is strategically motivated by idea generation incentives spilling over to idea evaluation, we would expect an increase in the probability to assign zero stars in contests with higher incentives (or more precisely, the prize spread between the contest winner and the contest loser(s); Lazear 1989, Harbring and Irlenbusch 2011). Conversely, if idea evaluation is not strategically motivated, we would expect no change or possibly even a decrease in the probability to rate zero stars after the incentive change because submission quality can be expected to increase as a result of increased effort due to the incentive effect (Jeppesen and Lakhani 2010). We leverage a platform change as a natural experiment—a change in prize money awarded to contest winners. The prize for winning the weekly contest doubled from \$500 to \$1,000 in 2005 (see Figure A.V in the Online Appendix). This doubling of the contest prize was announced on the company's blog on the day it took effect, and therefore, community members had no prior knowledge of it and were not able to withhold submissions or otherwise alter their behavior in anticipation of it.[11]

We use a difference-in-difference design to identify the causal effects of the incentive changes. We use a six-

Figure 1. (Color online) Strategic Behavior by Competitors of Heterogeneous Skill Levels



Notes. For these figures the mean submission quality is used as measure of skill. (a) (sabotage) Relative change in probability of rating zero stars when competing compared with not competing across skill levels (top right: positive values; bottom left: negative values). Outer rugs show distribution of data. There are 4,512 (1,091) observations for source (target) skill ≥ 3.5. (b) (self-promotion) Relative change in probability of rating five stars when rating own submission compared with submissions by others of same skill (error band is 95% confidence interval).

month time window before and after the incentive change. We estimate the same linear probability model for zero star ratings with individual- and submission-level fixed effects as before (Table 5). We find a significant 2.2% increase in the probability to rate zero stars when competing after the incentive change (Model 1: $\beta = 0.022$; $p < 0.001$). The effect is stronger (2.3%) in the first quarter after the incentive change and slightly weaker (2.0%) in the second quarter (Model 2: $\beta = 0.023$, $\beta = 0.020$; both $p < 0.001$). Furthermore, we find that submission quality (as determined by ratings of neutral outsiders who have no incentive to rate strategically) is significantly higher after the incentive change ($\mu_{before} = 1.18$, $\mu_{after} = 1.39$; t test, $p < 0.001$). This indicates that the incentive change, as intended, indeed increases the average quality in the idea generation phase but also increases sabotage in the idea evaluation phase. A placebo test in which we selected a fake date four weeks before the actual incentive change provides further evidence for the causal impact of the incentive change on the rating behavior (Model 3: $\beta = -0.009$; $p < 0.001$) as it rather indicates a negative time trend.

5.2. Self-Promotion Is Strategically Motivated and Not Only Overconfidence

Why do community members evaluate their own submissions so highly? High evaluations for one's own submission could reflect overconfidence of contestants in their own abilities (Camerer and Lovallo 1999, Benabou and Tirole 2002) or an increased preference fit of one's own creative work compared with the work of others (Franke et al. 2010, Berg 2016). To explore the mechanism behind the high self-evaluation, we leverage another

Table 5. Natural Experiments of Sabotage Behavior After an Incentive Changes from \$500 to \$1,000 Prize Money Using ±6 Months as Observation Windows

	Linear probability		
	Sabotage: Zero star rating		
Variables	Base (1)	Persistence (2)	Placebo test (3)
Submitted to same contest: Yes	−0.037***	−0.037***	−0.035***
	(0.000)	(0.000)	(0.000)
Rate own submission: Yes	−0.179***	−0.179***	−0.179***
	(0.000)	(0.000)	(0.000)
Submitted to same contest: Yes			
× *After*	0.022***		0.028***
	(0.000)		(0.000)
× *1st quarter after*		0.023***	
		(0.000)	
× *2nd quarter after*		0.020***	
		(0.000)	
× *Fake after*			−0.009***
			(0.000)
Individual	Fixed	Fixed	Fixed
Submission	Fixed	Fixed	Fixed
Adjusted R^2	0.390	0.390	0.390
No. of observations		825,504	

Note. Standard errors are in parentheses, clustered at the submission level.
***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

natural experiment around a change in the scoring mechanism.

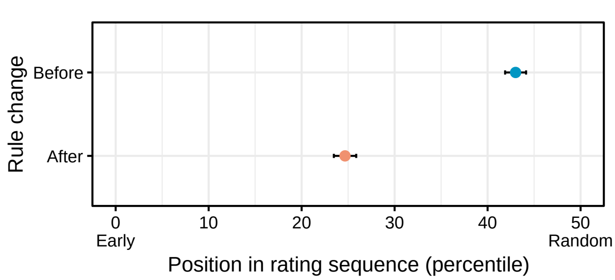
Before the change, every contest entry was posted on the website for the full seven-day rating period. After the change in 2005, contest entries with a rating below 1.5 stars were eliminated from the rating process before the seven-day rating period was complete (see Figure A.VI in the Online Appendix).[12] The change was announced and immediately implemented. If rating one's own submission is not intended strategically but only reflect biased perception and overconfidence, we should not see any change in rating behavior around this rule change. If, however, rating on one's own submission is intended strategically, such an evaluation should happen earlier to maximize the chance to progress past the 100 vote/1.5 stars cutoff.

We analyze the sequence of evaluations for each contest entry and test when in the sequence the self-promoting evaluation was cast (Figure 2). By focusing on the relative position in the evaluation sequence we control for any behavior change that might occur for community members evaluating the same contest entry. We focus our analysis on the 20-week period around the evaluation rule change (±10 weeks before/after the rule change). We find that before the rule change, self-evaluation happens roughly in the middle of the sequence (43rd percentile). That is, self-evaluations were cast roughly on a random day during the evaluation period. After the rule change, self-evaluations are cast significantly earlier, falling in the 25th percentile of the evaluation sequence. Our analysis of this natural experiment suggests that evaluating one's own submission has a strategic intention (make it past the 1.5 stars cutoff) rather than simply a reflection of overconfidence or preference for one's own submission.

5.3. Sabotage and Self-Promotion Are Mostly Used Together

Do individuals use sabotage and self-promotion together to maximize the impact of their strategic behavior or do they substitute one for the other, maybe to atone for moral cost? As noted before, investigating this question empirically is difficult as we can identify sabotage only probabilistically through the difference-in-difference

Figure 2. (Color online) Analysis of Rating Sequence



Note. Self-rating happens significantly earlier after the rule change.

approach. However, we can get some meaningful leverage to identify possible substitution by using self-promotion (or rather the lack thereof) as our starting point. Remember that self-promotion is very common: 86% of individuals who had a chance to self-promote (i.e., submitted to a contest and then rated at least some of their competitors) chose to do so and 97% of those self-ratings are five stars. Conversely, only 14% of individuals who had a chance to self-promote did not do so.

Using the same modeling approach with individual- and submission-level fixed effects as in the main analysis, we find (Table A.VI in the Online Appendix) that the likelihood to rate the competition with zero stars if one has also self-promoted in the same contest is significantly lower than for those who did not (Model 1, $\beta = -0.002; p < 0.01$; having self-promoted in a contest implies also having submitted to that contest). This suggests that some individuals do in fact sabotage somewhat less when they have self-promoted. However, this coefficient estimate captures the effect of marginal individuals: individuals who only sometimes self-promote but not always (as otherwise the individual fixed effect captures this time-invariant behavior). As such, the estimate may mask the more specific effect among that group of more strategic individuals who always self-promote and for which the model cannot pick up variation in behavior. To check this possibility, we estimate the same model, but without individual-level fixed effects (Model 2). Here we find a large positive effect: those who have self-promoted also have a much higher likelihood to sabotage ($\beta = 0.031; p < 0.001$). Therefore, although the small group of marginal raters may indeed commit less sabotage when they also self-promoted, this is not the case for large majority of individuals who tend to use sabotage and self-promotion together to maximize their likelihood of winning the contest.

5.4. Consequences of Strategic Behavior: Effect on Selection of Contest Winners

Does strategic behavior affect idea selection? To quantify the impact that strategic evaluations have on idea selection—compared with which ideas would be selected if all evaluations were truthful and followed the meritocratic ideal of the community—we perform two supplementary analyses. The first assesses the effect of self-promotion, the second of sabotage (see Online Appendix E). For self-promotion, keeping everything else constant, we find that the winner changes in 1.4% of contests (7 of 511), and in 5.5% (28 of 511) there is a change in at least one of the top three ranks. For sabotage, we find that the winner changes in 12% of the contests, and in 48% there is a change among the top three. The effect is especially pronounced in close contests where the contest winner would change in 25% of cases, whereas 65% would see a change in the top three.

6. Consequences of Sabotage on Long-Term Participation

This section investigates a key organizational-level outcome resulting from the emergence of strategic behavior: how does being the target of sabotage and leniency affect long-term participation and hence community structure? Retaining existing members is critical to sustaining the communities in which contests are embedded (Faraj et al. 2011, Ransbotham and Kane 2011). Although existing research has investigated the motivation to join communities (Lakhani and Wolf 2003), much less is known about individuals' progression and long-term participation in communities (Smirnova et al. 2022). There is a rich body of research on fair processes that would suggest that the targets of unfair behavior (such as sabotage) or those who experience random shocks to their performance evaluation, adjust their future effort and engagement (Gomez-Mejia and Balkin 1992, Balietti and Riedl 2021). Several studies that investigate sabotage in laboratory experiments find that being the target of sabotage increases the likelihood of dropout (Chen 2003, Charness et al. 2014, Balietti and Riedl 2021). Although we consider self-promotion of equal importance, this section focuses on the amount of sabotage and leniency received simply because the amount of self-promotion received is virtually identical across all individuals.

The analysis proceeds as follows. We analyze the likelihood to participate in next week's contest, based on the amount of sabotage and leniency an individual experienced on their current week's submission. For this analysis, it is important to recall that sabotage (leniency) can be identified only probabilistically. For example, a low-quality submission may deserve a zero star rating even if it comes from a competitor. Conversely, certain individuals may simply be harsh critics who rate most submissions with zero stars without any intention to sabotage them. To measure the amount of sabotage (leniency) that a submission received, we rely on our microlevel model of sabotage developed previously.

For each individual, we estimate the counterfactual likelihood to cast a zero stars evaluation of neutral outsider (using Model 1 from Table 4). We then compute the "residual" between the evaluation that was actually cast and a counterfactual without strategic behavior. That is, our microlevel model of sabotage allows us to estimate the degree to which any of the zero star ratings are likely due to sabotage (accounting for submission fixed effects, evaluator fixed effects, and time varying covariates like source and target skill). We also compute the amount of leniency a submission received as "residual" between a non–zero stars rating that should have been a zero star rating. We then aggregate the evaluation-level residual to the submission level by taking the average across all ratings. This then captures the level of sabotage (leniency) that a submission received and gives us a panel data set in which we have an estimate for the degree of sabotage (leniency) received by an individual for each contest. We then estimate a multiple event hazard model (Therneau and Grambsch 2000) to predict the likelihood of long-term participation.[13]

We find strong positive main effects of skill (Table 6, Model 1: $\beta = 0.165; p < 0.001$): Higher-skilled individuals are more likely to participate in future contests. There is a small marginally significant main effect of sabotage ($\beta = -0.017; p < 0.10$). We find a significant positive interaction between the level of sabotage received and skill ($\beta = 0.066; p < 0.001$). That is, high-skilled community members who received one standard deviation more sabotage are about 6.6% more likely to participate in the next contest compared with those who received only the average amount of sabotage. To better interpret the effect across skill levels, we convert skill to quintiles, using the middle quintile as our reference category (Model 2). Looking at the interaction between the level of sabotage received and skill, we find that individuals in first and second quintile react negatively to being targets of sabotage: their likelihood of participation in the next round drop significantly by 16% and 8%, respectively ($\beta = -0.158$; $p < 0.001$ and $\beta = -0.076; p = 0.003$). For high-skilled competitors, we find the opposite: They react positively to sabotage and their likelihood to participate increases. There is an 8% increase for individuals in the fourth skill quintile ($\beta = 0.075; p = 0.003$) and a 12% increase ($\beta = 0.119; p < 0.001$) for individuals in the highest skill quintile in the expected hazard to participate in the next round for a one unit change in received sabotage (one standard deviation).[14]

The pattern is reversed for the amount of leniency that a submission receives. The lowest skilled individuals increase their future participation by 17% (Model 4: $\beta = 0.166; p < 0.001$) when they receive an additional one standard deviation of leniency. The highest-skilled individuals on the other hand decrease their likelihood of future participation. For high-skilled individuals, we need to caution against placing too much weight on the coefficient for leniency: It is very rare that high-skilled individuals would submit a design of such low quality that it deserves a zero star rating, and hence there is virtually no room to show leniency toward high-skilled individuals.

7. Discussion

Crowdsourcing has evolved as an organizational approach to distributed problem solving and innovation. As contests are embedded in online communities and evaluation rights are assigned to the crowd, community members face a tension between competitive and collaborative participation motives. This tension tempts idea generators to violate community norms and evaluate their peers strategically. Using large-scale digital trace

Table 6. Analysis of Long-Term Participation

	Hazard: Participate in next round			
	Sabotage		Leniency	
Variables	Skill continuous (1)	Skill quintiles (2)	Skill continuous (3)	Skill quintiles (4)
Sabotage received	−0.017†	−0.017		
	(0.010)	(0.019)		
Sabotage received				
× *Skill*	0.066***			
	(0.009)			
× *Skill Q1*		−0.158***		
		(0.027)		
× *Skill Q2*		−0.076**		
		(0.027)		
× *Skill Q4*		0.075**		
		(0.026)		
× *Skill Q5*		0.119***		
		(0.026)		
Leniency received			−0.110***	−0.044
			(0.018)	(0.028)
Leniency received				
× *Skill*			−0.123***	
			(0.011)	
× *Skill Q1*				0.166***
				(0.028)
× *Skill Q2*				0.117***
				(0.029)
× *Skill Q4*				−0.208***
				(0.034)
× *Skill Q5*				−0.525***
				(0.046)
Controls				
Skill	0.165***		0.170***	
	(0.013)		(0.013)	
Skill Q1		−0.341***		−0.443***
		(0.035)		(0.035)
Skill Q2		−0.084*		−0.114***
		(0.032)		(0.033)
Skill Q4		0.122***		0.096**
		(0.035)		(0.035)
Skill Q5		0.225***		0.036
		(0.044)		(0.043)
Submission quality	0.976***	0.982***	0.821***	0.797***
	(0.027)	(0.028)	(0.034)	(0.036)
Submission popularity	0.067***	0.066***	0.071***	0.070***
	(0.008)	(0.008)	(0.008)	(0.008)
Competition size	−0.000	−0.000	−0.000†	−0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Average contest rating	0.953***	0.933***	0.845***	0.779***
	(0.053)	(0.053)	(0.056)	(0.056)
AIC	912,233.777	911,871.273	922,869.780	921,779.500
R^2	0.317	0.321	0.327	0.337
No. of observations	72,162	72,162	73,018	73,018
No. of events	48,052	48,052	48,639	48,639
PH test	0.000	0.000	0.000	0.000

Note. Multiple event hazard model of participation in the next week.
***$p < 0.001$; **$p < 0.01$; *$p < 0.05$; $^{\dagger}p < 0.1$.

data from Threadless, a prototypical crowdsourcing community (Majchrzak and Malhotra 2016), we answer the questions of how community members balance the competitive and collaborative motives as they evaluate their peers, and how the individual-level decisions they make change the structure of the community. We show that as their skill level increases, the competitive motive increasingly wins out: community members shift from using self-promotion to sabotaging their closest competitors (other high-skill individuals). However, community members also act in the collaborative spirit of the community and show leniency toward those who do not directly threaten their own chance of winning.

In addition to these immediate short-term effects, we also find surprising long-term effects of the fierce competition that ensues. Although low-skill targets of sabotage are less likely to participate in future contests, high-skill targets are more likely. Furthermore, low-skilled beneficiaries of leniency are encouraged and are more likely to participate in the future. This suggests a feedback loop between competitive evaluation behavior and future participation.

7.1. Theoretical Implications

Our findings have three important implications for the literature on the interplay between competitive and collaborative motives in crowdsourcing design, as well as the evolution and sustainability of crowdsourcing communities.

First, we extend prior research on the nature of competitive and collaborative behavior in crowdsourcing. The tension between competition and collaboration has been identified by several crowdsourcing researchers (Franke and Shah 2003, Bullinger et al. 2010, Adler and Chen 2011, Hutter et al. 2011, Majchrzak and Malhotra 2013, Boudreau and Lakhani 2015) and features prominently in the broader organizations literature (Deutsch 1949, Lado et al. 1997, Tsai 2002, Gallus et al. 2022). This work has shown that the level of knowledge sharing and mutual support decreases in communities when members are also competing (Franke and Shah 2003, Harhoff et al. 2003). We complement this research with insights into evaluation behavior. We show that the competitive incentives designed to encourage effort during idea generation spill over into the collaborative community, creating a mismatch between the collaborative organizational context and competitive incentives (Gallus et al. 2022). However, the resulting strategic behavior is not rampant and out of control as contest theory might predict. This suggests that embedding contests in communities drives up the cost of amoral strategic behavior even in a community without reputation costs (because peer ratings are anonymous) and low search costs (because sabotaging everyone would only cost a few mouse clicks). The tension between competitive and collaborative participation motives is felt most acutely by high-skilled individuals who are most likely to win (and thus have the most to gain). Hence, they are the most willing to accept the cost of violating community norms.

However, community members not only sabotage, but they also show leniency. We theorize that showing leniency is a crucial collaborative element that allows individuals to better justify violating community norms with competitive evaluations through a form of moral licensing (Blanken et al. 2015). Moral licensing describes behavior in which people who behave in a moral way can also display behaviors that are immoral, unethical, or otherwise problematic (Merritt et al. 2010). Specifically, we theorize that individuals accumulate credits on an invisible scorecard by promoting some low-skilled community members (giving them higher ratings than they deserve) and then spend these credits on self-interested strategic evaluations (giving some rivals lower ratings and self-promoting their own work). Often, the more individuals are invested in the community, the stronger their perceived need to atone for the violation of community norms may be (Ashforth and Johnson 2001). From the perspective of contest theory, leniency is unexpected collaborative behavior because promoting anyone other than oneself is counter to the self-interest of winning the contest. However, this leniency is not simply a collaborative act within the community. It has its own strategic nature as it is targeted toward those who least threaten one's own chance of winning. Recognizing that showing leniency toward low-skilled community members is strategically motivated has far-reaching implications for our understanding of interactions in communities (Wasko and Faraj 2005, Faraj and Johnson 2011, Ren et al. 2012). Future models of behavior in online communities need to consider not only interactions themselves but also how they are motivated. Thus, leniency appears to be a crucial collaborative element that allows individuals to better justify self-interested strategic evaluations. It alleviates the tension between the mismatched collaborative organization structure and the competitive incentive (Gallus et al. 2022).

Second, our dyad-level approach offers complementary insights to previous studies that examined participation motivations at the individual level (Hippel and Krogh 2003, Wasko and Faraj 2005). Prior work focuses on competitive and collaborative participation motives as individual-level traits (some individuals are competitive, whereas others are collaborative; Lakhani and Wolf 2003, Erat and Gneezy 2012, Belenzon and Schankerman 2015, Reuben et al. 2015) to explain why community members behave in certain ways. By contrast, we explain how the contextual factor of the competitive situation explains behavior. The same individual in the same contest may act competitively toward one individual yet show leniency toward another. Our theory and empirical results imply that within individual changes in skill and contextual factors (whether the individual is itself competing for the prize versus a neutral outsider, and who the target of the behavior is) make some interactions more (or less) competitive and explain the observed microlevel behavior. When community members are not themselves competing, they evaluate their peers fairly with equal treatment of everyone (i.e., according to equality matching; Fiske 1992). When they are competing, the evaluations depend on the competitiveness of the situation. The evaluation of those who do not threaten a community member's own chance of winning fall into a different category of relational behavior. Those evaluations are not governed by the competitive participation motivation (i.e., they are not based on a rational cost-benefit calculation of market pricing; Fiske 1992) and are instead evaluated under a collaborative scheme of kindness and selfless generosity (i.e., under communal sharing; Fiske 1992). As a result, skill (and social confirmation of skill from past success) is a key factor that shapes the strength and saliency of the self-interested motivation because it affects the gains individuals can expect from such behavior and can thus explain whether individuals wear a competitive or collaborative hat (Ashforth and Johnson 2001). Differentiating behavior based on the target allows individuals to be competitive and cooperative at the same time, thus alleviating the tension (Waldman et al. 2019). Individuals may have joined the community out of intrinsic motivation but once they have high skill (and have experienced success), the competitive motive takes over. This contributes to recent research which has started to acknowledge that there are changes and growth over individual "careers" in communities as they learn and improve their skill (Dahlander and O'Mahony 2011, Riedl and Seidel 2018, Soda et al. 2021, Smirnova et al. 2022).

Our dyad-level investigation also adds to the nascent crowdsourcing literature that has started to acknowledge the existence of strategic behaviors (Archak and Sundararajan 2009, Hutter et al. 2011, Liu et al. 2014, Hofstetter et al. 2018a, Chen et al. 2020, Deodhar et al. 2022, Klapper et al. 2024) by providing well-identified empirical evidence for both sabotage and self-promotion in a prominent crowdsourcing platform. In particular, our findings complement prior research on peer evaluations by highlighting the importance of competition and covertness of evaluations: Although Klapper et al. (2024) find that Wikipedia editors target negative evaluations at community members who are unlikely to retaliate, we show that when in direct competition—as it is the case in many crowdsourcing settings—negative evaluations are targeted toward the most prolific competitors. Furthermore, positive evaluations are more likely to be granted to those lower-skilled competitors who matter little for the overall outcome of the contest. Thus, the dynamic reverses compared with the findings of Klapper et al. (2024). This has important implications for crowdsourcing community organizers as it underscores the important roles of competition and transparency when designing the interactions of community members. Although nontransparent peer evaluations limit the possibility to use peer evaluations as a means to portray oneself, it does reduce the costs to engage in sabotage behavior toward competitors. For settings where competition is prevalent it also highlights the usefulness of game-theoretic considerations to assess heterogeneous evaluation patterns.

On a more general note, we add important field evidence to the existing economics literature on sabotage that has mostly relied on laboratory experiments (Lazear 1989, Konrad 2000, Harbring and Irlenbusch 2011, Charness et al. 2014) and found ambiguous evidence for the association between sabotage and skill (Harbring et al. 2007, Charness et al. 2014, Chambers and Baker 2020).

Third, our work contributes to our understanding of the evolution and sustainability of online communities by showing how individual-level strategic behavior affects the organizational-level social structure of communities (Faraj and Johnson 2011, Huang et al. 2014, Hofstetter et al. 2018b, Kim et al. 2018, Piezunka and Dahlander 2019). We document a self-reinforcing dynamic in which increased member skill not only leads to more strategic behavior but also increased future participation. This complements past studies that have investigated the fluidity of communities focused on aspects of self-selection and open boundaries—that is, dynamics *across* individuals—with insights of dynamics *within* individuals (Faraj et al. 2011, Felin et al. 2017). Contrary to the expectation that unfair behavior would reduce engagement (Gomez-Mejia and Balkin 1992, Franke et al. 2013, Faullant et al. 2017, Balietti and Riedl 2021), strategic behavior appears to facilitate future participation because it encourages low-skilled individuals with leniency and engages high-skilled community members in challenging competition through sabotage. Strategic behavior appears to play an important organizational role in stabilizing the core of a community by engaging members in intense competition.

Why does being a victim of sabotage make highly skilled community members more likely to participate in future contests? We theorize that the fierce competition among the high-skilled adds to their intrinsic motivation and promotion focus. Competition can be thrilling and exciting, even when outcomes are negative (Franken and Brown 1995). Intrinsic motivation and corresponding theories may provide an explanation. First, according to self-determination theory (Ryan and Deci 2000) skilled individuals may experience a sense of competence and autonomy that can enhance their motivation. Having been a victim of sabotage could serve as a confirmation of being considered a true competitor, which could increase motivation and thus the likelihood of future participation. Furthermore, the fierce competition among the high-skilled adds to the status incentive of the competition and may be perceived as a challenge to overcome, rather than a threat (To et al. 2020). Second, the challenge of competition is a trigger of a flow state (Csikszentmihalyi 1990). Higher-skilled individuals have a wider range in which flow can occur when the challenge-level rises compared with their lower-skilled counterparts. Being a victim of sabotage may thus increase the perceived challenge: Although low-skilled individuals perceive sabotage as a threat, high-skilled individuals perceive it as a challenge (promotion focus To et al. 2020), making their participation in future contests more likely.

Prior work has focused on collaborative behavior and explained why members in crowdsourcing act reciprocally toward others who have helped them in the past (Perry-Smith 2006, Jeppesen and Lakhani 2010, Dahlander et al. 2016, Safadi et al. 2021). By contrast, we explain why apparent reciprocity in terms of social network structure may also result from competitive behavior in which high-skilled members sabotage other high-skilled community members. We thus provide an additional explanation for the phenomenon of a consolidated community core of high-skill individuals that is a key characteristic of many crowdsourcing communities (Perry-Smith 2006, Jeppesen and Lakhani 2010, Dahlander et al. 2016, Safadi et al. 2021). Our contest theory model and empirical results imply that behaviors (and resulting network structures) that appear to be reciprocal and collaborative can in fact be deeply competitive. Together, this shows why strategic behavior must be incorporated in our theoretical understanding of crowdsourcing communities that have often been studied using structural social network approaches. The understanding of the implications of social structure in communities could therefore benefit from a more systematic integration of, and attention to, how actors' behavioral motivation guides their behavior.

With the proliferation of social network studies focused on static structural aspects (Perry-Smith 2006, Jeppesen and Lakhani 2010, Dahlander et al. 2016, Safadi et al. 2021), the feedback loops between dynamic interactions and their effect on structure suggest important questions for future research (Foley et al. 2021, Fulker et al. 2021, Soda et al. 2021). Our study is a first step in that direction. It provides an alternative explanation for why communities often have a core-periphery structure. Future work needs to consider not only the structure of community interaction itself but also the valence of interaction (e.g., are comments helpful or hurtful? Are ratings positive or negative?) and how the interactions are motivated. This may reveal more nuanced network structures in terms "negative" (competitive) ties (Labianca and Brass 2006), adding to our understanding of social networks from structural analysis that has focused on ties based on social relationships (e.g., friendship), communication, and information flow (Wasserman and Faust 1994, Borgatti et al. 2009).

7.2. Practical Implications

Our findings also have important practical implications. Counter to common sense, platform design that allows strategic behavior may not necessarily be bad. From an organizational perspective, strategic behavior is usually seen as destructive and thus undesirable. According to this view, organizations are well advised to limit its impact by designing incentives (and structures) that reduce strategic behavior. Our study reveals that this view is not always valid. The fierce competition among high-skilled community members which makes them targets of sabotage increases their future participation. It also leads to the encouragement of less skilled members through leniency. As a result, the competitive strategic behavior that ostensibly runs counter to the notion of a collaborative community does in fact have positive long-term effects on the community. Therefore, our study challenges conventional wisdom on strategic behavior: instead of eliminating strategic behavior entirely, the organizers of crowdsourcing communities may encourage it within controlled boundaries (e.g., anonymous voting).

7.3. Limitations, Generalizability, and Future Research

This paper is not without limitations. First, although Threadless is a prototypical crowdsourcing community and therefore highly suitable for generalization, there may still be some differences to other settings. For example, Threadless does not publish an overall ranking which may limit competitive motivation compared with other settings (e.g., Topcoder publishes public rankings that can spur rivalry; Grad et al. 2023). Future research could therefore investigate the patterns of strategic behaviors in settings that differ from ours. The contest theory model can serve as a starting point for that research. The findings may apply to other contexts beyond crowdsourcing. As the theoretical reasoning and empirical setting of this study is concerned with individuals anonymously evaluating their rivals in an otherwise collaborative environment, the study's results may generalize to contexts with similar characteristics such as peer evaluation in teams, academic publishing, grant awarding, or political races. Second, this paper focuses on two specific forms of strategic behavior on the individual level. The study of other forms of strategic behaviors that emerge between two or more participants over time, such as interpersonal rivalry and reciprocity, are exciting avenues for future research. Finally, our work opens new fields to investigate career dynamics in communities and the connection between individual behaviors and organizational-level outcomes like overall community structures.

Acknowledgments

The authors thank Jennifer Brown, Blair Davey, Imke Reimers, Ulrich Berger, and Ben Greiner, as well as participants at the National Bureau of Economic Research Summer Institute 2019 and participants of the Digital Innovation Workshop 2019 at Boston College, for helpful comments.

Endnotes

[1] In terms of rating behavior this corresponds to rate oneself up (self-promotion) and to rate competitors down (sabotage). There are other forms of strategic behavior such as strategic entry decisions

(Liu et al. 2014), cheap talk (Archak 2010), and reciprocal rating (Hutter et al. 2011).

[2] Skill in terms of generating high-quality contest entries not in terms of acting strategically, although the two may be correlated.

[3] Threadless lists several reasons for declining a submission at this stage: designs using copyrighted material, duplication of prior work, inappropriate content, technical errors with the image file such as low resolution, text only designs, designs that use too many colors, are too large or cannot be printed for other reasons. The rating period is seven days with the exception of the change in rules about dropping design earlier which we exploit in the natural experiment on self-promotion in Section 5.2.

[4] See Section C.2 in the Online Appendix for more details on the empirical patterns.

[5] Fake accounts are not of great concern to our analysis. Although fake accounts may exist on Threadless in principle, our analysis includes only ratings cast by designers—users who also make design submission. To qualify as "designer," a user needs to make a design submission that passes Threadless' basic muster of being a valid submission (a real t-shirt design on the Threadless template, cannot be blank, cannot be an obvious copy of other work, etc.) and have their design submission actually put up for voting.

[6] By using the two extreme outcomes, zero stars and five stars, respectively, we use a conservative measure of sabotage and self-promotion respectively, because (a) we only include the likely report of maximal lies (Gneezy et al. 2018), and (b) we make sure not to conflate the two behaviors. An alternative specification for sabotage would be, e.g., to use the probability that a design gets assigned a zero or one star rating. However, this would not just include sabotage (all designs that are of higher quality) but also self-promotion (all zero star designs that get assigned a one star rating) and therefore conflate the two behaviors. A similar argument can be made for self-promotion.

[7] We use the quality of the best previous submission as a robustness check that is consistent with our results, see Section C.4 in the Online Appendix.

[8] We cluster standard errors on the submission level to account for possible dependence of ratings of the same submissions and in the same contest. One may alternatively be concerned about capturing autocorrelation at the rater level. We examine the robustness of our specification in Online Appendix C.5. The clustered standard errors are virtually identical to robust standard errors. Furthermore, p-values are generally much smaller than 0.001 (i.e., $1e-9$) so that our substantial conclusions are not affected by the level of clustering of standard errors.

[9] This analysis relies on a reduced sample as we need to exclude (a) all ratings that an individual casts before making their first submission (because for those individuals we cannot compute idea generation skill), and (b) all ratings on every contestant's first submission (again, because idea generation skill is still unknown for those individuals). We check whether the use of subsamples drives our findings in Section C.4 in the Online Appendix but find no evidence of this.

[10] We perform three important robustness tests. First, we repeat the same analysis using the quality of the *best* design an individual submitted in the past (instead of the average of all past submissions) and find almost identical coefficients (Table A.III and Figure A.IV in the Online Appendix). Second, we explore alternative measures of skill: tenure, experience, and being a past winner (Table A.IV in the Online Appendix). We find no significant effect of tenure. We find mixed results for experience. We find a lower likelihood to sabotage for those with more evaluation experience but higher likelihood to sabotage for those with more submission experience. This is consistent with the interpretation that competitive motivation is the driver behind strategic behavior (i.e., those who are competitively motivated and make many submissions, sabotage more while those who may be more socially motivated and submit many evaluations). We find that past winners act significantly more strategically. They are both much more likely to sabotage, and other past winners are the targets. Being a past winner could be seen as socially validated skill. Third, we explore whether the reported effect of strategic behavior could be explained by dyadic rivalries instead (Kilduff et al. 2010, section D.2). We find that sabotage happens outside, and on top of, any pre-existing dyadic rivalries. Even within the same dyad, we still find a difference in individuals' rating behavior between competing for the same prize versus not competing which indicates strategic behavior. In summary, dyadic rivalry patterns are not sufficient to explain strategically motivated sabotage.

[11] This natural experiment also helps us address another potential alternative explanation. An alternative explanation for the observed negative rating behavior could be seen in the endogenous decision to enter contests (Bockstedt et al. 2016). Contestants may submit entries to contests where they expect that competition is weak and therefore deserve lower ratings. Following the argument of endogenous entry, we would expect either no change in zero-star ratings as contestants should not be able to time their contest entry around the incentive change.

[12] Specifically, contest entries that had received ratings from 100 different people, but the average of those rating was less than 1.5 stars, were eliminated. The average rating still was not visible during rating process but only to the platform owner. This rule change was introduced to help the platform owner to focus on a smaller set of contest entries by removing low quality submissions from the contests earlier.

[13] We estimate an interval Cox proportional hazard model for multiple events, controlling for individual hazard rate of participation, exploiting within-individual variation in received sabotage; robust standard errors clustered at the individual level. We include the key observables that a competitor observes after a contest: the rating their submission received, the number of ratings received, the level of competition (number of competitors and average quality of all submissions in the contest), and our focal variable of interest, the amount of sabotage received.

[14] We can also interpret the coefficients of our control variables. Having done well in a contest (high rating, many ratings) makes future participation significantly more likely. Competition has mixed effects (positive effect of high average rating across all contest submissions, no effect of competition size).

References

Aadland E, Cattani G, Ferriani S (2019) Friends, gifts, and cliques: Social proximity and recognition in peer-based tournament rituals. *Acad. Management J.* 62(3):883–917.

Abeler J, Nosenzo D, Raymond C (2019) Preferences for truth-telling. *Econometrica* 87(4):1115–1153.

Adler PS, Chen CX (2011) Combining creativity and control: Understanding individual motivation in large-scale collaborative creativity. *Accounting Organ. Soc.* 36(2):63–85.

Afuah A, Tucci CL (2023) Reflections on the 2022 AMR decade award: Crowdsourcing as a solution to distant search. *Acad. Management Rev.* 48(4):597–610.

Angrist JD (2001) Estimation of limited dependent variable models with dummy endogenous regressors. *J. Bus. Econom. Statist.* 19(1):2–28.

Archak N (2010) Money, glory and cheap talk: Analyzing strategic behavior of contestants in simultaneous crowdsourcing contests on TopCoder.com. *Proc. 19th Internat. Conf. World Wide Web* (Association for Computing Machinery, New York), 21–30.

Archak N, Sundararajan A (2009) Optimal design of crowdsourcing contests. *Proc. Internat. Conf. Inform. Systems* (Association for Information Systems, Atlanta, GA), 200.

Ashforth BE, Johnson SA (2001) Which hat to wear. *Social Identity Processes in Organizational Contexts* (Psychology Press, New York), 31–48.

Balietti S, Riedl C (2021) Incentives, competition, and inequality in markets for creative production. *Res. Policy* 50(4):104212.

Bauer J, Franke N, Tuertscher P (2016) Intellectual property norms in online communities: How user-organized intellectual property regulation supports innovation. *Inform. Systems Res.* 27(4): 724–750.

Belenzon S, Schankerman M (2015) Motivation and sorting of human capital in open innovation. *Strategic Management J.* 36(6): 795–820.

Benabou R, Tirole J (2002) Self-confidence and personal motivation. *Quart. J. Econom.* 117(3):871–915.

Berg JM (2016) Balancing on the creative highwire: Forecasting the success of novel ideas in organizations. *Admin. Sci. Quart.* 61(3): 433–468.

Blanken I, Van De Ven N, Zeelenberg M (2015) A meta-analytic review of moral licensing. *Personality Soc. Psych. Bull.* 41(4): 540–558.

Blohm I, Riedl C, Füller J, Leimeister JM (2016) Rate or trade? Identifying winning ideas in open idea sourcing. *Inform. Systems Res.* 27(1):27–48.

Bockstedt J, Druehl C, Mishra A (2016) Heterogeneous submission behavior and its implications for success in innovation contests with public submissions. *Production Oper. Management* 25(7): 1157–1176.

Borgatti SP, Mehra A, Brass DJ, Labianca G (2009) Network analysis in the social sciences. *Science* 323(5916):892–895.

Boudreau KJ, Lakhani KR (2015) "Open" disclosure of innovations, incentives and follow-on reuse: Theory on processes of cumulative innovation and a field experiment in computational biology. *Res. Policy* 44(1):4–19.

Boudreau KJ, Lakhani KR, Menietti M (2016) Performance responses to competition across skill levels in rank-order tournaments: Field evidence and implications for tournament design. *RAND J. Econom.* 47(1):140–165.

Brabham DC (2010) Moving the crowd at threadless. *Inform. Comm. Soc.* 13(8):1122–1145.

Bullinger AC, Neyer AK, Rass M, Moeslein KM (2010) Community-based innovation contests: Where competition meets cooperation. *Creative Innovation Management* 19(3):290–303.

Camerer C, Lovallo D (1999) Overconfidence and excess entry: An experimental approach. *Amer. Econom. Rev.* 89(1):306–318.

Chambers CR, Baker WE (2020) Robust systems of cooperation in the presence of rankings: How displaying prosocial contributions can offset the disruptive effects of performance rankings. *Organ. Sci.* 31(2):287–307.

Charness G, Masclet D, Villeval MC (2014) The dark side of competition for status. *Management Sci.* 60(1):38–55.

Chen KP (2003) Sabotage in promotion tournaments. *J. Law Econom. Organ.* 19(1):119–140.

Chen P, Sun H, Fang Y, Liu X (2020) Conan: A framework for detecting and handling collusion in crowdsourcing. *Inform. Sci.* 515(April):44–63.

Chiu CM, Huang HY, Cheng HL, Sun PC (2015) Understanding online community citizenship behaviors through social support and social identity. *Internat. J. Inform. Management* 35(4):504–519.

Cropanzano R, Goldman B, Folger R (2005) Self-interest: Defining and understanding a human motive. *J. Organ. Behav.* 26(8):985–991.

Csikszentmihalyi M (1990) *Flow: The Psychology of Optimal Experience* (Harper & Row, New York).

Dahlander L, O'Mahony S (2011) Progressing to the center: Coordinating project work. *Organ. Sci.* 22(4):961–979.

Dahlander L, Jeppesen LB, Piezunka H (2019) How organizations manage crowds: Define, broadcast, attract, and select. Sydow J, Berends H, eds. *Managing Inter-Organizational Collaborations: Process Views,* Research in the Sociology of Organizations, vol. 64 (Emerald, Leeds, UK), 239–270.

Dahlander L, O'Mahony S, Gann DM (2016) One foot in, one foot out: How does individuals' external search breadth affect innovation outcomes? *Strategic Management J.* 37(2):280–302.

Deodhar SJ, Babar Y, Burtch G (2022) The influence of status on evaluations: Evidence from online coding contests. *Management Inform. Systems Quart.* 46(4):2085–2110.

Deutsch M (1949) A theory of co-operation and competition. *Human Relations* 2(2):129–152.

Edelman B, Larkin I (2015) Social comparisons and deception across workplace hierarchies: Field and experimental evidence. *Organ. Sci.* 26(1):78–98.

Elster J (1989) Social norms and economic theory. *J. Econom. Perspective* 3(4):99–117.

Erat S, Gneezy U (2012) White lies. *Management Sci.* 58(4):723–733.

Faraj S, Johnson SL (2011) Network exchange patterns in online communities. *Organ. Sci.* 22(6):1464–1480.

Faraj S, Jarvenpaa SL, Majchrzak A (2011) Knowledge collaboration in online communities. *Organ. Sci.* 22(5):1224–1239.

Faullant R, Füller J, Hutter K (2017) Fair play: Perceived fairness in crowdsourcing competitions and the customer relationship-related consequences. *Management Decisions* 55(9):1924–1941.

Felin T, Lakhani KR, Tushman ML (2017) Firms, crowds, and innovation. *Strategic Organ.* 15(2):119–140.

Fiske AP (1992) The four elementary forms of sociality: Framework for a unified theory of social relations. *Psych. Rev.* 99(4):689–723.

Fjeldstad ØD, Snow CC, Miles RE, Lettl C (2012) The architecture of collaboration. *Strategic Management J.* 33(6):734–750.

Foley M, Smead R, Forber P, Riedl C (2021) Avoiding the bullies: The resilience of cooperation among unequals. *PLOS Comput. Biol.* 17(4):e1008847.

Franke N, Shah S (2003) How communities support innovative activities: An exploration of assistance and sharing among end-users. *Res. Policy* 32(1):157–178.

Franke N, Keinz P, Klausberger K (2013) "Does this sound like a fair deal?": Antecedents and consequences of fairness expectations in the individual's decision to participate in firm innovation. *Organ. Sci.* 24(5):1495–1516.

Franke N, Schreier M, Kaiser U (2010) The "i designed it myself" effect in mass customization. *Management Sci.* 56(1):125–140.

Franken RE, Brown DJ (1995) Why do people like competition? The motivation for winning, putting forth effort, improving one's performance, performing well, being instrumental, and expressing forceful/aggressive behavior. *Personality Individual Differences* 19(2):175–184.

Fulker Z, Forber P, Smead R, Riedl C (2021) Spite is contagious in dynamic networks. *Nature Comm.* 12(1):1–9.

Gallus J, Reiff J, Kamenica E, Fiske AP (2022) Relational incentives theory. *Psych. Rev.* 129(3):586–602.

Gaure S (2013) lfe: Linear group fixed effects. *R J.* 5(2):104–116.

Gebauer J, Füller J, Pezzei R (2013) The dark and the bright side of co-creation: Triggers of member behavior in online innovation communities. *J. Bus. Res.* 66(9):1516–1527.

Gneezy U, Kajackaite A, Sobel J (2018) Lying aversion and the size of the lie. *Amer. Econom. Rev.* 108(2):419–453.

Gomez-Mejia LR, Balkin DB (1992) Determinants of faculty pay: An agency theory perspective. *Acad. Management J.* 35(5):921–955.

Grad T, Riedl C, Kilduff GJ (2023) When rivalry backfires: How individual skill and risk of status loss moderate the effects of rivalry on performance. Working paper, Northeastern University, Boston.

Greene WWH (2012) *Econometric Analysis*, 6th ed. (Pearson, New York).

Harbring C, Irlenbusch B (2011) Sabotage in tournaments: Evidence from a laboratory experiment. *Management Sci.* 57(4):611–627.

Harbring C, Irlenbusch B, Kräkel M, Selten R (2007) Sabotage in corporate contests–An experimental analysis. *Internat. J. Econom. Bus.* 14(3):367–392.

Harhoff D, Henkel J, Von Hippel E (2003) Profiting from voluntary information spillovers: How users benefit by freely revealing their innovations. *Res. Policy* 32(10):1753–1769.

Hippel Ev, Krogh Gv (2003) Open source software and the "private-collective" innovation model: Issues for organization science. *Organ. Sci.* 14(2):209–223.

Hofstetter R, Aryobsei S, Herrmann A (2018a) Should you really produce what consumers like online? Empirical evidence for reciprocal voting in open innovation contests. *J. Production Innovation Management* 35(2):209–229.

Hofstetter R, Zhang JZ, Herrmann A (2018b) Successive open innovation contests and incentives: Winner-take-all or multiple prizes? *J. Production Innovation Management* 35(4):492–517.

Huang Y, Vir Singh P, Srinivasan K (2014) Crowdsourcing new product ideas under consumer learning. *Management Sci.* 60(9):2138–2159.

Hutter K, Hautz J, Füller J, Mueller J, Matzler K (2011) Communitition: The tension between competition and collaboration in community-based design contests. *Creative Innovative Management* 20(1):3–21.

Ivaturi K, Chua C (2019) Framing norms in online communities. *Inform. Management* 56(1):15–27.

Jeppesen LB, Lakhani KR (2010) Marginality and problem-solving effectiveness in broadcast search. *Organ. Sci.* 21(5):1016–1033.

Kilduff GJ, Elfenbein HA, Staw BM (2010) The psychology of rivalry: A relationally dependent analysis of competition. *Acad. Management J.* 53(5):943–969.

Kim Y, Jarvenpaa SL, Gu B (2018) External bridging and internal bonding: Unlocking the generative resources of member time and attention spent in online communities. *Management Inform. Systems Quart.* 42(1):265–283.

Klapper H, Piezunka H, Dahlander L (2024) Peer evaluations: Evaluating and being evaluated. *Organ. Sci.* Forthcoming.

Knudsen T, Levinthal AD, Puranam P (2019) Editorial: A model is a model. *Strategy Sci.* 4(1):1–3.

Konrad K (2000) Sabotage in rent-seeking contests. *J. Law Econom. Organ.* 16(1):155–165.

Konrad KA (2009) *Strategy and Dynamics in Contests* (Oxford University Press, Oxford, UK).

Labianca G, Brass DJ (2006) Exploring the social ledger: Negative relationships and negative asymmetry in social networks in organizations. *Acad. Management Rev.* 31(3):596–614.

Lado AA, Boyd NG, Hanlon SC (1997) Competition, cooperation, and the search for economic rents: A syncretic model. *Acad. Management Rev.* 22(1):110–141.

Lakhani KR, Kanji Z (2008) *Threadless: The Business of Community* (Harvard Business School, Cambridge, MA).

Lakhani KR, Wolf RG (2003) Why hackers do what they do: Understanding motivation and effort in free/open source software projects. Technical report, Open Source Software Projects.

Lazear EP (1989) Pay equality and industrial politics. *J. Political Econom.* 97(3):561.

Lazear EP, Rosen S (1981) Rank-order tournaments as optimum labor contracts. *J. Political Econom.* 89(5):841–864.

Lewis MW (2000) Exploring paradox: Toward a more comprehensive guide. *Acad. Management Rev.* 25(4):760–776.

Li H, Bingham JB, Umphress EE (2007) Fairness from the top: Perceived procedural justice and collaborative problem solving in new product development. *Organ. Sci.* 18(2):200–216.

Liu TX, Yang J, Adamic LA, Chen Y (2014) Crowdsourcing with all-pay auctions: A field experiment on taskcn. *Management Sci.* 60(8):2020–2037.

Magee JC, Galinsky AD (2008) Social hierarchy: The self-reinforcing nature of power and status. *Acad. Management Ann.* 2(1):351–398.

Majchrzak A, Malhotra A (2013) Toward an information systems perspective and research agenda on crowdsourcing for innovation. *J. Strategic Inform. Systems* 22(4):257–268.

Majchrzak A, Malhotra A (2016) Effect of knowledge-sharing trajectories on innovative outcomes in temporary online crowds. *Inform. Systems Res.* 27(4):685–703.

Merritt AC, Effron DA, Monin B (2010) Moral self-licensing: When being good frees us to be bad. *Soc. Personality Psych. Compass* 4(5):344–357.

Moffitt R (2001) Policy interventions, low-level equilibria, and social interactions. Durlauf S, Young P, eds. *Social Dynamics* (MIT Press, Cambridge, MA).

Münster J (2007) Selection tournaments, sabotage, and participation. *J. Econom. Management Strategy* 16(4):943–970.

Nambisan S, Baron RA (2010) Different roles, different strokes: Organizing virtual customer environments to promote two types of customer contributions. *Organ. Sci.* 21(2):554–572.

Nickell J (2010) *Threadless: Ten Years of T-Shirts from the World's Most Inspiring Online Design Community* (Abrams Image, New York).

Perry-Smith JE (2006) Social yet creative: The role of social relationships in facilitating individual creativity. *Acad. Management J.* 49(1):85–101.

Piezunka H, Dahlander L (2019) Idea rejected, tie formed: Organizations' feedback on crowdsourced ideas. *Acad. Management J.* 62(2):503–530.

Ransbotham S, Kane GC (2011) Membership turnover and collaboration success in online communities: Explaining rises and falls from grace in Wikipedia. *Management Inform. Systems Quart.* 35(3):613–627.

Ren Y, Harper FM, Drenner S, Terveen L, Kiesler S, Riedl J, Kraut RE (2012) Building member attachment in online communities: Applying theories of group identity and interpersonal bonds. *Management Inform. Systems Quart.* 36(3):841–864.

Reuben E, Sapienza P, Zingales L (2015) Taste for competition and the gender gap among young business professionals. Technical report 21695, National Bureau of Economic Research, Cambridge, MA.

Riedl C, Seidel VP (2018) Learning from mixed signals in online innovation communities. *Organ. Sci.* 29(6):1010–1032.

Riedl C, Hutter K, Füller J, Tellis G (2024) Cash or non-cash? Unveiling ideators' incentive preferences in crowdsourcing contest. *J. Management Inform. Systems* Forthcoming.

Roberts JA, Hann IH, Slaughter SA (2006) Understanding the motivations, participation, and performance of open source software developers: A longitudinal study of the Apache projects. *Management Sci.* 52(7):984–999.

Ryan RM, Deci EL (2000) Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Amer. Psych.* 55(1):68.

Safadi H, Johnson SL, Faraj S (2021) Who contributes knowledge? Core-periphery tension in online innovation communities. *Organ. Sci.* 32(3):752–775.

Smirnova I, Reitzig M, Sorenson O (2022) Building status in an online community. *Organ. Sci.* 33(6):2519–2540.

Smith A (2010) *The Theory of Moral Sentiments* (Penguin, London).

Soda G, Mannucci PV, Burt RS (2021) Networks, creativity, and time: Staying creative through brokerage and network rejuvenation. *Acad. Management J.* 64(4):1164–1190.

Therneau TM, Grambsch PM (2000) *Modeling Survival Data: Extending the Cox Model* (Springer, New York).

To C, Kilduff GJ, Rosikiewicz BL (2020) When interpersonal competition helps and when it harms: An integration via challenge and threat. *Acad. Management Ann.* 14(2):908–934.

Tsai W (2002) Social structure of "coopetition" within a multiunit organization: Coordination, competition, and intraorganizational knowledge sharing. *Organ. Sci.* 13(2):179–190.

Tullock G (1980) Efficient rent seeking. Buchanan J, Tollison R, Tullock G, eds. *Toward a Theory of the Rent-Seeking Society* (Texas A&M University Press, College Station, TX).

Waldman DA, Putnam LL, Miron-Spektor E, Siegel D (2019) The role of paradox theory in decision making and management research. *Organ. Behav. Human Decision Processes* 155(November): 1–6.

Wasko MM, Faraj S (2000) “It is what one does”: Why people participate and help others in electronic communities of practice. *J. Strategic Inform. Systems* 9(2–3):155–173.

Wasko MM, Faraj S (2005) Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *Management Inform. Systems Quart.* 29(1):35–57.

Wasserman S, Faust K (1994) *Social Network Analysis: Methods and Applications* (Cambridge University Press, New York).

Zaggl MA, Malhotra A, Alexy O, Majchrzak A (2023) Governing crowdsourcing for unconstrained innovation problems. *Strategic Management J.* 44(11):2783–2817.

Christoph Riedl is professor of information systems at the D’Amore-McKim School of Business at Northeastern University. He obtained his PhD from Technische Universität München. His research focuses on collective intelligence, crowdsourcing, and collaboration in human–artificial intelligence teams.

Tom Grad is an assistant professor of strategy and innovation at Copenhagen Business School. He obtained his PhD from the Vienna University of Economics and Business (WU Vienna). His research examines behavior in digital environments, with a special focus on platform design, crowdsourcing, and competition.

Christopher Lettl is professor of entrepreneurship and innovation at the Vienna University of Economics and Business. He obtained his PhD from the Hamburg University of Technology (TUHH). His research focuses on open and user innovation, new organizational forms, and crowdsourcing.