Dec 12th, 12:00 AM

Heavy Medal - The Consequences of Introducing Symbolic Awards on Contribution Behavior in Online Communities

Alexander Staub
Institute for Strategy, Technology and Organization, alexander.staub@wu.ac.at

Tom Grad
Copenhagen Business School, tg.si@cbs.dk

Christopher Lettl
Institute for Strategy, Technology and Organization, Christopher.Lettl@wu.ac.at

Heavy Medal—The Consequences of Introducing Symbolic Awards on Contribution Behavior in Online Communities

Completed Research Paper

Alexander Staub
WU Vienna University of Economics and Business
alexander.staub@wu.ac.at

Tom Grad
Copenhagen Business School
tg.si@cbs.dk

Christopher Lettl
WU Vienna University of Economics and Business
christopher.lettl@wu.ac.at

Abstract

Online communities, like Wikipedia and Stack Overflow, have made a vast repository of knowledge available as a public good. However, they suffer from under-contribution in terms of quantity and quality. To tackle this issue, online communities have increasingly been relying on gamification, the use of game elements in non-game settings, to incentivize their members. The consequences of introducing such features on members' behavior has remained elusive—partly due to the lack of controlled experiments. Herein, we take advantage of a natural experiment in which a technical online community introduced gamified rewards, which are awarded contingent on performance thresholds—termed performance contingent symbolic awards. Employing a difference-in-differences design using a comparable online community as a control group, we find that the introduction of performance contingent symbolic awards has a negative impact on the contribution behavior overall, and that experienced members reduce their contribution quantity while inexperienced members reduce their contribution quality.

Keywords: Online community, gamification, symbolic awards, user behavior, natural experiment

Introduction

Online communities (OC) have developed into a vital repository of publicly accessible knowledge and new organizational forms in their own right (Faraj et al. 2011; Fjeldstad et al. 2012). However, the reliance on voluntary contributions of their members leads to issues common to public goods settings (Mindel et al. 2018). First, the distribution of contributions by members follows a power law (Lerner and Tirole 2002; van Mierlo 2014), meaning that around 90-95% of users consist of lurkers or rare contributors who profit from the vast amount of knowledge contributed by less than 5% of members. Second, not only the quantity of contributions but also the quality of contributions is a rising concern for OCs (Lou et al. 2013; Wang et al. 2022). Finally, the population in communities is heterogeneous in motives, preferences, and experience (Arazy et al. 2016; von Krogh et al. 2003; Wiertz and de Ruyter 2007). Fostering only experienced members can result in the risk of organizational instability due to a dwindling flow of new talent and an eventual decline if core members decide to depart (Oh and Jeon 2007), while focusing on inexperienced members can alienate experienced members that were responsible for the initial growth of an OC (Kraut et al. 2012, 2011). Hence OCs are faced with the challenge of how to foster active engagement of their heterogeneous members, both in terms of quantity and quality.

To support member engagement, most online communities employ some form of non-monetary incentives associated with gamification (Friedrich et al. 2020; Liu et al. 2017). These often come in the form of virtual badges, trophies, and medals, or in other words symbolic awards (Gallus 2017). In particular, performance contingent symbolic awards (PCSA), which set out contribution behavior thresholds for the reception of the incentives upfront, are a common occurrence. IS research has studied the use of various manifestations of symbolic awards as gamification elements on behavior pre-reception (Anderson et al. 2013; Chen et al. 2018) or post-reception (Wang et al. 2021b) and has repeatedly found a positive relationship between earning such rewards and OC member engagement. However, there is a lack of understanding of how the *introduction* of commonly employed symbolic awards that are tied to OC relevant performance criteria like contribution quantity and quality affects the incentivized behavior. This introduction of PCSA creates an external intervention into a reward system, given that such awards serve as both intangible rewards tied to externally-determined goals and reputation signals to others due to their public visibility (Bénabou and Tirole 2006; Deci et al. 1999). Within online communities, intrinsic and extrinsic rewards systems tend to emerge implicitly by a shared understanding of community members von Krogh et al. 2012. This understanding varies among OC members, where those who are experienced have a higher identification with OC norms and values compared to their inexperienced counterparts. While the extant literature has highlighted the relevance of members and their contribution heterogeneity (Arazy et al. 2016; David and Rullani 2008; Wiertz and de Ruyter 2007), how the introduction of PCSA interacts with OC member experience and associated contribution behavior motives remains unclear. We aim to address these gaps in the literature by answering the following question:

How does the introduction of performance contingent symbolic awards affect online community members' contribution behavior 1) in terms of quantity and 2) in terms of quality, 3) and does the effect depend on members' experience within the online community?

Herein, we take advantage of a natural experiment that occurred in 2015, when a technical Q&A OC focused on electrical engineering introduced a set of tiered PCSA aimed at both contribution quantity and quality. Employing the Stack Exchange Electrical Engineering Q&A OC as a comparison group, we construct a matched panel dataset of 88.000 individual-week observations one year around treatment. We then employ a difference-in-differences (DiD) estimation method to estimate the Average Treatment Effect on the Treated (ATT) of the PCSA introduction. We find that introducing PCSA leads to an overall reduction in the average member contribution quantity, while the quality of contributions does not change significantly on average. However, by taking the experience of members into account, we find that experienced members significantly reduce their contribution quantity while their contribution quality remains constant, and inexperienced members do not display any change in contribution quantity but significantly reduce their contribution quality.

This paper makes two important contributions. First, we contribute to the literature on gamification (Liu et al. 2017; Riar 2020) and symbolic awards (Gallus 2017). Prior literature has mostly focused on the effect of *receiving* gamified rewards on contribution quantity (Anderson et al. 2013; Goes et al. 2016; Wang et al. 2021b), with some research differentiating the impact on the quantity and quality (Chen et al. 2010b; Hsieh et al. 2010; Yu et al. 2022). We extend this literature by providing causal evidence for both the contribution quantity and quality undermining potential of *introducing* PCSA. Furthermore, while prior literature has underlined the relevance of member heterogeneity for various OC engagement dimensions (Arazy et al. 2016; von Krogh et al. 2003; Wiertz and de Ruyter 2007), we unpack how members who vary in their within-OC experience differ in their behavioral reaction toward gamified rewards. In particular, by providing insights on how the level of experience moderates members' contribution behavior upon the introduction of PCSA, we answer the call to consider member heterogeneity in relation to symbolic awards (Gallus 2017). Second, we build on the literature on new forms of organizing in general and OCs in particular (Faraj et al. 2011; Fjeldstad et al. 2012; Puranam et al. 2014). By investigating how gamification elements that are tied to specific performance dimensions of OC members affect their contribution behavior, we offer insights into the connection between two universal problems of organizing–namely the self-selected task behavior of members and the provision of rewards (Puranam et al. 2014).

Related Literature

Contribution Behavior in Online Communities

Online communities provide widely popular knowledge repositories, like Wikipedia, Stack Exchange, Reddit, and Quora, that are based on the coordination of and voluntary contributions by individuals. However, the non-hierarchical, voluntary, and self-organized nature of OCs leads to well-known collective action problems like free riding and having the contributions of members resemble a power-law distribution (van Mierlo 2014), which in turn can lead to an over-dependence on key individuals and even result in the unravelling of the organization (Oh and Jeon 2007). In addition, not only the number of contributions but also the quality of member contributions is important to the health of an OC. What constitutes whether contributions are high quality differs between kinds of OCs—e.g., contributions to open-source software (von Krogh et al. 2012) vs. to Wikipedia articles (Stvilia et al. 2008)–and can take on multiple dimensions within any OC. Regardless of its definition or operationalization, issues around contribution quality are of growing concern. This has been highlighted by research on the perceived quality of questions and answers (Lappas et al. 2017; Lou et al. 2013) and the proliferation of misleading information or harmful conduct (Bachschi et al. 2020; Bowen et al. 2021).

The motives for voluntarily contributing range from extrinsic motivators like gaining reputation and status within any given OC (Wasko and Faraj 2005) to intrinsic factors like the joy of problem-solving and the drive to help others, as well as a desire to learn new skills or improve existing ones (Lerner and Tirole 2002; von Krogh et al. 2012). While the contribution quantity of members is an objective, countable statistic, contribution quality is more subjective and depends in part on community norms (Stvilia et al. 2008) and can thus vary greatly (Wiertz and de Ruyter 2007). Beyond the motives related to contribution quantity, contribution quality has been linked to higher levels of skill within, and knowledge of, a particular domain, familiarity with the norms and values of a given OC, as well as a willingness to provide higher levels of effort (Cerasoli et al. 2014; Stvilia et al. 2008). The latter of which is the dimension of quality we focus on in this study.

The heterogeneity of members in an online community is a further relevant factor to take into account when considering the contribution behavior of members. Past studies have differentiated OC members by their social network position (Wasko et al. 2009), activity levels (Leclercq et al. 2020; Yeow et al. 2006), or whether they are project starters or project joiners (von Krogh et al. 2003), to name but a few. Beyond offering various categorizations of members, research incorporating member heterogeneity has shown that member types associated with higher levels of experience tend to be more intrinsically motivated (Khern-am-nuai et al. 2018) as well as exhibit a higher level of identification with the community (Stvilia et al. 2008). On the other hand, members that are comparatively inexperienced exhibit a low natural propensity for engagement, and lack the intrinsic motivation, time, confidence, and/or extrinsic incentives to increase their activity levels. Simultaneously, they do not (yet) possess the skill or knowledge of the OC norms needed to reliably produce high-quality contributions (Kraut et al. 2012, 2011; Stvilia et al. 2008).

This in turn implies that members differing in their experience may react differently to externally-imposed governance or steering mechanisms. Thus, the question arises as to how extrinsic, gamified rewards, which are commonly-found incentives within OCs (Liu et al. 2017), affect OC members of varying experience levels.

Gamification and Symbolic Awards in Online Communities

While there are some early studies of online communities employing remuneration for contributions (Chen et al. 2010b; Hsieh et al. 2010), OCs have increasingly turned to non-monetary forms of rewards, commonly referred to as gamification (Deterding et al. 2011; Liu et al. 2017). Gamification elements, defined by Deterding (2009, p.9) as ”the use of game design elements in non-game contexts”, can include anything from a custom social presence to points and leaderboards (Friedrich et al. 2020; Liu et al. 2017; Mekler et al. 2017). One commonly encountered manifestation of gamified rewards are virtual, publicly displayed badges, medals, trophies, etc., which are also referred to as symbolic awards (Gallus 2017; Wang et al. 2021b). Incentives in general, and symbolic awards in particular, can be categorized as either discretionary and contingent or confirmatory (Deci et al. 1999; Gallus and Frey 2017). Specifically, tying the award reception to performance criteria, like contribution thresholds, is particularly popular within organizational knowledge management systems (Friedrich et al. 2020) and OCs (Anderson et al. 2013; Burtch et al. 2021; Chen et al. 2018). For OC community organizers, the benefit of employing such performance contingent symbolic awards (PCSA) is that they are comparatively cheap to implement (Gallus 2017), easy to distribute (Anderson et al. 2013), and they transparently set out extrinsic goals for OC members (Goes et al. 2016).

Several studies have found a positive relationship between the reception of symbolic awards on contribution quantity across the board, or more specifically for new (Burtch et al. 2021) or less active members (Chen et al. 2018; Leclercq et al. 2020). Their impact on contribution quality is less clear, with some studies finding a positive effect (see, e.g. Harper et al. 2008), while others find inconclusive results (see, e.g. Chen et al. 2010b; Hsieh et al. 2010). However, investigating the behavior related to the reception of rewards within an established incentive system does not necessarily allow us to determine the consequences of introducing a novel, performance-based reward system that effectively consists of externally-imposed extrinsic incentives. It has been shown that introducing gamification elements in the form of contingent awards can have a positive effect on member transactions and comments (Hamari 2017). However, the research on this was conducted on a platform for economic exchange which entails different kinds of a priori motives and a different type of content produced than that of an OC focused on the provision of knowledge as a public good. Hence it is not clear if and how the findings generalize to a setting in which intrinsic motivations for contributions are arguably more prevalent (Fjeldstad et al. 2012).

Despite the sizeable and growing body of literature on gamified rewards in OCs in general (see, e.g. Friedrich et al. 2020; Liu et al. 2017; Riar 2020, for reviews), and symbolic awards in particular, there are still important gaps in our understanding of how and why PCSA affect contribution behavior in OCs. It is not clear how the existing member population reacts to the imposition of extrinsic goals set forth by the introduction of PCSA–especially when such goals mesh quantity and quality objectives, two contribution behavior dimensions that are related but vary in knowledge and motivation requirements for members. Furthermore, we have little understanding if and how experience, acquired via engagement in an OC before the introduction of such PCSA, determines the behavioral reactions towards gamified rewards tied to extrinsically determined performance thresholds.

Hypotheses

In line with the economic literature on incentives (Lazear 1986) and behavioral theories (Eisenberg et al. 1996), rewards, both monetary and non-monetary, that are contingent on a recipient’s performance have been found to positively affect incentivized behavior. However, according to Self-Determination Theory (Deci and Ryan 1985) or Crowding Out Theory (Frey and Jegen 2001), rewards, especially contingent ones, have also been found to have performance-deteriorating effects by undermining intrinsic and/or prosocial motivation. These adverse effects were found to be particularly pronounced for activities that are intrinsically and prosocially rewarding, both of which are important drivers for OC contributions (von Krogh et al. 2012; Zhao et al. 2016). What’s more, the crowding out of intrinsic motivation tends to be particularly pronounced in cases where the award reception allows for the comparison among recipients and non-recipients, a factor that is highly salient for PCSA (Deci et al. 1999). In a similar vein, signalling theory suggests that providing public rewards can call into question the reasons for contributing to OCs by a relevant audience, thus decreasing the pro-social signalling value of contributing for members (Bénabou and Tirole 2006). This argument leads to the following hypothesis about the introduction of PCSA on OC member contributions:

H1: The introduction of PCSA will negatively affect online community members' contribution quantity.

Compared to the number of contributions, the relationship of extrinsic incentives to the quality of contributions is more ambiguous considering task performance in terms of quality in general (Byron and Khazanchi 2012) and contributions to OCs specifically (Chen et al. 2010b; Hsieh et al. 2010). In terms of antecedents, it has been shown that individual ability and effort driven by intrinsic motivation are closely related to performance in terms of quality on tasks in general (Cerasoli et al. 2014). This relationship is likely to hold within OCs, where contributions to a public good are made voluntarily and where superior quality requires increased effort (Chen et al. 2010b), a higher level of skill (Khern-am-nuai et al. 2018), expertise in the relevant topic(s) (Lou et al. 2013), familiarity with standards of excellence within an OC (Stvilia et al. 2008), or a combination of the above. The introduction of PCSA is unlikely to affect the majority of the aforementioned factors directly. However, the effort of individuals, which in absence of salient extrinsic incentives will have been driven to a degree by prosocial or intrinsic motivation, may be influenced negatively via a motivation-undermining effect of the PCSA introduction (Byron and Khazanchi 2012; Cerasoli et al. 2014). Thus, we would expect a similar outcome as in the case of contribution quantity:

H2: The introduction of PCSA will negatively affect online community members' contribution quality.

Across most OCs, members' propensity to be active within communities is closely related to differences in individual-level characteristics like self-efficacy (Lappas et al. 2017; Zhao et al. 2016), sources of motivation (Sun et al. 2012), and identification with a community's norms and values (Gallus 2017; Wiertz and de Ruyter 2007). While the above traits are hard to assess reliably, members' experience expressed as their accumulated record of activity within a particular community is indicative of their underlying characteristics (David and Rullani 2008; Kokkodis et al. 2020). More specifically, members that have acquired experience through prior knowledge provision within an OC tend to have done so out of own-use needs and joy of learning (Lerner and Tirole 2002). Inexperienced members can include newer and inactive members or lurkers (Chen et al. 2018; Kokkodis et al. 2020), all of whom are more opportunistically driven or extrinsically motivated to become active compared to experienced members (Leclercq et al. 2020; Zhao et al. 2016). Thus, given that experienced members are more prosocially and/or intrinsically motivated, and taking into account the potentially undermining effect of PCSA, the mere existence of such rewards may crowd out these members' intrinsic motivations to contribute to an OC.

In addition, the awarding criteria can be an important factor in determining how PCSA affect members of different experience levels. Discrete forms of rewards, which PCSA are an example of, require the definition of a threshold. This threshold is often defined as a compromise in light of heterogeneous ability profiles, which leads to inefficiencies in the chosen effort of individuals (Lazear 2000). Within OCs, this translates into a situation where the goals set out by the PCSA are set below the natural propensity to contribute of experienced individuals, leading to a so-called "done enough" effect (Gallus 2017) or a lack of perceived challenge (Leclercq et al. 2020). Conversely, members that exhibit contribution behavior far below the threshold may view PCSA as unattainable or irrelevant. Thus, according to Goal Setting Theory, while not necessarily leading to lower levels of contribution, inexperienced members may not be effectively incentivized to do more (Wang et al. 2021b). Hence, we posit:

H3: The introduction of PCSA will have a stronger negative effect on the contribution quantity of experienced members compared to inexperienced members.

Skill, knowledge, and familiarity with an OC context, all of which are relevant determinants of contribution quality as elaborated upon above, are developed over time with increasing engagement within a community

(Wiertz and de Ruyter 2007; Yeow et al. 2006). Thus, experienced OC members tend to be responsible for a higher average contribution quality. On the one hand, the increased effort necessary to produce contributions of high quality by experienced members as an expression of either intrinsic or prosocial motivation may be undermined by the PCSA introduction, in line with the argumentation for hypothesis 1. On the other hand, experienced users driven by the extrinsic motive of seeking reputation, which has been found to foster high-quality contributions (Stvilia et al. 2008), may be averse to risking reputation loss within OC contexts by reducing their contribution quality (Lappas et al. 2017). Moreover, it has been found that when OC organizers impose codified quality norms, longstanding members were less willing to adapt and tended to rely on heuristics and pragmatic decision processes in their contribution style (Danescu-Niculescu-Mizil et al. 2013; Stvilia et al. 2008).

Inexperienced members that do become or remain active upon PCSA introduction will not be as familiar with implicit quality standards within a given OC and are most likely not on the same skill and knowledge level as experienced members. In the absence of PCSA, even a small number of inexperienced members may engage with OC specific norms and values due to intrinsic, prosocial, or extrinsic reputation-seeking motives, thereby gradually improving the group's average contribution quality. However, the introduction of PCSA defines codified and transparent goals, even though these goals may be hard to reach. Thus, instead of gaining reputation via steadily improving one's contribution quality through learning, inexperienced users may opt for strategic, low-effort behavior to game the incentive system (Liu et al. 2017). What adds to the allure of engaging in opportunistic, low-effort contribution behavior is the lack of status or reputation to lose (Lappas et al. 2017), which in turn can lead to less thoughtful actions by these users as compared to their experienced counterparts (Kraut et al. 2012, 2011). In conclusion, we posit:

H4: The introduction of PCSA will have a stronger negative effect on the contribution quality of inexperienced members compared to experienced members.

Methods

Empirical Setting

To investigate our research question, we exploit a natural experiment in which a set of PCSA was introduced in April of 2015 to a technical Q&A OC focused on sharing knowledge on the topic of electrical engineering. It is among the oldest and largest OCs in the field, founded in 2009 and counting approximately 800.000 registered members at the time of data acquisition. While the site offers various means of knowledge sharing, we focus on the self-contained, Q&A-like discussion threads, which make up the bulk of the OC's activity (60% of all user-generated content).

The treatment is the introduction of a collection of awards made available to any registered user on the day of treatment. They signify the provision of publicly communicated achievement thresholds in the form of externally defined numbers of *correct* and/or *helpful* answers, as determined by other community members. Four sets of "expert" badges were introduced: One accrediting correct answers made to any discussion, and three specific to the topics of 3D printing focusing on Arduino and Raspberry Pi. The fifth badge relates to the provision of helpful answers, which is awarded upon receiving a specific amount of "helpful" marks by peers. All sets of badges have three tiers, with each having a higher threshold for reception in terms of the number of correct or helpful answers. Hence, they are contingent on both quantity and quality of contributions.

We use data from the Q&A OC Stack Exchange, specifically the Stack Exchange sub-community *Electrical Engineering*, as an off-platform control group. This group has been used frequently in IS research on user behavior in OCs (see, e.g. Anderson et al. 2013; Chen et al. 2018). Beyond the fact that the topical focus of both OCs is comparable, we were able to affirm their similarity by comparing the top 20 user-submitted tags (Stack Exchange) and sub-group names (treated OC). Both communities were created in 2009 and had a similar reputation system in place before treatment introduction. In addition, both communities had mechanisms in place for rating questions and answers, as well as selecting correct answers. Importantly for this research, the StackExchange OC did not experience any change to their incentive system 6 months before or after the treatment of relevance.

The data on the treatment community was acquired from publicly-available data using web scraping in September 2019. Data was gathered on the post and user level, ranging from two years before the treatment introduction to two years after. The data from StackExchange is publicly available and includes information on both the user and post level. The sample was reduced to users who were registered before the treatment introduction and were active at least once one year around the treatment date. This resulted in a sample of 931 users in the treated OC and 2536 in the control OC. Finally, as we are interested in the change in average user behavior, we created a balanced user-week panel for 48 weeks, resulting in a panel dataset of 166,000 observations.

Main Variables and Descriptive Statistics

We measure the quantity of contributions in terms of the weekly number of replies to questions posed by a user, which is a commonly-used operationalization of knowledge contributions to OCs (see, e.g. Chen et al. 2018; Li et al. 2012). Concerning quality, the operationalization is more challenging given that it is inherently hard to measure and context dependent (Stvilia et al. 2008). We employ the number of words per answer per week as a proxy for contribution quality, which was found to be an accurate predictor of contribution quality by multiple studies (Blumenstock 2008; Harper et al. 2008; Khern-am-nuai et al. 2018). This approach is supported in our data by finding a significant and positive correlation between word count and likes/score in both OCs. In addition, we are interested in heterogeneity across community members. In line with prior research (see, e.g. Gallus 2017; Khern-am-nuai et al. 2018; Wiertz and de Ruyter 2007), we use a dichotomous categorization of members according to their number of contributions before the window of observation. This moderator variable, termed experience, is created via a proportional split at the threshold of one contribution prior to the window of observation, leading to a 69:31 (71:29) split of inexperienced and experienced members in the treated (control) community. Changing the contribution cut-offs for experienced vs. inexperienced members to 0 or 2 instead of 1 does not change the significance or sign of our results. Table 1 shows descriptive statistics for the two communities based on a matched sample (the details of which are given below) and gives an overview of the main dependent variables of interest. For each group, summary statistics are provided for the moderator of interest (experience as a binary category).

	Control Community		**Treated Community**	
	inexperienced (N = 31,920)	**experienced** (N = 12,432)	**inexperienced** (N = 30,576)	**experienced** (N = 13,776)
answers per week				
Mean (SD)	0.06 (0.47)	0.19 (1.11)	0.06 (0.47)	0.28 (1.53)
Median [Min, Max]	0 [0, 22]	0 [0, 31]	0 [0, 23]	0 [0, 45]
words per asw per week				
Mean (SD)	98.20 (111)	109 (86.10)	71.70 (110)	71.00 (64.40)
Median [Min, Max]	69 [1, 1520]	84.50 [3, 674]	46 [1, 2250]	53 [2, 495]
likes/score per asw per week				
Mean (SD)	1.41 (2.32)	2.07 (3.23)	0.17 (0.39)	0.37 (0.64)
Median [Min, Max]	1 [-5, 22]	1 [-3, 54]	0 [0, 3]	0 [0, 6]

Table 1. Descriptive statistics based on sample selection conducted via 1:1 Nearest Neighbour PSM.

Identification and Estimation

We employ a difference-in-differences (DiD) design to identify the effects of PCSA introduction on OC member contribution behavior. The DiD estimator allows us to estimate the average treatment effect on the treated (ATT) and is commonly employed for causal inference using observational data (see, e.g. Barbosu and Gans 2022; Huang et al. 2017). We provide evidence that in our setup the parallel pre-trends assumption is not violated by running a dynamic DiD specification (see, e.g. Barbosu and Gans 2022; Huang et al. 2017), the results of which can be found in Figure A.1. in the appendix. To lend credibility that the stable unit treatment value assumption (SUTVA) holds in our setting, we perform username matching between the two communities and find only one user name that was active in both communities. Finally, we rely on the information provided to us by the community manager in charge of incentive design and implementation in a call in December 2019 to rule out any potential anticipatory effects. In addition to asserting that the assumptions necessary for DiD estimation are met, we implement propensity score matching (PSM) to address concerns around selection into treatment (Angrist and Pischke 2009). We adjust the sample for our main analysis based on 1:1 Nearest Neighbour PSM, in addition to conducting 1:3 Nearest Neighbour PSM with replacement for robustness purposes (Dehejia and Wahba 2002).

Our main estimating equations for the ATT across all users are shown below in equations 1 and 2. For both quantity and quality, we use panel data on the user-week level 24 weeks before and 24 weeks after treatment. i indexes individuals, g indexes the OC, and t indexes weeks. The $group$ variable represents a dummy variable indicating the treated OC (1) or the control OC (0). The $treatment$ variable takes the value 1 for the period after treatment introduction and 0 for the period before. Thus, the DiD coefficient of interest is represented by δ, capturing the effect of the PCSA introduction in the treated OC compared to a counterfactual scenario. We include individual-level fixed effects (α_i) to control for any time-invariant unobservable characteristics of OC members. Time (week) fixed effects (γ_t) are included to capture any events that might affect both groups. Estimators for the $treatment$ and $group$ variables are dropped, as they are perfectly co-linear with the fixed effects. β captures potential confounding effects due to differences in the supply of questions.

$$average\ answers_{igt} = \Phi(exp(\delta \times (group_g \times treatment_t) + \beta \times questions_{gt} + \alpha_i + \gamma_t + \epsilon_{igt})) \tag{1}$$

, where Φ is the Poisson distribution

$$\log(average\ words_{igt}) = \delta \times (group_g \times treatment_t) + \alpha_i + \gamma_t + \epsilon_{igt} \tag{2}$$

For the estimation of treatment effects on quantity, both in the estimation of average effects across all members (H1) and moderated by experience (H3), we employ a balanced panel, an observation taking the value 0 when a user did not provide any answers. As the dependent variable of interest is distributed as a count variable, in addition to being 0-inflated, we employ a Fixed-Effect Poisson regression estimation as our main specification, as well as providing results from an OLS estimation with fixed effects (Silva and Tenreyro 2011; Wooldridge 1999). To estimate the ATT for contribution quality, we only include user-week level observations conditional on contributing. Here, the main model is estimated using OLS regression with a logged DV. For the moderator analysis, we present the results in the form of separate, category-wise estimations for members of both the experienced and inexperienced groups according to Holgersson et al. 2014 and as employed in other studies (Barbosu and Gans 2022; Wang et al. 2021a). We refrain from using a continuous moderator variable, as this would provide biased or imprecise results in the presence of outliers (DeCoster et al. 2009), which occur naturally when studying online community member behavior (Anderson et al. 2013; Chen et al. 2018) and thus are not removed.

Results

In the following section, we will present the results of our DiD estimation to answer the question of how the introduction of PCSA affects OC member contribution behavior. Each analysis is based on user-week level panel data 24 weeks before and after the treatment introduction (approximately 6 months before and after the PCSA introduction). The DiD coefficient of relevance (δ) is represented by the interaction term, and we present results with and without user-week fixed effects.

Contribution Quantity

First, we look at how the introduction of PCSAs affects the number of answers contributed by OC members in the treated OC. In models 1 to 4 of Table 2, we see that the introduction of PCSAs does lead to a significant reduction of the contribution quantity. Our main model (Model 4) estimates an ATT of $\delta = -0.53$ ($p < 0.01$) via Fixed-Effects Poisson regression, which translates into a 41% reduction in the average weekly answering behavior. Using time and individual fixed effects allows us to control for time-invariant individual and time-specific unobservables. However, it leads to the loss of observations for individuals who were only active in one period. Hence, we also estimate the results without fixed effects (*see models 1 and 3*), which support our conclusions. In summary, the results from Table 2 provide support for our Hypothesis 1, which states that we expect a reduction in contribution behavior after the introduction of PCSAs.

	Main Effects				Moderator: Experience			
	OLS		Poisson					
DV = Quantity	1	2	3	4	5 inexp	6 exp	7 inexp	8 exp
Intercept	-0.14^{***}		-4.40^{***}		-4.94^{***}	-3.68^{***}		
	(0.04)		(0.14)		(0.23)	(0.19)		
Group	0.28^{***}		2.31^{***}		2.03^{***}	2.41^{***}		
	(0.03)		(0.12)		(0.19)	(0.15)		
Treated	-0.04^{***}		-0.43^{***}		-0.54^{***}	-0.36^{***}		
	(0.01)		(0.03)		(0.05)	(0.04)		
Q's per w.	0.00^{***}	0.00^{**}	0.01^{***}	0.00^{**}	0.01^{***}	0.01^{***}	0.01	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Group:Treated	-0.06^{***}	-0.06^{***}	-0.48^{***}	-0.53^{**}	-0.29^{***}	-0.60^{***}	-0.38	-0.68^{**}
	(0.01)	(0.02)	(0.05)	(0.19)	(0.08)	(0.06)	(0.32)	(0.26)
R^2	0.00	0.32						
Adj. R^2	0.00	0.31						
AIC			76679	49102	30640	41249	27261	21714
Observations	88704	88704	88704	88704	62496	26208	62496	26208
Week FE	No	Yes	No	Yes	No	No	Yes	Yes
Indiv. FE	No	Yes	No	Yes	No	No	Yes	Yes

Notes: SEs clustered at individual level for fixed effects models $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table 2. Results of DiD estimation - contribution quantity.

Regarding the effect of PCSA introduction conditional on the experience level of members registered before treatment, we estimate models via a proportional sample split (see models 5 to 8 in Table 2). In Model 8, within-individual estimated effects suggest that the average contribution volume of experienced members is significantly reduced ($\delta = -0.68; p < 0.01$). In terms of effect sizes, both in the model estimated with and without fixed effects, the decrease in contribution quantity for the experienced users is significant at the 0.01 level and of high magnitude (*Model* 6: $\delta = -0.60$, *Model* 8 $\delta = -0.68$). This translates into a decrease in the average number of answers per week in the experienced group of 49% according to our preferred specification (models 7 and 8). For the inexperienced users, the estimated change is not statistically different to 0 in Model 7 with fixed effects, and the estimated change is below that of experienced users in Model 5. Hence, we find evidence in support of Hypothesis 3, stating that we expect the decrease in contribution behavior to be higher for experienced than for inexperienced members.

Contribution Quality

To investigate the impact of treatment on the quality of contributions, we start again with the estimation of the ATT of PCSA introduction on all members of the treated community. We can see that the DiD coefficients of the OLS models in Table 3 do estimate a significant negative effect of treatment introduction on average weekly contribution quality, in terms of the length of answers (*Model* 1: δ = -18.62; *Model* 2: δ = -14.06). However, standard errors based on the models with an untransformed dependent variable may be estimated too liberally given the presence of several significant outliers and the highly right-skewed distribution of the dependent variable. The results from our main specification, the OLS model with a logged DV (Models 3 and 4), provide statistically insignificant estimates for our DiD coefficients (*Model* 3: δ = -0.10; *Model* 4: δ = -0.12). Hence, the results of models 1 to 4 in Table 3 do not offer strong support in favor of our Hypothesis 2, stating that the introduction of PCSA leads to an overall reduction in contribution quality.

	Main Effects				Moderator: Experience			
	OLS		OLS log(DV)					
DV = Quality	1	2	3	4	5 inexp	6 exp	7 inexp	8 exp
Intercept	90.96***		4.22***		4.07***	4.41***		
	(2.78)		(0.03)		(0.04)	(0.04)		
Group	−22.03***		−0.44***		−0.38***	−0.55***		
	(3.68)		(0.04)		(0.05)	(0.05)		
Treated	28.11***		0.23***		0.37***	0.05		
	(4.20)		(0.04)		(0.06)	(0.06)		
Group:Treated	−18.62**	−14.06*	−0.10	−0.12	−0.21*	0.02	−0.54***	−0.07
	(6.40)	(6.34)	(0.06)	(0.07)	(0.10)	(0.08)	(0.14)	(0.08)
R^2	0.04	0.72	0.08	0.72	0.08	0.10	0.85	0.59
Adj. R^2	0.04	0.47	0.08	0.47	0.08	0.10	0.50	0.43
Observations	4017	4017	4017	4017	1940	2077	1940	2077
Week FE	No	Yes	No	Yes	No	No	Yes	Yes
Indiv. FE	No	Yes	No	Yes	No	No	Yes	Yes

Notes: SEs clustered at individual level for fixed effects models $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table 3. Results of DiD estimation - contribution quality.

Models 5 to 8 in Table 3 show the results of estimating the ATT of PCSA introduction on members' contribution quality conditional on their experience. Compared to the results on the quantity of contributions, we notice an opposite dynamic when quality is the dependent variable of interest. In our main specification (Model 7), we estimate a large and significant drop in contribution quality of 54% ($\delta = -0.54$, $p < 0.001$) for inexperienced members. We also show a model without fixed effects— i.e., the between-estimator—in Model 5 ($\delta = -0.21$; $p < 0.05$) to account for the loss in observations when including individual fixed effects. The effect is of the same sign as the within-estimator for inexperienced members, albeit with a lower effect size and statistical significance level. The change in average contribution quality of experienced members is estimated to not be significantly different to 0 in either model, with a point estimate that is meaningfully below the estimate for inexperienced members. This provides support for our Hypothesis 4, which states that inexperienced members are expected to decrease their contribution quality more than experienced members upon PCSA introduction.

Robustness Checks

To support our findings, we conducted a number of robustness checks. First, we employ user-submitted signals as an alternative operationalization to measure changes in average contribution quality. These take on the form of "likes" in the treated community and "score" in the control community. Results when using the peer signals as an alternative specification for quality lead to the same conclusions as with our main specification, as can be seen in models 1 to 3 of Table 4. Second, there might be concerns that our results are driven by how we operationalize member experience as a moderator. As can be seen in models 4 to 7 in Table 4, our main results are confirmed when using a median split of the days a member was registered as a community member up to the PCSA introduction.

Third, we investigate whether our choice of matching procedure affects our results. In particular, we run our models using a sample based on 1:3 nearest neighbour matching with replacement. The results of the DiD estimations with the alternative matching procedure echo the results of the models based on the 1:1 matching procedure in both sign and significance levels and can be found in Table A.1. in the appendix. Finally, our

	likes/score as quality			days registered as moderator			
	Poisson			(quant) Poisson		(qual) log(DV)	
Models:	1	2	3	4	5	6	7
Member types:	all	inexp	exp	inexp	exp	inexp	exp
Avg. q. views	0.00^{***}	0.00^{**}	0.00^{**}				
	(0.00)	(0.00)	(0.00)				
Q's per w.				0.00	0.01^{*}		
				(0.00)	(0.00)		
Group:Treated	0.19	-0.75^{*}	0.29	0.13	-0.86^{***}	-0.48^{***}	-0.07
	(0.15)	(0.36)	(0.16)	(0.33)	(0.22)	(0.13)	(0.08)
R^2						0.85	0.66
Adj. R^2						0.54	0.45
AIC	9072.41	3874.42	5205.83	20413.32	28294.70		
Observations	2810	1075	1735	44400	44304	1433	2584
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Indiv. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: SEs clustered at individual level for fixed effects models $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table 4. Robustness checks for contribution quality and experience.

results may not be based on the introduction of PCSAs, but due to a combination of the time period we chose and different contribution dynamics within the OCs. Therefore, we ran placebo checks estimating DiD coefficients for any change in contribution behavior with an artificial treatment date 1 year prior to PCSA introduction, shown in Table A.2. in the appendix. Both for quantity and quality, we find no indication of our main results being driven by our sample selection procedure when using our original matched sample or when using an unmatched sample based on users registered before the artificial treatment date.

Discussion and Conclusion

In this study, we investigate the effects of introducing PCSA on the contribution behavior of OC members, contingent on whether they are experienced or inexperienced. We find that the contribution quantity of OC members is significantly reduced after PCSA introduction, while contribution quality, on average, is not significantly reduced. By segmenting OC members by their experience, the study reveals that after the introduction of PCSA, experienced members significantly reduce their contribution quantity, while their inexperienced counterparts do not. We observe the opposite effect with regards to quality. Here, experienced members retain their average contribution quality after PCSA introduction, while inexperienced members reduce their average quality of contributions. In addition, when taking results from the event study into account, as seen in Figure A1 in the appendix, it can be seen that these effects only become significant 3 months after the introduction of the awards. This points towards a learning effect, in that it takes some time before members become acquainted with the new incentives. Taken together, we show that the introduction of PCSA harms the OC member contribution behavior in terms of both quantity and quality. However, heterogeneity in member experience within the community determines which dimension of contribution behavior is compromised.

With this paper, we contribute to the literature on gamification (Liu et al. 2017; Riar 2020) and symbolic awards as a specific manifestation of gamified reward elements (Friedrich et al. 2020; Gallus 2017). In particular, we highlight the negative effects of gamification features on both quantity and quality of contributions when they are tied to performance thresholds. While past research has pointed towards the potential positive effects of introducing PCSAs (Hamari 2017), we show that such extrinsic reward elements, as defined by their externally-determined goals and public visibility (Deci et al. 1999; Deci and Ryan 1985; Frey and Jegen 2001), can in fact affect behavior in ways that suggest a motivation undermining effect. This is in line with studies that hint at the potential negative effects on online community engagement of rewards that set out criteria for reception transparently upfront (Leclercq et al. 2020) vs. rewards that are unexpected and discretionary (Gallus 2017). In addition, we offer more nuanced insights by unpacking the relevance of experience on the relevant behavior dimensions, as suggested by Gallus (2017). Finally, we extend prior research on new organizational forms by investigating the role of gamification elements in the form of extrinsic incentives on contribution behavior. By studying the introduction of PCSA into an OC, we provide insight into the interplay of two universal problem-solving approaches of new forms of organizing–namely, how extrinsic reward provision affects self-selected task behavior (Puranam et al. 2014).

Our findings also have important implications for practitioners with regard to gamification feature design for OCs. We highlight the importance of considering which kinds of members as well as which performance dimensions are targeted when deciding to introduce gamified rewards in general and symbolic awards in particular. Among other things, we identify an effect similar to traditional discrete performance contingent monetary rewards, which were shown to be inefficient in light of heterogeneity in organizational member ability (Lazear 2000), or when highly active reviewers are informed about their relative performance compared to their peers (Chen et al. 2010a). Thus, our findings offer empirical support for recommendations that the reward elements within online communities should be personalized to heterogeneous member needs and abilities (Leclercq et al. 2020; Ping 2008) . Furthermore, we find that the quality of a contribution is a dimension that may be hard to encourage with gamified rewards that target extrinsic motivation, but is potentially easy to hamper. This is in line with research that finds that performance in terms of quality is related more strongly to interest and intrinsic motivation (Cerasoli et al. 2014) and learned community standards of quality and excellence (Stvilia et al. 2008) rather than the prospect of receiving extrinsic rewards.

Our study has some limitations of note, which in turn offer interesting avenues for future research. First, we are not able to assess to perfect certainty whether members truly noticed the PCSA introduction. However, based on extensive background research, anecdotal evidence from the community manager, pre-trend analyses, and robustness checks, we are confident that the effects we measure are not caused by an alternative shock. Nonetheless, future research that investigates the impact of the introduction of PCSA via randomized field experiments may be able to incorporate measures of member attention to the intervention. Second, we are not able to directly assess the motivations of individuals, and thus the mechanism underlying our hypothesized effect of crowding out. Hence, assessing the motivation of OC members, specifically intrinsic versus extrinsic dimensions, would be of value in future studies. Third, we opt for a very specific operationalization of both quality and experience that may not reflect the full depth and breadth of information quality (Stvilia et al. 2008) or the heterogeneity of OC members (Arazy et al. 2016; von Krogh et al. 2003; Wasko et al. 2009) respectively. Our robustness checks for our quality as well as experience variables give us confidence that our results are not spurious. However, future research is well advised to investigate the impact of awards on other relevant qualitative factors like the perceived trustworthiness of the information. Finally, we investigate both a very specific kind of gamification element, PCSA, in a particular kind of OC, a technical QA OC. Therefore, one must be careful in generalizing these results to the impact of gamification in general, or applying them to all forms of OCs and knowledge-sharing platforms. Indeed, prior studies pointing out both positive (see, e.g. Hamari 2017) and negative effects of gamified rewards (see, e.g. Leclercq et al. 2020) highlight the relevance of heterogeneity in both, context and gamification features (Friedrich et al. 2020). Hence, future research may want to compare the effect of PCSA in different kinds of communities or investigate the impact of various gamified reward elements within a technical QA OC.

References

Anderson, A., Huttenlocher, D., Kleinberg, J., and Leskovec, J. 2013. "Steering User Behavior with Badges∗: May 13 - 17, 2013, Rio de Janeiro, Brazil". In: *Proceedings of the 22nd International Conference on the World Wide Web*.

Angrist, J. D. and Pischke, J.-S. 2009. *Mostly harmless econometrics: An empiricist's companion / Joshua D. Angrist & Jörn-Steffen Pischke*. Princeton, N.J. and Oxford: Princeton University Press.

Arazy, O., Daxenberger, J., Lifshitz-Assaf, H., Nov, O., and Gurevych, I. 2016. "Turbulent Stability of Emergent Roles: The Dualistic Nature of Self-Organizing Knowledge Coproduction". *Information Systems Research* (27:4), pp. 792–812.

Bachschi, T., Hannak, A., Lemmerich, F., and Wachs, J. 2020. "From Asking to Answering: Getting More Involved on Stack Overflow". *CoRR* (abs/2010.04025). arXiv: 2010.04025.

Barbosu, S. and Gans, J. S. 2022. "Storm crowds: Evidence from Zooniverse on crowd contribution design". *Research Policy* (51:1), p. 104414.

Bénabou, R. and Tirole, J. 2006. "Incentives and Prosocial Behavior". *American Economic Review* (96:5), pp. 1652–1678.

Blumenstock, J. E. 2008. "Size matters: word count as a measure of quality on wikipedia". In: *Proceedings of the 17th International Conference on World Wide Web, April 21 - 25, 2008, Beijing, China*. Ed. by J. Huai, R. Chen, H.-W. Hon, Y. Liu, W.-Y. Ma, A. Tomkins, and X. Zhang. New York: ACM, pp. 1095–1096.

Bowen, R., Dmitrev, D., and Galperti, S. 2021. "Learning from Shared News: When Abundant Information Leads to Belief Polarization". *NBER Working Paper* (w28465).

Burtch, G., He, Q., Hong, Y., and Lee, D. 2021. "How Do Peer Awards Motivate Creative Content? Experimental Evidence from Reddit". *Management Science* (forthcoming).

Byron, K. and Khazanchi, S. 2012. "Rewards and creative performance: a meta-analytic test of theoretically derived hypotheses." *Psychological bulletin* (138:4), p. 809.

Cerasoli, C. P., Nicklin, J. M., and Ford, M. T. 2014. "Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis". *Psychological Bulletin* (140:4), pp. 980–1008.

Chen, W., Wei, X., and Xiaoguo Zhu, K. 2018. "Engaging Voluntary Contributions in Online Communities: A Hidden Markov Model". *MIS Quarterly* (42:1), 83-A8–A8.

Chen, Y., Harper, F. M., Konstan, J., and Li, S. X. 2010a. "Social comparisons and contributions to online communities: A field experiment on movielens". *American Economic Review* (100:4), pp. 1358–98.

Chen, Y., Ho, T.-H., and Kim, Y.-M. 2010b. "Knowledge Market Design: A Field Experiment at Google Answers". *Journal of Public Economic Theory* (12:4), pp. 641–664.

Danescu-Niculescu-Mizil, C., West, R., Jurafsky, D., Leskovec, J., and Potts, C. 2013. "No country for old members: User lifecycle and linguistic change in online communities". In: *Proceedings of the 22nd international conference on World Wide Web*, pp. 307–318.

David, P. A. and Rullani, F. 2008. "Dynamics of innovation in an "open source" collaboration environment: lurking, laboring, and launching FLOSS projects on SourceForge". *Industrial and Corporate Change* (17:4), pp. 647–710.

Deci, E. L., Koestner, R., and Ryan, R. M. 1999. "A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation". *Psychological Bulletin* (125:6), pp. 627–668.

Deci, E. L. and Ryan, R. M. 1985. *Intrinsic motivation and self-determination in human behavior*. New York: Springer Science+Business Media.

DeCoster, J., Iselin, A.-M. R., and Gallucci, M. 2009. "A conceptual and empirical examination of justifications for dichotomization." *Psychological methods* (14:4), p. 349.

Dehejia, R. H. and Wahba, S. 2002. "Propensity Score-Matching Methods for Nonexperimental Causal Studies". *The Review of Economics and Statistics* (84:1), pp. 151–161.

Deterding, S., Dixon, D., Khaled, R., and Nacke, L. 2011. "From game design elements to gamefulness: defining" gamification"". In: *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*, pp. 9–15.

Eisenberg, N., Fabes, R. A., Karbon, M., Murphy, B. C., Wosinski, M., Polazzi, L., Carlo, G., and Juhnke, C. 1996. "The relations of children's dispositional prosocial behavior to emotionality, regulation, and social functioning". *Child development* (67:3), pp. 974–992.

Faraj, S., Jarvenpaa, S. L., and Majchrzak, A. 2011. "Knowledge Collaboration in Online Communities". *Organization Science* (22:5), pp. 1224–1239.

Fjeldstad, Ø. D., Snow, C. C., Miles, R. E., and Lettl, C. 2012. "The architecture of collaboration". *Strategic Management Journal* (33:6), pp. 734–750.

Frey, B. S. and Jegen, R. 2001. "Motivation Crowding Theory". *Journal of Economic Surveys* (15:5), pp. 589–611.

Friedrich, J., Becker, M., Kramer, F., Wirth, M., and Schneider, M. 2020. "Incentive design and gamification for knowledge management". *Journal of Business Research* (106), pp. 341–352.

Gallus, J. 2017. "Fostering Public Good Contributions with Symbolic Awards: A Large-Scale Natural Field Experiment at Wikipedia". *Management Science* (63:12), pp. 3999–4015.

Gallus, J. and Frey, B. S. 2017. "Awards as Strategic Signals". *Journal of Management Inquiry* (26:1), pp. 76–85.

Goes, P. B., Guo, C., and Lin, M. 2016. "Do Incentive Hierarchies Induce User Effort? Evidence from an Online Knowledge Exchange". *Information Systems Research* (27:3), pp. 497–516.

Hamari, J. 2017. "Do badges increase user activity? A field experiment on the effects of gamification". *Computers in Human Behavior* (71), pp. 469–478.

Harper, F. M., Raban, D., Rafaeli, S., and Konstan, J. A. 2008. "Predictors of answer quality in online Q&A sites". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Ed. by M. Czerwinski. ACM Digital Library. New York, NY: ACM, p. 865.

Holgersson, H., Nordström, L., and Öner, Ö. 2014. "Dummy variables vs. category-wise models". *Journal of Applied Statistics* (41:2), pp. 233–241.

Hsieh, G., Kraut, R. E., and Hudson, S. E. 2010. "Why pay? : Exploring How Financial Incentives are Used for Question & Answer". In: *Mynatt, Schoner et al - Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, p. 305.

Huang, N., Hong, Y., and Burtch, G. 2017. "Social Network Integration and User Content Generation: Evidence from Natural Experiments". *MIS Quarterly* (41:4), pp. 1035–1058.

Khern-am-nuai, W., Kannan, K., and Ghasemkhani, H. 2018. "Extrinsic versus Intrinsic Rewards for Contributing Reviews in an Online Platform". *Information Systems Research* (29:4), pp. 871–892.

Kokkodis, M., Lappas, T., and Ransbotham, S. 2020. "From Lurkers to Workers: Predicting Voluntary Contribution and Community Welfare". *Information Systems Research* (31:2), pp. 607–626.

Kraut, R. E., Resnick, P., and Kiesler, S. 2012, 2011. *Building successful online communities: Evidence-based social design*. Cambridge (Mass.) and London: MIT.

Lappas, T., Dellarocas, C., and Derakhshani, N. 2017. "Reputation and contribution in online question-answering communities". *Available at SSRN: https://ssrn.com/abstract=2918913* ().

Lazear, E. P. 1986. "Salaries and Piece Rates". *The Journal of Business* (59), pp. 405–431.

Lazear, E. P. 2000. "The Power of Incentives". *American Economic Review* (90:2), pp. 410–414.

Leclercq, T., Poncin, I., Hammedi, W., Kullak, A., and Hollebeek, L. D. 2020. "When gamification backfires: the impact of perceived justice on online community contributions". *Journal of Marketing Management* (36:5-6), pp. 550–577. eprint: https://doi.org/10.1080/0267257X.2020.1736604.

Lerner, J. and Tirole, J. 2002. "Some Simple Economics of Open Source". *Journal of Industrial Economics* (50:2), p. 197.

Li, Z., Huang, K.-w., and Cavusoglu, H. 2012. "Quantifying the Impact of Badges on User Engagement in Online Q&A Communities". In: *ICIS 2012 Proceedings*. 74.

Liu, D., Santhanam, R., and Webster, J. 2017. "Toward meaningful engagement: a framework for design and research of gamified information systems". *MIS Quarterly* (41:4), pp. 1011–1034.

Lou, J., Fang, Y., Lim, K. H., and Peng, J. Z. 2013. "Contributing high quantity and quality knowledge to online Q&A communities". *Journal of the Association of Information System* (64:2), pp. 356–371.

Mekler, E. D., Brühlmann, F., Tuch, A. N., and Opwis, K. 2017. "Towards understanding the effects of individual gamification elements on intrinsic motivation and performance". *Computers in Human Behavior* (71), pp. 525–534.

Mindel, V., Mathiassen, L., and Rai, A. 2018. "The sustainability of polycentric information commons". *MIS Quarterly* (42:2), pp. 607–632.

Oh, W. and Jeon, S. 2007. "Membership Herding and Network Stability in the Open Source Community The Ising Perspective". *Management Science* (53:7), pp. 1086–1101.

Ping, Z. 2008. “Motivational affordances: Fundamental reasons for ICT design and use”. *Communications of the ACM (CACM)* (51:11).
Puranam, P., Alexy, O., and Reitzig, M. 2014. “What’s ”New” About New Forms of Organizing”. *Academy of Management Review* (39:2), pp. 162–180.
Riar, M. 2020. “Using Gamification to Motivate Cooperation: A Review”. In: *ICIS 2020 Proceedings*. 12.
Silva, J. S. and Tenreyro, S. 2011. “Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator”. *Economics Letters* (112:2), pp. 220–222.
Stvilia, B., Twidale, M. B., Smith, L. C., and Gasser, L. 2008. “Information quality work organization in wikipedia”. *Journal of the American Society for Information Science and Technology* (59:6), pp. 983–1001.
Sun, Y., Fang, Y., and Lim, K. H. 2012. “Understanding sustained participation in transactional virtual communities”. *Decision Support Systems* (53:1), pp. 12–22.
van Mierlo, T. 2014. “The 1% rule in four digital health social networks: an observational study”. *Journal of medical Internet research* (16:2), e33.
von Krogh, G., Haefliger, Spaeth, and Wallin 2012. “Carrots and Rainbows: Motivation and Social Practice in Open Source Software Development”. *MIS Quarterly* (36:2), p. 649.
von Krogh, G., Spaeth, S., and Lakhani, K. R. 2003. “Community, joining, and specialization in open source software innovation: a case study”. *Research Policy* (32:7), pp. 1217–1241.
Wang, J., Li, G., and Hui, K.-L. 2021a. “Monetary incentives and knowledge spillover: evidence from a natural experiment”. *Management Science* (forthcoming).
Wang, N., Liu, Y., and Xiao, S. 2022. “Which feedback matters? The role of expressions and valence in continuous high-quality knowledge contribution in the online QA community”. *Decision Support Systems* (156), p. 113750.
Wang, X., Sanders, S. P., and Sanders, G. L. 2021b. “Examining the Impact of Yelp’s Elite Squad on Users’ Following Contribution”. In: *ICIS 2021 Proceedings*. 23.
Wasko, M. M. and Faraj, S. 2005. “Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice”. *Management Information Systems Quarterly* (29:1), pp. 35–57.
Wasko, M. M., Teigland, R., and Faraj, S. 2009. “The provision of online public goods: Examining social structure in an electronic network of practice”. *Decision Support Systems* (47:3), pp. 254–265.
Wiertz, C. and de Ruyter, K. 2007. “Beyond the Call of Duty: Why Customers Contribute to Firm-hosted Commercial Online Communities”. *Organization Studies* (28:3), pp. 347–376.
Wooldridge, J. M. 1999. “Distribution-free estimation of some nonlinear panel data models”. *Journal of Econometrics* (90:1), pp. 77–97.
Yeow, A., Johnson, S., and Faraj, S. 2006. “Lurking: Legitimate or Illegitimate Peripheral Participation?” In: *ICIS 2006 Proceedings*. 62.
Yu, Y., Khern-am-nuai, W., and Pinsonneault, A. 2022. “When paying for reviews pays off: The case of performance-contingent monetary rewards”. *MIS Quarterly* (46:1), pp. 609–626.
Zhao, L., Detlor, B., and Connelly, C. E. 2016. “Sharing Knowledge in Social Q&A Sites: The Unintended Consequences of Extrinsic Motivation”. *Journal of Management Information Systems* (33:1), pp. 70–100.

Appendix

Parallel Pretrends Investigation

With Figure A.1., we aim to provide support for assumptions that our comparison OC provides a suitable approximation of a counterfactual for our treated OC. This is accomplished by testing whether pre-trends are considered parallel between our treated context and the control OC by estimating a dynamic DiD model (see, e.g. Barbosu and Gans 2022; Huang et al. 2017). To do this, we interact the group dummy variable with a dummy variable for each month, omitting the month of the first period, illustrated in Equation 3 below:

$$y_{igt} = \delta \times (group_g \times monthdummy_t) + X \times \beta + \alpha_i + \gamma_t + \epsilon_{igt} \quad (3)$$

Y_{igt} represents our dependent variable of interest, X represents the covariates of relevance, α and γ are the fixed effects, and ϵ is our individually clustered error term. The coefficients of interest are all the δ coefficients. As in our main specification, the model for our quantity variable is estimated via a Fixed Effects Poisson regression, while we estimate coefficients of the quality analysis via OLS with a logged dependent variable. We then plot the δ estimates with error bars based on the 95% confidence interval. In the event study below, we find no indication of a pre-trend that may explain the results we find in our main analysis. We do see that the effect of award introduction is not immediately significant. This makes sense, as it would take some time for OC members to become aware of and acquainted with the PCSA.

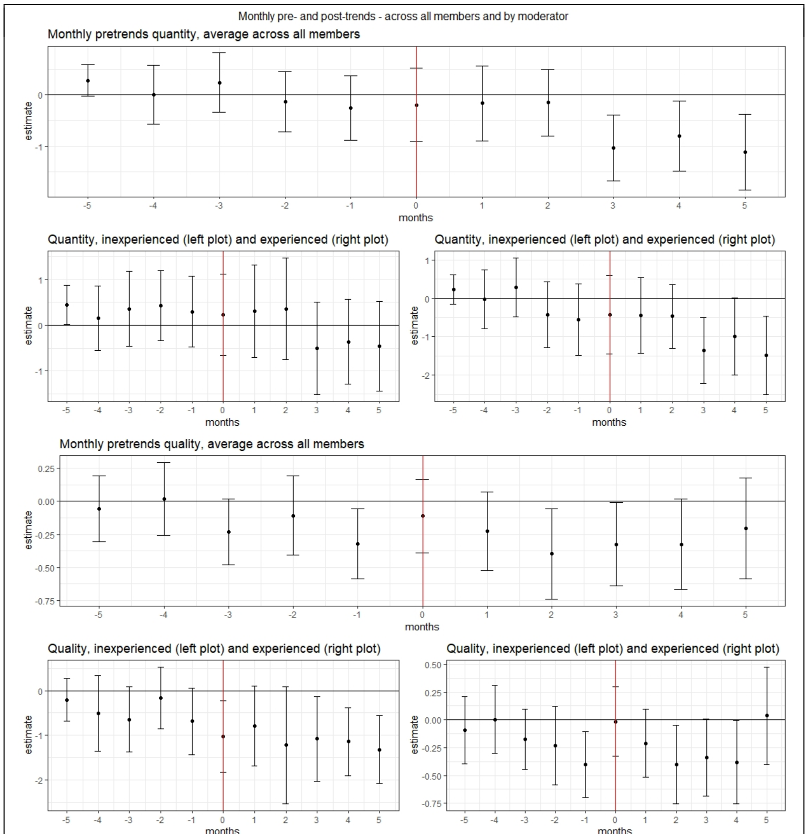

Figure A1. Event Study showing pre- and post-trends

Robustness Check - 1:3 NN Matching with Replacement

Table A.1. repeats the estimation procedures from our main analyses, using a sample based on 1:3 nearest neighbour PSM matching with replacement as a robustness check. For the 1:3 nearest neighbour PSM matching with replacement, members from the control group were only included once, even if they were matched to multiple treatment units. Coefficients are estimated via OLS without fixed effects (Models 1 and 5) as well as via our main estimation models–Fixed Effects Poisson for contribution quantity and OLS with a logged DV for contribution quality. As can be seen in Table A.1., the results of the regressions have the same sign and statistical significance levels as in our main specification.

Robustness Check - Placebo Check

To test if our results are truly due to the introduction of the PCSA, or whether they simply reflect underlying differences in contribution dynamics between the two communities, we ran placebo checks shown in Table A.2. below. Herein, we estimate the DiD coefficients for any change in contribution behavior with an arti-

Models	Contribution quantity OLS 1	Poisson 2	3 inexp	4 exp	Contribution quality (# of words) OLS 5	OLS Log(DV) 6	7 inexp	8 exp
Intercept	-0.12^{***}				107.61^{***}			
	(0.03)				(2.50)			
Group	0.27^{***}				-38.67^{***}			
	(0.03)				(3.70)			
Treated	-0.03^{***}				29.74^{***}			
	(0.01)				(3.81)			
q's per week	0.00^{***}	0.01^{***}	0.01^{**}	0.00				
	(0.00)	(0.00)	(0.00)	(0.00)				
Group:Treated	-0.07^{***}	-0.64^{***}	-0.35	-0.92^{***}	-20.25^{**}	-0.10	-0.49^{**}	-0.07
	(0.01)	(0.18)	(0.33)	(0.22)	(6.68)	(0.07)	(0.15)	(0.07)
R^2	0.00				0.06	0.73	0.83	0.62
Adj. R^2	0.00				0.06	0.50	0.50	0.48
Num. obs.	109248	109248	77328	31920	5163	5163	2522	2641
AIC		61210.48	33932.39	26983.97				
Indiv. + week FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes

Notes: SEs clustered at individual level for fixed effects models $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table A.1. Robustness checks using sample based on 1:3 NN Propensity Score Matching.

Models:	Quantity - num. of answers Poisson 1 no match t-1	2 matched t-1	3 Main spec	Qual - num. of words OLS log(DV + 1) 4 no match t-1	5 matched t-1	6 Main spec
questions per week	0.00	-0.00	0.00^{**}			
	(0.00)	(0.00)	(0.00)			
Group:Treated	0.22	0.28	-0.53^{**}	0.10	0.03	-0.12
	(0.16)	(0.33)	(0.19)	(0.05)	(0.09)	(0.07)
R^2				0.70	0.49	0.72
Adj. R^2				0.55	0.36	0.47
AIC	96615.69	18727.56	49102.88			
Observations	139104	13680	88704	8531	1627	4017
Indivi. + week FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: SEs clustered at individual level for fixed effects models $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table A.2. Placebo checks for the period one year before window of observation.

ficial treatment date 1 year prior to PCSA introduction. When using our original matched sample (models 2 and 5) or the unmatched sample based on users registered before the artificial treatment date (models 1 and 4), we now find significant treatment effects. Thus, both for quantity and quality, we find no indication of our main results being driven by our sample selection procedure.