*Information Systems Research*
Vol. 27, No. 3, September 2016, pp. 497–516
ISSN 1047-7047 (print) | ISSN 1526-5536 (online)
http://dx.doi.org/10.1287/isre.2016.0635


# Do Incentive Hierarchies Induce User Effort? Evidence from an Online Knowledge Exchange

Paulo B. Goes
Dean and Halle Chair in Leadership, Eller College of Management, University of Arizona, Tucson, Arizona 85721,
pgoes@eller.arizona.edu

Chenhui Guo, Mingfeng Lin
Department of Management Information Systems, Eller College of Management, University of Arizona, Tucson, Arizona 85721
{chguo@email.arizona.edu, mingfeng@eller.arizona.edu}

To motivate user contributions, user-generated content sites routinely deploy incentive hierarchies, where users achieve increasingly higher statuses in the community after achieving increasingly more difficult goals. Yet the existing empirical literature remains largely unclear whether such hierarchies are indeed effective in inducing user contributions. We gather data from a large online crowd-based knowledge exchange to answer this question, and draw on goal setting and status hierarchy theories to study users' contributions before and after they reach consecutive ranks on a vertical incentive hierarchy. We find evidence that even though these glory-based incentives may motivate users to contribute more before the goals are reached, user contribution levels drop significantly after that. The positive effect on user contribution appears only temporary. Moreover, such impacts are increasingly smaller for higher ranks. Our results highlight some unintended and heretofore undocumented effects of incentive hierarchies, and have important implications for business models that rely on user contributions, such as knowledge exchange and crowdsourcing, as well as the broader phenomenon of gamification in other contexts.

*Keywords*: online knowledge exchange; motivation; status; incentive hierarchy; goals; effort; user-generated content
*History*: Chris Forman, Senior Editor; Param Vir Singh, Associate Editor. This paper was received on March 24, 2014, and was with the authors 14 months and 10 days for 2 revisions. Published online in *Articles in Advance* June 23, 2016.

## 1. Introduction

From e-commerce websites to online communities, many websites today rely heavily on user-generated content (henceforth UGC). They regularly draw on voluntary contributions from a fluid membership to generate sufficient contents and rely on these contents to attract new members and retain old ones. The reach of such sites is only limited by imagination, ranging from wikis, blogs, social networking sites, consumer review sites, and online games, to question-and-answer and other crowdsourcing platforms. Regardless of their nature, a fundamental issue for such websites is virtually the same: How do we motivate these users to continue participating and voluntarily contribute new contents?

This issue is challenging due to the fundamental public goods problem inherent in UGC sites: Users need not exert any effort to enjoy the contributions of others. A classic example is product reviews (Chen et al. 2010), since review writers expend the costs (time and effort) but may not fully reap the benefits of providing the reviews. Such "public goods" nature of UGCs will, as economic theories predict, lead to an *undersupply* of such contents due to insufficient incentives. While the long memory of the Internet can capture and accumulate some whimsical moments of kindness—thereby allowing more people to benefit from someone else's efforts—the provision of timely information or contents can be highly challenging to sustaining the growth of UGC sites.

One possible solution to this problem is to increase the rewards for users who contribute, so that they *internalize* more benefits from their efforts. Sometimes such internalization can occur endogenously. In open-source software development, some users can improve their offline employment prospects because of their online reputation (Roberts et al. 2006, Hann et al. 2013). Unfortunately, such internalization may not be applicable to all UGC sites, and it is only realistic for some extremely well-known members. More importantly for website managers, this is beyond their control.

For these reasons, more and more UGC sites adopt "incentive hierarchies," which allow users to receive points by making contributions and automatically grant them various accolades as their cumulative points reach increasingly higher thresholds. Such mechanisms intend to help users internalize more benefits of providing public goods by recognizing them in front of their peers and bestowing glory, honor, or bragging rights. It appears obvious that these accolades should incentivize users to contribute more, which could explain why incentive hierarchies have become so popular.

However, there remains a significant gap in our understanding of how such incentive hierarchies actually affect user behaviors. In particular, these incentive hierarchies are very different from the traditional—and mostly successful—loyalty rewards programs such as those used in marketing (e.g., Kivetz et al. 2006[1]), despite some apparent similarities.

First, the "effort" in a traditional rewards program is rarely about a user's voluntary contribution. Rewards programs in marketing allow users to accumulate points from their *purchases*, such as buying coffee or airline tickets. The points users receive on UGC sites do not come from purchases, but rather voluntary contributions. An important impetus for voluntary contributions is intrinsic motivations (Ryan and Deci 2000), the inherent enjoyment of helping others. Such motivations are generally irrelevant for purchases. In addition, while extrinsic rewards such as free coffee may encourage purchase behaviors, they will negatively affect intrinsic motivations (Lepper et al. 1973). Therefore, the apparent similarity between incentive hierarchies and rewards programs does not guarantee that incentive hierarchies will be effective in UGC contexts.

Second, the "goals" in traditional rewards programs are typically meant for a user's private consumption with no publicity. By contrast, an incentive hierarchy consists of increasingly rarer symbols of statuses, and by definition, statuses make sense only when they are public. This difference will impact user behaviors in very different ways. For example, whereas users in traditional rewards programs tend to significantly decrease their activities upon reaching their goals (e.g., earning free coffee), this may not be true in a UGC context because a drop in a user's activity level can be easily seen by the whole community.[2] The user therefore may be less likely to decrease their efforts after goal attainment. Again, what we know from traditional rewards programs does not necessarily apply.

Third, an incentive hierarchy is uniquely comprised of progressively more challenging goals as well as increasingly higher statuses. A higher rank can only be attempted when all previous ranks have been reached. By contrast, marketing rewards programs repeat the same reward for the same milestones (e.g., always the free 11th cup of coffee), and the badge systems on some sites provide different badges for various aspects of a user's activities[3] (e.g., Anderson et al. 2013). Therefore, findings from studies of those systems or programs do not necessarily hold true for incentive hierarchies.

For these reasons, we cannot rely on what we know from studies of badges or traditional rewards systems to understand user behaviors in an incentive hierarchy. In fact, incentive hierarchies could be completely ineffective in a UGC context. The glory and recognition associated with each rank on an incentive hierarchy is in essence an *extrinsic* reward that users cannot opt out of. Yet at least some users participate in UGC communities for *intrinsic* reasons or the mere joy of helping others. A long stream of research (e.g., Lepper et al. 1973) has documented that giving explicit and *expected* rewards can in fact dampen those users' motivations to contribute. Therefore, incentive hierarchies may not necessarily serve to motivate user efforts at all because of the nature of UGC sites.

Given the popularity of incentive hierarchies in UGC contexts, therefore, it is urgent that we systematically examine whether and how incentive hierarchies actually influence user contributions. Such an investigation will not only enrich the UGC literature but also provide significant value for practitioners. For example, if our study shows that incentive hierarchies do provide the expected benefits, then it will provide helpful validation for this common practice. Yet if we find that the effects are small or short-lived, then our study will bring attention to an important but underappreciated issue and initiate discussions on how to improve incentive hierarchies or develop better ways of motivating users.

Our research context is a leading online knowledge exchange where users can ask technology-related questions and seek answers from peers. The members who ask (henceforth "*askers*") can choose one or more answers as the solution, and those who provide the answers (henceforth "*answerers*") receive points based on the points that the asker assigned to the question and the asker's evaluation of the answer. Once the answerers accumulate a sufficient number of points, their rank on the site will be elevated from "regular member" to increasingly higher statuses such as "master" or other designations. We investigate whether this meritocracy type of incentive hierarchy actually induces users to increase their contributions to the community by answering more

[1] An example of the marketing rewards program used in Kivetz et al. (2006) is a free cup of coffee every time the customer pays for 10 cups.

[2] We thank an anonymous reviewer for this suggestion.

[3] For example, there may be one type of badge for logging in for the first time, another badge for making a referral, and so on.

questions posted by peers. Since incentive hierarchies feature increasingly more challenging goals associated with increasingly higher statuses, we draw on both the *goal-setting* and *status hierarchy* theories to examine the following research questions:

1. *How do users' contribution levels change* before *they reach goals*?
2. *How do users' contribution levels change* after *they reach goals*?
3. *Do different ranks on the incentive hierarchy affect user behaviors differently*?

We track a random sample of users on the website over time and construct a panel data set to study these questions. We find evidence that even though different ranks motivate users to contribute more *before* the goals are reached, their effort levels drop significantly upon goal attainment. After that, as users make progress toward the next rank, their effort only slowly ramps up, at a speed much slower than before goal attainment. Hence, the positive effect of goals in the incentive hierarchies appears only temporary. Furthermore, we find that in general, the higher the rank of the hierarchy, the smaller the impact on user behaviors.

Our study is one of the first to apply goal-setting and status hierarchy theories to empirically studying the effect of incentive hierarchies in a UGC site, and to document several important findings. Our findings challenge the common wisdom that incentive hierarchies are effective in inducing user contributions; rather, after goals are reached, users are slow to resume their effort level. These findings have important implications not only for knowledge exchanges but also for other contexts where similar incentive hierarchies are present, such as online games or online education. Our study also contributes to the empirical goal-setting literature by examining the effects of consecutive and increasingly higher goals, rather than a single goal.

## 2. Literature and Hypotheses

### 2.1. Related Literature

The incentive hierarchy we study consists of predefined, consecutive goals. To motivate our tests, we first review the literature on *motivations*. In UGC communities, the most prevalent incentive is status. Therefore, we next review literature on *status*, especially *status hierarchies* because of the hierarchical nature of our context. Finally, statuses or ranks along the hierarchy represent predefined *goals* for members of the community, so our literature review will not be complete without considering the literature on *goals*, especially how goals induce goal-pursuit behaviors.

**2.1.1. Motivation, Status, and Goals.** Social psychologists broadly classify motivations into two categories: *intrinsic motivations* and *extrinsic motivations* (Ryan and Deci 2000). *Intrinsic* motivations refer to the innate desire to perform an action, such as contributing to a community, because the action itself provides utility or enjoyment for that person. Such motivations lead to actions not because of the appeal of a reward or the threat of a punishment (Ryan and Deci 2000). *Extrinsic* motivations are the exact opposite: An act is performed not because it is enjoyable, but because there are external reasons to do so, such as potential rewards or punishments. Financial motivations naturally fall into the category of extrinsic motivators, such as explicit payments (Roberts et al. 2006) or the prospect of improved employability (Hann et al. 2013). However, such financial benefits are rare in online communities and will accrue only to a few elite members. Nonfinancial extrinsic rewards are much more prevalent. Prior studies have documented that factors such as status or reputation among peers (Roberts et al. 2006, Wasko and Faraj 2005), perceived image (Jabr et al. 2014), community response (Zhang et al. 2013), identification with a group (Ma and Agarwal 2007, Bagozzi and Dholakia 2006), or social comparison (Chen et al. 2010) can all extrinsically motivate users to contribute contents on UGC sites.

Status among peers (Anderson et al. 2006, 2015) is a powerful extrinsic motivator. It is directly related to the incentive hierarchies we study, because each level of the hierarchy defines a milestone in a member's status trajectory in the community. The status literature has consistently found that people pay close attention to status, even when it is represented by seemingly trivial *symbols* (Anderson et al. 2015). A high status commands respect and admiration from peers. Accordingly, the pursuit and preservation of status can serve as a powerful impetus for certain behaviors (Anderson and Brown 2010, Sheldon 2011, Baumeister and Leary 1995), especially those that provide *value* to a community (Anderson et al. 2015). Such "value" is consistent with the concept of public goods we discussed. For example, Goes et al. (2014) find that *popularity*—a concept directly related to one's perceived status within a community—leads users to contribute more in a UGC site.

Status is not binary but often manifests itself in the form of a *status hierarchy*, a sequence of status levels or a pecking order among members. Status hierarchy is almost universal and often essential in organizations (Bavelas 1950), but particularly so when the hierarchy is based on technical capabilities (Anderson and Kilduff 2009). Despite a much more fluid membership, online communities are still organizations of people. Hence, it seems plausible that incentive hierarchies are implemented to establish status hierarchies. Furthermore, organizations with steeper status hierarchies—where there is "larger asymmetry in member's power, status and influence" (Anderson and Brown 2010, p. 6)[4]—often perform better than those with flatter ones. This appears consistent with the typical implementation of incentive hierarchies (i.e., many goals, ever more challenging, with very few members at the top ranks).

Despite their apparent similarities, incentive hierarchies are distinct from status hierarchies, so incentive hierarchies are not necessarily as effective as status hierarchies documented in the literature. Most importantly, status hierarchies can naturally emerge because of varying contributions from members (e.g., the popularity of users; see Goes et al. 2014), whereas incentive hierarchies are designated by the website with clear criterion for each rank. Status can emerge even if there is no incentive hierarchy or similar ranking mechanisms. By contrast, incentive hierarchies institute predefined (therefore anticipated), highly visible milestones of a member's status trajectory. Even though members' inherent pursuit of status can serve as goals and lead to goal-directed behaviors (Sheldon 2011, Baumeister and Leary 1995), ranks on incentive hierarchies provide much more concrete, salient, and uniform reference points—goals—for all members of a community. Given the delicate nature of human motivations (Ryan and Deci 2000), the *explicit* incentive hierarchies in UGC contexts are not necessarily as effective as status hierarchies that would have naturally emerged. The value of incentive hierarchies, therefore, should be empirically established rather than presumed.

**2.1.2. Pursuit of Goals.** Psychologists have long been interested in goal-pursuit behaviors (e.g., Hull 1932, Miller 1944, Earley et al. 1989, Louro et al. 2007, Koo and Fishbach 2008, Liberman and Förster 2008, Bonezzi et al. 2011). Hull (1932) found that rats ran faster when they were closer to the food, providing one of the earliest pieces of evidence that distance to goals affects effort. Such goal-pursuit behaviors readily extend to human beings, since we are able to anticipate consequences of goals and exert effort to achieve them (Lewin et al. 1944). Interestingly, many later studies in this literature focus on the presence or absence of goals, rather than the distance from them. For example, Locke's goal-setting theory (1967, 1968) suggested that individuals achieve higher performance in the presence of goals (see Locke and Latham 2002 for a review). Research consistently shows that "specific, challenging goals" have a stronger effect than "do-your-best" goals. In fact, the difficulty of goals can positively influence effort and performance (Locke and Latham 1990, 2002; Mento et al. 1987; Tubbs 1986).

The *goal-gradient hypothesis* (Hull 1932) in this literature is particularly relevant since we are interested in the distance from goals. Researchers have empirically examined the effect of goal proximity in several contexts, such as loyalty rewards programs in marketing (Kivetz et al. 2006, Cheema and Bagchi 2011, Zhang and Huang 2010, Huang and Zhang 2011). However, most studies in this literature, especially the empirical ones, are based on contexts very different from UGC sites. For example, "rewards" in these programs are mostly for consumers' private consumption with no implications of "glory." By contrast, each rank in an incentive hierarchy signifies a milestone in a member's *status*, and status is meaningful only socially (in the eyes of peers). In addition, most studies focus on individual behaviors prior to goal attainment, but very few, if any, consider behaviors *after* goal attainment, most likely because there is typically only one goal in those studies. Last but not least, almost all empirical studies in this literature are about a single goal (even if it can be repeated), rather than consecutive and ever more challenging goals. In rewards programs for instance, while a consumer may receive every 11th coffee for free, receiving the second free coffee in these programs is very different from reaching the second level in an incentive hierarchy. Because of these differences, existing research is not yet conclusive on whether incentive hierarchies are effective in inducing user efforts in a UGC context. Incentive hierarchy represents a new phenomenon that requires new theorizing and systematic empirical tests.

**2.1.3. Related Studies in the UGC Context.** We now review a small but growing literature on how UGC-website users respond to website-designated goals. Two studies presented in recent conferences are directly related to our study, although they both use data from StackOverflow.com, a free question-and-answer website different from our context. On StackOverflow.com, users can earn various *badges* on different aspects of their activities, and the effect of badges is the focus of both studies. Specifically, Li et al. (2012) identified a short-term positive effect of winning new badges (i.e., users contributed more if they obtained new badges in the previous period). Another study (Anderson et al. 2013) argued that users increase efforts to obtain badges and plotted user activities before and after that as empirical evidence.

Compared to the rewards programs discussed earlier, the badges in these two studies are more similar to the incentive hierarchies we study, because badges carry psychological rewards and signify a status. However, badges are *horizontal*: There are many different badges for a variety of activities. In addition,

[4] In a steep hierarchy, the number of people with higher status is much smaller than the rest, therefore more asymmetric in status.

each badge represents a highly simplified incentive hierarchy because there is only one level (users either have a badge or do not have a badge). Incentive hierarchies, by contrast, are vertical and comprised of multiple consecutive ranks. Furthermore, whereas incentive hierarchies provide a uniform ranking system against which all members can easily compare with each other, badges do not allow such comparisons: One particular badge does not necessarily provide more glory than another. Empirically, the presence of different badges, and the fact that one action may simultaneously contribute to multiple badges, make it difficult to disentangle the effect of each. By contrast, we study the effect of a unified, multilevel hierarchy, in a context with no multiple horizontal badges. Moreover, the business model of our context is highly different from the open and free websites that they study. Hence, findings derived in these studies do not necessarily apply, and our research further contributes to this growing literature.

In summary, we situate our study in a new context that is more amenable to empirically disentangling the effects of multiple, consecutive status goals, and we try to provide more systematic empirical evidence for this important but still underdeveloped literature.

### 2.2. Hypotheses Development

We now formally develop a series of hypotheses on users' behaviors as they progress along an incentive hierarchy. For ease of exposition, we refer to different ranks on the incentive hierarchy as goals, though it should be noted that these are platform-designated goals. Following the process through which users reach ranks, we start with the effect on user behavior as users *approach* goals—taking into account the distance from goals—followed by the instantaneous effect upon goal attainment. We next discuss the "recovery" of user contributions after reaching goals, due to the influence of the next goal. Subsequently, we contrast the effect of different levels of goals on user efforts.

**2.2.1. Approaching Goals: The Effect of Distance.** Ranks on an incentive hierarchy, once obtained, will be prominently displayed on a user's public profile. They serve as symbols of the user's status in the community and are prime examples of extrinsic rewards because even trivial symbols can influence behaviors (Anderson et al. 2015). Furthermore, the psychology literature has long demonstrated that in a phenomenon known as "goal gradient," the motivating effect increases as the distance to a goal decreases (Heath et al. 1999, Hull 1932, Kivetz et al. 2006). Therefore, these ranks should provide strong impetus for users to contribute more, especially as users approach the goals and the goals appear more and more within reach.

The motivational effect of goals can be absent when the goal is excessively challenging, especially when a user is not committed to that goal (Locke 1968, Hollenbeck and Klein 1987). However, such inhibition effects should be increasingly less likely to occur as a user approaches goals. Therefore, the net effect of goal proximity on users should still follow the prediction of the goal gradient argument.[5] This general result should hold even in the presence of multiple sequential goals in the hierarchy. As users approach one rank, they are also making progress toward ranks after that. The effect of those future goals—if any—will affect user behaviors in the same direction. Therefore, we hypothesize that:

Hypothesis 1 (H1). *Users increasingly accelerate their efforts before goal attainment. For a given length of time, they contribute increasingly* more *contents as their distance to the next goal decreases.*

**2.2.2. Attaining Goals: The Instantaneous Effect of Goal Attainment.** As discussed, incentive hierarchy ranks can serve as goals for users. While such goals induce user efforts as users approach goals, the literature also suggests that there will be a drop in efforts upon goal attainment. For example, the Drive–Reduction Theory proposed by Clark Hull (see Dewey 2007 or Hilgard and Bower 1975) suggests that satisfaction from achieving a goal leads to a drop in efforts when the goal is reached. Additionally or alternatively, the attainment of the status symbol—an extrinsic reward—could even dampen intrinsic motivations (Lepper et al. 1973). This literature suggests that we should expect user efforts to drop as soon as a goal is reached.

There are two potential counterarguments to this prediction in the UGC context. First, since UGC sites are public and user activities are readily observable by peers, users who hope to ultimately obtain a higher status in the community may feel embarrassed if their contribution drops abruptly,[6] so they may have an incentive *not* to reduce efforts upon goal attainment.

[5] It is possible that some users may be completely oblivious to the presence of incentive hierarchies and remain solely driven by intrinsic motivations. The presence of such users, however, will not change our theoretical prediction, because these users' effort level should be invariant regardless of their distance from goals. Therefore, when we consider all users, as long as not everyone is oblivious to the ranks, our Hypothesis 1 holds. In addition, studies such as Lepper et al. (1973) show that intrinsic motivations are often dampened, if not eliminated, when extrinsic rewards can be anticipated. The goals on an incentive hierarchy are such extrinsic rewards, and they are not only visible but advertised to all members of a UGC site to engage them. Members also cannot choose not to have their status information displayed on the site. Therefore, users who completely ignore the incentive hierarchy are likely to be small in number even if they exist.

[6] We thank an anonymous reviewer for this observation.

This is different from marketing rewards programs because those rewards are for private consumptions only (Kivetz et al. 2006). Second, empirical support for the Drive–Reduction Theory is typically obtained from lab experiments where there is only one goal. The presence of future goals could eliminate the drop in user efforts when users reach a goal, because the attainment of one goal suggests that the user is now closer than ever to the next rank, which will offer an even higher status and more recognition. Users who are sufficiently motivated by the lure of the next goal, then, will continue to exert effort.

The net effect of these forces, therefore, will depend on how far the next goal is from the goal just accomplished. If it is within reach, then we should not expect to see the drop, because some users may continue to maintain their effort level due to the observability of efforts, and some users may even increase their efforts due to the next goal. Yet if it is very far away and the user needs to put in a significant amount of effort to reach the next goal, then immediately after goal attainment there *will* be a drop in contributions because the lure of the next goal will not kick in. For many incentive hierarchies, including the one we study, the distance between ranks is increasingly larger (the distance between levels 2 and 3 is larger than that between 1 and 2, and so on). Because of the sheer distance, future ranks will appear virtually unattainable to a user who just obtained a rank and therefore will not motivate them (Hollenbeck and Klein 1987). Hence, the positive effect of future ranks is unlikely to overcome the drop in contributions due to the satisfaction, or complacency, from the completion of the current rank. We therefore hypothesize a drop upon goal attainment:

Hypothesis 2 (H2). *Users will reduce their efforts once they reach goals in a hierarchy. In other words, users' effort levels should be* higher *before they reach goals than after.*

**2.2.3. After Goal Attainment: The Lure of the Next Goal.** Once a goal is reached, its motivational effect—if any—is gone, because the desire for the status associated with that goal has been satisfied. A natural and important question at this point for website administrators is, given the drop hypothesized in H2, how quickly can the user recover from the drop and return to a high level of contribution? In other words, since the users are now faced with and hopefully intrigued by the next rank, how will the increase of their efforts compare to that right before the attainment of the previous rank? As discussed, higher ranks are increasingly more prominent and confer increasingly higher statuses. Perhaps the lure of the next rank is so strong that the users immediately recover from any reduction of effort (if any) and strive toward the next goal.

However, this is an unlikely scenario. As mentioned earlier, in virtually all incentive hierarchies, the distance between each pair of consecutive ranks is increasingly higher. The additional amount of contribution to the community required to achieve the second rank after achieving the first, for instance, will be significantly *higher* than the amount of contribution required to achieve the first rank. The third rank will require even more. Since the *distance* between consecutive goals is increasingly higher, the lure of the next rank will be highly unlikely to influence users who just obtained the previous rank. In fact, most existing research models the effort trajectory toward a goal as a convex relationship: A goal only starts to affect users when the users are within a certain distance, analogous to a "gravitational field" of a goal's influence. For this reason, when the users have just obtained a rank, even if the next rank may prompt them to recover their efforts, the speed of the recovery will be much lower than the acceleration toward the previous goal. Therefore, when we consider the users' effort level at a given distance *after* goal attainment, it will be much lower compared to their effort level at the same distance *prior to* the attainment of that goal. We hence hypothesize that:

Hypothesis 3 (H3). *Users will gradually recover their efforts after the drop at goal attainment, due to the presence of the next goal. However, quantitatively the recovery speed will be* lower *than the acceleration prior to the attainment of the goal.*

**2.2.4. Hierarchical Effect: Comparison Across Ranks.** The presence of consecutive ranks is a unique feature of an incentive hierarchy. Since these ranks are vertical (users cannot reach a higher rank unless they attain a lower rank first), any user at any given time will only be faced with *one* immediate next rank. Hence, our hypotheses so far (H1–H3) should apply to the effect of all ranks—we should observe that users increase their efforts as they approach a rank threshold, drop when they reach it, and then recover (at a lower speed) after that. However, the *magnitude* of this effect may not be the same for lower and higher ranks. This is an important question to consider. If this effect is increasingly larger as the user moves up the incentive hierarchy, then there should not be a top rank in the hierarchy, unless that rank is simply impossible to achieve. Yet if this effect is increasingly smaller as we move up the hierarchy, then the ranks are perhaps only meaningful (in terms of inducing efforts) up to a certain level; after that, higher ranks may not have any substantial impact on user contributions even if they exist. Therefore, understanding the varying impact of different ranks has important implications for the design of incentive hierarchies.

To understand how different ranks affect users in different ways, we must first understand what distinguishes a higher rank from a lower one: (1) the status of the ranks is increasingly higher; (2) the amount of additional effort required to reach each rank (i.e., from start to level 1, from level 1 to level 2, and so on) is also increasingly larger; and (3) importantly, the *subset of users* who are able to progress close enough to a rank and therefore be subject to its influences is different. The third point is especially important. If we compare the pools of users approaching each rank, they are different not just in how many they are, but more importantly in their *composition* in terms of engagement with the community, or what motivates them to contribute.

As mentioned earlier, motivations are broadly classified into extrinsic and intrinsic motivations, and this taxonomy can help us conceptualize the heterogeneity among users. It is useful to think of the extreme cases first. In the UGC context, some users participate because they truly enjoy the act of helping others, even if there is no explicit financial or nonfinancial benefits to do so; that is, they have strong intrinsic motivations. These users will contribute even when they are far away from the thresholds of ranks (i.e., when they are not in the goals' "gravitational" field of influence that we discussed). By contrast, some users participate solely to derive benefits, such as the glory associated with ranks of an incentive hierarchy; that is, they are to a great extent only motivated extrinsically. These users will be very reluctant to contribute when the goals are so far away, because those goals have no perceivable influence on them. Importantly, these two types of users represent the theoretical extremes, and actual users on the site have varying degrees of combination in terms of their motivations.

The vertical and "nested" nature of ranks on an incentive hierarchy, and the goal gradient from each rank (H1), suggest that the composition of users who will be approaching each rank will vary. Such compositional differences will lead to differences in what we observe or measure as the ranks' effectiveness in inducing efforts. For simplicity, let us compare the impact of the first and second rank. After reaching the first rank, users who participate mostly because they enjoy helping others (more intrinsically driven than extrinsically) will continue to participate, and because of that they will continue to accumulate points and gradually approach the second rank as the gradient effect of the second goal kicks in (H1). By contrast, for users who are largely driven by extrinsic reasons, such as the glory with the status, their contributions will mostly depend on the lure of the next rank. As discussed in H3, the second rank is so far away that its impact can hardly be felt when these users just achieve the first rank. Therefore, users who are primarily extrinsically motivated will be less likely to continue participating in the community. In other words, the chances of these users dropping out of the community after achieving the first rank will be higher than users who are more intrinsically driven.

For these reasons, the proportion of users who are more intrinsically driven will be *higher* among the group of users that we will observe approaching the second rank than those approaching the first rank. More generally, for any pair of ranks, the higher rank is more likely to find in its field of influence users who are more intrinsically driven than the lower rank, because the gradient effect of goals and an ever increasing distance to the next goal create a "Darwinian" process that filters out extrinsically driven users.

Ranks are extrinsic rewards by design. Therefore, they will be more influential when there are more users who care about those ranks (i.e., more extrinsically motivated). To put it differently, a rank's overall impact will be higher when more users in its "scope of influence" care about the glory associated with it. Because higher ranks will find in its scope of influence a lower proportion of users who are extrinsically motivated, we should observe that the impact of ranks hypothesized so far (H1–H3) should be *smaller* for higher level ranks than the lower level ones: The increase, drop, and recovery in efforts should be increasingly smaller for higher ranks.[7] We therefore hypothesize that:

Hypothesis 4 (H4). *In an incentive hierarchy, higher level ranks will show smaller impacts on user behaviors—as hypothesized in H1–H3—than lower level ranks.*

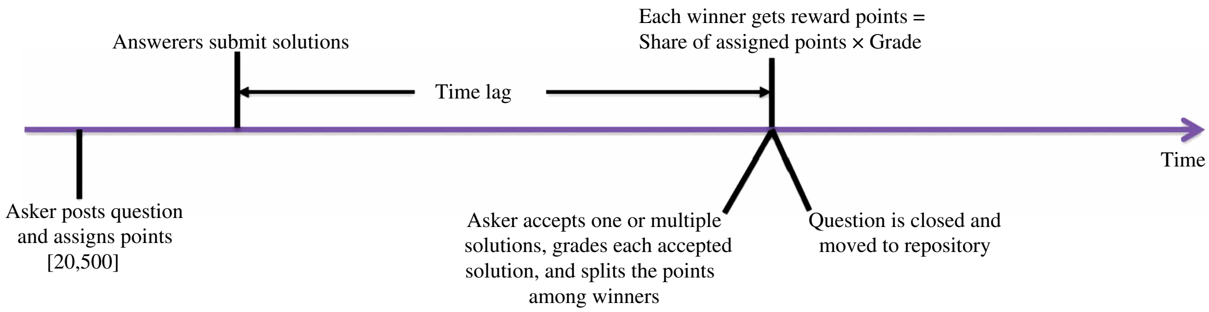
## 3. Research Context

Our research context is a popular online knowledge exchange, where members ask and answer technical questions. Launched in the late 1990s, this website is a global community built around information technology (IT)-related questions. As of 2012, more than 100,000 users had contributed to the platform's knowledge repository (archived questions and answers) by providing detailed solutions to more than three million questions.

In a typical question-answering process on this site (illustrated in Figure 1), an asker posts a question and assigns points to it (between 20 and 500 points) based

[7] One possible counterargument is that, if a user is (extrinsically) driven by the highest rank available, and only by that goal, then because of the gradient effect (H1), at a given distance from each of the lower ranks the effect should be increasingly higher across the ranks (counter to our prediction). However, the highest rank is extremely far away for a new user, so the impact is extremely unlikely to reach users who have yet to gain the first rank. In addition, if this were true, then we should find no support for H2 or H3 other than the top rank. This did not turn out to be the case.

**Figure 1** **(Color online) Timeline of Question Answering**




on the question's level of difficulty and urgency. After answerers submit their responses or comments, the asker can accept one or more as solutions and grade each solution as "A," "B," or "C." Reward points will then be accordingly awarded to answerers. The number of points that each answerer earns is determined by the share of points for the answer they provide, as well as the grades received.[8]

Users accumulate these reward points to reach ranks in an incentive hierarchy designated by the site.[9] When users' accumulated points exceed predefined thresholds, they will be automatically granted increasingly higher rank, such as "master" at the first level (50,000 points), "guru" at the second level (150,000 points), "wizard" at the third level (300,000 points), and more glorious titles at increasingly larger intervals. These ranks, or titles, are prominently displayed on the users' profile avatars, which appear in all messages that they post on the website next to their user name. It is also conspicuously displayed on the user's profile webpage and signifies the user's status in the community.[10]

## 4. Data and Method

### 4.1. Data

The platform provides a list of all users who have earned reward points on the site. We obtained the complete list, which contains 117,174 users who have at least one reward point as of March 26, 2012. From this list, we drew a random sample of 2,000 individuals and collected their complete activity history between their first day on the site (when they registered) and March 26, 2012. We then construct an unbalanced panel data set from these answerers and their activities, where each observation records activities of one user in one week. As is common in online communities, not all users are constantly active. For those who stopped contributing, their total points no longer change, but the website retained their records. To account for this, we drop individual-period pairs four weeks[11] after each individual's last observed action (either answering or asking). This mitigates possible estimation bias from inactive periods. In our main analysis, we include only individual-period pairs that remained active after the introduction of the incentive hierarchy in 2007, since users who became inactive before that do not generate variations relevant for the incentive hierarchy.[12]

### 4.2. Main Variables and Summary Statistics

*Dependent variables.* The main outcome we are interested in is the level of effort that users exert, and we study how users' distance from hierarchical thresholds affects their efforts. We measure user effort using the number of questions that a user answers in each period (*Number of Questions Attempted*), no matter whether the answers are accepted as correct or not. The platform allows answerers to return to question pages and make additional comments before the asker closes the question. It is reasonable to assume that the cost of submitting the first comment is the highest, because it involves the cost of reading and understanding the question. We consider additional

[8] Award points for "A" will be multiplied by 4, "B" by 3, and "C" by 2. For example, if an answerer gets an 80% share from a 500-point question with a grade of "A," she earns $500 \times 0.8 \times 4 = 1{,}600$ points.

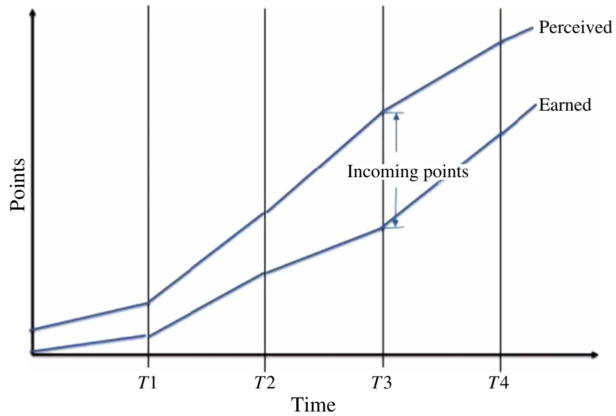
[9] In addition to answering questions, users may also write articles to earn points; however, such activities are extremely rare, and the points awarded in that manner are negligible. We therefore focus on questions and answers.

[10] The website also has a topic-specific hierarchy for each topic (e.g., SQL Server DBMS). This hierarchy is also based on rewards points but only those on a specific topic. However points and ranks within areas are only visible on the user's profile and are much less conspicuous than the overall hierarchy. The topics are also subject to change over time, as old technologies become obsolete and new technologies emerge. We focus on the overall hierarchy because it is more stable and much more likely to convey "status" and provide extrinsic motivations to contribute.

[11] Using a longer time period, such as eight weeks rather than four, yields consistent results.

[12] In a robustness test, we focused only on users who signed up before the introduction of the incentive hierarchy and obtained highly similar results. This suggests that users who signed up later are not inherently different than those who signed up earlier.

**Figure 2** (Color online) Illustration of Total Points and Incoming Points




comments as follow-up discussions for the same solution, so our outcome variable does not include questions that users had already answered in previous periods.[13] In our robustness tests, we explore alternative outcome variables.

*Independent variables.* The key independent variable that we are interested in is the user's distance from rank thresholds. A natural way of measuring this distance is the difference between the number of points required for each rank (such as 50,000 points for the first rank) and the total points that each user had earned at the beginning of each period (*TotalPoints*). However, this metric does not reflect an important feature of this website: There can be significant delays between the time that answerers submit their answers and the time that askers accept them. To meaningfully calculate the distance from goals as a motivator (or demotivator) on user contribution, we need to include the points that the users can *reasonably expect* to receive from the solutions that they submitted previously, even if their solutions have not yet been accepted, in addition to *TotalPoints*.[14] Figure 2 visualizes the incoming points due to the time delay between submission and acceptance of solutions.

To further illustrate this, suppose an answerer had 40,000 points at the beginning of period $t$, and suppose the threshold for the first level in the incentive hierarchy is 50,000 points. A naïve measurement of distance would be $40{,}000 - 50{,}000 = -10{,}000$ points. However, the user had submitted solutions to five questions in the previous period $t-1$, but the askers of those questions had not yet formally accepted any solutions. From their past experience on this site, the answerer expected that at least half of the answers they provided would solve the problems for the askers. This is possible not only because of the answerer's experience but also due to the technical nature of the site. Hence, the "*perceived*" distance that would affect the answerer's behavior in period ($t$) was not the naïve $-10{,}000$ points, but in fact shorter, because more points (which we call "incoming" points) would come from the solutions that she had already provided in the previous period. For simplicity, let us assume a 50% historical success rate for the answerer and a maximum 2,000 points per question; the distance that would more likely affect the answerer's behaviors should be ($-10{,}000 + 2{,}000 \times 5 \times 1/2 = -5{,}000$) points. Therefore, we measure the answerer's distance from thresholds as a *Modified Distance*

$$\begin{aligned} &\mathit{ModifiedDistance}\\ &\quad = \mathit{ModifiedTotalPoints} - \mathit{Threshold}\\ &\quad = \mathit{TotalPoints} + \mathit{IncomingPoints} - \mathit{Threshold}. \end{aligned} \tag{1}$$

More formally, we calculate the incoming points by multiplying the total number of possible points with the historical "success" rate of the answerer (in terms of probabilities of being selected as the right answer) and "quality" of the answer (in terms of the share of assigned points and grade received)

$$\begin{aligned} &\mathit{IncomingPoint}_{it}\\ &\quad = \mathit{SuccessRate}_{it} \times \mathit{Quality}_{it}\\ &\qquad \times \sum_{\{j\in J \mid \text{attempted by } i \text{ till } t\}} \mathit{Open}_{jt} \times \mathit{AssignedPoint}_j, \end{aligned} \tag{2}$$

$$\mathit{Open}_{jt} = \begin{cases} 1 & \textit{if } j \textit{ is solved by } i \textit{ and } \mathit{ClosureTime}_j > t,\\ 0 & \textit{if } j \textit{ is solved by } i \textit{ and } \mathit{ClosureTime}_j \le t,\\ \mathrm{Prob}(\mathit{ClosureTime}_j > t), & \\ \quad \textit{if } j \textit{ is not solved by } i, & \end{cases} \tag{3}$$

$$\begin{aligned} \mathrm{Prob}(\mathit{ClosureTime}_j > t) &= 1 - F(t - \mathit{AttemptTime}_{ij})\\ &= \exp(-\lambda(t - \mathit{AttemptTime}_{ij})). \end{aligned} \tag{4}$$

Here, $\mathit{SuccessRate}_{it}$ denotes user $i$'s perception of success rate at time $t$, which is updated by the user at each period. The term

$$\mathit{Quality}_{it} = \frac{1}{J} \sum_{\{j\in J \mid \textit{solved by } i \textit{ till } t\}} \mathit{Share}_{ij} \times \mathit{Grade}_{ij}$$

captures the variation in the share of assigned points and grades of previously answered questions, which also varies across user and time. The term

$$\sum_{\{j\in J \mid \textit{attempted by } i \textit{ till } t\}} \mathit{Open}_{jt} \times \mathit{AssignedPoint}_j$$

[13] This is not a restrictive assumption; results are highly consistent when we consider all comments as efforts.

[14] When we discard these "incoming" points, we find that users' effort levels start to decrease prior to reaching the goal, which lends further support to using this modified distance metric by taking into account these incoming points. Results of this test are reported in Online Appendix A (available as supplemental material at http://dx.doi.org/isre.2016.0635).

**Table 1 Description of Main Variables**

| Variable | Description | Mean | S.D. | Min | Max | Skewness |
|---|---|---|---|---|---|---|
| $NumAttempt_{it}$ | Number of questions participated (no matter solved) by user *i* at time *t* | 0.356 | 2.443 | 0.000 | 205.000 | 25.256 |
| $TotalPoint_{it}$ | Total point value (in thousands) user *i* has accumulated at time *t* | 17.824 | 43.196 | 0.000 | 299.916 | 3.713 |
| $IncomingPoint_{it}$ | Expected incoming points (in thousands) for user *i* at time *t* | 0.578 | 2.561 | 0.000 | 95.869 | 11.145 |
| *FirstGoal* | An indicator that equals to 1 if user *i* reaches first level goal at time *t* | 0.105 | 0.307 | 0.000 | 1.000 | 2.576 |
| $SecondGoal_{it}$ | An indicator that equals to 1 if user *i* reaches second level goal at time *t* | 0.031 | 0.174 | 0.000 | 1.000 | 5.378 |
| *AskCount* | Number of questions asked by user *i* at time *t* | 0.080 | 0.452 | 0.000 | 13.000 | 10.114 |
| $Tenure_{it}$ | Number of weeks user *i* has been in market since registration at time *t* | 215.016 | 148.027 | 0.000 | 791.000 | 0.651 |

*Note.* Number of individuals is 1,001 and number of observations is 103,129.

denotes the sum of points assigned by askers to user $i$'s open questions at time $t$, where $Open_{jt}$ is an indicator of whether question $j$ is still open at time $t$. If we observe that ultimately user $i$'s solution is chosen as correct, then this indicator is deterministic. Otherwise, we need to estimate how soon that question will be resolved (i.e., how soon the asker will choose one or more answers as correct). To account for this probability, we assume an exponential distribution with cumulative distribution function $F(\cdot)$ for the delay between the solution submission time ($AttemptTime_{ij}$) and the question closure time ($CloseTime_j$), with a parameter $\lambda$. Thus, we have $1-F(t-AttemptTime_{ij})$ as the probability of question $j$ being resolved after time $t$, which is equivalent to the probability of question $j$ being open at time $t$. Parameter $\lambda$ is calibrated using the empirical distribution of time delay from solved questions.[15] Our calculation of incoming points captures users' perception of their success rate, as well as the expected timing of closure for the unsolved questions. Hence, the *Modified Distance* metric better reflects the user's perception of relative distance to the rank thresholds.

We also include in our models a number of control variables such as *Tenure* (number of weeks since registration) and *Ask Count* (number of questions asked). The descriptions and summary statistics of the main variables used in our analysis are in Table 1. It shows that 1,001 individuals are used for the main analysis, and the unbalanced panel data set is comprised of 103,129 individual-period pairs. The mean weekly number of questions attempted by users is 0.356. Overall, 10.5% of individual-period pairs reached the first level goal, while 3.1% reached the second level. This indicates that the incentive hierarchy is indeed increasingly challenging. The average incoming point is 578.

### 4.3. Model Specification and Estimation

To test our hypotheses, we simultaneously investigate—in one model—changes in user effort levels as the user approach the goal (H1), at the time they reach the goal (H2), and after they have attained the goal (H3). In this model, the outcome variable is a function of the distance to the goals as well as other covariates. Inspired by the regression discontinuity design (Thistlethwaite and Campbell 1960), we use the following parametric polynomial model for direct effects of goals:

$$\begin{aligned} y_{it} = {} & (1-Goal_{it}) \times \sum_{p=1}^{\bar{p}} \beta_{1,p} Distance_{it}^{p} \\ & + Goal_{it} \times \sum_{p=1}^{\bar{p}} \beta_{2,p} Distance_{it}^{p} + \beta_3 Goal_{it} \\ & + \gamma_1 LogAskCount_{it} + \gamma_2 LogTenure_{it} \\ & + \delta\, MonthDummies + \alpha_i + \varepsilon_{it}. \end{aligned} \tag{5}$$

The outcome variable $y_{it}$ is the number of questions attempted by user $i$ at time $t$. The variable $Goal_{it}$ is an indicator of whether user $i$ reaches the first or second rank at time $t$, and its coefficient estimate tests H2. To test H1 and H3, we focus on the distance to rank thresholds (goals). The variable $Distance_{it}$ captures the distance to the focal goal along the incentive hierarchy for user $i$ at time $t$. Our specification allows for different impacts of distance on either side of the threshold, consistent with our hypotheses. If an individual has not reached the goal, $Goal_{it} \times \sum_{p=1}^{\bar{p}} \beta_{2,p} Distance_{it}^{p}$ equals zero, while the term $(1-Goal_{it}) \times \sum_{p=1}^{\bar{p}} \beta_{1,p} Distance_{it}^{p}$ appears; and vice versa. Thus, parameters of these terms represent user behaviors on either side of goal attainment. We estimate up to the third polynomial of distance, though we focus more on the linear and quadratic terms of distance for the following reasons. First, results from cubic or higher order models are very difficult to interpret intuitively. Second, as we will show, effect sizes associated with higher order terms are generally miniscule. Third, most papers in the goal setting literature use only linear terms, and a few use quadratic models (Kivetz et al. 2006). We do not know of papers using higher order polynomials. In fact, theoretical and empirical work related to *any* distance typically uses linear and quadratic terms of distance (e.g., Akerlof 1997, p. 1011), likely due

[15] Average delay for the solved questions is 33 days, therefore we use the maximum likelihood estimate $\lambda = 1/33$.

to their roots in the physics literature on the *law of gravitation*—which states that the amount of acceleration is proportional to the inverse of the square of distance between objects (Feynman 1963). The way that different ranks on an incentive hierarchy "accelerate" user efforts is highly analogous to the gravitational pull of an astronomical object. We therefore focus on linear and quadratic terms of distance.

We incorporate $LogTenure_{it}$ and $LogAskPoint_{it}$ as controls, as well as calendar month dummies to control for unobserved time trend and seasonal variations. Individual fixed effects are further included to control for time-invariant individual heterogeneity, and we assume that $\varepsilon_{it}$ is normally distributed with zero mean for estimation purposes.

Because the outcome variable can be expressed as a linear function of observables given parameters $\beta$, $\gamma$, and $\delta$, and an additively separable disturbance, all parameters can be obtained by applying the within estimator that differences out individual fixed effects. To illustrate the hypothesized goal effect graphically, we calculate the conditional mean function of the outcome variable given different distances from the goal threshold. We use the delta method to derive the standard errors for the 95% confidence intervals. The conditional mean of outcome given distance from goal is

$$\begin{aligned} E[y \mid Distance] &= Constant + (1 - Goal) \times \sum_{p=1}^{\bar{p}} \beta_{1,p} Distance^p \\ &\quad + Goal \times \sum_{p=1}^{\bar{p}} \beta_{2,p} Distance^p + \beta_3 Goal. \end{aligned} \tag{6}$$

To contrast the effects of different levels of ranks on the incentive hierarchy, we extend our main model to account for multiple ranks at the same time. We note that these goals are *nested*, meaning that what happens *after* the first goal is precisely what happens *before* the second goal. Therefore, we separately compare across ranks the three predictions of H4: What happens as users approach goals, upon goal attainment, and after goal attainment—and how each of these differs across ranks. Because of data availability (too few users passed the third rank), we focus on the first two levels only. More specifically, we consider the interval from 0 to 50,000 as "region 1," the interval from 50,000 to 150,000 as "region 2," and the interval between 150,000 and 300,000 as "region 3." We create two indicator variables for regions 1 and 3, respectively. Then we extend our main model to test the predictions on user behavior as they approach goals using Equation (7) and those after goal attainment using Equation (8). Our predictions on the contrast across ranks can be captured by the coefficient of the interaction term between region indicators and distance metrics

$$\begin{aligned} y_{it} &= \sum_{p=1}^{\bar{p}} \beta_{1,p} DistanceBefore_{it}^p \\ &\quad + \beta_2 Region1_{it} \times DistanceBefore_{it} \\ &\quad + \beta_3 Region3_{it} \times DistanceBefore_{it} \\ &\quad + \beta_4 Region1_{it} + \beta_5 Region3_{it} \\ &\quad + \gamma_1 LogAskCount_{it} + \gamma_2 LogTenure_{it} \\ &\quad + \delta MonthDummies + \alpha_i + \varepsilon_{it}, \end{aligned} \tag{7}$$

$$\begin{aligned} y_{it} &= \sum_{p=1}^{\bar{p}} \beta_{1,p} DistanceAfter_{it}^p \\ &\quad + \beta_2 Region1_{it} \times DistanceAfter_{it} \\ &\quad + \beta_3 Region3_{it} \times DistanceAfter_{it} \\ &\quad + \beta_4 Region1_{it} + \beta_5 Region3_{it} + \gamma_1 LogAskCount_{it} \\ &\quad + \gamma_2 LogTenure_{it} + \delta MonthDummies + \alpha_i + \varepsilon_{it}. \end{aligned} \tag{8}$$

To measure the *instantaneous* "drop" effect upon goal attainment and compare across ranks, we have to focus on a very short interval around goal attainment. The interval's length should approach zero to reflect the instantaneous change. We therefore use a weighted regression. Specifically, we use the region indicators mentioned above and add an indicator to capture whether the goal being attained is the second level goal. We then estimate a regression model of user efforts, treating this new indicator variable as a moderating variable. Most importantly, we estimate weighted regression, where the weight is the inverse of absolute distance (larger weights for observations close to the goal threshold, thereby capturing the *instantaneous* effect). The coefficient of the interaction term between this variable and the regions indicator will reveal whether the magnitude of drop is statistically different across these two ranks

$$\begin{aligned} y_{it} &= (1 - Goal_{it}) \times \sum_{p=1}^{2} \beta_{1,p} Distance_{it}^p \\ &\quad + Goal_{it} \times \sum_{p=1}^{2} \beta_{2,p} Distance_{it}^p \\ &\quad + \beta_3 Goal_{it} + \beta_4 Second_{it} \\ &\quad + \beta_5 Second_{it} \times (1 - Goal_{it}) \times Distance_{it} \\ &\quad + \beta_6 Second_{it} \times Goal_{it} \times Distance_{it} \\ &\quad + \beta_7 Second_{it} Goal_{it} + \gamma_1 LogAskCount_{it} \\ &\quad + \gamma_2 LogTenure_{it} + \delta MonthDummies \\ &\quad + \alpha_i + \varepsilon_{it}. \end{aligned} \tag{9}$$

# 5. Results

## 5.1. Effect of the First Rank

We first analyze the effect of the first goal using individual-period pairs that have not reached the second goal, so as to avoid confounding influences from the second goal. In addition, we use the modified distance metrics to incorporate incoming points. There are 994 individuals forming an unbalanced panel data with 99,895 individual-period pairs. Results are reported in Table 2.

The coefficient of the binary goal indicator provides an estimate for the effect of reaching the first level goal. In the linear, quadratic, and cubic specifications, the coefficients of *FirstGoal* are significantly negative, suggesting that the instantaneous effect of goal attainment is negative, consistent with H2. In terms of practical significance, on average and across specifications, each user reduces their effort by at least 1.79 questions per week upon reaching the first rank. Hence, H2 is supported.

We next turn to the distance measures. In the linear distance model, the acceleration rate for users before reaching the first goal is 0.04 ($p < 0.001$) questions per week per 1,000 extra points; thus, H1 is supported. Moreover, the acceleration rate after reaching the first goal is 0.02 ($p < 0.001$) questions per week per 1,000 extra points, which is smaller than that before goal attainment. We further conduct a Wald test to compare the coefficients, and it shows that the coefficient of *before* is statistically larger than that of *after* ($F = 95.75$). Therefore, H3 is also supported. The quadratic and cubic models yield consistent results. To illustrate the above results, we plot the conditional mean function of outcome at different levels of distance with 95% point-wise confidence intervals on the upper left panel of Figure 3. The vertical dotted green line denotes the rank threshold.

It can be seen that the conditional mean of outcome exhibits a significant increase from 0.0 to 2.0 (questions per week) before goal attainment. However, it drops to about 0 once the goal is reached, and then it increases at a lower rate till reaching the second goal. These results support H1 and H3. Findings on other covariates such as *LogAskCount* and *LogTenure* also offer some interesting insights into user behaviors. All else equal, users who ask more questions in a given period are also likely to answer more in that period, and users who are on the site longer tend to exert lower levels of effort per week. This may reflect user attrition or a more selective behavior in answering questions, or both.

In summary, our analysis of the incentive hierarchy's first rank demonstrates that users accelerate efforts before reaching the threshold. However, they significantly reduce their efforts once they obtain the rank. After that, working toward the next goal, they slowly recover their efforts but at a much slower rate. These findings are all consistent with our hypotheses.

**Table 2 Estimation of Panel Data Model (First Rank)**

| Distance fun. | None | Linear | Quadratic | Cubic |
|---|---|---|---|---|
| *FirstGoal* | | $-1.7922^{***}$ | $-2.0005^{***}$ | $-2.4250^{***}$ |
| | | (0.0712) | (0.0949) | (0.1108) |
| *BeforeFirst* | | $0.0362^{***}$ | $0.0500^{***}$ | $0.2492^{***}$ |
| | | (0.0013) | (0.0044) | (0.0074) |
| *BeforeFirst*$^2$ | | | $0.0002^{***}$ | $0.0113^{***}$ |
| | | | (7.49e−5) | (0.0003) |
| *BeforeFirst*$^3$ | | | | $0.0001^{***}$ |
| | | | | (4.26e−6) |
| *AfterFirst* | | $0.0185^{***}$ | $0.0253^{***}$ | $0.0578^{***}$ |
| | | (0.0012) | (0.0039) | (0.0079) |
| *AfterFirst*$^2$ | | | $-6.73\text{e}{-5}^{\dagger}$ | $-0.0008^{***}$ |
| | | | (3.71e−5) | (0.0002) |
| *AfterFirst*$^3$ | | | | $4.23\text{e}{-6}^{***}$ |
| | | | | (9.33e−7) |
| *LogAskCount* | $1.6613^{***}$ | $1.6337^{***}$ | $1.6356^{***}$ | $1.6272^{***}$ |
| | (0.0352) | (0.0351) | (0.0351) | (0.0349) |
| *LogTenure* | $-0.4113^{***}$ | $-0.4812^{***}$ | $-0.4760^{***}$ | $-0.5120^{***}$ |
| | (0.0145) | (0.0147) | (0.0148) | (0.0148) |
| AIC | 409,790.1 | 408,778.9 | 408,768.9 | 407,613.7 |
| BIC | 410,398.9 | 409,416.2 | 409,425.2 | 408,289.1 |

*Notes.* Individual fixed effects are included. The number of observations is 99,895 and the number of individuals is 994. Calendar month dummies are included, but their coefficients are not reported for brevity.

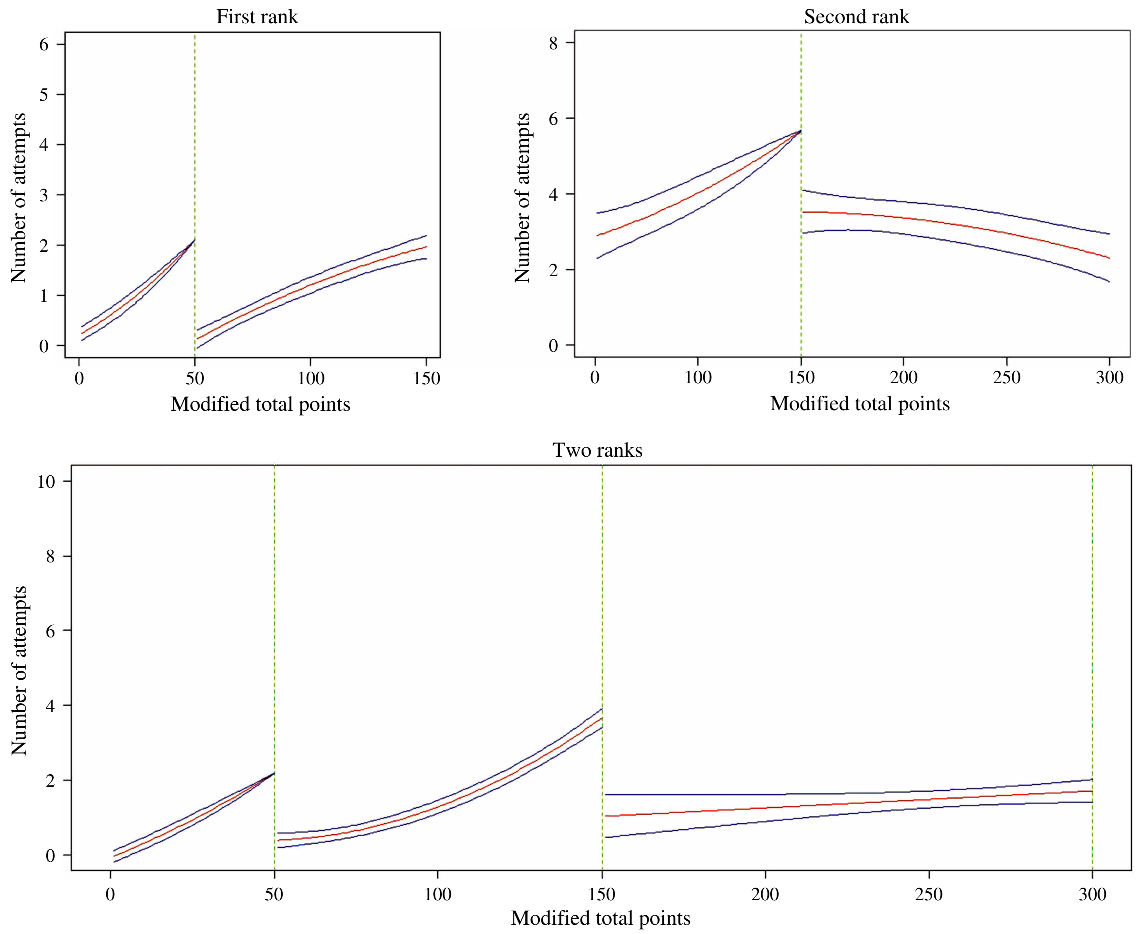
$^{\dagger}$Significant at 0.1; $^{***}$significant at 0.001.

## 5.2. Effect of the Second Rank

We next consider the effect of the second rank, focusing on users who have obtained the first rank but not the third. We similarly construct a panel data set with 136 individuals and 10,833 individual-period pairs. Users who never reached the first rank are automatically excluded, as no observation is available on their behaviors in this range. We estimate a similar set of models and present the results in Table 3.

The coefficient for the binary goal indicator in all specifications verifies that H2 is still supported, that users also reduce efforts upon obtaining the second rank. For the coefficient of distance in the linear distance specification, users increase effort level by 0.03 ($p < 0.001$) questions per week per 1,000 extra points before reaching the second goal, thus H1 is supported. Interestingly, users do not accelerate effort, but even slowly *decelerate* effort by 0.01 ($p < 0.001$) questions per week per 1,000 extra points after reaching the goal. The Wald test also shows that the coefficient for *Before* is statistically larger than that of *After* ($F = 109.76$). This is still qualitatively consistent with H3, but more importantly, it appears that the third rank is perhaps too far away to motivate users to recover their efforts. Results from the quadratic

**Figure 3** **(Color online) Conditional Mean Functions: Main Effects**




and cubic models are highly consistent. We further use the quadratic specification to graphically illustrate the results by calculating the conditional mean of outcome variables, as depicted in the upper right panel of Figure 3. This provides additional support for our hypotheses. The conditional mean increases from 3.0 to 5.5 questions per week before the goal is reached, while it dramatically decreases to 3.5 questions per week upon goal attainment. After that, users would even *decrease* effort level to 2.0 questions per week as they progress toward the third rank. Therefore, H1–H3 are supported for the second level goal as well.

### 5.3. Contrasting Effects of Two Sequential Goals

H4 contrasts the effect of different levels of goals on user contributions. We hypothesized that for higher ranks, users' increase in efforts while approaching the goal should be slower, and they will also be slower to recover after goal attainment. In Section 4.3 we described our empirical methods to test H4. Since only 44 users are above the third rank, we focus on comparing the effects of the first and second ranks. Results are reported in Tables 4 and 5, and graphically illustrated in the bottom panel of Figure 3.

We first compare the goal-gradient effect of the two ranks. Results are consistent whether we use the linear or the quadratic model. For simplicity, we focus on the linear one. We find that the acceleration rate toward the second rank is 0.01 ($p < 0.001$) fewer questions per week per 1,000 extra points (in distance) than the first rank, and for recovery it is 0.03 ($p < 0.001$) fewer for the second rank. These results are qualitatively consistent with H4. We then compare the instantaneous "drop" effect of goal attainment for these two goals. The coefficient of the interaction term is insignificant in both linear and quadratic distance specifications, suggesting that the amount of drop in efforts does not differ significantly between the higher and lower goals. It is the increase before goal attainment and the recovery after that show significant difference. The overall pattern, however, remains consistent with H4: the overall impact on user efforts is *smaller* for higher ranks.

Table 6 summarizes our main findings and how they correspond to the hypotheses.

**Table 3** **Estimation of Panel Data Model (Second Rank)**

| Distance fun. | None | Linear | Quadratic | Cubic |
|---|---|---|---|---|
| *SecondGoal* | | −1.8387*** | −2.1466*** | −1.8573*** |
| | | (0.2270) | (0.2947) | (0.3458) |
| *BeforeSecond* | | 0.0292*** | 0.0380*** | 0.0896*** |
| | | (0.0030) | (0.0079) | (0.0099) |
| *BeforeSecond*$^2$ | | | 0.0001 | 0.0022*** |
| | | | (8.22e−5) | (0.0003) |
| *BeforeSecond*$^3$ | | | | 1.75e−6*** |
| | | | | (2.01e−6) |
| *AfterSecond* | | −0.0086*** | −0.0008 | 0.0017 |
| | | (0.0024) | (0.0071) | (0.0154) |
| *AfterSecond*$^2$ | | | −4.88e−5 | −9.23e−5 |
| | | | (4.20e−5) | (0.0002) |
| *AfterSecond*$^3$ | | | | 2.37e−7 |
| | | | | (8.43e−7) |
| *LogAskCount* | 2.8266*** | 2.7373*** | 2.7300*** | 2.6782*** |
| | (0.2006) | (0.1992) | (0.1993) | (0.1988) |
| *LogTenure* | −2.5787*** | −2.2951*** | −2.2724*** | −2.1636*** |
| | (0.1646) | (0.1774) | (0.1779) | (0.1791) |
| AIC | 58,646.1 | 58,465.6 | 58,466.6 | 58,393.2 |
| BIC | 59,112.7 | 58,954.1 | 58,969.6 | 58,910.9 |

*Notes*. Individual fixed effects are included. The number of observations is 10,833 and the number of individuals is 136. Calendar month dummies are included, but their coefficients are not reported.
***Significant at 0.001.

**Table 4** **Estimation of Panel Data Model (Two Level Goals, Pre-Post Goal Attainment)**

| Variable | Pre-goal | | Variable | Post-goal | |
|---|---|---|---|---|---|
| Distance fun. | Linear | Quadratic | Distance fun. | Linear | Quadratic |
| *DistanceBefore* | 0.0330*** | 0.0423*** | *DistanceAfter* | 0.0330*** | 0.0246*** |
| | (0.0015) | (0.0024) | | (0.0013) | (0.0024) |
| *DistanceBefore*$^2$ | | 9.15e−5*** | *DistanceAfter*$^2$ | | 8.57e−5*** |
| | | (2.02e−5) | | | (2.08e−5) |
| *DistanceBefore* × *Region*1 | 0.0105*** | 0.0064** | *DistanceAfter* × *Region*1 | 0.0105*** | 0.0189*** |
| | (0.0020) | (0.0022) | | (0.0020) | (0.0028) |
| *DistanceBefore* × *Region*3 | −0.0263*** | −0.0228*** | *DistanceAfter* × *Region*3 | −0.0263*** | −0.0316*** |
| | (0.0018) | (0.0020) | | (0.0018) | (0.0022) |
| *Region*1 | −1.1690*** | −1.2674*** | *Region*1 | −0.0440 | −0.1621* |
| | (0.1096) | (0.1117) | | (0.0746) | (0.0799) |
| *Region*3 | −1.4838*** | −1.3877*** | *Region*3 | 0.8159*** | 1.0432*** |
| | (0.1459) | (0.1424) | | (0.1216) | (0.1335) |
| *LogAskCount* | 1.7498*** | 1.7506*** | *LogAskCount* | 1.7498*** | 1.7500*** |
| | (0.0383) | (0.0383) | | (0.0383) | (0.0383) |
| *LogTenure* | −0.5875*** | −0.5845*** | *LogTenure* | −0.5875*** | −0.5866*** |
| | (0.0163) | (0.0163) | | (0.0163) | (0.0163) |

*Notes.* Individual fixed effects are included. The number of observations is 103,129 and the number of individuals is 1,001. Calendar month dummies are included, but their coefficients are not reported.
*Significant at 0.05; **significant at 0.01; ***significant at 0.001.

## 6. The Effect of Introducing an Incentive Hierarchy

Our results so far show that users' efforts noticeably ramp up as they approach goals but also drop significantly upon goal attainment. A natural follow-up question is: What is the "net" effect on user efforts, then, of having an incentive hierarchy? This is a challenging question, and a full investigation would likely require a separate paper; but given how closely it is related to our main analyses above, we exploit the introduction of the incentive hierarchy on this website to provide some preliminary answers.

**Table 5 Estimation of Panel Data Model (Two Level Goal, Weighted)**

| Distance fun. | Linear | Quadratic |
|---|---|---|
| *Goal* | −1.7792*** | −1.2006*** |
| | (0.0681) | (0.0721) |
| *Before* | 0.1517*** | 0.2095*** |
| | (0.0019) | (0.0023) |
| *Before*$^2$ | | 0.0015*** |
| | | (3.48e–5) |
| *After* | −0.0065*** | −0.0474*** |
| | (0.00210) | (0.0036) |
| *After*$^2$ | | 0.0003*** |
| | | (3.95e–5) |
| *Second* | 0.4320*** | −0.0183 |
| | (0.0698) | (0.0702) |
| *Second* × *Goal* | 0.0718 | −0.0278 |
| | (0.1097) | (0.1108) |
| *Second* × *Before* | −0.1074*** | −0.0626*** |
| | (0.0028) | (0.0032) |
| *Second* × *After* | −0.0133*** | −0.0052 |
| | (0.0028) | (0.0032) |
| *LogAskCount* | 3.2810*** | 3.3136*** |
| | (0.0718) | (0.0712) |
| *LogTenure* | −1.6329*** | −1.4981*** |
| | (0.0368) | (0.0366) |

*Notes.* Weighted regression model with individual fixed effects. The number of observations is 110,728 and the number of individuals is 1,004. Calendar month dummies are included, but their coefficients are not reported.
***Significant at 0.001.

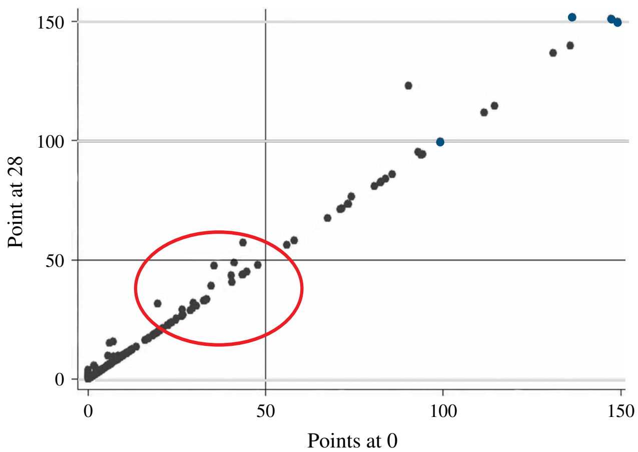
**Table 6 Summary of Hypotheses Test Results**

| Summary of hypothesis | Hypothesized parameter signs | Finding |
|---|---|---|
| H1: Before goal attainment: Closer to goal, more effort (tested separately for the first and second ranks) | $\beta_{11} > 0$ in Equation (5) | Supported (Tables 2 and 3) |
| H2: Upon goal attainment: Drop in efforts (tested separately for the first and second ranks) | $\beta_3 < 0$ in Equation (5) | Supported (Tables 2 and 3) |
| H3: After goal attainment: Effort recovers at a lower rate than acceleration prior to goal attainment (tested separately for the first and second ranks) | $\beta_{21} < \beta_{11}$ in Equation (5) | Supported (Tables 2 and 3) |
| H4: The higher the rank, the slower the user accelerates prior to goal attainment | $\beta_2 > 0$, $\beta_3 < 0$ in Equation (7) | Supported (Table 4) |
| The higher the rank, the smaller the drop upon goal attainment | $\beta_7 > 0$ in Equation (9) | Not supported (Table 5) |
| The higher the rank, the slower the user recovers their effort after goal attainment | $\beta_3 < 0$ in Equation (8) | Supported (Table 4) |

Prior to 2007, the website did not have an incentive hierarchy, even though it did have the point-based system. At the introduction of the incentive hierarchy, users who already had more than 50,000 points were automatically granted the title of the first rank—almost as a surprise. Those who had less than 50,000 points did not receive the title, but users now would know how far they were from the threshold. Because of the small number of users whose points were around the threshold, we are unable to conduct full-scale statistical analyses. Nonetheless, we exploit this event in several ways.

We first nonparametrically examine users' activities immediately before and after the incentive hierarchy was introduced. We plot each user's accumulated points at the time of introduction and their points as of 28 days after that (Figure 4). Each dot represents a user (possibly overlapping). If a user did not receive any points during that time, the corresponding dot will be on the 45-degree line. We find that users who were *just below* the threshold of 50,000 at the time of introduction *increased* their efforts and received rewards points. By contrast, users who were either far below the threshold or already above it at

**Figure 4** (Color online) Total Points at Day 0 and 28 of Incentive Hierarchy Introduction




the time of introduction barely received any rewards points during that time window. In fact, 63% of users who had between 50,000 and 150,000 points at the time of introduction did not answer any questions during the 28 days after they received their new title. This finding, while only descriptive, is consistent with our previous findings. More important, it shows that introducing the incentive hierarchy only impacted a small number of users.

We next examine the time series of several variables that reflect the vitality of the community (i.e., the number of questions attempted by all members and the average number of comments per question) around the time that the incentive hierarchy was introduced. The time series model is

$$\begin{aligned} y_t = \alpha + \beta_1\, AfterShock_t + \beta_2\, TimeTrend \\ + \beta_3\, TimeTrend \times AfterShock_t \\ + \beta_4 y_{t-1} + \beta_5 y_{t-2} + \varepsilon_t. \end{aligned} \tag{10}$$

We focus on a 200-day window (100 days before and after the introduction, respectively) for this analysis and present the results in Table 7. Across different specifications, the total number of attempts (comments) from all members *decreased* by at least 50 comments per week and about 1 comment per question. The introduction of incentive hierarchy did not increase user efforts, at least in the short term.[16]

To further delve into individual user efforts, we turn to a parametric test using individual-week observations. We are interested in how users' behaviors change as a result of the incentive hierarchy introduction. We estimate the following model using observations 20 weeks before and after the introduction and with less than 150,000 points[17]

$$\begin{aligned} y_{it} = \alpha + \beta_1\, AfterShock_t + \gamma_1\, LogAskCount_{it} \\ + \gamma_2\, LogTenure_{it} + \gamma_3\, Time + \varepsilon_{it}. \end{aligned} \tag{11}$$

Here, $AfterShock_t$ captures how the introduction of the hierarchy affects the number of attempts that a user makes. For robustness, we estimate this model both with and without individual fixed effects, and also try 20, 10, or 5 weeks before and after for the sampling time frame. As can be seen across different specifications in Table 8, the coefficient of $AfterShock_t$ is either statistically insignificant or significantly negative but with small effect size. The introduction of the hierarchy, at least in the short time windows we study, was largely ineffective in inducing user efforts. This is also consistent with the previous analyses and suggests that despite the popularity of incentive hierarchies, their effectiveness may not be as strong as website administrators would hope.

Some caveats should be noted when interpreting these results. First, our analysis focuses on the short term, yet there might be longer-term benefits of introducing this hierarchy. Second, although this exogenous change offers an opportunity for empirical tests, the thresholds were not randomly determined but decided by the platform, likely by considering the distribution of users' contribution levels around each cutoff. Even though such deliberations should have led to a more positive outcome and our estimates are likely conservative, a full investigation of the effect of incentive hierarchy warrants a separate paper. Our results therefore should be taken as preliminary evidence that incentive hierarchies may not be as effective in inducing efforts as their popularity might suggest.

## 7. Robustness

### 7.1. Alternative Independent Variable: Time-Based Distance Metric

In our main analysis, we measure users' distance to goals using the difference between their modified total points and the threshold specified by the incentive hierarchy. An alternative measure, used in Anderson et al. (2013), is the number of time periods before or after achieving goals (i.e., distance in time). This approach does not require the calculation of expected incoming points. To test whether our

[16] One potential confounding factor is that the website also changed its user interface at that time. However, the aesthetic improvement should have improved user participation. Our estimates are therefore conservative. In addition, replicating the analysis on the *asking* behavior of users reveals that the number of questions *asked* per time period did not change with the introduction.

[17] Very few users pass the second threshold at that time; therefore, there are not enough observations to estimate the effect of the second goal.

**Table 7 The Overall Effect of the Incentive Hierarchy (Answering Behaviors)**

| Dependent var. | Total number of comments from all members | | | Average number of comments per question | | |
|---|---|---|---|---|---|---|
| Spec. | 1 | 2 | 3 | 4 | 5 | 6 |
| $AfterShock_t$ | −72.1545** | −50.6589* | −55.2139** | −0.9772** | −0.9535** | −0.9158* |
| | (23.6881) | (22.7820) | (22.7390) | (0.3595) | (0.3683) | (0.3767) |
| $TimeTrend$ | 0.1239 | 0.0226 | 0.1227 | −0.0023 | −0.0030 | −0.0036 |
| | (0.2865) | (0.2748) | (0.2767) | (0.0043) | (0.0044) | (0.0045) |
| $TimeTrend \times AfterShock_t$ | −0.1487 | 0.0123 | −0.0994 | 0.0035 | 0.0041 | 0.0047 |
| | (0.4082) | (0.3890) | (0.3889) | (0.0062) | (0.0063) | (0.0063) |
| $Lag1$ | | 0.3193*** | 0.3589*** | | 0.0007 | −0.0030 |
| | | (0.0671) | (0.0710) | | (0.0715) | (0.0719) |
| $Lag2$ | | | −0.0702 | | | 0.0216 |
| | | | (0.0703) | | | (0.0717) |
| Constant | 221.3454*** | 149.3293*** | 159.1095*** | 5.5444*** | 5.5174*** | 5.3972*** |
| | (16.5825) | (21.4524) | (23.5532) | (0.2517) | (0.4722) | (0.6199) |
| $N$ | 201 | 200 | 199 | 201 | 200 | 199 |
| Adjusted $R^2$ | 0.1282 | 0.2199 | 0.2261 | 0.1315 | 0.1301 | 0.1285 |

*Note.* We run a linear regression model with lagged dependent variables.
*Significant at 0.05; **significant at 0.01; ***significant at 0.001.

**Table 8 Effect of Incentive Hierarchy Introduction on the Number of Attempts**

| Time window (weeks) | [−20, 20] | | [−10, 10] | | [−5, 5] | |
|---|---|---|---|---|---|---|
| Specification | OLS | FE | OLS | FE | OLS | FE |
| $AfterShock$ | −0.0920* | −0.0550 | −0.1046* | −0.0703 | −0.0801 | −0.0686 |
| | (0.0361) | (0.0363) | (0.0520) | (0.0646) | (0.0667) | (0.0611) |
| $LogAskCount$ | 1.5247*** | 1.3694*** | 1.6427*** | 1.4104*** | 1.6157*** | 1.3264*** |
| | (0.0789) | (0.0974) | (0.1289) | (0.1365) | (0.1701) | (0.1746) |
| $LogTenure$ | −0.1953*** | −0.9678*** | −0.1611*** | −0.9618*** | −0.2000*** | −1.9097*** |
| | (0.0292) | (0.3057) | (0.0239) | (0.2177) | (0.0385) | (0.4619) |
| $TimeTrend$ | −0.0020 | 0.0099*** | 0.0042 | 0.0128* | 0.0001 | 0.0204† |
| | (0.0015) | (0.0031) | (0.0045) | (0.0056) | (0.0112) | (0.0116) |
| Individual fixed effects | No | Yes | No | Yes | No | Yes |
| Constant | 1.1514*** | | 0.9755*** | | 1.1695*** | |
| | (0.1543) | | (0.1346) | | (0.2089) | |
| $N$ (No. of users) | 20,373 | 20,373 | 10,443 | 10,443 | 5,467 | 5,467 |
| | (568) | (568) | (532) | (532) | (513) | (513) |

*Notes.* We run linear models on $NumAttempt_{it}$. Standard errors are shown in parentheses.
†Significant at 0.1; *significant at 0.05; ***significant at 0.001.

results are robust under this distance measure, we modify our model in the following way:

$$\begin{aligned} y_{it} = {} & (1 - Goal_{it}) \times \sum_{p=1}^{\bar{p}} \beta_{1,p}\, TimeDistance_{it}^{p} \\ & + Goal_{it} \times \sum_{p=1}^{\bar{p}} \beta_{2,p}\, TimeDistance_{it}^{p} \\ & + \beta_3\, Goal_{it} + \gamma_1\, LogAskCount_{it} \\ & + \gamma_2\, LogTenure_{it} + \delta\, MonthDummies \\ & + \alpha_i + \varepsilon_{it}. \end{aligned} \tag{12}$$

Although the results (reported in Online Appendix C) are qualitatively consistent with our main findings, we retain the point-based distance as the main specification for two reasons. First, the time-based distance metric assumes that users can *perfectly* anticipate the exact time that they will reach goals, including whether and when askers will accept their answers, the share of assigned points, the grade on their answers, the "supply" of questions in future periods, and their competition outcome against other answerers. These are hard to justify. Second, the time-based distance metric cannot be computed for users who were eventually unable to attain the goals, so this specification reduces the statistical power of our tests. For these important reasons, we retain the point-based distance in our main analysis.

### 7.2. Alternative Dependent Variables: Number of Points Attempted and Number of Questions Solved

In our main specifications we use the *Number of Questions Attempted* as the outcome variable. Alternatively, we could measure user efforts by the *Number of Points Attempted*, the point value of questions attempted, or the *Number of Questions Solved* by the users. We replicate the main analysis with these alternative outcome variables, and the results are highly consistent (see Online Appendix D).

## 8. Discussions, Managerial Implications, and Future Research

This paper draws on goal-setting and status hierarchy theories to examine the effect of a platform-designated, status-based incentive hierarchy on user contributions in an online knowledge exchange. Our results indicate that individuals exert more effort before reaching goals; the closer they are to the goal, the more effort we observe. However, their efforts significantly drop upon goal attainment. In addition, the higher the rank to be achieved, the *less* the influence on contributions. Furthermore, by analyzing data around the time that the website implemented the incentive hierarchy, we find that the overall impact of incentive hierarchy on user efforts appears almost negligible. Our study is among the first to empirically examine such status-based incentive hierarchies in the context of online UGCs. It contributes to the goal-setting literature by identifying several important boundary conditions, and by extending it to a setting with consecutive and increasingly more challenging goals. It also contributes to the large but still growing information systems literature on user motivations and behaviors in online communities.

Our study has some data limitations that should be acknowledged. Although our original random sample contains 2,000 users drawn from a master list of answerers, the effective number of users in the estimations is often much smaller. This is because empirical tests on the effect of a rank can only be carried out on users who were able to obtain that rank, and that number is very small at the third rank and above. These small numbers of users do not allow us to estimate more flexible models, or to examine and contrast the effect of higher levels of ranks beyond the second rank. Nonetheless, the methods in this paper can be easily replicated when larger data sets are available.

Despite these limitations, our study has important managerial implications. First and foremost, the popularity of incentive hierarchies does not mean that they are necessarily effective or that they should absolutely be implemented. Our main empirical analyses show that users increase their efforts as they approach goals but that that impetus is almost gone when they reach the goals. Our analysis of the incentive hierarchy's introduction also shows that the overall effect of introducing a hierarchy appears very small. While additional research is needed to address the limitations of our study, our findings are consistent with existing studies, such as Lepper et al. (1973), that caution against the detrimental effect of extrinsic rewards, of which the "glory" bestowed by the ranks on an incentive hierarchy is an excellent example.

Second, if a website is to implement an incentive hierarchy, then our finding on the relative impact of different levels should be taken into account. While our data set only allows us to compare the first two levels of ranks, the pattern is quite clear that higher ranks show less impact than lower ranks. That is, not only are the motivational effect of goals only visible prior to goal attainment, even that temporary effect dwindles as we move up the hierarchy. This finding suggests that a larger number of levels is not necessarily more effective. In addition, there remain many open questions regarding the "optimal" design of incentive hierarchies. For example, should incentive hierarchies be mandated on the whole site in the first place? Should users be allowed to "opt out" in terms of showing the number of points or the titles? Is there an optimal distance from the start to the first rank, and then to the second rank, and so on? Is there an optimal total number of ranks? What happens when many users already achieve the top honor—should new ranks be introduced at that point, essentially further increasing the "tallness" of the hierarchy? Should there be additional accolades such as "top contributors" that provide even greater visibility and recognition than ranks?[18] In addition to future academic research, website administrators can also engage their members and conduct experiments to identify the best implementation of incentive hierarchies.

Third, while previous research has shown that having a hierarchy of statuses can be beneficial (Anderson et al. 2015), incentive hierarchy is just one way to create a status hierarchy. It is a very mechanical one and turns out not to be very effective. It is mechanical because the website administrators *have to* "promote" a member's status once a threshold is reached, no matter how they get there, how long it takes the member to get there, what they do before that, and what they do after. Because of that, these thresholds create what the literature calls "expected" extrinsic rewards (Lepper et al. 1973), which have been shown to dampen intrinsic motivations. There are many other ways that will lead to a pecking order of statuses

[18] We thank the associate editor and an anonymous reviewer for this suggestion.

among members, such as a bottom-up voting mechanism among regular members of a site, or even a top-down approach such as commendations from website administrators. Website administrators can test these or other novel methods to create more flexible, less "expected," and more effective status hierarchies that induce efforts in a more sustainable way.

As is well known, UGC sites and knowledge exchanges are ultimately based on voluntary contributions, so fostering a sense of community and nurturing members' intrinsic motivations should perhaps take a higher priority over mechanisms that cater to extrinsic motivations. More broadly, our study also informs the practice of other forms of user and customer interactions, such as crowdsourcing, open innovations, and the recent trend of "gamification" in various contexts. Although the "public goods" problem of user contribution is less salient in some of those contexts since users are able to *internalize* some of their efforts, how incentive hierarchies interact with user motivations to induce their efforts can be a fertile area for future research. Ultimately, a well designed incentive system of a UGC community should be able to, first and foremost, cultivate and preserve intrinsic motivations of its members, and then if needed, institute extrinsic rewards that do not interfere with or contaminate those intrinsic motivations.

## Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/isre.2016.0635.

## Acknowledgments

The authors thank the senior editor, the associate editor, and the three anonymous reviewers for their constructive comments. The authors also thank seminar participants at the University of Arizona, University of Pennsylvania, University of Texas at Austin, Michigan State University, and the University of Copenhagen for their suggestions. All errors remain our own.

## References

Akerlof GA (1997) Social distance and social decisions. *Econometrica* 65(5):1005–1027.

Anderson A, Huttenlocher D, Kleinberg J, Leskovec J (2013) Steering user behavior with badges. *Proc. 22nd Internat. Conf. World Wide Web* (ACM, New York), 95–106.

Anderson C, Brown CE (2010) The functions and dysfunctions of hierarchy. *Res. Organ. Behav.* 30:55–89.

Anderson C, Kilduff GJ (2009) Why do dominant personalities attain influence in face-to-face groups? The competence-signaling effects of trait dominance. *J. Personality Soc. Psych.* 96(2):491–503.

Anderson C, Hildreth JAD, Howland L (2015) Is the desire for status a fundamental human motive? A review of the empirical literature. *Psych. Bull.* 141(3):574–601.

Anderson C, Srivastava S, Beer JS, Spataro SE, Chatman JA (2006) Knowing your place: Self-perceptions of status in face-to-face groups. *J. Personality Soc. Psych.* 91(6):1094–1110.

Bagozzi RP, Dholakia UM (2006) Open source software user communities: A study of participation in Linux user groups. *Management Sci.* 52(7):1099–1115.

Baumeister RF, Leary MR (1995) The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psych. Bull.* 117(3):497–529.

Bavelas A (1950) Communication patterns in task oriented groups. *J. Acoustical Soc. Amer.* 22(6):725–730.

Bonezzi A, Brendl CM, DeAngelis M (2011) Stuck in the middle: The psychophysics of goal pursuit. *Psych. Sci.* 22(5):607–612.

Cheema A, Bagchi R (2011) The effect of goal visualization on goal pursuit: Implications for consumers and managers. *J. Marketing* 75(2):109–123.

Chen Y, Harper FM, Konstan J, Li SX (2010) Social comparisons and contributions to online communities: A field experiment on MovieLens. *Amer. Econom. Rev.* 100(4):1358–1398.

Dewey R (2007) *Psychology: An Introduction*. http://www.intropsych.com/index.html.

Earley PC, Connolly T, Ekegren G (1989) Goal, strategy development, and task performance: Some limits on the efficacy of goal setting. *J. Appl. Psych.* 74(1):24–33.

Feynman RP (1963) *The Feynman Lectures on Physics* (Addison Wesley, Reading, MA).

Goes PB, Lin M, Au Yeung CM (2014) "Popularity effect" in user-generated content: Evidence from online product reviews. *Inform. Systems Res.* 25(2):222–238.

Hann IH, Roberts JA, Slaughter SA (2013) All are not equal: An examination of the economic returns to different forms of participation in open source software communities. *Inform. Systems Res.* 24(3):520–538.

Heath C, Larrick RP, Wu G (1999) Goals as reference points. *Cognitive Psych.* 38(1):79–109.

Hilgard ER, Bower GH (1975) *Theories of Learning*, 4th ed. (Prentice Hall, Englewood Cliffs, NJ).

Hollenbeck J, Klein HJ (1987) Goal commitment and the goal-setting process: Problems, prospects, and proposals for future research. *J. Appl. Psych.* 72(2):212–220.

Huang SC, Zhang Y (2011) Motivational consequences of perceived velocity in consumer goal pursuit. *J. Marketing Res.* 48(6):1045–1056.

Hull CL (1932) The goal-gradient hypothesis and maze learning. *Psych. Rev.* 39(1):25–43.

Jabr W, Mookerjee R, Tan Y, Mookerjee VS (2014) Leveraging philanthropic behavior for customer support: The case of user support forums. *MIS Quart.* 38(1):187–208.

Kivetz R, Urminsky O, Zheng Y (2006) The goal-gradient hypothesis resurrected: Purchase acceleration, illusionary goal progress, and customer retention. *J. Marketing Res.* 43(1):39–58.

Koo M, Fishbach A (2008) Dynamics of self-regulation: How (un)accomplished goal actions affect motivation. *J. Personality Soc. Psych.* 94(2):183–195.

Li Z, Huang KW, Cavusoglu H (2012) Can we gamify voluntary contributions to online Q&A communities? Quantifying the impact of badges on user engagement. *Proc.* 2012 *Workshop Inform. Systems Econom.* (*WISE* 2012), *Orlando, FL.*

Liberman N, Förster J (2008) Expectancy, value and psychological distance: A new look at goal gradients. *Soc. Cognition* 26(5):515–533.

Lepper MR, Greene D, Nisbett RE (1973) Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *J. Personality Soc. Psych.* 28(1):129–137.

Lewin K, Dembo T, Festinger L, Sears PS (1944) Level of aspiration. Hunt JM, ed. *Personality and the Behavior Disorders*, Vol. 1 (Ronald Press, New York), 333–378.

Locke EA (1967) Motivational effects of knowledge of results: Knowledge or goal setting? *J. Appl. Psych.* 51(4):324–329.

Locke EA (1968) Toward a theory of task motivation and incentives. *Organ. Behav. Human Performance* 3(2):157–189.

Locke EA, Latham GP (1990) *A Theory of Goal Setting and Task Performance* (Prentice Hall, Englewood Cliffs, NJ).

Locke EA, Latham GP (2002) Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *Amer. Psychologist* 57(9):705–717.

Louro MJ, Pieters R, Zeelenberg M (2007) Dynamics of multiple-goal pursuit. *J. Personality Soc. Psych.* 93(2):174–193.

Ma M, Agarwal R (2007) Through a glass darkly: Information technology design, identity verification, and knowledge contribution in online communities. *Inform. Systems Res.* 18(1):42–67.

Mento AJ, Steel RP, Karren RJ (1987) A meta-analytic study of the effects of goal setting on task performance: 1966–1984. *Organ. Behav. Human Decision Processes* 39(1):52–83.

Miller NE (1944) Experimental studies of conflict. Hunt JM, ed. *Personality and the Behavior Disorders*, Vol. 1 (Ronald Press, New York), 431–465.

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) Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Ed. Psych.* 25(1):54–67.

Sheldon KM (2011) Integrating behavioral-motive and experiential requirement perspectives on psychological needs: A two-process model. *Psych. Rev.* 118(4):552–569.

Thistlethwaite D, Campbell D (1960) Regression-discontinuity analysis: An alternative to the ex-post facto experiment. *J. Ed. Psych.* 51(6):309–317.

Tubbs ME (1986) Goal setting: A meta-analytic examination of the empirical evidence. *J. Appl. Psych.* 71(3):474–483.

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

Zhang C, Hahn J, De P (2013) Continued participation in online innovation communities: Does community response matter equally for everyone? *Inform. Systems Res.* 24(4):1112–1130.

Zhang Y, Huang SC (2010) How endowed versus earned progress affects consumer goal commitment and motivation. *J. Consumer Res.* 37(4):641–654.