User Motivation and the Effects of Multitasking: An Analysis of an Online CQA Forum

Jed DeVaro* Jin-Hyuk Kim† Liad Wagman‡ Ran Wolff§

February 24, 2015

Abstract

Large-scale knowledge sharing has become a ubiquitous online phenomenon. Collaborative Question Answering (CQA) forums often depend on volunteers' provision of informational content. We examine a rich dataset of user contributions to study user motivation and its effects on multitasking performance. We find that most users on the platform we study behave in a manner consistent with the incentives to win contests rather than to make public-good contributions. Further, for those users who are relatively more contest-motivated, multitasking across subject categories is associated with lower performance levels, and more so when answering activities are spread over time.

Keywords: User-generated content; online forum; user motivation; multitasking; crowdsourcing

*California State University, East Bay. E-mail: jed.devaro@csueastbay.edu
†University of Colorado at Boulder. E-mail: jinhyuk.kim@colorado.edu
‡Illinois Institute of Technology. E-mail: lwagman@stuart.iit.edu
§Yahoo Labs, Haifa. E-mail: ranw@yahoo-inc.com

1 Introduction

Much of online content is generated and provided to websites by user-contributors, who are not employed by the firms running those sites. Such users' contributions of time and content can be important determinants of a firm's value (Brynjolfsson et al., 2014). A number of authors have analyzed data from Wikipedia, the online user-contributed encyclopedia, by bringing in economic perspectives; however, there are many other knowledge-sharing forums and there is a paucity of quantitative, economic analysis that studies them. This paper aims to understand some aspects of user contribution and performance within an economic framework, by analyzing a large set of data from an online question-answering forum.

The forum we analyze is *Yahoo Answers*. It is the second most popular such website in terms of traffic according to the measurement of Alexa.com.[1] On this forum, the number of open questions awaiting answers is vast at any moment, facilitating the continuous exchange of large amounts of information.[2] The data set we extracted is detailed and at the user level. Specifically, our data set tracks a cohort of nascent (new) users over a 12-month period and records all of their answering activities and the timing of those submissions. We first reconcile the pattern of user contributions with two well-known economic theories, and then account for the effect of user motivation on multitasking performance.

The first theory we consider is the framework proposed by Bergstrom et al. (1986) concerning the private provision of public goods. That work and the voluminous literature

[1]The first-ranked website in the 'Reference' category is Stackoverflow.com. *Stackoverflow* specializes in questions and answers for programmers while *Yahoo Answers* spans a full spectrum of general subject categories.

[2]This contrasts with Wikipedia, where articles tend to stabilize and mature leaving less opportunity to contribute after much of the low-hanging fruit has been picked. Although there may be an uncharted terrain on Wikipedia where users might contribute, identifying the subset of that terrain that matches a user's knowledge set involves relatively high search costs.

that followed it suggests that it is possible to overcome the free-riding problem when making voluntary contributions to public goods in a dynamic environment. For instance, Marx and Matthews (2000) show the existence of an equilibrium that achieves an asymptotically efficient outcome, where a person's response to previous contributions is affected by the cumulative level of contributions. Although these models focus on monetary contributions, their implications for user-generated content are quite plausible because the contributed information is a public good generating positive externalities for those who search for an answer to the same question (e.g., Jian and MacKie-Mason, 2012).

The second theory is that answering questions in a public forum resembles a contest (e.g., Morgan and Wang, 2010). Indeed, creating the perception that answering questions is a contest is arguably one of the reasons that *Yahoo Answers* became a successful forum. Promoting the idea that user contribution is a game was achieved by creating a "winner" in the sense that each question is organized as a contest where volunteer answerers compete to have their answers selected as the "best answer" by the person who originally posted the question (henceforth, "the asker"), or by a community vote. Those who win best answer are rewarded with virtual points, and a user's accumulated points and level (defined by ranges of points) are a publicly visible indicator of their social status on the forum. This contest motivation may be a source of user motivation to post answers.

We proxy user "effort" by the length of the answers, which we measure in two ways (number of characters, and number of words) that yield very similar results. Particular cases in which posting length is a poor proxy for effort can be easily imagined. For instance, a well thought out answer and a poorly thought out one may be of the same length. But in a typical case, it seems reasonable to assume that a longer response consumes more time and hence effort. To be clear, we are not trying to capture answer "quality" with our measure of effort.[3] Our purpose is to understand whether the data is broadly consistent with either one of the two aforementioned economic theories.

The two theories of user motivation have opposite predictions concerning a user's contribution patterns, which suggests a simple empirical specification that nests both theories and that allows us to see which theory better describes the data. Of course, we do not argue that these two theories are the only, literal interpretation of a user's inner thought process. Further, a user's browsing behavior might account for data patterns, but the data generating process we refer to is based on theories rather than individual behavior. In fact, surveys that directly ask the users for reasons for their online activities might reveal individual-level, real-life motivations. Nonetheless, the two theories on which we focus are and have been of particular interest to economists, as they contrast the general ideas of altruism and selfishness of economic agents.

The empirical specification captures the best-response function of a given user in terms of effort (proxied by posting lengths) to that of the collective effort already posted in a question thread at the time of the given user's submission. The contributions of users who are likely motivated by public-good provisions (contests) are decreasing (increasing) in the collective efforts already invested by others. We find that the data is consistent with most users having a contest motivation rather than a public-good motivation. Further, we show that users with a public-good motivation cease their answering activities (and hence drop out of the contest) earlier than those who are more contest motivated, and the estimated

[3]That is, the relation between text content and answer quality is not the topic of our paper; we are also unable to objectively evaluate millions of answers by hiring independent third-parties due to budget constraints.

preference parameter capturing this behavior seems stable over time among users. With the point system, the platform seems to attract mostly contest-motivated contributors.

We essentially treat this coefficient estimate at the individual level, which predicts different observable posting behaviors, as the user's preference parameter. One might object that we do not have a plausible identification strategy for this parameter. Yet, our belief is that the endogeneity is less likely to be a concern in our analysis because of the sequential nature of posting activities. That is, the amount of existing contributions are predetermined from the standpoint of potential contributors. We also think that the selection (participation) into answering a particular question is plausibly exogenous holding constant the user's domain expertise. This is because the website lists the most recently posted (open) questions at the top, so that it takes time for users to browse questions far down the list. Thus, users arrive and only consider a subset of questions that are exogenous.

One of the well-known behavioral patterns of online user activities is multitasking (e.g., contributing answers across multiple subject categories in a given time period). The literature has noted pros and cons associated with multitasking; however, existing evidence of the effects of multitasking on performance is mostly based on lab experiments. In our context, the incentive mechanism is the "best answer" point system, and we hypothesize that the effects of multitasking on performance would be stronger for those who are contest motivated than for those users who are motivated by public-good provision. A key feature of the platform we study is that the reward (ten points for winning "best answer") is the same across subject categories. This is convenient for us because it means that the coefficient on multitasking will not reflect different underlying incentive schemes.[4]

[4]If the "best answer" award can be a different number of points for different categories or over time,

We measure multitasking in two dimensions over the history of users' activities: across subject categories and across time. In our data, contest-motivated users tend to have lower performance metrics. That is, the number of completed answers (or best answers) for a given posting length is lower for contest-motivated users than for public-goods-motivated users. Consistent with our predictions, this performance gap is larger at the high levels of multitasking across subject categories, and this is particularly so when users spread their answering activities more equally over time. The former result indicates that when users multitask they become jacks-of-all-trades, and this is a problem for contest-motivated users. The latter implies that multitasking further lowers contest-motivated users' performance when they engage in answering activities in a ubiquitous fashion.

2 Related Literature

A large literature in computer science examines user behavior in online conversational spaces such as Usenet, where users read messages from and post messages to threaded discussions. This literature has uncovered a rich pattern of user interactivity. For instance, Whittaker et al. (1998) analyze millions of messages collected from newsgroups to find the relationship between newsgroup characteristics (e.g., number of messages and participants) and conversational characteristics (e.g., message length and thread depth). More recently, researchers have used various visualization techniques to study the relationship between role behavior and social-network characteristics (Welser et al., 2007) as well as to cluster forum categories

then the effect of multitasking on performance would need to control for the different incentives each user might be facing. Hence, it is beneficial not to have this variation created by differential incentives in our cross-sectional analysis.

according to content characteristics and patterns of user interactions (Adamic et al., 2008).

The paper closest to ours in this literature is Backstrom et al. (2008), which examines user groups of varying degrees of engagement. They employ a large data set from an online message board, classify users by the level of patronage and contribution, and study their respective behaviors. Their main finding is that the higher the number of groups to which a user belongs, the less likely it is that the user can be heavily engaged in all of them; moreover, new users who will become active members of a group tend to receive preferential treatment early on from other active group members, possibly because of their personal relationships and/or fit with the group. While our analysis examines a similar context of online user contribution, the difference is that we focus on understanding some of these patterns by applying economic theories.

We aim to build a bridge between this literature and economics. The two economic theories underlying our analysis (i.e. contests and public-goods provision) connect broadly to multiple literatures in economics and beyond. For instance, the tournament theory (e.g., Lazear and Rosen, 1981) relates to our setting as multiple players compete for a smaller number of prizes and outcomes are determined by the players' relative performances. One study that is closest to ours in the literature on information technology and labor productivity is Aral et al. (2012), where the effects of multitasking on employee productivity are mediated by knowledge networks. Specifically, the authors find that the workers who have access to a more evenly distributed expertise in the company are more productive when juggling multiple projects.

The theory of public-goods provision naturally relates to intrinsic motivation (e.g., Kreps, 1997 and Besley and Ghatak, 2005, explore theoretically the idea that this intrinsic motivation in nonprofits can substitute for explicit rewards). For instance, in the nonprofit sector, many individuals volunteer to work for lower pay than they would receive elsewhere. There is also an experimental literature in psychology establishing that explicit reward can crowd out intrinsic motivation, and in economics, Frey (1997) identifies conditions under which explicit rewards crowd out intrinsic motivation. Thus, if rewards such as "best answer" crowd out intrinsic motivation, then the approach may backfire, because users who would otherwise be driven by the provision of a public good may lose that motivation and drop out early.

Recent empirical studies have focused on identification of group size effects using exogenous variations. For instance, Zhang and Zhu (2011) show that the free-riding hypothesis is inconsistent with the evidence following the block of Chinese Wikipedia and conclude that contributors must receive some other (to which they refer as "social") benefits. Like *Yahoo Answers*, Wikipedia offers social recognition. It publicly acknowledges the highest-performing few thousand contributors. Given the vast number of contributors to Wikipedia, however, the online social status of being named to this elite list would be extremely difficult to achieve, whereas in *Yahoo Answers*, each new posted question offers a level playing field and a chance to win the "best answer" and receive some points. Hence, the status-seeking motivation can be stronger in *Yahoo Answers*.

Existing evidence on online user motivation is mostly based on the open-source software literature, where researchers directly survey the developers (e.g., Lerner and Tirole, 2002; Lakhani and von Hippel, 2003; Roberts et al., 2006). This literature often focuses on whether extrinsic motivation (such as future career) substitutes or complements intrinsic motivation, where the consensus appears that the latter incentive is at least as important as the former. To our knowledge, Nam et al. (2009) is the only study that deals with the motivation for participating in a question answering forum, based on telephone interviews with a small group of answerers. They find that a point system similar to the one we consider here works well in motivating answerers; in particular, questions that are assigned more points in their study attract more answerers.

3 User Motivation

3.1 Theoretical Predictions

The first of two theoretical perspectives we examine is public-goods provision theory. Sequential contribution models are particularly suitable to our context because on the *Yahoo Answers* platform users post their answers sequentially, observing all previously posted answers for a given question, within four (or eight, if the asker requests an extension) days after the question is posted. Marx and Matthews (2000) consider a model where an agent's best response is to contribute either the full shortfall (i.e., "finishing the job" by filling the gap between the target level of the public good and the current cumulative contribution) or nothing. Further, a player's payoff depends on the contributions of the others only through their sum, and in equilibrium individual contributions tend to decrease over time as the collective contribution increases. This implies that holding constant a player's sequence of moves, a user's contribution would be decreasing in the sum of existing contributions by others.

The reason for the negative correlation between own and others' contributions is that an early mover's contribution crowds out the contribution from the late movers. Consider a canonical model (e.g., Varian, 1994) where two players divide their free time t_i, $i \in \{1, 2\}$, between private consumption, $x_i \geq 0$, and a contribution to the public good, $g_i \geq 0$. The total amount of the public good is $g_i + g_j$ for $i \neq j$. Agent i's utility function is quasilinear in his time, concave and increasing in the public good, and specified by $A_i u_i(g_i + g_j) + t_i - g_i$, where the parameter $A_i > 0$ scales the relative weight on the public good provision. Let $\bar{g}_i$ denote the amount of the public good that maximizes agent i's utility when the other agent contributes zero; that is, $\bar{g}_i$ satisfies $A_i u_i'(\bar{g}_i) = 1$. Suppose agent i observes agent j's contribution prior to deciding his own. The best-response function of agent i, denoted by $BR_i(g_j)$, satisfies $A_i u_i'(BR_i(g_j) + g_j) = 1$, so that $BR_i(g_j) = \max\{\bar{g}_i - g_j, 0\}$. Figure 1 depicts $BR_i(g_j)$ for three different levels of A_i.

The remaining issue is whether it is possible for agent i to contribute a positive amount, that is, whether data will be observed in the downward-sloping portion of the curve. As Varian (1994) shows, this is possible if the first-mover j has a belief distribution F on the second-mover's valuation A_i (or equivalently $\bar{g}_i$).[5] This seems plausible in our context because millions of users frequent the *Yahoo Answers* platform, so that a user would only form a belief distribution over other contributors' public-good preferences. Figure 1 illustrates possible data points with dots, where agent i's contribution is non-zero. The figure also shows that the slope of the function does not depend on the parameter A_i as long as A_i has a multiplicative effect on utility. Therefore, our first hypothesis is that if users are motivated by public-good provisions, then their contributions would be decreasing in existing levels of contributions with roughly the same sensitivities, holding constant a player's position in the overall sequence of moves.

[5]It can readily be shown that the first-mover's contribution g_j^* satisfies $A_j u_j'(g_j^*)F(g_j^*) = 1$, whereby the first-mover would contribute less than what he would have contributed under certainty (i.e., $g_j^* < \bar{g}_j$).

The second theory we examine is the theory of contests (e.g., Tullock, 1980), where two or more rent-seekers simultaneously determine their outlays to win a prize according to some "contest success function." While the literature that follows Tullock's seminal paper is large, an extension of this basic framework that is particularly relevant to our study is a sequential contest with asymmetric contestants. Specifically, Dixit (1987) and Morgan (2003) show that the "favorite" contestant (e.g., the one who has a higher valuation for the prize) has an incentive to overcommit an outlay compared to the Nash equilibrium level; and the opposite holds if the "underdog" contestant (e.g., the one who has a lower valuation for the prize) has an opportunity to move first. Baik and Shogren (1992) show that given endogenous order of moves, the favorite finds it advantageous to wait until the underdog moves and, thus, at the equilibrium levels of contributions the second mover's best response is increasing in the first mover's contribution.

This leads to our second hypothesis that if players are motivated by the contest prize, then their outlays would be increasing in the existing levels of outlays. Consider a standard contest model involving two players (i and j) who choose effort levels, $e_i \geq 0$. Let A_i represent the relative value that player i places on the contest reward. Player i's expected payoff is $A_i p(e_i, e_j) - c(e_i)$, where $p(e_i, e_j)$ is the probability that player i will win, and $c(e_i)$ represents effort cost. For the typical contest success functions p (such as the ratio form), it is well known that player i's best-response function, $BR_i(e_j)$, is inverse-U shaped. Further, under some mild conditions (such as the concavity of p in e_i) the slope of $BR_i(e_j)$ also increases in A_i.[6] Let player j be the underdog and player i be the favorite in the contest ($A_i > A_j$),

[6]For instance, suppose $p = e_i/(e_i + e_j)$, and $c(e_i) = e_i$. Then it is straightforward to derive agent i's best response function as $BR_i(e_j) = \sqrt{A_i e_j} - e_j$ and show that $\partial^2 BR_i(e_j)/\partial A_i \partial e_j = 1/4\sqrt{A_i e_j} > 0$.

so that player j is the first-mover. It can be shown that the equilibrium intersection of the two best-response functions occurs on the downward-sloping segment of BR_j and the upward-sloping segment of BR_i. Figure 2 depicts player i's best-response curve for three different levels of A_i, where possible data points are indicated with dots.

It seems reasonable to assume that on the *Yahoo Answers* platform users know whether they are an underdog or a favorite; that is, whether they have more than the average preference for the contest reward (hence more than one-half chance to win the contest if they were to simultaneously compete with a randomly drawn user). Given that users have access to all the archival answers posted on the platform, they should have an idea of how much they care to write lengthy answers to win "best answer" points. If the first-mover j has a belief distribution F on the second-mover's valuation A_i, then j will choose the level of effort e_j that maximizes his expected utility, which will fall somewhere between zero and the value that maximizes player i's utility. The equilibrium occurs on the upward-sloping segment of i's best-response function, because j reduces his effort relative to the Nash equilibrium. Thus, a user's effort would increase in the amount of existing contributions, and the more so for those users having stronger preferences for contest rewards.

3.2 Empirical Evidence

Our analysis is made possible by the granular level of data in the activity traces on *Yahoo Answers*. We follow a cohort of users who began their *Yahoo Answers* activities during the months of January, February, or March of 2013, extracting all of their answering activities for the ensuing twelve months. For each answer of a given user in our sample, we have a snapshot of the Unix time (in seconds) at which the answer was given and the question was asked, the number of answers already given (by others) at the time of the answer submission, the number of total characters and terms in the existing answers for the question, the same count measures for the user's own answer, the eventual number of answers provided for the question, and the names of the subject category and subcategory in which the question was posted. We also record the user's cumulative profile at each snapshot including the number of total answers provided and the number of total best answers.

We use a large fraction of these data (to protect the business information), randomly chosen at the user level (hence containing all answer sequences of any chosen user) for this study.[7] There are a total of 236,531 unique user identifiers in our dataset who self-reported the U.S. and English as their country of residence and primary language, respectively, and a total of 5,900,247 answers (see Table 1 for summary statistics). The average answer in the sample contains 240 characters or 43 words, and at the time of submission there are, on average, three answers already posted by others, which sum to roughly 600 characters or 110 words. There is a rapid response to questions on this forum, with modal time-to-answer being less than ten minutes from the initial posting of the question. This is likely due to the user interface, as previously mentioned, which lists the most recently posted questions at the top, whereby it takes longer to find previously posted questions.

Our aim in this section is to show whether user behavior is, on average, consistent with the data generating process predicted by the theoretical model of dynamic public-goods provision or sequential contest. If the data are consistent with either prediction, then we can refer that users are "public-good motivated" or "contest motivated," respectively. (However, to

[7] There can be some randomly missing answer sequences for a given user that we believe to be due to removed content that violated the terms of service. We believe that these should have little effect on our results.

be precise, it should be recognized that this inference may only be a "labelling" of underlying process, elicited here from the activity data, and not about self-stated motivation as in a survey.) Both theories seem plausible to us, *a priori*. That is, some socially-motivated users might be motivated by providing an incremental answer to the question, hoping that others follow suit, regardless of whether their answers are chosen as the best answer or not. Others might be motivated by the social comparison and status in the community, which is maintained by accumulating recognition points when their answers are chosen as the best one.

Our two hypotheses are that, holding constant a user's position in the contribution sequence and the total number of answers for a given question, the user's own contribution should be negatively correlated with the sum of existing contributions under the public goods hypothesis and positively correlated under the sequential contest hypothesis. We consider the following regression equation:

$$\begin{aligned} OwnContrib_{ik} &= \beta_0 + \beta_1 ExistingContrib_{ik} + \beta_2 NumofExistingAns_{ik} \\ &\quad + \beta_3 NumofEventualAns_{ik} + \sum_{h=1}^{23} HourofDay(h)_{ik} \\ &\quad + \sum_{d=1}^{6} DayofWeek(d)_{ik} + \sum_{w=1}^{65} Weekly(w)_{ik} + \zeta_i + \epsilon_{ik}, \end{aligned} \tag{1}$$

where users and answers are indexed by i and k, respectively.

User i's own contribution (the dependent variable) is measured either as the number of characters in a posted answer or as the number of words. As previously mentioned, these are proxies for user effort in the theory and not the answer quality. We estimate the effort regression (1) using both measures, defining the sum of existing contributions (the first independent variable) in the corresponding units. The specification includes twenty-three indicators for the hour of the day and six indicators for the day of the week to control for factors associated with hourly or daily routines and also user fixed effects and weekly dummies to absorb unobserved heterogeneity. The error term is independent across users but not necessarily within a user.

Estimation is by ordinary least squares, and the results are presented in Table 2. The signs of the coefficients on the existing contributions are consistently and statistically significantly positive throughout the specifications. The coefficient decreases by about one third when user fixed effects are included in columns (3) and (6), compared to the immediately preceding two columns. The coefficient is very similar regardless of whether contributions are measured by using the number of characters or the number of words. Thus, we conclude that, on average, the users posting answers in this forum are better described by the contest theory. It should be understood that this result does not imply the absence of the public-goods motivation. It is an average effect across all users, some of whom may be contest motivated. Users may also have a blend of both types of motivation.

We conjecture that those users who are more public-goods motivated may have lower levels of interest in answering questions because of the presence of those who are contest motivated, so that they drop out earlier than the point gatherers. To investigate this possibility, we divide the sample of users (at the user level) according to the number of total active days (i.e., days in which there is at least one answering activity) into four segments of data with roughly equal proportions. We then re-estimate specification (1) separately for each subgroup. Table 3 shows that indeed the estimated coefficient $\hat{\beta}_1$ is higher for the group of users who are active on the forum over more days. In particular, the coefficient is much higher for the last subgroup of 90+ active days, which is consistent with the conjecture that those users who are more public-goods motivated are more likely to quit early.

An alternative possibility is that (even initially public-goods-motivated) users may become contest-motivated as they experience the point system. We find some support for this possibility, but this effect is not large. We excluded from the sample all users who were active for less than 91 days. That is, we now only use the fourth quartile of the previously defined subgroup of users. Then we divide the data by the cumulative number of active days and re-estimate specification (1) by experience level. That is, for a given user, activities in the earlier and later periods are grouped into different bins. Table 4 shows the results, where the coefficient estimate, $\hat{\beta}_1$, is increasing slightly over the experience bins but relatively little in the first three categories, which suggests that the dominance of contest motivated users is due to drop-outs rather than due to changing preferences. The coefficient increases in the highest experience category, but this is a confounded estimate because it only comprises of longer-running users.

Estimates so far describe user behavior at the conditional means. However, it has long been recognized in the literature that there is considerable individual heterogeneity, and big data might offer an opportunity to measure and incorporate such heterogeneity (Einav and Levin, 2013). Thus, we estimate the specification in (1) for each of the users who are observed with enough activities (i.e., more than 30 answering events following a rule of thumb), using a simple regression with only the first three explanatory variables ($ExistingContrib_{ik}$, $NumofExistingAns_{ik}$, $NumofEventualAns_{ik}$). This allows us to investigate how the motivation parameter, $\hat{\beta}_1$, varies across the individual users. Figures 3 and 4 display the distribution of this parameter using variables measured in characters and those measured in words, respectively. The histograms are plotted on the range $[-0.5, 1]$, excluding 0.65% of the sample which falls outside of this range. The two distributions are very similar, and about two thirds are between 0 and 0.2.

4 Effects of Multitasking

4.1 Theoretical Predictions

We assume that the parameter β_1 measures how much more, in relative terms, a user cares about incentive rewards (i.e., points). In this section, we illustrate how this imputed preference parameter can be useful in understanding user performance at the aggregate level. Specifically, we investigate the effect of multitasking on performance. Multitasking means that a user is equipped to engage in multiple tasks and that the user regularly performs all of those tasks.[8] Making the notion of "regularly performs" precise in the empirical counterpart requires stating a specific timeframe (e.g. a minute, an hour, or a day) during which to record which activities a user performs. That is, whether or not a user is said to be "multitasking" hinges on which timeframe is chosen as well as the granularity of those activities (see Circella et al., 2012).

In most settings, the degree of multitasking decreases as the time unit of reference decreases to zero.[9] In our context, a user can only answer one question from a single category

[8]A recent theoretical exploration of the tradeoff between multitasking and the alternative (i.e. specialization) is DeVaro and Gürtler (2015).

[9]Some basic functions of the brain may operate simultaneously even in a relatively small unit of time. For instance, one may be concurrently typing at the keyboard, reading on a screen what is being typed, and thinking about what to type next. We abstract from this finer level of multitasking.

at a time, and thus would always be engaged in a single activity if a sufficiently narrow timeframe is considered, although the user might switch tasks *across* such timeframes. In our application, the choice of any finite time scale that is uniformly applied to every user is problematic — this is due to the large variation in the frequencies with which users make contributions to the forum. Thus, in what follows, to define a user's level of multitasking, the timeframe we choose is the entire period during which the user's answering activities are observed.[10] We believe this leads to a better measurement while we hold constant the level of detail characterizing a task or activity.

Both positive and negative effects of multitasking have been documented. There is a long tradition in experimental psychology showing that multitasking can have detrimental effects on performance if subjects are required to switch between multiple tasks within a given time period than if only a single task is performed, due to the limited capacity of the human mind (e.g., Koechlin and Hyafil, 2007; Just et al., 2008; Kiesel et al., 2010; Buser and Peter, 2012). On the positive side, multitasking have been sometimes linked to various types of innovation in other strands of literature (e.g., Aoki, 1986; Morita, 2005; Gibbs et al., 2007), which in our context would mean users who are actively participating in multiple subject categories might see linkages that more specialized users might miss and that could result in more creative, and better answers.[11]

We attempt to link these two strands of literature by proposing that the effects of multitasking on performance may be mediated by the user's motivation. To be more specific,

[10] For instance, if a user's activities are observed over a year, and during that time they post one answer in category A and all other answers in category B, the user is defined to be multitasking although we use a continuous measure of multitasking to address these issues.

[11] For further information on the positives and negatives associated with multitasking and multiskilling at both the individual and organization levels, see Carmichael and McLeod (1993), Owan (2011), DeVaro and Farnham (2011), Farnham and Hutchinson (2011), and Kato and Owan (2012).

we hypothesize that the effects of multitasking would be more pronounced for those who respond relatively more strongly to the point system (i.e., contest motivated users) than for those whose behavior is relatively more consistent with the public-goods theory. Our assumption is that the latter group does not care much about winning more points by specializing in certain categories, so that their performance depends little on the observed extent of multitasking. We test this correlative effect in a cross-section of users, by investigating whether the performance gap between the two differentially motivated groups tends to widen or shrink as the level of multitasking increases in data.

Notice that the user's expertise and subject knowledge are not observable to the platform, so the platform does not personalize the point system to each individual user. Thus, if users are solely interested in gaining "best answer" points, they should focus all their efforts on a single (or few) subject categories representing their area(s) of greatest knowledge and expertise, which should yields the greatest number of points earned for a given investment of effort. This assumption is similar to and consistent with the prediction and evidence in the incentive theory literature that high-powered incentive schemes tend to make workers neglect quality (e.g., Hong et al., 2013; Bartel et al., 2013). Since incentive schemes typically do not adjust for quality (as in our case of *Yahoo Answers*), workers tend to focus on one "task" (such as quantity of items produced).

Finally, in addition to choosing a subject category, users on the forum also decide when to contribute answers, if at all. That is, there is also a time dimension in multitasking between answering activities and other time-use activities (that are not observed in our data). The experimental psychology literature indicates that when people try to do multiple things simultaneously, it impairs the speed and/or quality of their work because the competition for the allocation of brain resources tends to increase switching costs between different activities. Thus, we predict that the performance of the users who post answers to multiple subject categories tends to be lower when they also frequently switch in and out of the answering activity among daily/weekly routines. By the same logic as above, the effects on performance would bear on those who are contest motivated.

4.2 Empirical Evidence

The analysis in this section is at the aggregate (user) level. We are interested in determining how the effect of multitasking on performance is moderated by the user's motivation parameter imputed from the previous section. The user-level estimation of equation (1) has three explanatory variables, and we required $N \geq 30$ for each user to remain in the data set. This reduces the sample size from the previous 5,900,247 answers to 4,597,747, and from 236,531 unique user identifiers to 32,135. The latter is unsurprising because most users engage in activities on the platform once or twice and never return. From the previous section, we know that the early dropouts are more likely to be public-goods motivated; hence, dropping these users would only be biased against our prediction, which focus on the most contest-motivated group's performance.

The summary statistics are shown in Table 5. $Coeff1$ is the individual response parameter β_1 estimated from above using variables measured in characters, and $Coeff2$ is the corresponding estimate when they are measured in words. The correlation coefficient between these two variables is 0.95, and it affects little the grouping of users into quintiles that we use in our analysis below. Thus, to save space, we only present results using $Coeff1$, though the results are virtually the same if we instead use $Coeff2$. The table shows that a typical user answers questions in nine subject categories, or 20 subcategories, during the active period. *Duration* is the number of days between the first and the last observed answering activity in the data.[12] *Activedays* counts only the number of days when a user posted at least one answer, which averages about 20 days.

The middle panel in Table 5 lists our measures of multitasking with respect to subject and time. We use Shannon entropy, $-\sum_{i=1}^{n} p(x_i)\log(p(x_i))$, where $p(x_i)$ is the probably mass function of task x_i, $i \in \{1, ..., n\}$. The measure is standard in the information science literature, and it quantifies the amount of uncertainty associated with a realization of n outcomes. We use this entropy measure as a proxy for a user's level of multitasking because it measures how evenly a user's activities are spread over tasks. That is, simply counting the number of active subject categories would ignore information on the respective amounts of contributions made to each category. The Shannon entropy is maximal when activities are equally distributed over tasks (all of the outcomes are equally likely), and it is also additively separable if activities are divided into sub-tasks.[13]

The first two entropies are measured with respect to subject categories and subcategories, respectively, to which there is a unique assignment from each question. There are 26 *Yahoo Answers* subject categories and 272 subcategories in our data. The next two entropies are measured with respect to the timing of answering activities. Specifically, we classify each answer's submission time into (modulus of) 7 days of the week as well as 24 hours of the day. Then the entropies based on the modulo number can capture the user's time-use pattern over

[12] The data are right-truncated. That is, some 1300 users (out of 32,135) still are observed posting answers after one year from their first activity, so they are likely to continue their activities beyond the time period of our data. Further, a dormant user may return in the future.

[13] When the realizations are equally spread over n tasks, then the Shannon entropy is $\log(n)$, which means that the entropy is in fact an increasing function of the number of active categories. Thus, we control for the number of active (sub)categories in our regressions.

personal routines in addition to the online activity. Hence, if a user posts answers online only at a certain hour of the day during the entire history of online activities, then the hour entropy would be zero, whereas if the user has posted answers on many different hours of the day, then the entropy would be high. As previously discussed, frequent switching between activities over time as well as across subjects tends to lower performance.

The bottom panel in Table 5 lists our four "performance" measures, intended to check the robustness of our results. We posit that performance of an answerer is defined by how many questions the user serves for a given effort. This can be operationalized by dividing the total number of answers by the sum of total number of characters (or terms) typed by the user. We believe this metric captures the commonly accepted definition of performance, that is, the accomplishment of tasks measured against some known standards or costs. However, this captures the quantity rather than the quality of answers for a given amount of efforts. We partly address this issue by measuring the same performance metrics, where the numerator counts only the total number of best answers a user has provided. Thus, under this definition, if an answer was not selected as a best answer, then it does not count at all. Although rather extreme, we show that our results are robust to these specifications.[14]

[14] Hence, our performance measures are observed under two "virtual" incentive schemes: One is to get two points for each answer a user posts, and the other is to get ten points only when the answer is selected as a best answer by the asker (or by community vote).

The performance regression equation is

$$Perf_i = \tag{2}$$

$$\beta_0 + \beta_1 Coeff_i + \beta_2 Ent(subject)_i + \beta_3 Ent(time)_i + \beta_4(CoeffQ5_i \times Ent(subject)Q5_i)$$

$$+\beta_5(CoeffQ5_i \times Ent(time)Q5_i) + \beta_6(CoeffQ5_i \times Ent(subject)Q5_i \times Ent(time)Q5_i)$$

$$+\beta_7 CoeffQ5_i + \beta_8 Ent(subject)Q5_i + \beta_8 Ent(time)Q5_i + \mathbf{X}_i\beta + \epsilon_i,$$

where $\mathbf{X}_i$ includes duration (in days), number of active days (or hours), and number of active (sub)categories, and the error term is i.i.d. across user i.

Table 6 presents the estimation results with robust standard errors, where the dependent variable is the total number of answers divided by the sum of efforts measured in characters. The first three columns (1)-(3) pair the two upper-level entropies, namely, across subject categories and days of week; the next three columns (4)-(6) pair the other lower-level entropies across subject sub-categories and hours of day. The specification of each column is maintained for the remaining Tables 7-9, where the only difference is the left-hand side variable. Column (1) shows only direct correlative effects of the motivation parameter, and subject and time entropies on performance. We find that all three variables are significantly negatively associated with the performance measure. That is, it appears that more contest-motivated users have lower performance levels, and multitasking across either subjects or time is also negatively correlated with performance.

Column (2) tests the main prediction that the negative effects of multitasking on performance would be stronger for contest-motivated users compared to that of public-goods motivated users. To highlight this result, we classify users into quintiles based on the motivation parameter, subject entropy, and time entropy, each separately, and then we employ an indicator variable for the fifth quintile (labeled with $Q5$). The prediction is supported by the statistically significant negative coefficient on the interaction term, $CoeffQ5 \times Ent(subject)Q5$, in Column (2). Another prediction was that this moderating effect would be even larger for the users who tend to spread over time due to switching costs. Column (3) shows that this prediction is also supported by the statistically-significant negative coefficient on the triple interaction term, $CoeffQ5 \times Ent(subject)Q5 \times Ent(time)Q5$. Results in columns (4)-(6) confirm these findings with a pair of lower-level entropies.

Table 7 shows the same estimation results when we divide the total number of answers by the sum of posting length measured in words rather than characters. The results are essentially identical to those in Table 6. We also find that more contest-motivated users tend to perform better when their answers are relatively more spread over time (holding constant multitasking across subjects). An interpretation of these findings is that contest-motivated users tend to access the forum and post their answers throughout a day (or a week), rather than posting an answer regularly at a certain time of usage. Finally, Tables 8 and 9 show the estimation results using performance based only on the number of best answers provided by the user. Except for one case (the insignificant coefficient on the triple interactive term in column (3) of Table 8), the above predictions are strongly confirmed even when we use quality-adjusted performance metrics. Thus, we conclude that our hypotheses are supported by the data.

5 Conclusion

Our results have a number of implications. One is that instituting a recognition-incentive system, such as rewarding virtual points for best answers, can have a significant effect on the composition of the user population active on the forum. Specifically, we found that from a given cohort of nascent users, those who are more drawn to the contest aspects of winning best answers tend to be active for more days than those who might care more about the public-good-provision aspect of their contributions. Thus, if overcoming the free-rider problem and maximizing the time spent on the platform is the objective of the operating platform, then a recognition system based on a stream of contests can be a useful means to achieve this goal and will cost little to the firm or any other users involved.

Another implication is that there are statistically significant effects from spreading contributions more evenly over multiple subjects for those who are drawn to such an incentive system. This effect is more pronounced when the users spread their answering activities more evenly over time (within a day or a week). Thus, the platform operator may want to discourage contest-motivated users from multitasking across subjects and time. The current system awards the same number of points across all categories and time, but equipped with information about imputed user types, the platform operator can identify those users for whom multitasking is an issue, and could, for instance, target personalized point offers to direct their focus to fewer categories and induce their participation.

Some limitations of our analysis are also worth noting. First, it is unclear to what extent our findings can be generalized beyond *Yahoo Answers*. However, we believe that our methodology can be easily adapted and expanded when analyzing user-generated content from other knowledge exchange forums where there is some contest-based recognition system. Second, perhaps more importantly, this paper focused only on the turnover of user-generated content, and we do not address the issue of answer quality. In fact, there is no guarantee that the chosen best answer is even correct. Although we do not claim here that the recognition system or multitasking improves answer quality, other researchers could examine it in future work by analyzing an answer's popularity.

References

[1] Adamic, L., J. Zhang, E. Bakshy, and M. Ackerman. 2008. "Knowledge Sharing and Yahoo Answers: Everyone Knows Something." WWW '08 Proceedings of the 17th International Conference on World Wide Web: 665–674.

[2] Aoki, M. 1986. "Horizontal vs. Vertical Information Structure of the Firm." *American Economic Review* 76(5): 971-983.

[3] Aral, S., E. Brynjolfsson, and M. Van Alstyne. 2012. "Information, Technology, and Information Worker Productivity." *Information Systems Research* 23: 849–867.

[4] Backstrom, L., R. Kumar, C. Marlow, J. Novak, and A. Tomkins. 2008. "Preferential Behavior in Online Groups." WSDM '08 Proceedings of the 2008 International Conference on Web Search and Data Mining: 117–128.

[5] Baik, K., and J. Shogren. 1992. "Strategic Behavior in Contests: Comment." *American Economic Review* 82: 359–362.

[6] Bartel, A., B. Cardiff-Hicks, and K. Shaw. 2013. "Compensation Matters: Incentives for Multitasking in a Law Firm." NBER Working Paper No. 19412.

[7] Bergstrom, T., H. Varian, and L. Blume. 1986. "On the Private Provision of Public Goods." *Journal of Public Economics* 29: 25–49.

[8] Besley, T., and M. Ghatak. 2005. "Competition and Incentives with Motivated Agents." *American Economic Review*, 95: 616–636.

[9] Brynjolfsson, E., S. Kim, and J. Oh. 2014. "User-Generated Capital and Firm Value: Theory and Evidence from Internet Firms." Mimeo.

[10] Buser, T., and N. Peter. 2012. "Multitasking." *Experimental Economics* 15: 641–655.

[11] Carmichael, H. & MacLeod, W. 1993. "Multiskilling, Technical Change and the Japanese Firm." *Economic Journal*, 103, 142-160.

[12] Circella, G., P. Mokhtarian, and L. Poff. 2012. "A Conceptual Typology of Multitasking Behavior and Polychronicity Preferences." *International Journal of Time Use Research* 9: 59–107.

[13] DeVaro, J., and M. Farnham. 2011. "Two Perspectives on Multiskilling and Product Market Volatility" *Labour Economics* 18, 862-871.

[14] DeVaro, J., and O. Gürtler. 2015. "Strategic Shirking: A Theoretical Analysis of Multitasking and Specialization" Forthcoming in *International Economic Review*.

[15] Dixit, A. 1987. "Strategic Behavior in Contests." *American Economic Review* 77: 891–898.

[16] Einav, L., and J. Levin. 2013. "The Data Revolution and Economic Analysis." NBER Working Paper No. 19035.

[17] Farnham, M., and E. Hutchinson. 2011. "The Effect of Multiskilling on Labor Productivity, Product Quality, and Financial Performance." *Advances in the Economics of Participatory and Labor Managed Firms*, 12.

[18] Frey, B. 1997. "On the Relationship Between Intrinsic and Extrinsic Work Motivation." *International Journal of Industrial Organization*, 15: 427–439.

[19] Gibbs, M., Levenson, A., and Zoghi, C. 2010. "Why Are Jobs Designed the Way They Are?" *Research in Labor Economics* 30, 107-154.

[20] Hong, F., T. Hossain, J. List, and M. Tanaka. 2013. "Testing the Theory of Multitasking: Evidence from a Natural Field Experiment in Chinese Factories." NBER Working Paper No. 19660.

[21] Jian, L., and J. MacKie-Mason. 2012. "Incentive-Centered Design for User-Contributed Content." In M. Peitz and J. Waldfogel (Eds.), *Oxford Handbook of The Digital Economy*, New York: Oxford University Press.

[22] Just, M., T. Keller, and J. Cynkar. 2008. "A Decrease in Brain Activation Associated with Driving When Listening to Someone Speak." *Brain Research* 1205, 70-80.

[23] Kato, T., and H. Owan. 2012. "Market characteristics, intra-firm coordination, and the choice of human resource management systems: evidence from new Japanese data." *Journal of Economic Behavior and Organization*.

[24] Kiesel, A., M. Wendt, K. Jost, M. Steinhauser, M. Falkenstein, A. Philipp, and I. Koch. 2010. "Control and Interference in Task Switching – A Review." *Psychological Bulletin* 136: 849–874.

[25] Koechlin, E., and A. Hyafil. 2007. "Anterior Prefrontal Function and the Limits of Human Decision-Making." *Science*, 3181(5850): 594-598.

[26] Kreps, D. 1997. "Intrinsic Motivation and Extrinsic Incentives." *American Economic Review*, 87: 359–364.

[27] Lakhani, K., and E. von Hippel. 2003. "How Open Source Software Works: "Free" User-to-User Assistance." *Research Policy* 32: 923–943.

[28] Lazear, E., and S. Rosen. 1981. "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy*, 89: 841–864.

[29] Lerner, J., and J. Tirole. 2002. "Some Simple Economics of Open Source." *Journal of Industrial Economics* 50: 197–234.

[30] Marx, L., and S. Matthews. 2000. "A Dynamic Voluntary Contribution to a Public Project." *Review of Economic Studies* 67: 327–358.

[31] Morgan, J. 2003. "Sequential Contests." *Public Choice* 116: 1–18.

[32] Morgan, J., and R. Wang. 2010. "Tournaments for Ideas." *California Management Review* 52: 77–97.

[33] Morita, H. 2005. "Multi-skilling, Delegation and Continuous Process Improvement: A Comparative Analysis of US-Japanese Work Organizations." *Economica*, 72: 69-93.

[34] Nam, K., M. Ackerman, and L. Adamic. 2009. “Questions in, Knowledge iN? A Study of Naver’s Question Answering Community.” SIGCHI ’09 Proceedings of the 2009 Conference on Human Factors in Computing Systems: 779–788.

[35] Owan, H. 2011. “Specialization, Multiskilling, and Allocation of Decision Rights.” Advances in the Economic Analysis of Participatory and Labor-Managed Firms, 12: 3-34.

[36] Ransbotham, S., G. Kane, and N. Lurie. 2012. “Network Characteristics and the Value of Collaborative User-Generated Content.” *Marketing Science* 31: 387–405.

[37] Roberts, J., I.-H. Hann, and S. Slaughter. 2006. “Understanding the Motivations, Participation, and Performance of Open Source Software Developers: A Longitudinal Study of the Apache Projects.” *Management Science* 52: 984–999.

[38] Tullock, G. 1980. “Efficient Rent-Seeking.” In J Buchanan et al. (Eds.), *Toward a Theory of the Rent-Seeking Society*, College Station: Texas A&M Press.

[39] Varian, H. 1994. “Sequential Contributions to Public Goods.” *Journal of Public Economics* 53: 165–186.

[40] Welser, H., E. Gleave, D. Fisher, and M. Smith. 2007. “Visualizing the Signatures of Social Roles in Online Discussion Groups.” *Journal of Social Structure* 8: No. 2.

[41] Whittaker, S., L. Terveen, W. Hill, and L. Cherny. 1998. “The Dynamics of Mass Interaction.” Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work: 257–264.

[42] Zhang, X., and F. Zhu. 2011. “Group Size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia.” *American Economic Review* 101: 1601–1615.

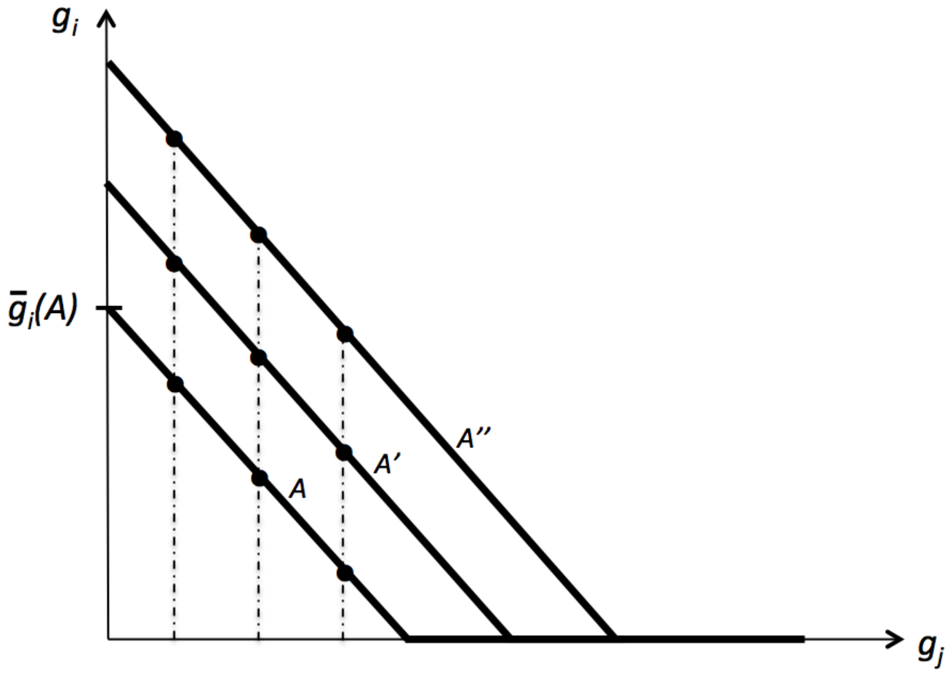

Figure 1: Public-good contributions best-response curves, with $A < A' < A''$.

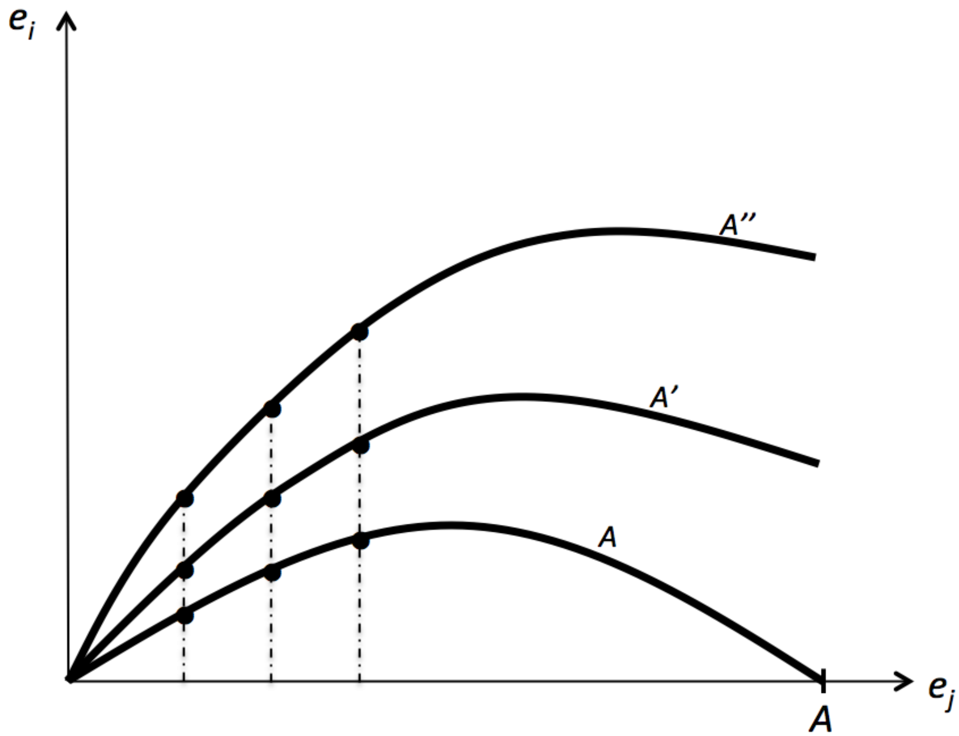

Figure 2: Tullock contest best-response curves, with $A < A' < A''$.

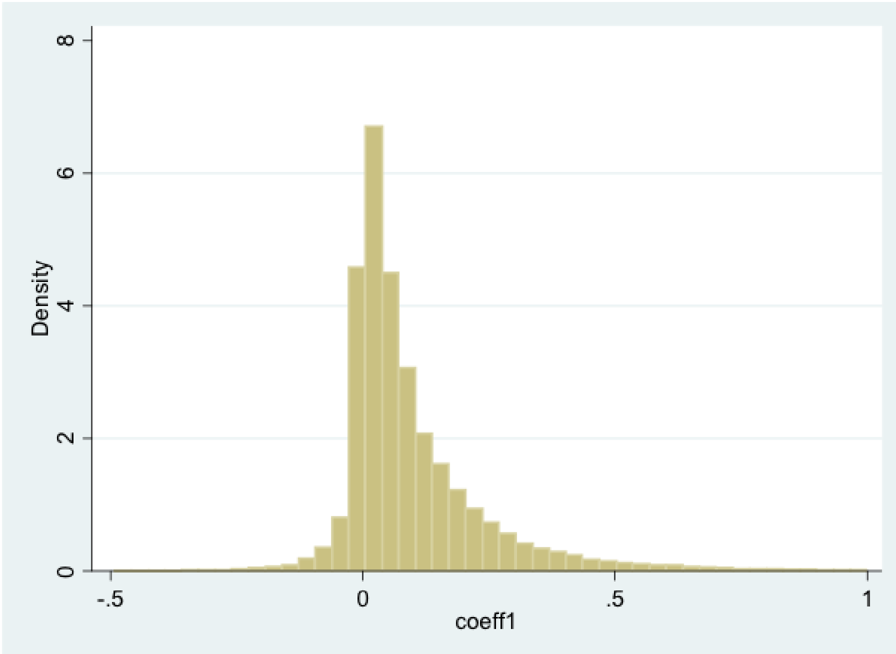

Figure 3: Histogram of coefficients using a measurement of characters used. The histogram excludes 0.65% of the sample which falls outside the given range. About two thirds of the density is between 0 and 0.2.

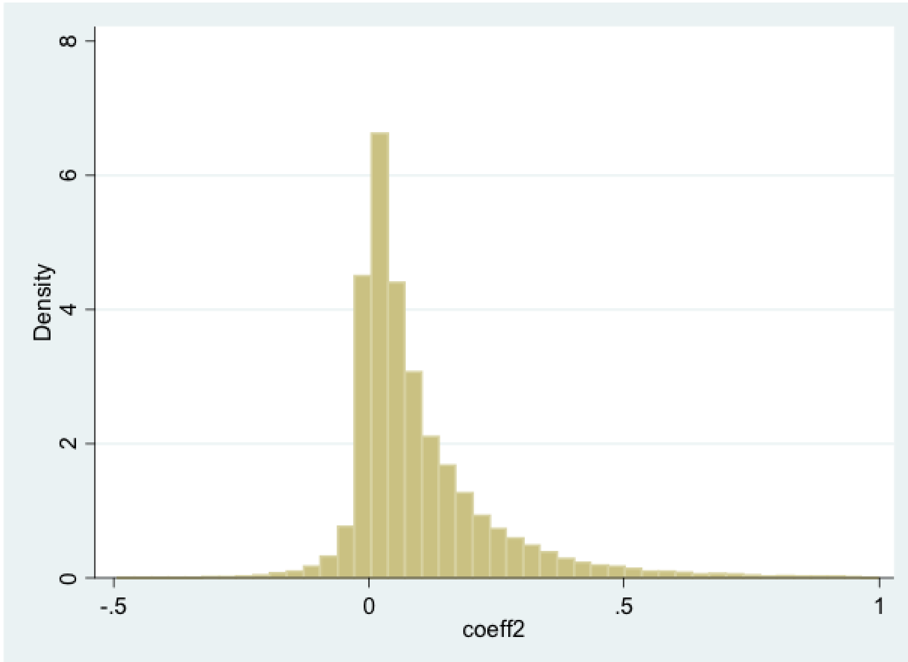

Figure 4: Histogram of coefficients using a measurement of terms used. The histogram excludes 0.65% of the sample which falls outside the given range. About two thirds of the density is between 0 and 0.2.

Variable	Mean	Std. dev.	Min	Max	N
Own contribution (in characters)	239.6146	394.7037	2	11378	5900247
Own contribution (in terms)	43.17206	69.11497	1	1888	5900247
Existing contributions (in characters)	613.306	3225.927	0	450152	5900247
Existing contributions (in terms)	108.9812	583.2364	0	83418	5900247
Number of existing answers	3.381745	12.6679	0	1854	5900247
Number of eventual answers	8.005107	21.04932	1	1856	5900247

Table 1: Summary statistics for answer-level data set

Term(=words) includes the articles, "a(n)" and "the."

	Panel A: Y = Contributions (in char.)			Panel B: Y = Contributions (in terms)		
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Exiting contributions (in char.)	0.0330 (0.0012)***	0.0329 (0.0012)***	0.0206 (0.0007)***			
Exiting contributions (in terms)				0.0315 (0.0011)***	0.0314 (0.0011)***	0.0200 (0.0007)***
Number of existing answers	-4.0539 (0.2272)***	-4.0372 (0.2251)***	-3.0480 (0.1494)***	-0.7031 (0.0378)***	-0.7024 (0.0376)***	-0.5327 (0.0249)***
Number of eventual answers	-2.2897 (0.1169)***	-2.2923 (0.1175)***	-1.0171 (0.0522)***	-0.3949 (0.0203)***	-0.3953 (0.0203)***	-0.1782 (0.0090)***
Hour of day	No	Yes	Yes	No	Yes	Yes
Day of week	No	Yes	Yes	No	Yes	Yes
Week FE	No	Yes	Yes	No	Yes	Yes
User FE	No	No	Yes	No	No	Yes
R^2	0.0198	0.0214	0.4149	0.019	0.0201	0.4119
N	5900247	5900247	5900247	5900247	5900247	5900247

Table 2: Effort regressions (using the whole sample)

User-clustered standard errors are in the parenthesis; *** 1%, ** 5%, * 10%

	Panel A: Y = Contributions (in char.)							
Variable	Active = 1-7 day		Active = 8-30 days		Active = 31-90 days		Active = 90+ days	
Exiting contributions (in char.)	0.0193 (0.0007) ***	0.0120 (0.0006) ***	0.0288 (0.0017) ***	0.0179 (0.0009) ***	0.0379 (0.0026) ***	0.0235 (0.0014) ***	0.0708 (0.0060) ***	0.0405 (0.0032) ***
Number of existing answers	-2.9489 (0.1906) ***	-1.7922 (0.1475) ***	-3.5362 (0.3191) ***	-2.6428 (0.2002) ***	-4.5815 (0.5115) ***	-3.5102 (0.3023) ***	-7.7226 (0.9120) ***	-5.3809 (0.5129) ***
Number of eventual answers	-1.3237 (0.0859) ***	-0.8194 (0.0606) ***	-1.8620 (0.1579) ***	-0.8600 (0.0735) ***	-2.3912 (0.2220) ***	-1.0011 (0.1044) ***	-4.1952 (0.5317) ***	-1.5408 (0.1773) ***
Hour of day	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day of week	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	-	-	Yes	Yes	Yes	Yes	Yes	Yes
User FE	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.011	0.5487	0.0178	0.386	0.0249	0.3533	0.0505	0.3834
N	1612387	1612387	1582275	1582275	1277090	1277090	1428495	1428495

	Panel B: Y = Contributions (in terms)							
Variable	Active = 1-7 day		Active = 8-30 days		Active = 31-90 days		Active = 90+ days	
Exiting contributions (in terms)	0.0189 (0.0007) ***	0.0121 (0.0006) ***	0.0281 (0.0017) ***	0.0178 (0.0009) ***	0.0365 (0.0024) ***	0.0229 (0.0014) ***	0.0677 (0.0056) ***	0.0387 (0.0029) ***
Number of existing answers	-0.5353 (0.0334) ***	-0.3304 (0.0266) ***	-0.6351 (0.0551) ***	-0.4765 (0.0352) ***	-0.7968 (0.0874) ***	-0.6178 (0.0524) ***	-1.3321 (0.1518) ***	-0.9209 (0.0837) ***
Number of eventual answers	-0.2290 (0.0151) ***	-0.1484 (0.0111) ***	-0.3211 (0.0272) ***	-0.1517 (0.0132) ***	-0.4149 (0.0385) ***	-0.1768 (0.0182) ***	-0.7181 (0.0923) ***	-0.2643 (0.0296) ***
Hour of day	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day of week	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	-	-	Yes	Yes	Yes	Yes	Yes	Yes
User FE	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.0109	0.5473	0.0172	0.3743	0.0239	0.3354	0.0477	0.3922
N	1612387	1612387	1582275	1582275	1277090	1277090	1428495	1428495

Table 3: Effort regressions (using all user samples) by total active days

User-clustered standard errors are in the parenthesis; *** 1%, ** 5%, * 10%

	Panel A: Y = Contributions (in char.)							
Variable	Active = 1-39 day		Active = 40-78 days		Active = 79-132 days		Active = 133+ days	
Exiting contributions (in char.)	0.0576 (0.0059) ***	0.0329 (0.0034) ***	0.0619 (0.0064) ***	0.0334 (0.0032) ***	0.0685 (0.0092) ***	0.0356 (0.0047) ***	0.0969 (0.0101) ***	0.0573 (0.0055) ***
Number of existing answers	-6.5527 (1.2365) ***	-4.4927 (0.7356) ***	-7.2178 (1.1359) ***	-4.4202 (0.6014) ***	-6.7969 (1.4292) ***	-4.7381 (0.6568) ***	-7.3234 (1.5410) ***	-6.4528 (1.0508) ***
Number of eventual answers	-3.7904 (0.6405) ***	-1.5256 (0.2618) ***	-2.9739 (0.6170) ***	-1.1013 (0.2099) ***	-4.4465 (0.8886) ***	-1.2594 (0.2632) ***	-7.6949 (1.1687) ***	-2.2970 (0.4984) ***
Hour of day	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day of week	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
User FE	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.0431	0.4148	0.0433	0.4421	0.0485	0.4537	0.0772	0.3374
N	364262	364262	350491	350491	357147	357147	356595	356595

	Panel B: Y = Contributions (in terms)							
Variable	Active = 1-39 day		Active = 40-78 days		Active = 79-132 days		Active = 133+ days	
Exiting contributions (in terms)	0.0557 (0.0058) ***	0.0321 (0.0033) ***	0.0607 (0.0064) ***	0.0331 (0.0031) ***	0.0667 (0.0087) ***	0.0349 (0.0044) ***	0.0911 (0.0094) ***	0.0528 (0.0047) ***
Number of existing answers	-1.1529 (0.2141) ***	-0.7945 (0.1276) ***	-1.2944 (0.1949) ***	-0.8008 (0.1032) ***	-1.1875 (0.2384) ***	-0.8296 (0.1090) ***	-1.2375 (0.2618) ***	-1.0327 (0.1683) ***
Number of eventual answers	-0.6642 (0.1135) ***	-0.2687 (0.0459) ***	-0.5178 (0.1109) ***	-0.1921 (0.0367) ***	-0.7723 (0.1564) ***	-0.2213 (0.0454) ***	-1.2372 (0.1712) ***	-0.3717 (0.0791) ***
Hour of day	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day of week	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
User FE	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.0418	0.4162	0.0416	0.4543	0.0472	0.4709	0.0736	0.3354
N	364262	364262	350491	350491	357147	357147	356595	356595

Table 4: Effort regressions (using only users with more than 90 active days) by total active days

User-clustered standard errors are in the parenthesis; *** 1%, ** 5%, * 10%

Variable	Mean	Std. dev.	Min	Max	N
Coeff1 (first-stage parameter in char.)	0.100566	0.276102	-33.2217	8.255607	32135
Coeff2 (first-stage parameter in terms)	0.103762	0.230992	-23.8581	6.569354	32135
Number of active (main) categories	9.306364	5.768125	1	26	32135
Number of active subcategories	20.22922	17.95117	1	207	32135
Number of total active days	20.77567	29.47209	1	387	32135
Number of total active hours	45.37921	109.1597	2	3441	32135
Total duration (in days)	143.0513	122.7452	1	396	32135
Entropy over main categories	1.329898	0.81183	0	3.100776	32135
Entropy over subcategories	1.947031	1.03329	0	4.580973	32135
Entropy over hours of day	2.085094	0.482003	0	3.163832	32135
Entropy over days of week	1.510556	0.340179	0	1.94503	32135
Number of total answers	166.5196	496.0233	17	22001	32135
Number of best answers	30.34713	117.3526	0	9384	32135
Total contributions (in char.)	33807.44	155471.8	33	11800000	32135
Total contributions (in terms)	6061.71	26530.19	60	2250493	32135
#Best answers / ΣContrib. (in char.)	1.377749	8.835681	0	1424.46	32135
#Best answers / ΣContrib. (in terms)	6.987059	12.10376	0	712.2302	32135
#Total answers / ΣContrib. (in char.)	10.54363	21.93251	0.24218	1436.306	32135
#Total answers / ΣContrib. (in terms)	51.98561	60.10996	1.176514	893	32135

Table 5: Summary statistics for user-level data set

The four ratio measures are multiplied by 1,000.

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Coeff	-6.1853	-2.2285	-2.1970	-5.6795	-1.9755	-1.8984
	(2.6962)**	(1.1587)*	(1.1488)*	(2.5690)**	(1.1374)*	(1.1223)*
Ent(main category)	-1.7982	-2.4125	-2.2320			
	(0.2914)***	(0.2906)***	(0.2963)***			
Ent(day of week)	-5.4340	-5.2119	-6.0246			
	(0.4792)***	(0.4720)***	(0.5217)***			
Ent(subcategory)				-1.1951	-1.7323	-1.5162
				(0.2501)***	(0.2547)***	(0.2556)***
Ent (hour of day)				-6.6494	-6.4034	-7.7927
				(0.3351)***	(0.3257)***	(0.4194)***
CoeffQ5 x Ent(m)Q5		-3.8621	-3.5337			
		(0.7771)***	(0.8342)***			
CoeffQ5 x Ent(d)Q5			1.2142			
			(0.3531)***			
CoeffQ5 x Ent(m)Q5 x			-2.3973			
Ent(d)Q5			(0.7746)***			
CoeffQ5 x Ent(s)Q5					-3.4914	-3.1318
					(0.7847)***	(0.8709)***
CoeffQ5 x Ent(h)Q5						1.4985
						(0.2873)***
CoeffQ5 x Ent(s)Q5 x						-2.5581
Ent(h)Q5						(0.7505)***
CoeffQ5		-5.0521	-5.2337		-4.8701	-5.0085
		(0.3847)***	(0.4051)***		(0.3810)***	(0.3958)***
Ent(m)Q5		7.0558	7.0477			
		(0.4790)***	(0.4784)***			
Ent(d)Q5			1.7771			
			(0.4078)***			
Ent(s)Q5					7.1291	7.1534
					(0.5073)***	(0.5061)***
Ent(h)Q5						2.9533
						(0.3572)***
Total duration (in days)	-0.0194	-0.0175	-0.0173	-0.0152	-0.0137	-0.0135
	(0.0010)***	(0.0010)***	(0.0010)***	(0.0009)***	(0.0009)***	(0.0009)***
Total active days	0.0013	0.0041	-0.0043			
	(0.0028)	(0.0028)	(0.0031)			
Total active hours				0.0001	0.0021	0.0007
				(0.0007)	(0.0007)***	(0.0007)
No. of active	0.5215	0.2912	0.2678			
categories	(0.0362)***	(0.0329)***	(0.0325)***			
No. of active				0.1514	0.0695	0.0554
subcategories				(0.0106)***	(0.0093)***	(0.0094)***
R^2	0.031	0.0468	0.0476	0.0407	0.0556	0.0575
N	32135	32135	32135	32135	32135	32135

Table 6: Performance regressions (Y = #Total answers / ΣContrib. in char.)
Heteroscedasticity-consistent standard errors are in the parenthesis; *** 1%, ** 5%, * 10%

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Coeff	-29.3676 (12.4000)**	-11.3427 (5.4028)**	-11.1791 (5.3420)**	-27.0043 (11.8045)**	-10.1830 (5.3047)*	-9.8594 (5.2328)*
Ent(main category)	-11.6443 (0.9228)***	-13.9454 (0.8461)***	-13.4036 (0.8528)***			
Ent(day of week)	-23.4010 (1.4609)***	-22.5047 (1.4112)***	-25.0765 (1.5448)***			
Ent(subcategory)				-5.9904 (0.5966)***	-7.8785 (0.5277)***	-7.0879 (0.5310)***
Ent (hour of day)				-30.4286 (1.0450)***	-29.3554 (0.9927)***	-34.5020 (1.2327)***
CoeffQ5 x Ent(m)Q5		-22.3108 (1.5996)***	-20.8935 (1.6057)***			
CoeffQ5 x Ent(d)Q5			6.5975 (1.1656)***			
CoeffQ5 x Ent(m)Q5 x Ent(d)Q5			-10.2875 (2.2161)***			
CoeffQ5 x Ent(s)Q5					-20.8313 (1.5403)***	-18.8323 (1.5642)***
CoeffQ5 x Ent(h)Q5						7.8233 (1.0768)***
CoeffQ5 x Ent(s)Q5 x Ent(h)Q5						-13.8406 (1.9359)***
CoeffQ5		-21.8691 (1.7301)***	-22.9046 (1.7893)***		-20.9631 (1.7185)***	-21.8396 (1.7713)***
Ent(m)Q5		30.2966 (1.3871)***	30.2561 (1.3843)***			
Ent(d)Q5			5.2053 (1.0085)***			
Ent(s)Q5					29.3873 (1.4277)***	29.4470 (1.4231)***
Ent(h)Q5						10.7348 (1.0398)***
Total duration (in days)	-0.0863 (0.0027)***	-0.0784 (0.0025)***	-0.0777 (0.0025)***	-0.0687 (0.0024)***	-0.0629 (0.0023)***	-0.0622 (0.0023)***
Total active days	-0.0447 (0.0114)***	-0.0332 (0.0109)***	-0.0592 (0.0112)***			
Total active hours				-0.0069 (0.0030)**	0.0011 (0.0027)	-0.0043 (0.0027)
No. of active categories	3.2632 (0.1486)***	2.2858 (0.1301)***	2.2183 (0.1317)***			
No. of active subcategories				0.8978 (0.0417)***	0.5612 (0.0361)***	0.5118 (0.0367)***
R^2	0.0953	0.1367	0.1379	0.1238	0.1612	0.1649
N	32135	32135	32135	32135	32135	32135

Table 7: Performance regressions (Y = #Total answers / ΣContrib. in terms)
Heteroscedasticity-consistent standard errors are in the parenthesis; *** 1%, ** 5%, * 10%

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Coeff	-0.5693	-0.2346	-0.2331	-0.5402	-0.2174	-0.2133
	(0.2393)**	(0.1076)**	(0.1068)**	(0.2310)**	(0.1057)**	(0.1043)**
Ent(main category)	-0.2739	-0.3177	-0.3128			
	(0.1425)*	(0.1582)**	(0.1629)*			
Ent(day of week)	0.0583	0.0749	0.0515			
	(0.0873)	(0.0857)	(0.0941)			
Ent(subcategory)				-0.2008	-0.2441	-0.2370
				(0.1519)	(0.1650)	(0.1639)
Ent (hour of day)				-0.2745	-0.2535	-0.3008
				(0.1346)**	(0.1323)*	(0.1444)**
CoeffQ5 x Ent(m)Q5		-0.2974	-0.2816			
		(0.0891)***	(0.0878)***			
CoeffQ5 x Ent(d)Q5			0.0664			
			(0.0804)			
CoeffQ5 x Ent(m)Q5 x			-0.1160			
Ent(d)Q5			(0.0772)			
CoeffQ5 x Ent(s)Q5					-0.3155	-0.2842
					(0.0888)***	(0.0867)***
CoeffQ5 x Ent(h)Q5						0.1273
						(0.0651)*
CoeffQ5 x Ent(s)Q5 x						-0.2138
Ent(h)Q5						(0.0662)***
CoeffQ5		-0.4338	-0.4442		-0.4215	-0.4386
		(0.0706)***	(0.0819)***		(0.0600)***	(0.0675)***
Ent(m)Q5		0.5405	0.5401			
		(0.1183)***	(0.1180)***			
Ent(d)Q5			0.0470			
			(0.0730)			
Ent(s)Q5					0.5962	0.5962
					(0.1178)***	(0.1182)***
Ent(h)Q5						0.0929
						(0.0536)*
Total duration (in days)	-0.0015	-0.0013	-0.0013	-0.0010	-0.0009	-0.0009
	(0.0003)***	(0.0003)***	(0.0003)***	(0.0002)***	(0.0002)***	(0.0002)***
Total active days	-0.0012	-0.0010	-0.0013			
	(0.0009)	(0.0009)	(0.0007)*			
Total active hours				-0.0004	-0.0002	-0.0003
				(0.0003)	(0.0002)	(0.0002)
No. of active	0.0565	0.0384	0.0378			
categories	(0.0088)***	(0.0064)***	(0.0068)***			
No. of active				0.0176	0.0107	0.0103
subcategories				(0.0052)***	(0.0042)**	(0.0041)**
R^2	0.001	0.0016	0.0016	0.0011	0.0018	0.0018
N	32135	32135	32135	32135	32135	32135

Table 8: Performance regressions (Y = #Best answers / ΣContrib. in char.)
Heteroscedasticity-consistent standard errors are in the parenthesis; *** 1%, ** 5%, * 10%

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Coeff	-2.5381	-1.0513	-1.0470	-2.4663	-1.0038	-0.9884
	(0.9878)**	(0.4060)**	(0.4049)**	(0.9839)**	(0.4122)**	(0.4055)**
Ent(main category)	-0.9825	-1.1189	-1.1109			
	(0.1867)***	(0.1921)***	(0.1942)***			
Ent(day of week)	0.6904	0.7527	0.7036			
	(0.2880)**	(0.2845)***	(0.3111)**			
Ent(subcategory)				-0.2642	-0.4069	-0.4036
				(0.1438)*	(0.1481)***	(0.1479)***
Ent (hour of day)				-0.8742	-0.7789	-0.8257
				(0.2116)***	(0.2078)***	(0.2533)***
CoeffQ5 x Ent(m)Q5		-2.6028	-2.4712			
		(0.2820)***	(0.2825)***			
CoeffQ5 x Ent(d)Q5			0.3287			
			(0.2217)			
CoeffQ5 x Ent(m)Q5 x Ent(d)Q5			-0.9848			
			(0.3554)***			
CoeffQ5 x Ent(s)Q5					-2.7473	-2.4766
					(0.2869)***	(0.2902)***
CoeffQ5 x Ent(h)Q5						0.8649
						(0.2059)***
CoeffQ5 x Ent(s)Q5 x Ent(h)Q5						-1.7665
						(0.3165)***
CoeffQ5		-1.6205	-1.6719		-1.6146	-1.7491
		(0.1572)***	(0.1686)***		(0.1570)***	(0.1652)***
Ent(m)Q5		2.3624	2.3614			
		(0.2677)***	(0.2675)***			
Ent(d)Q5			0.0942			
			(0.1943)			
Ent(s)Q5					2.6339	2.6230
					(0.2896)***	(0.2902)***
Ent(h)Q5						0.0301
						(0.1965)
Total duration (in days)	-0.0063	-0.0057	-0.0057	-0.0047	-0.0042	-0.0042
	(0.0005)***	(0.0005)***	(0.0005)***	(0.0004)***	(0.0004)***	(0.0004)***
Total active days	-0.0079	-0.0070	-0.0076			
	(0.0019)***	(0.0019)***	(0.0019)***			
Total active hours				-0.0016	-0.0009	-0.0009
				(0.0005)***	(0.0005)*	(0.0004)**
No. of active categories	0.3498	0.2764	0.2758			
	(0.0258)***	(0.0247)***	(0.0250)***			
No. of active subcategories				0.0887	0.0598	0.0604
				(0.0073)***	(0.0070)***	(0.0072)***
R^2	0.0182	0.025	0.0251	0.0171	0.0245	0.0247
N	32135	32135	32135	32135	32135	32135

Table 9: Performance regressions (Y = #Best answers / ΣContrib. in terms)
Heteroscedasticity-consistent standard errors are in the parenthesis; *** 1%, ** 5%, * 10%