Chat More and Contribute Better: An Empirical Study of a Knowledge-Sharing Community*

Xiaomeng Chen
Dyson School of Applied Economics and Management
Cornell University
137 Reservoir Avenue
Ithaca, NY 14853
USA
xc382@cornell.edu

Chris Forman
Dyson School of Applied Economics and Management
Cornell University
137 Reservoir Avenue
Ithaca, NY 14853
USA
chris.forman@cornell.edu

Michael Kummer
School of Economics
2.25R REG, University of East Anglia
Norwich Research Park
Norwich NR4 TJ7
UK
M.Kummer@uea.ac.uk

Abstract

We analyze whether an informal second channel for communication can improve the efficiency of knowledge transfer in an electronic network of practice. We explore this question by analyzing the effect of chat rooms in the well-known Q&A forum Stack Overflow. We identify the causal effect using a difference-in-differences approach which exploits a feed functionality that non-selectively pushed all questions from the Q&A into the relevant chat rooms. We report two main findings: First, chat rooms reduced the time until a question in the main Q&A received a satisfactory answer. Second, chat rooms disproportionally benefited new users who asked lower quality questions. Our study has clear managerial implications: A second channel for communication can complement the main channel in online communities to enhance both efficiency and inclusion.

Keywords: Online community, knowledge sharing, user contribution

* We thank seminar participants at Cornell University, the Workshop on Information Systems and Economics, the Conference on Information Systems and Technology, and the INFORMS annual meeting for helpful comments and suggestions. We thank Stack Overflow for providing essential data. All opinions and errors are ours alone.

1. Introduction

In the digital age, the sharing of practice-related knowledge such as programming or legal advice often comes through computer- and internet-based communication technologies known as electronic networks of practice (Wasko and Faraj 2005). These networks of practice have been found to be an effective means of transferring knowledge across users (Boudreau and Lakhani 2009) and, as a result, have grown rapidly in recent years. For example, the discussion forum Stack Overflow has over 100 million users (Fullerton 2019) and a recent survey of that site shows that over 85% of its users visit the site multiple times per week (Stack Overflow 2019); Quora reached 300 million monthly visitors in 2018 (Browne 2018); and Zhihu, the largest Chinese Q&A community exceeded 220 million monthly visitors by 2018 (Wikipedia 2018).

Given the importance of these sites, recent research has sought to identify and improve design features that will improve interactions within them. Much of this line of work has sought to identify design choices that will improve user contributions (e.g., Wasko and Faraj 2005; Jabr et al 2014; Goes, Guo, and Lin 2016). However, while stimulating user contributions is important to the success of knowledge-sharing platforms, insufficient contributions are not the only barrier to effective knowledge transfer. In particular, many networks of practice incorporate specific conventions, such as gamification, that encourage contributions. However, these conventions may at the same time create barriers to community participation among certain groups, such as new users (Ford, Lustig, Banks, and Parnin 2018; Ford, Smith, Guo, and Parnin 2016; Valisescu, Capiluppo, and Serebrenik 2014) when members enforce community norms (Yang 2020).

Online communities have recently explored adding new virtual spaces that, while lacking well-known features like gamification that encourage contributions, may facilitate informal conversations among users. These spaces, sometimes termed "third places" because of their similarity to offline community meeting places like cafes and coffee shops (Oldenburg 1999), are becoming increasingly popular. For example, Wikipedia Talk pages provide a place for editors to interact informally and discuss edits to Wikipedia pages and the group chat feature in Reddit. These pages provide a place for users to interact with fewer formal norms, and in doing so may improve outcomes in the core community. However, to our knowledge there has been little formal assessment on the impact of these new spaces on networks of practice. This is a significant gap in understanding. If such spaces improve outcomes, then networks of practice should consider them as a potential additional channel for users to interact, particularly if they help new users who may be unaware of norms in the core community.

In this paper we examine how the introduction of a chat functionality affected efficiency of knowledge exchange on Stack Overflow (SO), the largest online question and answer (Q&A) community for programming. We measure the changes to the efficiency of the knowledge exchange on SO as the probability with which questions receive an accepted answer within 1 hour, 2 hours, 4 hours or 8 hours. Ideally we would like to compare the efficiency of questions that are discussed in chat rooms with the efficiency of questions that do not appear in a chat room. However, because questions that appear in chat rooms might be fundamentally different from question that do not appear in a chat room, a direct comparison of questions that do vs. do not appear in chat rooms may be contaminated by selection bias.

To overcome concerns of selection bias and identify the effect of interest, we exploit the feed function in the Stack Overflow chat rooms. Whenever a user turns on the feed function of the chat room, the feed will push all newly generated questions that are related into the chat room. We build a difference in difference (DID) approach around the activation of a feed to identify the effect of chat rooms on the speed with which questions receive an acceptable answer.

Our findings suggest that chat rooms increase the efficiency of the knowledge exchange in the main Q&A community: for example, when a question is pushed into the chat room of a Q&A community, the probability of the question getting an accepted answer within 2 hour increases by 3.2 percentage points. A causal interpretation of this estimate is supported by a range of robustness analyses, including a pretrend analysis and exploration of alternative control groups.

The implications of being pushed into chat rooms are not the same for all questions and question askers, however. They are strongest for questions that are less likely to receive an answer in the Q&A site, namely those that are asked by new users and those that are of low quality. For example, one set of estimates show that users with zero or lower reputation are 14.6 percentage points less likely than users with positive reputation to receive an answer within 2 hours in the Q&A site. However, when questions from these users are pushed into the chat room, the likelihood of receiving an answer within 2 hours increases by an additional 5.9 percentage points. We further decompose these effects based on user reputation and question quality, finding that the effects of the feed are strongest among questions of low quality and who are asked by new (low reputation) users.

Our findings have several managerial implications, because they highlight the ability of less formal and complementary communication channels to support the transfer of knowledge in communities of practice. First, these additional channels can increase the overall efficiency. Second, they can complement an efficiency-focused main channel by providing a space in which disadvantaged groups, such as newcomers, can benefit from a "more welcoming" environment which facilitates improved access to community knowledge.

2. Related Literature.

Our research furthers prior work that has investigated the implications of platform governance choices on networks of practice where sharing of practice-related knowledge is mediated through online channels, labeled electronic networks of practice (Wasko and Faraj 2005). One line of literature has investigated platform actions that increase the knowledge *contribution* of users (Raban 2008, Wasko and Faraj 2005, Wiertz and de Ruyter 2007 and Jabr, Mookerjee, Tan, and Mookerjee 2014) and their *commitment* to the community (Moon and Sproull 2008). One particular area of focus has been an exploration of the implications of gamification, the process of adding game-like features to something to encourage participation. Previous studies have shown that the addition of gamification incentives motivate user contributions (Jabr, Mookerjee, Tan, and Mookerjee 2014; Liu, Santhanam, and Webster 2017). Gamified elements such as badges (Anderson et al 2013), status (Liu, Santhanam, and Webster 2017; Goes, Guo and Lin 2019) and feedback (Jabr, Mookerjee, Tan, and Mookerjee 2014; Moon and Sproull 2008) can effectively motivate users to engage in online systems.

While increases in user contributions will increase the likelihood of knowledge transfer between network participants, they do not ensure it. It has long been known in other settings of computer-mediated communications that knowledge can be communicated without being absorbed by the intended receiver of the communication (Alavi and Leidner 2001). A selected few studies discussed the knowledge transfer and exchange activities among users (Wicks et al 2012, Curran et al 2009, Dinh et al 2011), but they were unable to observe to which extent the transmission of knowledge is efficient or successful. Increases in contributions may not lead to an improvement in outcomes, however. Our research contributes to two lines of work that have explored different aspects of this problem.

First, incentives to contribute may indirectly affect outcomes for some community members. Within the context of research on electronic networks of practice like Stack Overflow, the existence of barriers to participation from groups such as newcomers and females is an active area of work (Ford et al 2016; Ford et al 2018; Nian et al. 2019). We contribute to a line of literature that has explored issues around how to best integrate newcomers into online communities. Early research in this area is summarized in Kraut et al (2012). For example, since newcomers are especially likely to leave the community if their first contribution is rejected (Halfaker, Kittur, and Riedl 2011, Zhang and Zhu 2006), online communities such as Wikipedia employ “don’t bite the newcomer” policies to lessen the negative impact of entry barriers for newcomers. More recent research has investigated how peer awards can disproportionately increase the engagement of users in communities (Burtch et al. 2020). The present paper contributes to this line of research by showing the implications of the introduction of a supportive channel on outcomes for new users relative to the rest of the community.

Second, it is possible that answers provided are not understood or absorbed by the original asker of the question. A long line of literature in information systems has explored the properties of different communication channels and their efficacy in different offline contexts (e.g., Daft and Lengel 1986; Daft et al. 1987; Dennis et al. 2008; George, Carlson, and Valacich 2013). One takeaway from this literature is that communication channels may show differential efficacy in different contexts. In online spaces, prior work has investigated the impact of introducing new communication channels in specific settings such as open source development and education. For example, Shihab et al (2009) and Elliott and Scacchi (2003) documented that the introduction of a supportive channel based on chat increased contributor retention and resolved conflicts in an open source project. Prior work has also investigated the implications of new online communication channels such as chat on learning outcomes (e.g., Coetzee, Fox, and Hartmann 2014). We further this line of work by identifying the benefits of a supportive channel in a popular electronic network of practice.

3. Research Context

Stack Overflow is a large programming Q&A community where users can ask and answer questions. In this section we provide a brief overview of how the primary Stack Overflow site functions, and then discuss the introduction of chat rooms. Last we discuss the introduction of the feed function, which will play a critical role in our identification strategy.

3.1 Stack Overflow Q&A community.

Users can ask and answer questions related to programming in the Stack Overflow Q&A community. The user who asks the question ("askers") can post a question and other users can answer the question ("answerers"). We summarize key features that are important to our analysis.

Questions, answers and accepted answers. A question can have multiple answers and all users can vote up or vote down for an answer or a question, which will be summarized as the "score" of the question or answer (up votes minus down votes). A user who asks the question can choose to select only one answer as the accepted answer if he thinks the answer is working well and sometimes the first answer may not be the accepted answer.

Tags. For each question, the asker can assign tags indicating the programming language, framework or the related module about the question. For example, the set of tags for one question in our data includes "c#" and "asp.net-mvc", indicating that solving this question requires a user who uses the c# language and who is familiar with the asp.net-mvc framework. Users can assign multiple tags to one question to help answerers to quickly identify the expertise needed to answer the question.

User reputation system. Stack Overflow uses a reputation system and badges to motivate users to contribute. Users can gain reputation points from the upvotes and downvotes of their questions or answers.

These votes are assigned from the community members. For example, users will earn 10 points for receiving an upvote, lose 2 points for receiving a downvote and gain an extra 15 points if their answers are "accepted."

3.2 Stack Overflow Chat rooms.

In October 15 2010 Stack Overflow launched a chat room feature to allow users to interact in a manner that is different from the main Q&A community (Atwood 2010).[1] There are two major types of chat rooms, chat rooms for interest-groups and discussion chat rooms. Discussion chat rooms are created on an ad hoc basis by users to discuss relatively narrow issues, such as a particular question on the Q&A forum. In contrast, chat rooms for an interest-group serve as a forum for a large number of users that share a particular interest, usually based upon a programming language. For example, "c#" "JavaScript", "PHP" "Python" and "Lounge<C++>" are all examples of typical chat rooms for interest-groups. In this paper, we focus on the interest-group chat rooms instead of the discussion chat rooms in this paper.

Questions in the main question and answer (Q&A) site can be added to the chat room through different means. For one, they can be pushed into the chat rooms by the askers themselves, or by other users. Questions can also be pushed into the chat rooms via feeds, i.e. bots that automatically push questions from the Q&A site to the chat rooms.

Chat room feeds can be turned on and off by users. Once a feed is turned on for a chat room, it pushes all questions containing a certain tag into the associated chat room. Figures 1a and 1b provide an example of how the feed works: Figure 1a shows a question that was asked on Dec 21st, 2013, and that contains the tag "Perl." Figure 1b shows the question after it was pushed into the Perl chat room immediately after being asked, together with all other newly generated questions tagged with "Perl." The question did not receive any response for 4 hours until the asker discusses the question with a potential answerer in the chat room and provides some clarification (seen in Figure 1b).

[Insert Figure 1a, Figure 1b here]

---

[1] https://stackoverflow.blog/2010/04/29/do-trilogy-sites-need-a-third-place/

4. Theoretical Motivation: Benefits of chat rooms

In this section we will discuss the theoretical motivation for our hypotheses. We start by discussing differences in the operational characteristics of online forums and chat rooms in online communities, simultaneously discussing their implications for user behavior. We will then examine the implications of these differences specifically in the context of an electronic network of practice like Stack Overflow.

Differences in the characteristics of online forums and chat rooms

Earlier Q&A communities utilized relatively unstructured forms of online communication such as electronic mail (e.g., Ahuja, Galletta, and Carley 2003) or online bulletin boards (Wasko and Faraj 2005). More recently, electronic networks of practice have increasingly relied upon more structured forms of interaction in which gamification incentives are provided to users to encourage contributions (Jabr, Mookerjee, Tan, and Mookerjee 2014; Liu, Santhanam, and Webster 2017). Gamification takes elements from game designs to make tasks more engaging for users to, within the context of online forums, generate additional content (Liu, Santhanam, and Webster 2017; Goes, Guo, and Lin 2016).

Gamification: A variety of studies have reviewed the potential benefits of gamification (Jabr, Mookerjee, Tan, and Mookerjee 2014; Liu, Santhanam, and Webster 2017). For example, gamification has been shown to increase user contribution (Goes, Guo and Lin 2016, Cavusoglu et al 2015), user retention (Moon and Sproull 2008), quality of contributions (Moon and Sproull 2008) and user coordination (Forte et al 2012). While the benefits of gamification are well-established, recent literature has sought to understand how the design of gamification should fit the task to which it is intended. Motivated from the literature on task-technology fit (TTF), game design elements should match the intended purpose of a system (Goodhue 1995; Goodhue and Thompson 1995; Liu, Santhanam, and Webster 2017). This same work suggests environments in which gamification may be incongruent with the underlying task. We highlight certain gamification decisions that could impact its efficacy in certain types of environments.

For example, gamification often engenders competition among users (Morschheuser et al 2016, Massung et al 2013). For example, if a question answerer arrives at the correct solution for a problem, this may foreclose opportunities for others to contribute. Thus, users are competing against one another to be one of the first to answer a question. However, while in some cases competition is desired or at least neutral, in others it may be desirable to obtain cooperation between users. This may particularly be the case when users need to work together to solve a problem.

Further, in some cases the underlying task must be changed so that gamification is possible. For example, elements such as points or badges may not be suitable for work that lacks quantifiable performance measures (Liu, Santhanam, and Webster 2017).[2] For these reasons settings with gamification may choose incentivize certain types of (quantifiable) effort over others, or even restrict certain types of communication between users. For example, in a question and answer setting that rewards quality contributions arising from questions and quality answers, users may be discouraged from posting comments that are not aligned with the central purpose of the platform and for which it is difficult to objectively measure quality.

Many electronic networks of practice that employ gamification are characterized by asynchronous communication. One advantage of such asynchronous postings is that they allow users additional time to compose messages (Dennis, Fuller, and Valacich 2008), a characteristic that may be particularly valuable in an environment where messages may be evaluated by other users. However, synchronous communications may be preferred in some environments.

Synchronicity: One way to think through the effect of different communication methods on outcomes is through media synchronicity theory (MST), developed by Dennis et al. (2008). MST argues that communication is composed of two primary properties, conveyance and convergence (Dennis et al. 2008; George, Carlson, and Valacich 2013). Conveyance focuses on the transmission of a diversity of information, while convergence focuses on clarifying the meaning of information that has already been exchanged or shared. Without convergence, users will not reach shared understanding. Convergence often requires rapid back-and-forth transmission of small quantities of information (Dennis et al. 2008; George, Carlson, and Valacich 2013), the type used in synchronous communication. MST argues that for

[2] Although for a recent exception see Burtch, He, Hong, and Lee (2020).

conveyance processes, asynchronous communication like asynchronous forums is preferred. However, for convergence processes, synchronous communication like synchronous chat is better. Therefore, in an online question and answer forum, transmission of messages through asynchronous forums may be appropriate in most cases to facilitate conveyance. However, in settings where it is difficult to achieve shared understanding – as when, for example, a question is unclear – synchronous communication may be particularly valuable.

The heterogeneous effects of synchronous and asynchronous communication tools has been examined within the context of online learning contexts (e.g., Coetzee, Fox, and Hartmann 2014; Johnson 2006; Rovai 2001; Oztok, Zingaro, Brett, and Hewitt 2013) and open source projects (Shihab et al 2009; Elliott and Scacchi 2003). In those contexts, (synchronous) chat has been argued to provide faster responses to questions in MOOCs than (asynchronous) forums (Coetzee, Fox, Hearst, and Hartmann 2014) and help to resolve conflicts in open source software communities (Elliott and Scacchi 2003).

New communication methods can also bring about changes to the social structure of the community which can affect productivity and shape user contribution behaviors (Singh, Tan and Mookerjee 2011, Singh, Tan and Youn 2011). Such interactive environments may enhance network ties among users, which can have positive effects for complex knowledge exchange (Coleman 1988 and Walker et al. 1997). Studies also show that people are more willing to share knowledge when there is mutual knowledge and expectations of ongoing reciprocity (Athey and Ellison 2014), which may be more feasible in environments in which rapid communication is more feasible. For example, chat can more effectively (than forums) encourage community building and help to facilitate the forming of relationships (Coetzee, Fox, and Hartmann 2014; Rovai 2001). The role of free form discussion in facilitating community building has been highlighted in prior research around online learning and educational technology (Johnson 2006; Branon and Essex 2001). For example, comparing synchronous and asynchronous discussion, Dede and Kremer (1999) note that asynchronous discussion provided richer exchange, however required more time and engendered less social interaction than synchronous chat.

Implications for performance on Stack Overflow

In the prior section we discussed how two features, gamification and synchronicity of communications, could potentially impact behavior in an electronic network of practice. To test these assertions formally, we will examine the implications of the introduction of a supportive channel, or "third space" (Oldenburg 1999; Steinkuehler and Williams 2006), on interactions in an electronic network of practice. Consistent with the goals of a third space to create an environment that promotes conversation and in which rank and status do not play a role, we examine the implications of the introduction of a new supportive channel that includes synchronous communications and in which gamification incentives do not play a role on user behavior. Our interest is on the impact of third spaces on user behavior in general. However, the characteristics of a third space will likely vary across different electronic networks of practice. Given the presence of these differences and the impact they may have on user behavior, rather than examine the implications of the introduction of a third space across multiple platforms, we instead examine the implications in one particular environment, Stack Overflow.

The third space in Stack Overflow takes the form of a chat room (Stack Overflow 2020). Chat rooms in Stack Overflow function like discussion boards found elsewhere and there are several differences from the traditional Q&A forum. First, there are no explicit incentives to post into the chat room, unlike the Stack Exchange Q&A forums where users receive points for asking, answering, and editing questions and answers (among other things). That is, there are no gamification incentives in the chat rooms. Relatedly, there are fewer explicit rules on what can be posted in chat rooms. Lastly, communication in chat rooms occurs synchronously, with users responding immediately to the posts of one another.

As a result of synchronicity and the negligible role of gamification incentives, users may behave differently in chat rooms and in the general Q&A site. These differences may mean that chat rooms have advantages for answering certain types of questions. We focus on two particular differences: user reputation and question quality.

In the Stack Overflow Q&A site there are well-established norms of behavior; as noted above, these norms help to facilitate gamification. Active community members are concerned with preserving the platform's norms, and may behave in a hostile manner to users who do not follow them (Ford, Smith, Guo, and Parnin 2016, 2018; Valisescu, Capiluppo, and Serebrenik 2014). Moreover, for newcomers, there may be barriers to learning these norms (Steinmacher et al. 2015). As a result, some users may decide not to ask or answer a question due to fear of negative feedback (Ford et al. 2016). In chat rooms, norms of behavior may be less well established and so users may be less apprehensive about contributing. Moreover, given the more flexible and permissive behavioral norms with which users can contribute to the community, existing and experienced users may be able to mentor newcomers on behavioral norms to help them translate questions into a format that can better be answered by the community. These features may be particularly helpful in engaging with new users and their questions.

Gamification in the Stack Overflow forum may encourage competition between users in ways that make it more difficult to answer certain types of questions. While Stack Overflow offers a mechanism to add comments to an answer, they cannot be threaded, and incentives are geared toward being the first to answer a question (Bosu et al. 2013; Zagalskey et al. 2017). As a result, answers on Stack Overflow are geared more toward "crowd" knowledge construction that is not cumulative as opposed to "participatory" knowledge construction that involves a collaborative process of building upon prior knowledge (Zagalsky et al. 2017). The lack of threaded conversations and the incentives to be the first to answer may make it harder for the community to achieve shared understanding on questions that are unclear or initially appear to be low quality. Such questions are also more likely to be asked by new members of the community (Ahn et al. 2013).

As noted above, differences in communication patterns can engender differences to network structure which can influence outcomes (Rice 1994, Butler 2001). First, users interact only under one post in QA forum while in chat rooms they can interact more frequently within a stream of posts. Therefore, they could form stronger social ties by communicating with each other in chat environment. Stronger social tie can positively affect productivity in collective actions (Ren et al 2007, Marwell and Oliver 1988, Krackhardt et al 2003) such as responding to questions. Second, the chat room is often built by leading users and enables them to have social exchange with other users. It enhances the individual centrality, the extent to which the individual is linked to others (Ahuja et al 2003), for the leading users. Such individual centrality can motivate leading users to contribute (Wasko and Faraj 2005, Grewal, Lilien and Mallaprogada 2006) and advocate other users in the subcommunity to contribute.

The **synchronicity** of chat rooms may also help to facilitate shared understanding between the question asker and answerer. Even if a question is answered online, the question asker may not be able to absorb the knowledge effectively (King and Lakhani 2011); some authors have argued that in the Stack Overflow Q&A forum a certain level of experience is required to understand answers (Zagalsky et al 2017). That is, shared understanding or convergence may not be achieved. The synchronous communications available through chat rooms can help to facilitate convergence, particularly in environments in which shared understanding is difficult to achieve, as when questions are unclear or asked by new users who have less understanding of the norms of the community.

In this section we have provided an overview of how chat rooms can facilitate getting an answer earlier for questions that are posted on the Stack Overflow Q&A site. However, the benefits of chat rooms are not equal across questions. For many questions, posting on the Q&A site will be sufficient. We have identified two conditions for which the value of chat rooms are likely to be high: questions that are asked by new users and those that are not clear.

Research Questions: In future sections, we assess the salience of these observations in our data. In particular, *we examine whether, other things equal, adding a discussion chat room will decrease the time it takes for related questions that are posted in the chat room to receive an accepted answer on the Q&A site*. Given the discussion above, we then explore heterogeneity in the above treatment effect based upon characteristics of the question asker and question, *examining differences in the effects of adding a chat room on outcomes when the question asker is new to the community and when the question is low quality*. In the next section we discuss our empirical approach for answering these questions.

5. Identification and Data Strategy and Data

5.1 Identification of the Main Effect

We seek to understand how adding a question to the chat room impacts the time it takes for that question to receive an accepted answer on the Stack Overflow Q&A site. We face an identification challenge because many reasons could drive a user to push a question into a chat room. Such questions might be harder to answer, or they might be more interesting, important, or more urgent, etc. Therefore, a direct comparison between questions being pushed into the chat rooms and questions not being pushed into the chat rooms could lead to a selection bias in the results: For example, if users tend to push more difficult questions into the chat rooms which generally take longer to get an answer, the selection would lead to a negative bias in our estimates of the effects of chat rooms.

To address potential identification concerns arising from user-level choice to push questions into chat rooms, we exploit the feed function that pushes questions automatically into the chat rooms and use a difference in differences (DID) approach to identify the causal impact of moving a question into the chat room. Specifically, our primary estimation approach will compare outcomes of questions related to tags that have been "treated by the feed" (i.e., a user has specified that all questions with this tag will be pushed into a specific chat room) relative to a comparison group of questions that have not been so treated. Figure 2 illustrates examples of treated tags for which the associated chat rooms turned on the feed between October 2010 and December 2016, the period over which chat rooms and feeds were features that were incorporated into Stack Overflow until the end of our sample period.

Our approach of using treatment by the feed mitigates concerns about user level selection on unobservables that might occur if we used user decisions to move questions into the chat room as our treatment. However, there are several things to note about our approach. First, feeds are turned on over a period of time in our sample, so treatment occurs over a period of time (examples are shown in Figure 2). Second, the decision to turn on the feed is made by users of the chat rooms. While we do not have specific evidence for the factors shaping user decisions to turn on the feed, it is possible that these are shaped by activity in the Q&A site, activity in the associated chat room, and the likelihood of getting an answer in the associated tag in the Q&A site. If trends in our treated tags are shaped by these factors in a way that is the systematically different from control tags, they have the potential to shape our results. To address this issue, we do two things. First, we use matching techniques to identify control tags with similar characteristics. Second, we focus upon a short time window before and after the tag is treated by the feed, identifying the control group based on characteristics during this short period and examining outcomes over this same period. This will mitigate concerns about differences in unobservable trends at the tag level potentially shaping our results.

We use propensity score matching (PSM) to identify control tags that are similar to the treated tags but have not yet been treated. After identifying the control group, we construct a panel data set to run a difference in difference (DID) regression. For each incidence of a treated tag in our data, we identify suitable control tags as described below and then include in our sample all questions from treated and control tags generated in a 4-week time window: 2 weeks before the associated chat room turns on the feed and 2 weeks after. Below is our baseline empirical specification:

$$Y_{ijt} = \beta_0 + \beta_1 TurnOnFeed_{ijt} + \alpha_j + WeekDay_t + Week_t + \varepsilon_{ijt} \quad (1)$$

Y_{ijt} indicates the outcome variable of question i in tag j generated on time t. For treated tag j, $TurnOnFeed_{ijt} = 1$ when a chat room associated with tag j turns on the feed at time t and $TurnOnFeed_{ijt} = 0$ before and after the feed is turned on. For all control tag units, $TurnOnFeed_{ijt} = 0$ during the 4-week time window. We also include tag level fixed effects α_j, weekly dummies ($Week_t$) and weekday dummies ($WeekDay_t$).

Our primary outcome variables Y_{ijt} denotes whether the focal question has an accepted answer within 1, 2, 4, and 8 hours. Thus, our outcome variable is a binary variable and our estimation approach is a linear probability model. Our use of a linear probability model reflects several considerations. First, the linear probability model will provide consistent estimates of the parameters of interest. A major concern is the existence of heteroskedastic standard errors, which we adjust for using robust standard errors. Second, the linear model allows for a more straightforward interpretation of the implied marginal effects from our parameter estimates. This is particularly the case when we explore heterogeneity in our treatment effect based on question and user characteristics in which we will interact the variable $TurnOnFeed_{ijt}$ with other variables: In these models, the interaction term in nonlinear models will not identify the cross partial derivative (Ai and Norton 2003).

We estimate this regression using the within estimator. This method means that the calculated "within" R-squared values do not take into account the explanatory power of the fixed effects. Therefore, for our main results, we also estimated an equivalent, though computationally inefficient, "one-way fixed effects" estimator in order to calculate R-squared values that include the fixed effects.

5.2 Data

Our primary source of data is publicly available through the Stack Exchange API. We supplement the publicly available information with additional data scraped from chat.stackoverflow.com, and user reputation data obtained through agreement with Stack Overflow.[3]

To assemble our data we identify questions that are pushed into an associated chat room by a feed. We identify treated questions as those for which one chat room that is associated with a tag turns on a feed. We refer to a "tag episode" as a period over which a tag is treated by the feed. For example, Figure 2 lists six such tag episodes. Our data set consists of a set of tag episodes in addition to a set of control tags that are used in each tag episode. We next discuss our procedure for identifying tag episodes and their associated control tags.

As discussed above, our identification strategy requires that we use feeds that push *all* questions into the associated chat room without discrimination. Our matching strategy also requires all treated tag units to have at least one week of activity in the associated chat room before the feed was turned on. From

[3] No user identities were revealed through this process.

October 2010 to December 2016, we found 26 treated tag episodes that satisfy these criteria.[4] Each of these tags are unique except for python, which is treated twice in our data.[5]

There is great heterogeneity in the intensity of activity in chat rooms in Stack Overflow. Some tags have thousands of posts every week, while others go weeks at a time without activity. This heterogeneity creates problems for our estimation approach; chat rooms with very little activity are unlikely to influence the outcomes of questions that are pushed into them. For the propensity score matching to be able to identify appropriate control tags, we restricted our sample to tags (control and treated) that are active for more than 80% of the observable weeks during our sample period. This restriction was necessary to reduce heterogeneity across tags in the PSM procedure. Note, however, that the results are robust to using alternative cutoffs. We show the results using a 50% threshold in Table A4 and Table A5 in the Appendix. This procedure results in a dataset that contains 18 treated tags and 1012 potential control tag episodes before applying the PSM procedure.

We then run a logit regression where the unit of observation is tag-week and in which the dependent variable is whether the tag is associated with a chat room that turns on the feed during the week. As matching covariates we use weekly aggregations of tag-level characteristics (question count, asker count, answer count, answerer count, and number of answerers per question), chat room level characteristics (message count, user count), lagged values (t-1) of those characteristics, and weekly growth rate of question count as matching covariates. We then use the estimated coefficients to calculate a propensity score for tag-week observations. For each treated tag unit, we calculated the difference between its propensity score in week immediately before it is treated (days -7 to day -1 prior to treatment) and the propensity score of all the potential control tags that are active in the same week.

[4] Of 73 tags for which a feed was activated, 24 applied some selective filter such as feeding only questions with bounties (extra reputation reward provided by the askers). An additional 23 started the feed within one week of the associated chat room opening.

[5] The tag "python" has two associated chat rooms that turned on the feed function in January 2013 and September 2013, respectively. We include both episodes of treatment in our data.

For each treated tag episode, the PSM procedure pre-selected 5 control tags that have the closest squared difference in propensity score. Of those suggestions, we dropped control tag-episodes that had no questions during the tag episode and eliminated instances of duplicate use of the same control tag during the same time period, as might happen if the same control tag was identified for two tags that were treated at the same time.[6] The final matched data set includes 166,435 questions from 18 treated tag episodes and 83 control tags episodes. We observe treatment for 1,085 questions.

Table 1 and Table 2 show the balance check on matching covariates after all of these changes, before and after PSM. Appendix Table A0 provides a description of each of the matching covariates, and Appendix Table A1 provides the results of the PSM model estimation. Before PSM, the differences in means between the treatment and control are insignificant except for log(AnswerCount), log(AnswerCount(t-1)) and Answerer per question. Although the differences in means between the treatment and control group are small even before PSM, we choose to use control units selected by PSM in our baseline regressions because use of the PSM in general reduces these differences even further. The main results stay robust using control tags without being selected by PSM (Tables A2 and A3 in the Appendix).

[Insert Table 1, Table 2 here]

Table 3 provides summary statistics for the outcome variables in our regressions. We focus on time to receiving an accepted answer rather than time to receiving an answer because our analysis focuses on whether chat rooms facilitate the conveyance of an answer that addresses the question asker's need and for which the question asker and answerer have reached a shared understanding. Askers have the incentive to accept useful answers since the site provides a reputation reward for accepting an answer. 50.7% of questions receive an accepted answer within 8 hours (Accept8Hour) and only 9.9% (60.6%-50.7%) of

[6] For example, tag "java" was selected as control tag for both treated tag "mercurial" and "c++". And since the treatment time of "mercurial" and "c++" were very close, the 4-week time window of these two treated tags were overlapped. And we dropped "java" as control tags for "c++".

questions receive an accepted answer after 8 hours. So we focus on the likelihood of a question receiving an accepted answer within 1, 2, 4, and 8 hours as our outcome variables.

We use answer score to measure answer quality. The answer score is measured by the total upvotes minus downvotes assigned to the answer from users on the platform (excluding the question asker him/herself). It is an indicator of the collective evaluation of answers from users in the community (Burghardt et al 2016). Since the score of an accepted answer can only be observed for questions that have an accepted score, to incorporate the possibility that appearing in the chat room might change the distribution of quality across all answers we also include the average score of all the answers for the focal question as an alternative measure of question quality.

[Insert Table 3 here]

6. Results

6.1 Main effect of the feed function.

Tables 4 and 5 present our main results. Table 4 shows that feeding a question into a chat room increases the likelihood of getting an accepted answer within 2, 4, and 8 hours. Column 2 shows that after feeding a question into the chat rooms, the likelihood of getting an accepted answer within 2 hours increases by 3.2 percentage points, or 7% (0.032/0.46). Results are qualitatively similar for 4 hours and 8 hours. However, while the point estimate for the effect of feeding a question on the likelihood of receiving an accepted answer in one hour is positive, it is smaller than other time intervals and not statistically significant. We believe there are several reasons for the non-significant result over the shorter time interval. First, there is often a short lag—less than one hour – between the time when a question is posed on the Q&A site and when it appears in a chat room. Second, even after the question appears in the chat room, it takes some time for answerers to view the question and identify an answer, and then for the asker to test and accept a promising answer.

In Table 5 we examine the effects of treatment on answer quality. Our measure of answer quality is based upon the score of the answer. As noted above, this is a commonly used measure of question quality

(Baltadzhieva and Chrupala 2015, Ravi et al 2014; Arora et al 2015). The average score of answers that are treated by the feed increases by 0.3 or 5.9% (Columns 2 and 4). However, this is likely driven by the increase in views received by questions that are treated by the feed; the number of question views increases by 17.7% (Column 5). Moreover, treatment by the feed neither increases nor decreases the score of the accepted answer (Columns 1 and 3). Note that the number of observations in these tables vary based upon whether the question receives an answer or accepted answer.

Table 6 shows that the main results are driven by the tags associated with big chat rooms that are more active. We define big chat rooms as those that have more than 140 messages per week and 11 treated tags are classified as big chat rooms. Column 2 shows that while treatment has no effect (statistically or economically) on the likelihood of receiving an accepted answer in 2 hours for small chat rooms, for questions that are pushed into large chat rooms the likelihood of receiving an accepted answer increases by 7.9 percentage points. Moreover, Column 1 shows that for big chats, the feed significantly increases the likelihood of getting an accepted answer even within 1 hour.

6.2 Robustness checks

Our identification relies on the assumption that treated tag questions have similar time trends to accepted answer as control tag questions but for the incidence of treatment. To explore the validity of this assumption, we include leads and lags to illustrate the pre-trend and post-trend between control and treated units.

We present the pre- and post-trend analysis in Table 7. We generated dummy variables for questions that are affiliated with treated tags, based upon the timing of when the question was asked relative to treatment. To generate pre-trend variables, we created dummy variables that were equal to one for questions that are affiliated with treated tags but which were asked between 6 and 4 days before treatment (day - 6 to day – 4 in Table 7) and also for questions that were asked between 3 and 1 days prior to treatment (day - 3 to day – 1 in Table 7). We similarly generated dummy variables that are equal to one for treated questions based upon the number of days the questions were asked after to the initial switch in treatment status (i.e., the first day the questions with that tag began to be pushed into the chat room).[7] The reference group in these questions are questions generated on day -7 to day -14.

Our research design of using the feed to identify the effects of chat rooms brings a useful source of exogenous variation to identify our parameter estimates. However, because we use a fairly specific source of variation that exists for only a subset of questions, it limits the statistical power for our pretrend analysis. There are only 1085 questions feeded in the post-period for 18 treated tag episodes and the decomposition into days reduces the statistic power relative to our earlier results. The individual parameter estimates are identified off of questions that are asked in each specific time window and, as noted earlier, for some tags treatment does not last the entire post period.

Columns 1 to 4 show that there is no evidence of a statistically significant effect of being affiliated with a treated tag prior to treatment. Due to the statistical power issues mentioned above, the coefficients in the post period are economically significant but only weakly statistically significant. However, the coefficients after treatment are still in the 3-5% range as in the main results.

Because of the differences in results for questions pushed into big and small chat rooms shown in Table 6, Figure 3 presents a day-by-day pre-trend analysis for estimates using all chat rooms, big chat rooms only, and small chat rooms only. Figures 3(a) and 3(d) pool the estimates for big chat rooms and small chat rooms and so are consistent with Table 7, showing that the differences in the likelihood of getting an accepted answer between treated and control tags are insignificant before the feed is turned on in the treated tags. When we break the sample between big chat rooms (Figure 3(b) and 3(e)) and small chat rooms (Figure 3(c) and 3(f)), the results show that there is no evidence of a pre-trend but evidence of a statistically and economically significant impact of being pushed into the chat rooms after treatment. However, as expected, the parameter estimates for small chat rooms are noisy and show little evidence of a discernible pattern before or after treatment.

[7] We dropped questions generated on the day when feed is turned on because questions generated on that day can be either feeded questions (after the feed is on) or non-feeded questions (before feed is on).

Our identification strategy relies on comparing questions that are treated by the feed to those that are not treated by the feed. However, there are other means by which questions may appear in chat rooms. Users with sufficient reputation can push questions into chat rooms themselves, and other users may also push questions into chat rooms. In our main analysis we include questions that are untreated by the feed but pushed into chat rooms by other means in our control group. In Table 8 we include these as a separate treated category. The results show that our main results of being treated by the feed remain unchanged. Further, while we are unable to treat the estimates causally because they are determined by the choices of community members, the parameter estimates for the variables that indicate treatment by other means are interesting. In particular, questions that are pushed into chat rooms by askers are less likely to be answered, other things equal; these results may reflect selection on unobservables related to question quality.

We present two more robustness checks for two different samples in the Appendix. First, we analyze a "no PSM" sample, which compares the treated group against questions from *all* control tags without any matching (Table A2, Table A3, Figure A1(a) and (c)). Second, in the "50%" sample, we run the PSM logit using tags (control and treated) that are active for more than 50% of the observable weeks during our sample period, rather than 80% in our baseline analysis (Table A4, Table A5, Figure A1(b) and (d)). Both the main results and the pre-trend analysis are robust when we use these alternative samples.

6.3 Heterogeneous effect of the feed.

In section 4 we discussed how the effect of chat rooms might be greater among particular types of users and questions that might be less likely to receive an answer on the Q&A site. In particular, we discussed whether chat rooms could disproportionately influence outcomes among questions asked by low reputation users and among questions that are lower quality. In this section we examine whether the empirical evidence is consistent with these assertions.

First, in Table 9, we investigate heterogeneity in the effects of chat rooms based upon the reputation of users. We obtained reputation data from Stack Overflow and built a variable that indicates questions that are asked by users with 0 or negative reputation, *LowRep*.[8] We similarly create a variable that indicates whether users have reputation greater than 0, *HighRep*. We interact both of these variables with the treatment variable *TurnOnFeed.* In this sample, we dropped 12,146 questions that are from users who have no recorded reputation (this is because these users deleted their Stack Overflow account). Moreover, to keep the sample consistent with the later results in Table 10 and Table 11, we also dropped 360 questions for which we are unable to compute question quality (we do not observe question quality for these questions because they were later deleted by users). We checked the baseline specification for this sample and the results (see Appendix, Table A6) are consistent with those in table 4.

The questions of users with no reputation are less likely to receive an accepted answer. For example, in column 2, the coefficient on the variable *HighRep* is 0.1458, indicating that users with zero or lower reputation are 14.6 percentage points less likely to receive an accepted answer in two hours (31.9% =0.1458/0.4567). However, this asymmetry of the forum is mitigated by the chat room which significantly increases the probability of receiving an accepted answer for users with zero or negative reputation, but does not influence outcomes for users with an established positive reputation. For example, while in column 2 we estimate positive reputation users to be 14.6 percentage points more likely to receive an accepted answer within two hours than other users, this difference is estimated to be only 10.4 percentage points (=0.1458 + 0.0179 – 0.0594) when the chat room function is turned on. The results in Table 9 are robust to using a different threshold to identify low reputation (low reputation when reputation<=20 rather than reputation<=0; results in Appendix Table A7) and are also robust to controlling for a question's quality (see Table A8).

In the next Table (10), we analyze whether chat rooms can help users to figure out solutions to questions that are difficult to understand. Following prior literature (Ignatova et al 2008, Agichtein et al 2008 and Liu et al 2008), we measure question quality based upon the incidence of misspelled words. We didn't choose non-textual features such as score of questions and number of edits by other users because

[8] Users can lose reputation by receiving down votes from answers or questions.

those features could potentially be affected by treatment. We use Hunspell (Wikipedia 2017)[9] to identify the misspelled words in question text and compute percentage of misspelled words (number of misspelled words/ number of total words) as the measure for question quality.

As before, we create separate indicator variables for low quality and high quality questions. Low quality questions are defined as questions with quality lower than the bottom 25th percentile (*LowQuality*) and high quality questions are defined as questions with quality higher than the bottom 25th percentile (*HighQuality*). The table shows that questions of low quality are generally less likely to receive any answers. For example, in column 2 we find that low quality questions are 2.0 percentage points (4.3% = 0.0197/0.4567) less likely to be answered than other questions. Moreover, unlike higher quality questions, we find that low quality questions do not benefit from being forwarded to a chat room. Instead, the coefficient estimate of chat rooms for such low-quality questions cannot be discerned from zero.

In section 4 we asserted that chat rooms would have a stronger effect on time to accepted answer for questions that are unclear. Thus, these results are inconsistent with that assertion. While chat rooms may improve convergence for questions that are difficult to read, they will also be harder to answer on average and so even in the chat room users may still choose not to answer them. Because the propensity of potential question answerers to attempt to answer such questions may differ for questions asked by new and experienced users, in our next analysis we examine the joint effects of user reputation, question quality, and their interaction.

In Table 11 we combine question quality and user reputation into one single specification with a 3-way interaction. This allows us to present all puzzle pieces in one consolidated picture. The table includes three dummies to indicate when the feed was turned on (*TurnOnFeed*), whether a question is relatively low quality (*LowQuality*), and whether the question asker is inexperienced (*LowRep*). Moreover, we included all bilateral interactions and a 3-way interaction term to disentangle the interplay of quality and reputation when analyzing how chat rooms affect inexperienced users. Unlike prior tables, to simplify exposition we

---

[9] Details are available at https://hunspell.github.io/.

do not include the counterpart of these conditions (*HighQuality* and *HighRep*) however instead will provide a table that directly shows the marginal effect of being exposed to the feed under different conditions. The underlying specification of this regression is:

$$Y_{ijt} = \beta_0 + \beta_1 TurnOnFeed_{ijt} + \beta_2 LowRep_{ijt} + \beta_3 LowQuality_{ijt} + \beta_{32} LowQuality_{ijt} * LowRep_{ijt} + \beta_{31} LowQuality_{ijt} * TurnOnFeed_{ijt} + \beta_{12} TurnOnFeed_{ijt} * LowRep_{ijt} + \beta_{123} TurnOnFeed_{ijt} * LowRep_{ijt} * LowQuality_{ijt} + \alpha_j + WeekDay_t + Week_t + \varepsilon_{ijt} (2)$$

In this specification Y_{ijt} indicates the outcome variable of question i in tag j generated on time t. $TurnOnFeed_{ijt}$ indicates whether question i was pushed into the chat room by the feed function at time t. $LowRep_{ijt}$ indicates whether the reputation of the asker of question i is equal to or below zero. $LowQuality_{ijt}$ indicates whether the quality of question i is below the 25 percentile. We also include tag-chat level fixed effects α_j, weekly dummies ($Week_t$) and weekday dummies ($WeekDay_t$).

Table 11 shows the regression results of estimating equation (2). Table 12 provides marginal effects under different combinations of user reputation and question quality for the dependent variable Accept2Hour. Additional marginal effects calculations are included in Table A9. For simplicity, we focus our discussion on the results in column 2 (Accept2Hour) and the accompanying discussion in Table 12, though other results are similar.

The main coefficient on *TurnOnFeed* (which identifies the effects of the feed for high quality questions asked by higher reputation users) is similar to the estimates in the main specification (0.0426), but the interaction with *LowQuality* (-0.1097) suggests that the feed does not help questions of low quality. In fact, the results suggest that for low quality questions, the effects of the feed are negative for high reputation users (the marginal effect is -0.0671[10]) and statistically significant at the 10% level, although this result weakens in importance statistically and economically for longer time windows.

[10] This can be seen by adding up the coefficients to obtain: 0.0426 – 0.1097 = -0.0671

Our main focus lies on the coefficient of the 3-way interaction, which is estimated to be significantly positive and large (0.1788). The marginal effect of the feed for inexperienced users is greater for low quality questions that are asked by low reputation users. This finding is remarkable, especially against the backdrop of the negative coefficients for inexperienced users (in general) and for low-quality questions (in chat rooms). We explore the nuances of this result in several ways.

The changes in the likelihood of receiving an answer for low reputation users in the Q&A forum and when the questions are pushed into the chat room is particularly informative. In the Q&A site (i.e., excluding the effects of the feed), low reputation users (*LowRep*) asking low quality questions (*LowQuality*) are 17.6 percentage points less likely to receive an answer in two hours than high reputation/high quality users when questions are not pushed into the chat room.[11] This improves only slightly to a 13.6 percentage point relative deficit when low reputation users ask high quality questions. However, chat rooms help to mitigate these effects. Low reputation users asking low quality questions experience the greatest benefits from questions appearing in the chat room (an increase of 10.8 percentage points),[12] much greater than low reputation users asking high quality questions (an increase of 3.9 percentage points).[13] This helps to partially offset the disadvantages of being a low reputation user on Stack Overflow. In this context, the positive coefficient on the 3-way interaction highlights that chat rooms provide help to inexperienced users when they ask low quality questions, and are more tolerant towards low quality questions when they come from inexperienced users. In contrast, high reputation users receive few benefits when their questions appear in the chat room.

7. Discussion and Conclusions

We examine whether an informal new channel for communication can improve the efficiency of knowledge transfer in a network of practice. We analyze this question using the chat room functionality in the well-

[11] This can be seen by adding up the coefficients to obtain: - 0.1360 – 0.0070 - 0.0329= -0.176.
[12] This can be seen by adding up the coefficients to obtain: 0.0426 – 0.0041– 0.1097 + 0.1788 = 0.1076
[13] This can be seen by adding up the coefficients to obtain: 0.0426 – 0.0041 = 0.0385

known Q&A forum Stack Overflow. We identify the causal effect of chat rooms by exploiting a feed function which non-selectively pushed all questions from the Q&A forum into the relevant chat rooms that use this feature. We report two main findings: First, the adoption of a chat room reduces the time it takes for a question in the main Q&A to receive a satisfactory answer. The second channel thus increases the efficiency of the forum and it does so without any negative effects on answer quality. Second, we show that chat rooms disproportionately benefit user groups that face reduced access to the community's knowledge on the main channel. Specifically, inexperienced users that asked less well phrased questions benefitted most from the chat room.

This paper has several limitations to be acknowledged. First, our primary identification strategy relies on the feed function that pushed questions into the chat rooms. This means that our identification is based on a relatively small number of questions that are pushed into chat rooms by a feed. However, there are other means by which questions may appear in chat rooms. Users can push questions into chat rooms themselves, and other users may also push questions into chat rooms. While we document that questions are more likely to benefit when they are pushed into the chat room by users other than the asker, we are unable to draw causal inference on those questions, because they will reflect selection on unobservables such as question difficulty.

Our identification strategy requires that questions which are associated with tags that are treated by the feed are not changing over time in ways that are different from the control group generated by our propensity score estimator. To generate a valid control group we use propensity score matching. While balance checks support this approach and our main findings are unchanged if we use all tags, as in any matching estimator this represents a particular specification choice.

Our findings have important managerial implications. In particular, our results regarding low reputation and low quality questions showed that the chat room provides the structured main channel of the network of practice with a "friendlier space," that can facilitate newcomers' entry into the community. These findings suggest that opening a third space like chat rooms can be an effective managerial tool to support the main community by partially resolving inefficiencies in knowledge sharing as well as improving conditions for new users. A third space can thus alleviate inequalities in the access to the community knowledge that might arise from employing efficiency enhancing mechanisms, such as gamification, on the main channel. Hence, incorporating a third space into the network of practice might be an attractive option for managers who consider using efficiency enhancing mechanisms, but are worried about the inclusion of inexperienced users.

Our research contributes to a growing literature related to understanding platform decisions in online communities of practice. However, we depart from existing literature in important ways by focusing on a phenomenon that is growing in importance in these communities yet has so far remained underexplored in existing literature, the creation and emergence of a third space. Our research represents a first step in addressing these issues, yet many questions remain. Further research should examine the robustness of our results to other communities that have introduced third spaces, such as Wikipedia, and explore the features of such spaces that most effectively improve outcomes. Further, our work has focused upon identifying the effects of the treatment on the treated, namely those who have introduced chat rooms. Future would should investigate how the introduction of these spaces impact other parts of the community that have remained untreated. We hope our research spurs additional work in this important area.

References

Agichtein E, Castillo C, Donato D, Gionis A, Mishne, G (2008, February) Finding high-quality content in social media. *In Proceedings of the 2008 international conference on web search and data mining* (pp. 183-194).

Ahn J, Butler BS, Weng C, Webster S (2013) Learning to be a better Q'er in social Q&A sites: social norms and information artifacts. *Proceedings of the American Society for Information Science and Technology*, 50(1), 1-10.

Ahuja MK, Galletta DF, Carley KM (2003) Individual centrality and performance in virtual R&D groups: An empirical study. *Management science*. 49(1): 21-38.

Ai C, Norton EC (2003) Interaction terms in logit and probit models. *Economics letters*. 80(1):123-129.

Alavi M., Leidner DE (2001) Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS quarterly*. 107-136.

Arora P, Ganguly D, Jones GJ (2015, August) The good, the bad and their kins: Identifying questions with negative scores in stackoverflow. *In 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (ASONAM) (pp. 1232-1239). IEEE.

Athey S, Ellison G (2014) Dynamics of open source movements. *Journal of Economics & Management Strategy*. 23(2):294-316.

Atwood J (2010) Do Trilogy Sites Need a "Third Place"? Accessed December 29, 2020. https://stackoverflow.blog/2010/04/29/do-trilogy-sites-need-a-third-place/

Baltadzhieva A, Chrupała, G (2015, September) Predicting the quality of questions on stackoverflow. *In Proceedings of the international conference recent advances in natural language processing* (pp. 32-40).

Nonnecke B, Andrews D, Preece J (2006) Non-public and public online community participation: Needs, attitudes and behavior. *Electronic Commerce Research*. *6*(1): 7-20.

Bosu A, Corley CS, Heaton D, Chatterji D, Carver JC, Kraft NA (2013) Building reputation in stackoverflow: An empirical investigation. *In Proceedings of the 10th International Conference on Mining Software Repositories*, MSR '13, pp 89–92

Boudreau K, Lakhani K (2009) How to manage outside innovation. *MIT Sloan management review*. 50(4): 69.

Brown J S, Duguid P (2001) Knowledge and organization: A social-practice perspective. *Organization science*. 12(2):198-213.

Browne R (2018) Reach 300+ million monthly unique visitors on Quora. Accessed in December 29, 2020. https://www.quora.com/q/quoraforbusiness/Reach-300-million-monthly-unique-visitors-on-Quora

Branon RF, Essex C (2001) Synchronous and asynchronous communication tools in distance education. *TechTrends*. *45*(1):36.

Brynjolfsson E, McAfee A (2014) *The second machine age: Work, progress, and prosperity in a time of brilliant technologies* (WW Norton & Company)

Burtch G, He Q, Hong Y, Lee D (2020) Peer Awards Retain New Users and Encourage Exploitation in Users' Production of Creative UGC. Available at https://ssrn.com/abstract=3701259.

Burt RS (1992) *Structural Holes* (Cambridge, MA: Harvard University Press)

Butler BS (2001) Membership size, communication activity, and sustainability: A resource-based model of online social structures. *Information systems research*. 12(4):346-362.

Cavusoglu H, Li Z, Huang KW (2015, February) Can gamification motivate voluntary contributions? The case of Stack Overflow Q&A community. *In Proceedings of the 18th ACM conference companion on computer supported cooperative work & social computing* (pp. 171-174).

Chiu C M, Hsu MH, Wang ET (2006) Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. *Decision support systems*. 42(3):1872-1888.

Coetzee D, Fox A, Hearst MA, Hartmann B (2014, March) Chatrooms in MOOCs: all talk and no action. *In Proceedings of the first ACM conference on Learning@ scale conference* (pp. 127-136).

Cosley D, Frankowski D, Terveen L, Riedl J (2007, January) SuggestBot: using intelligent task routing to help people find work in wikipedia. *In Proceedings of the 12th international conference on Intelligent user interfaces* (pp. 32-41).

Coleman JS (1988) Social capital in the creation of human capital. *American journal of sociology*. 94:S95-S120.

Curran JA, Murphy AL, Abidi SR, Sinclair D, McGrath PJ (2009) Bridging the gap: knowledge seeking and sharing in a virtual community of emergency practice. *Evaluation & the health professions*. 32(3):314-327.

Daft RL, Lengel RH (1986) Organizational information requirements, media richness and structural design. *Management science*. *32*(5):554-571.

Daft RL, Lengel RH, Trevino LK (1987) Message equivocality, media selection, and manager performance: Implications for information systems. *MIS quarterly*. 355-366.

Dede C, Kremer A (1999) Increasing students' participation via multiple interactive media. *Invention: Creative Thinking about Learning and Teaching*. *1*(1):7.

Dennis AR, Fuller RM, Valacich JS (2008) Media, tasks, and communication processes: A theory of media synchronicity. *MIS quarterly*. *32*(3):575-600.

Ford D, Parnin C (2015, May) Exploring causes of frustration for software developers. In *2015 IEEE/ACM 8th International Workshop on Cooperative and Human Aspects of Software Engineering* (pp. 115-116). IEEE.

Ford D, Lustig K, Banks J, Parnin C (2018, April) "We Don't Do That Here" How Collaborative Editing with Mentors Improves Engagement in Social Q&A Communities. In *Proceedings of the 2018 CHI conference on human factors in computing systems* (pp. 1-12).

Ford D, Smith J, Guo PJ, Parnin C (2016, November) Paradise unplugged: Identifying barriers for female participation on stack overflow. In *Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering* (pp. 846-857).

Dinh M, Tan T, Bein, K, Hayman J, Wong YK, Dinh D (2011) Emergency department knowledge management in the age of Web 2.0: evaluation of a new concept. *Emergency Medicine Australasia*. 23(1):46-53.

Elliott MS, Scacchi W (2003, November) Free software developers as an occupational community: resolving conflicts and fostering collaboration. *In Proceedings of the 2003 international ACM SIGGROUP conference on Supporting group work* (pp. 21-30).

Ford D, Lustig K, Banks J, Parnin C (2018, April) " We Don't Do That Here" How Collaborative Editing with Mentors Improves Engagement in Social Q&A Communities. *In Proceedings of the 2018 CHI conference on human factors in computing systems* (pp. 1-12).

Ford D, Smith J, Guo PJ, Parnin C (2016, November) Paradise unplugged: Identifying barriers for female participation on stack overflow. *In Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering* (pp. 846-857).

Forte A, Kittur N, Larco V, Zhu H, Bruckman A, Kraut RE (2012, February) Coordination and beyond: social functions of groups in open content production. *In Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (pp. 417-426).

Fullerton D (2019) State of the Stack 2019: A Year in Review. Accessed December 29, 2020, https://stackoverflow.blog/2019/01/18/state-of-the-stack-2019-a-year-in-review/

George JF, Carlson JR, Valacich JS (2013) Media selection as a strategic component of communication. *MIS Quarterly*. 1233-1251.

Goodhue DL (1995) Understanding User Evaluations of Information Systems. *Management Science* 41(12):1827-1844.

Goodhue DL, Thompson RL (1995) Task-technology fit and individual performance. *MIS quarterly*. 213-236.

Goes PB, Guo C, Lin M (2016) Do incentive hierarchies induce user effort? Evidence from an online knowledge exchange. *Information Systems Research*. *27*(3):497-516.

Grewal R, Lilien GL, Mallapragada G (2006) Location, location, location: How network embeddedness affects project success in open source systems. *Management science*. 52(7):1043-1056.

Halfaker A, Kittur A, Riedl J (2011, October) Don't bite the newbies: how reverts affect the quantity and quality of Wikipedia work. *In Proceedings of the 7th international symposium on wikis and open collaboration* (pp. 163-172).

Harper FM, Moy D, Konstan JA (2009, April) Facts or friends? Distinguishing informational and conversational questions in social Q&A sites. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 759-768).

Harper FM, Sen S, Frankowski D (2007, October) Supporting social recommendations with activity-balanced clustering. In *Proceedings of the 2007 ACM conference on Recommender systems* (pp. 165-168).

Hines RA, Pearl CE (2004) Increasing interaction in web-based instruction: Using synchronous chats and asynchronous discussions. *Rural special education Quarterly*. *23*(2):33-36.

Howe, D. (2008). The emotionally intelligent social worker. Macmillan International Higher Education.

Ignatova K, Bernhard D, Gurevych I (2008, September) Generating high quality questions from low quality questions. In *Proceedings of Workshop on the Question Generation Shared Task and Evaluation Challenge*.

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

Johnson GM (2006) Synchronous and asynchronous text-based CMC in educational contexts: A review of recent research. *TechTrends*. *50*(4):46-53.

King AA, Lakhani KR (2011) *The contingent effect of absorptive capacity: an open innovation analysis*. Boston: Harvard Business School.

Krackhardt D, Nohria N, Eccles B (2003) *Networks in the knowledge economy* (Oxford University Press), 82-88.

Kraut RE, Resnick P (2012) *Building successful online communities: Evidence-based social design* (Mit Press), 21–76, 80–97, 145–147, 179–230 pages.

Liu D, Li X, Santhanam R (2013) Digital games and beyond: What happens when players compete? *Mis Quarterly*. 111-124.

Liu D, Santhanam R, Webster J (2017) Toward Meaningful Engagement: A Framework for Design and Research of Gamified Information Systems. *MIS quarterly*. *41*(4).

Liu Y, Bian J, Agichtein E (2008, July) Predicting information seeker satisfaction in community question answering. *In Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 483-490).

Marwell G, Oliver PE, Prahl R (1988) Social networks and collective action: A theory of the critical mass. III. *American journal of Sociology*. 94(3):502-534.

Massung E, Coyle D, Cater KF, Jay M, Preist C (2013, April) Using crowdsourcing to support pro-environmental community activism. *In Proceedings of the SIGCHI Conference on human factors in Computing systems* (pp. 371-380).

Mokyr J (2002) *The gifts of Athena: Historical origins of the knowledge economy* (Princeton University Press).

Moon JY, Sproull LS (2008) The role of feedback in managing the Internet-based volunteer work force. *Information Systems Research.* 19(4):494-515.

Morschheuser B, Hamari J, Koivisto J (2016, January) Gamification in crowdsourcing: a review. *In 2016 49th Hawaii International Conference on System Sciences (HICSS)* (pp. 4375-4384). IEEE.

Nian T, Lei X, Sanjeev D (2019) Can She Code? Gender Bias in an Open Source Software Community. Working Paper, University of California, Irvine.

Oztok M, Zingaro D, Brett C, Hewitt J (2013) Exploring asynchronous and synchronous tool use in online courses. *Computers & Education*. *60*(1):87-94.

Oldenburg R (1999) *The great good place: Cafes, coffee shops, bookstores, bars, hair salons, and other hangouts at the heart of a community* (Da Capo Press)

Ravi S, Pang B, Rastogi V, Kumar R (2014) Great question! question quality in community q&a. *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*, 426-435.

Raban DR (2008) The incentive structure in an online information market. *Journal of the American Society for Information Science and Technology*. 59(14):2284-2295.

Ren Y, Kraut R, Kiesler S (2007) Applying common identity and bond theory to design of online communities. *Organization studies*. 28(3):377-408.

Rice ER (1994) *Advances in social network analysis: Research in the social and behavioral sciences* (SAGE Publications), 167-190.

Rovai AP (2001) Building classroom community at a distance: A case study. *Educational technology research and development*. *49*(4):33.

Santhanam R, Liu D, Shen CM (2016) Research Note—Gamification of technology-mediated training: Not all competitions are the same. *Information systems research*. *27*(2):453-465.

Sambamurthy V, Subramani M (2005) Special issue on information technologies and knowledge management. *MIS quarterly*. 1-7.

Shihab E, Jiang ZM, Hassan AE (2009, May) On the use of Internet Relay Chat (IRC) meetings by developers of the GNOME GTK+ project. *In 2009 6th IEEE International Working Conference on Mining Software Repositories* (pp. 107-110). IEEE.

Singh PV, Tan Y, Mookerjee V (2011) Network effects: The influence of structural capital on open source project success. *Mis Quarterly*. 813-829.

Singh PV, Tan Y, Youn N (2011) A hidden Markov model of developer learning dynamics in open source software projects. *Information Systems Research*. 22(4):790-807.

Squire M (2015) Should we move to Stack Overflow: measuring the utility of social media for developer support. *37th International Conference on Software Engineering*. pp 219–228.

Stack Overflow (2019) Developer Survey Results 2019. Accessed December 29, 2020. https://insights.stackoverflow.com/survey/2019

Stack Overflow (2020) Talk in chat. Accessed December 29, 2020. https://stackoverflow.com/help/privileges/chat

Steinmacher I, Conte T, Gerosa MA, Redmiles D (2015, February) Social barriers faced by newcomers placing their first contribution in open source software projects. *In Proceedings of the 18th ACM conference on Computer supported cooperative work & social computing*, pp. 1379-1392.

Steinkuehler CA, Williams D (2006) Where everybody knows your (screen) name: Online games as “third places”. *Journal of computer-mediated communication*. 11(4):885-909.

Szulanski G (1996) Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic management journal*. 17(S2):27-43.

Steinmacher I, Conte T, Gerosa MA, Redmiles D (2015, February) Social barriers faced by newcomers placing their first contribution in open source software projects. In *Proceedings of the 18th ACM conference on Computer supported cooperative work & social computing* (pp. 1379-1392).

Vasilescu BN (2014) Social aspects of collaboration in online software communities. PhD thesis.

Vasilescu BN, Capiluppi A, Serebrenik A (2014) Gender, representation and online participation: A quantitative study. *Interacting with Computers*. 26(5):488-511.

Walker G, Kogut B, Shan W (1997) Social capital, structural holes and the formation of an industry network. *Organization science*. 8(2):109-125.

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

Wiertz C, de Ruyter K (2007) Beyond the call of duty: Why customers contribute to firm-hosted commercial online communities. *Organization studies*. 28(3):347-376.

Wicks P, Keininger DL, Massagli MP, de la Loge C, Brownstein C, Isojärvi J, Heywood J (2012) Perceived benefits of sharing health data between people with epilepsy on an online platform. *Epilepsy & Behavior*. 23(1):16-23.

Wikipedia (2018) Zhihu. Accessed December 29, 2020. https://en.wikipedia.org/wiki/Zhihu#cite_note-5

Yang X (2020) Effects of Peer Penalty on Online Community Participation. *Proceedings of the Forty-First ICIS*.

Zagalsky A, Teshima CG, German DM, Storey MA, Poo-Caamaño G (2016, May) How the R community creates and curates knowledge: a comparative study of stack overflow and mailing lists. In *Proceedings of the 13th International Conference on Mining Software Repositories* (pp. 441-451).

Zhang XM, Zhu F (2011) Group size and incentives to contribute: A natural experiment at Chinese Wikipedia. *American Economic Review*. 101(4):1601-15.

Tables and Figures

Figure 1a

Figure 1b

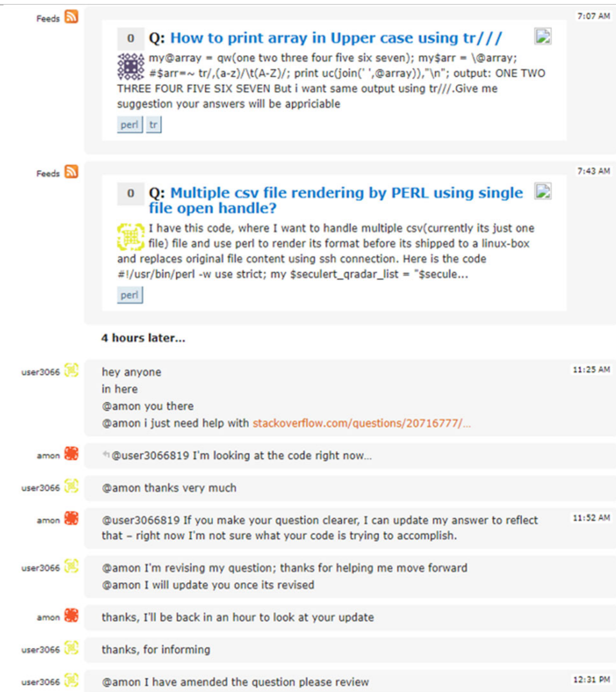

Figure 2

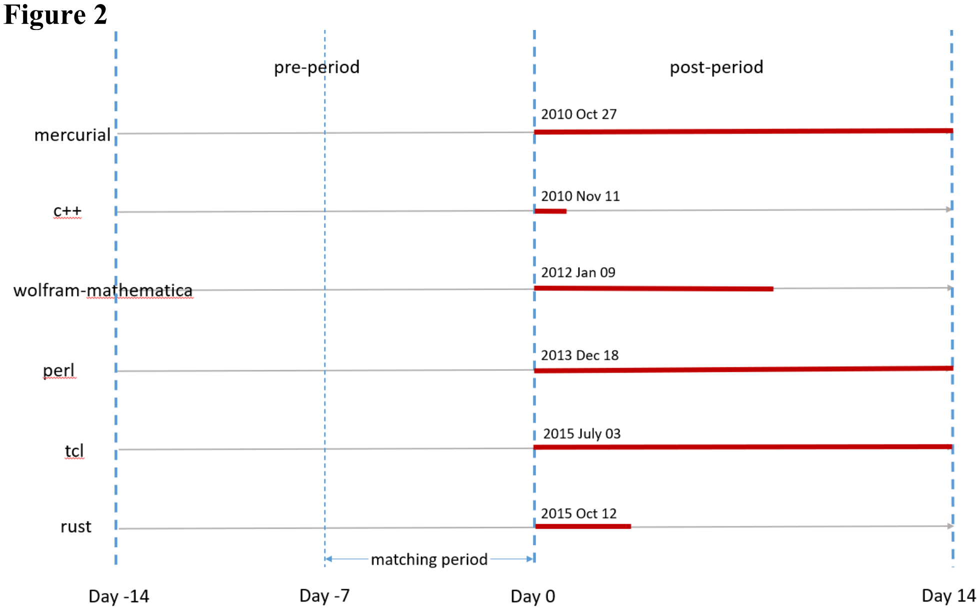

Table 1: Balance check before PSM

	treat	control	Difference	se	(p-value)
Estimated propensity score	0.03	0.01	0.02	0.01	0.00
log(QuestionCount)	4.90	4.23	0.67	0.44	0.13
log(AskerCount)	4.69	4.10	0.59	0.43	0.17
log(AnswerCount)	5.90	4.99	0.92	0.52	0.08
log(AnswererCount)	4.55	4.08	0.47	0.43	0.27
log(MessageCount)	3.74	3.43	0.32	0.58	0.59
log(UserCount)	1.63	1.57	0.05	0.27	0.84
log(QuestionCount(t-1))	4.92	4.21	0.72	0.45	0.11
log(AskerCount(t-1))	4.72	4.08	0.64	0.44	0.14
log(AnswerCount(t-1))	6.00	4.97	1.03	0.53	0.05
log(AnswererCount(t-1))	4.60	4.06	0.54	0.44	0.22
log(MessageCount(t-1))	3.25	3.61	-0.36	0.55	0.51
log(UserCount(t-1))	1.47	1.65	-0.19	0.26	0.47
GrowthRate	-0.00	0.09	-0.10	0.12	0.43
Answerer per question	0.73	0.92	-0.19	0.10	0.05
N. of tag episode	18	1008			

Note: The column “treat” represents the average value of variables in the first column across treated tag-episodes in the week before turning on the feed (matching period) and the column “control” represents the average value of variables in the first column across control tag-episodes in the week before turning on the feed (matching period). The column “Difference” represents the sample difference between treated tag-episodes and control tag-episodes.

Table 2: Balance check after PSM

	treat	control	Difference	se	(p-value)
Estimated propensity score	0.03	0.01	0.02	0.01	0.06
log(QuestionCount)	4.90	4.46	0.43	0.54	0.42
log(AskerCount)	4.69	4.33	0.36	0.52	0.49
log(AnswerCount)	5.90	5.59	0.31	0.59	0.60
log(AnswererCount)	4.55	4.44	0.11	0.49	0.83
log(MessageCount)	3.74	3.85	-0.11	0.59	0.85
log(UserCount)	1.63	1.73	-0.11	0.27	0.69
log(QuestionCount(t-1))	4.92	4.49	0.43	0.53	0.42
log(AskerCount(t-1))	4.72	4.36	0.36	0.51	0.48
log(AnswerCount(t-1))	6.00	5.54	0.46	0.60	0.44
log(AnswererCount(t-1))	4.60	4.42	0.18	0.50	0.72
log(MessageCount(t-1))	3.25	4.02	-0.77	0.57	0.19
log(UserCount(t-1))	1.47	1.88	-0.42	0.27	0.12
GrowthRate	-0.00	0.01	-0.01	0.08	0.85
Answerer per question	0.73	1.10	-0.37	0.13	0.01
N. of tag episode	18	83			

Note: The column “treat” represents the average value of variables in the first column across treated tag-episodes in the week before turning on the feed (matching period) and the column “control” represents the average value of variables in the first column across control tag-episodes in the week before turning on the feed (matching period). The column “Difference” represents the sample difference between treated tag-episodes and control tag-episodes.

Table 3 Summary statistics

	N	Mean	St.Dev	min	max
AcceptHour	166435	0.41	0.49	0	1
Accept2Hour	166435	0.46	0.50	0	1
Accept4Hour	166435	0.49	0.50	0	1
Accept8Hour	166435	0.51	0.50	0	1
AcceptScore	100869	4.63	19.48	-20	1983
AveScore	153935	2.14	5.44	-7	362.50
log(AcceptScore)	81687	0.96	0.98	0	7.59
log(AveScore)	113827	0.52	0.91	-2.20	5.89
log(ViewCount)	166435	6.35	1.66	1.95	14.62

Table 4 The effect of the chat rooms on the efficiency of knowledge exchange

	AcceptHour (1)	Accept2Hour (2)	Accept4Hour (3)	Accept8Hour (4)
TurnOnFeed	0.0174	0.0320**	0.0259**	0.0335**
	(0.0107)	(0.0144)	(0.0120)	(0.0132)
R^2-within	0.0007	0.0007	0.0007	0.0007
R^2-total	0.0295	0.0286	0.0283	0.0286
N	166,435	166,435	166,435	166,435
N. of tag-episode	101	101	101	101
Mean	0.41	0.46	0.49	0.51

* p<0.1; ** p<0.05; *** p<0.01
Heteroskedasticity robust standard error clustered at tag-episode level in parentheses
Including tag-episode level fixed effect, weekly dummies and weekday dummies

Table 5 The effect of the chat rooms on answer quality

	AcceptScore (1)	AveScore (2)	ln(AcceptScore) (3)	lnAveScore (4)	ln(ViewCount) (5)
TurnOnFeed	0.0099	0.3067**	0.0401	0.0572*	0.1634***
	(0.4458)	(0.1217)	(0.0316)	(0.0292)	(0.0617)
R^2-within	0.0007	0.0007	0.0007	0.0007	0.0006
R^2-total	0.0130	0.0236	0.0540	0.0483	0.1101
N	100,869	153,935	81,687	113,827	166,435
N. of tag-episode	101	101	101	101	101
Mean	4.63	2.14	0.96	0.52	6.35

* p<0.1; ** p<0.05; *** p<0.01
Heteroskedasticity robust standard error clustered at tag-episode level in parentheses
Including tag-episode level fixed effect, weekly dummies and weekday dummies

Table 6 Big chat rooms benefit more from the feed function

	AcceptHour (1)	Accept2Hour (2)	Accept4Hour (3)	Accept8Hour (4)
TurnOnFeed*smallChat	-0.0043	0.0043	0.0044	0.0140
	(0.0161)	(0.0152)	(0.0129)	(0.0147)
TurnOnFeed*bigChat	0.0542***	0.0791**	0.0622***	0.0664***
	(0.0166)	(0.0307)	(0.0201)	(0.0249)
R^2-within	0.0007	0.0007	0.0007	0.0007
R^2-total	0.0295	0.0286	0.0284	0.0286
N	166,435	166,435	166,435	166,435
N. of tag-episode	101	101	101	101
Diff: bigChat -smallChat	0.0585**	0.0748**	0.0577**	0.0525*
	(0.0231)	(0.0342)	(0.0238)	(0.0289)
Mean	0.41	0.46	0.49	0.51

* p<0.1; ** p<0.05; *** p<0.01
Heteroskedasticity robust standard error clustered at tag-episode level in parentheses
Including tag-episode level fixed effect, weekly dummies and weekday dummies

Table 7 Pre-trend analysis

	AcceptHour (1)	Accept2Hour (2)	Accept4Hour (3)	Accept8Hour (4)
Treat*(Day -6 to -4)	0.0012	0.0026	0.0064	0.0105
	(0.0101)	(0.0110)	(0.0120)	(0.0095)
Treat*(Day -3 to -1)	0.0109	0.0104	0.0090	0.0095
	(0.0079)	(0.0084)	(0.0088)	(0.0090)
TurnOnFeed * (Day 1 to 3)	0.0157	0.0335	0.0269	0.0489
	(0.0477)	(0.0416)	(0.0493)	(0.0507)
TurnOnFeed * (Day 4 to 6)	0.0290	0.0412	0.0378*	0.0479***
	(0.0218)	(0.0251)	(0.0206)	(0.0181)
TurnOnFeed * (After day 6)	0.0170	0.0192	0.0109	0.0182
	(0.0253)	(0.0252)	(0.0258)	(0.0296)
R^2-within	0.0007	0.0007	0.0007	0.0007
R^2-total	0.0295	0.0286	0.0284	0.0286
N	165,784	165,784	165,784	165,784
N. of tag-episode	101	101	101	101
Mean	0.41	0.46	0.49	0.51

* p<0.1; ** p<0.05; *** p<0.01
Heteroskedasticity robust standard error clustered at tag-episode level in parentheses
Including tag-episode level fixed effect, weekly dummies and weekday dummies

Figure 3: Estimated impact of chat rooms on the likelihood of getting an accepted answer within 2 hours and 4 hours

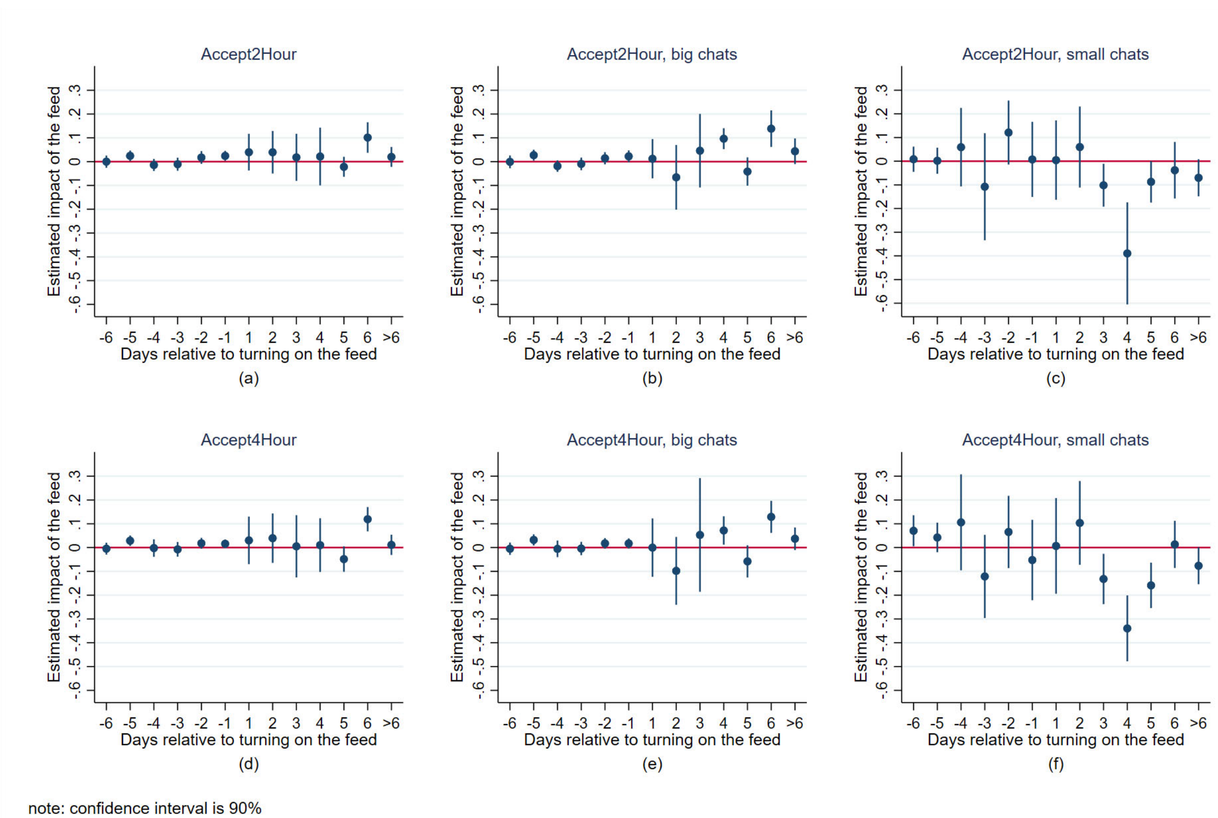

Table 8 Robustness check: Control for questions treated by askers and other users

	AcceptHour (1)	Accept2Hour (2)	Accept4Hour (3)	Accept8Hour (4)
TurnOnFeed	0.0233**	0.0384**	0.0268**	0.0318**
	(0.0108)	(0.0147)	(0.0124)	(0.0142)
treatasker	-0.1164***	-0.1060***	-0.0873***	-0.0901***
	(0.0198)	(0.0245)	(0.0264)	(0.0284)
treatother	0.0590***	0.0625***	0.0607***	0.0630***
	(0.0119)	(0.0122)	(0.0136)	(0.0134)
R^2	0.0010	0.0010	0.0009	0.0010
N	166,435	166,435	166,435	166,435
N. of tag-episode	101	101	101	101
Mean	0.41	0.46	0.49	0.51

* p<0.1; ** p<0.05; *** p<0.01
Heteroskedasticity robust standard error clustered at tag-episode level in parentheses
Including tag-episode level fixed effect, weekly dummies and weekday dummies

Table 9 Low reputation users benefit more from the chat rooms

	AcceptHour (1)	Accept2Hour (2)	Accept4Hour (3)	Accept8Hour (4)
TurnOnFeed*LowRep	0.0511***	0.0594**	0.0647**	0.0661**
	(0.0183)	(0.0252)	(0.0323)	(0.0306)
TurnOnFeed*HighRep	-0.0013	0.0179	0.0085	0.0187
	(0.0128)	(0.0157)	(0.0132)	(0.0159)
HighRep	0.1331***	0.1458***	0.1547***	0.1620***
	(0.0091)	(0.0094)	(0.0099)	(0.0102)
R^2	0.0125	0.0146	0.0162	0.0178
N	153,929	153,929	153,929	153,929
N. of tag-episode	93	93	93	93
Diff (marginal): HighRep - LowRep	-0.0525***	-0.0415	-0.0563	-0.0474
	(0.0188)	(0.0276)	(0.0383)	(0.0385)
Diff (total): HighRep - LowRep	0.1037***	0.1043***	0.1055***	0.1061***
	(0.0258)	(0.0258)	(0.0257)	(0.0258)
Mean	0.4147	0.4567	0.4868	0.5081

* p<0.1; ** p<0.05; *** p<0.01
Heteroskedasticity robust standard error clustered at tag-episode level in parentheses
Including tag-episode level fixed effect, weekly dummies and weekday dummies

Table 10 Questions with lower quality benefit less from the chat rooms

	AcceptHour (1)	Accept2Hour (2)	Accept4Hour (3)	Accept8Hour (4)
TurnOnFeed*HighQuality	0.0254**	0.0476***	0.0418***	0.0395***
	(0.0116)	(0.0146)	(0.0110)	(0.0131)
TurnOnFeed*LowQuality	-0.0258	-0.0220	-0.0305	0.0134
	(0.0258)	(0.0269)	(0.0284)	(0.0279)
LowQuality	-0.0173***	-0.0197***	-0.0199***	-0.0213***
	(0.0054)	(0.0051)	(0.0052)	(0.0051)
R^2	0.0010	0.0011	0.0011	0.0012
N	153,929	153,929	153,929	153,929
N. of tag-episode	93	93	93	93
Diff: LowQuality - HighQuality	-0.0512*	-0.0695***	-0.0723**	-0.0261
	(0.0268)	(0.0254)	(0.0289)	(0.0262)
Mean	0.4147	0.4567	0.4868	0.5081

* p<0.1; ** p<0.05; *** p<0.01
Heteroskedasticity robust standard error clustered at tag-episode level in parentheses
Including tag-episode level fixed effect, weekly dummies and weekday dummies

Table 11 Three-way interaction

	AcceptHour (1)	Accept2Hour (2)	Accept4Hour (3)	Accept8Hour (4)
TurnOnFeed	0.0193	0.0426**	0.0309**	0.0302*
	(0.0137)	(0.0165)	(0.0122)	(0.0162)
TurnOnFeed*LowRep	0.0061	-0.0041	0.0267	0.0155
	(0.0225)	(0.0283)	(0.0393)	(0.0441)
TurnOnFeed*LowRep*LowQuality	0.1769***	0.1788***	0.1234*	0.1204*
	(0.0450)	(0.0633)	(0.0626)	(0.0622)
LowQuality*LowRep	-0.0296***	-0.0329***	-0.0336***	-0.0336***
	(0.0082)	(0.0094)	(0.0090)	(0.0093)
TurnOnFeed * LowQuality	-0.0917***	-0.1097***	-0.1004***	-0.0528
	(0.0318)	(0.0365)	(0.0376)	(0.0393)
LowQuality	-0.0058	-0.0070	-0.0067	-0.0078
	(0.0050)	(0.0050)	(0.0050)	(0.0049)
LowRep	-0.1243***	-0.1360***	-0.1447***	-0.1520***
	(0.0095)	(0.0098)	(0.0105)	(0.0110)
R^2	0.0128	0.0150	0.0166	0.0181
N	153,929	153,929	153,929	153,929
N. of tag-episode	93	93	93	93
Mean	0.4147	0.4567	0.4868	0.5081

* p<0.1; ** p<0.05; *** p<0.01
Heteroskedasticity robust standard error clustered at tag-episode level in parentheses
Including tag-episode level fixed effect, weekly dummies and weekday dummies

Table 12 Effect of the feed on the likelihood for receiving an answer within 2 hours for different levels of reputation and quality

Panel A: Marginal Effects of the Feed		
	Low reputation	High reputation
Low quality	0.1076	-0.0671
High quality	0.0385	0.0426
Panel B: Relative likelihood of receiving an answer within 2 hours for questions that are not treated by the feed		
	Low reputation	High reputation
Low quality	-0.1759	-0.0070
High quality	-0.1360	0
Panel C: Total relative magnitude of likelihood of receiving an answer within 2 hours if pushed into chat room		
	Low reputation	High reputation
Low quality	-0.0683	-0.0741
High quality	-0.0975	0.0426
Panel D: Total probability of receiving an answer within 2 hours if pushed into chat room		
	Low reputation	High reputation
Low quality	0.389	0.383
High quality	0.359	0.500

Notes: The table analyzes the effect of the feed on the likelihood of receiving an answer (within 2 hours) for questions of different combinations of reputation and quality. Panel A shows the marginal effects of the feed and how it differs based on reputation and quality. Panel B shows the likelihood a question will be answered based on quality and reputation if it is not treated by the feed. Panel C summarizes panels A and B by combining their effects to show the total relative magnitudes of the likelihood that a question will receive an answer within 2 hours when pushed into a chat room. Panel D sums the effects in Panel C with the average likelihood that a question will receive and answer in the Q&A forum. More details providing marginal effects for other dependent variables are provided in Table A9.

Appendix

Table A0 PSM: matching covariates

log(QuestionCount)	Log of question count under associated tag in each week
log(AskerCount)	Log of asker count under associated tag in each week
log(AnswerCount)	Log of answer count under associated tag in each week
log(AnswererCount)	Log of answerer count under associated tag in each week
log(MessageCount)	Log of message count under associated chat room in each week
log(UserCount)	Log of unique users count under associated chat room in each week
log(QuestionCount(t-1))	Log of question count under associated tag in previous week
log(AskerCount(t-1))	Log of asker count under associated tag in previous week
log(AnswerCount(t-1))	Log of answer count under associated tag in previous week
log(AnswererCount(t-1))	Log of answerer count under associated tag in previous week
log(MessageCount(t-1))	Log of message count under associated chat room in previous week
log(UserCount(t-1))	Log of unique users count under associated chat room in previous week
GrowthRate	Growth rate in question count by week
Answerer per question	Number of answerers per question

Table A1 PSM: logit regression

	TurnOnFeed
log(QuestionCount)	-4.9149***
	(1.1889)
log(AskerCount)	3.8173***
	(1.1232)
log(AnswerCount)	2.4174***
	(0.2622)
log(AnswererCount)	-1.9448***
	(0.5124)
log(MessageCount)	0.6104***
	(0.0937)
log(UserCount)	-1.4057***
	(0.2927)
log(QuestionCount(t-1))	-4.6931***
	(1.0580)
log(AskerCount(t-1))	4.3533***
	(1.1108)
log(AnswerCount(t-1))	2.4208***
	(0.2551)
log(AnswererCount(t-1))	-2.4967***
	(0.4451)
log(MessageCount(t-1))	0.3385***
	(0.0939)
log(UserCount(t-1))	-0.6666**
	(0.2904)
GrowthRate	0.0020
	(0.0039)
Answerer per question	-0.2738
	(0.2731)
Pseudo R^2	0.3290
N. of tag-weeks	46,528

Notes: N. of treated tags 18; N. of treated tags 706; Robust standard error in parentheses. * p<0.1; ** p<0.05; *** p<0.01

Table A2 Robustness: Effect of chat rooms on the efficiency of knowledge exchange (before PSM)

	AcceptHour (1)	Accept2Hour (2)	Accept4Hour (3)	Accept8Hour (4)
TurnOnFeed	0.0218**	0.0332***	0.0273***	0.0361***
	(0.0108)	(0.0108)	(0.0081)	(0.0094)
R^2	0.0002	0.0002	0.0002	0.0002
N	1,097,468	1,097,468	1,097,468	1,097,468
N. of tag-episode	1,026	1,026	1,026	1,026
Mean	0.39	0.43	0.46	0.48

* p<0.1; ** p<0.05; *** p<0.01
Heteroskedasticity robust standard error clustered at tag-episode level in parentheses. Including tag-episode level fixed effect, weekly dummies and weekday dummies
We keep tag episodes that have observations in matching covariates for at least 80% of the weeks and include all 1012 tags without selecting on control tags by PSM.
18 treated tags and 1008 control tags are included in the final sample because 4 control tags didn't have questions during the 4 week time window.

Table A3 Robustness: Effect of chat rooms on answer quality (before PSM)

	AcceptScore (1)	aveScore (2)	Ln(AcceptScore) (3)	Ln(AveScore) (4)	Ln(ViewCount) (5)
TurnOnFeed	-0.2049	0.2147**	0.0358	0.0632**	0.1261**
	(0.3334)	(0.1039)	(0.0281)	(0.0268)	(0.0570)
R^2	0.0002	0.0002	0.0006	0.0004	0.0006
N	639,621	995,547	504,446	703,975	1,097,468
N. of tag-episode	1,024	1,026	1,022	1,025	1,026
Mean	3.77	1.84	0.86	0.46	6.11

* p<0.1; ** p<0.05; *** p<0.01
Heteroskedasticity robust standard error clustered at tag-episode level in parentheses. Including tag-episode level fixed effect, weekly dummies and weekday dummies
We keep tags that have observations in matching covariates for at least 80% of the weeks and include all 1012 tags without selecting on control tags by PSM. 18 treated tags and 1008 control tags are included in the final sample because 4 control tags didn't have questions during the 4 week time window.

Table A4 Robustness: Effect of chat rooms on efficiency of knowledge exchange (50% active chat rooms)

	AcceptHour (1)	Accept2Hour (2)	Accept4Hour (3)	Accept8Hour (4)
TurnOnFeed	0.0177*	0.0299**	0.0233***	0.0315***
	(0.0101)	(0.0114)	(0.0088)	(0.0109)
R^2	0.0009	0.0009	0.0009	0.0010
N	140,629	140,629	140,629	140,629
N. of tag-episode	103	103	103	103
Mean	0.43	0.47	0.50	0.52

* p<0.1; ** p<0.05; *** p<0.01
Heteroskedasticity robust standard error clustered at tag-episode level in parentheses. Including tag-episode level fixed effect, weekly dummies and weekday dummies
We keep tags that have observations in matching covariates for at least 50% of the weeks and include tags selected by PSM. 19 treated tags and 84 control tags are included in the sample.

Table A5 Robustness: Effect of chat rooms on answer quality (50% active chat rooms)

	AcceptScore (1)	aveScore (2)	Ln(AcceptScore) (3)	Ln(AveScore) (4)	Ln(ViewCount) (5)
TurnOnFeed	-0.0142	0.1692	0.0415	0.0433	0.1538**
	(0.5219)	(0.1750)	(0.0300)	(0.0299)	(0.0590)
R^2	0.0008	0.0008	0.0012	0.0012	0.0009
N	88,327	131,835	72,858	100,584	140,629
N. of tag-episode	102	102	102	102	103
Mean	5.01	2.31	1.02	0.56	6.51

* p<0.1; ** p<0.05; *** p<0.01
Heteroskedasticity robust standard error clustered at tag-episode level in parentheses. Including tag-episode level fixed effect, weekly dummies and weekday dummies
We keep tags that have observations in matching covariates for at least 50% of the weeks and include tags selected by PSM. 19 treated tags and 84 control tags are included in the sample.

Figure A1 Robustness check: The estimated impact of the chat rooms on the likelihood of getting an accepted answer within 2 hours and 4 hours

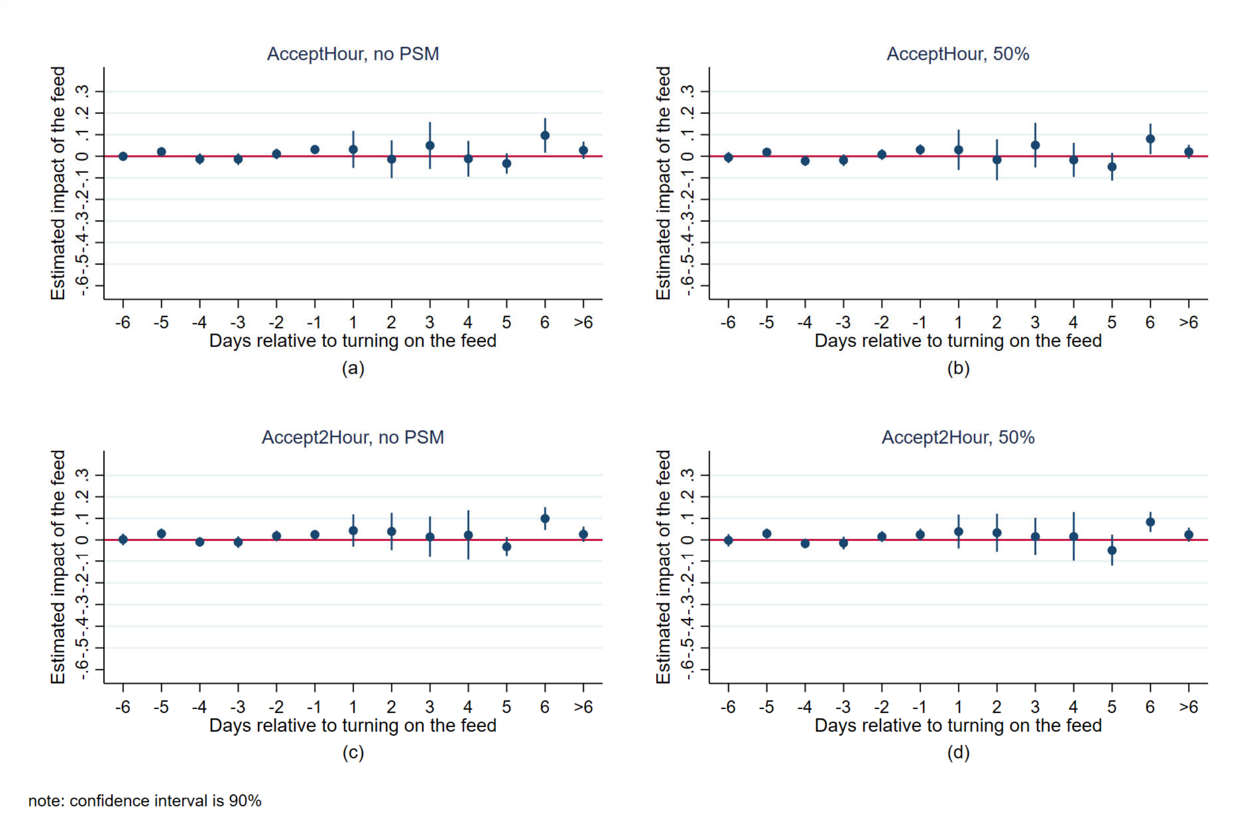

Table A6 Main effect on heterogeneity analysis sample: user with reputation record and questions with quality measure

	AcceptHour	Accept2Hour	Accept4Hour	Accept8Hour
	(1)	(2)	(3)	(4)
TurnOnFeed	0.0132	0.0310**	0.0247**	0.0335**
	(0.0117)	(0.0150)	(0.0120)	(0.0141)
R^2	0.0007	0.0008	0.0007	0.0008
N	153,929	153,929	153,929	153,929
N. of tag-episode	93	93	93	93
Mean	0.41	0.46	0.49	0.51

* p<0.1; ** p<0.05; *** p<0.01
Heteroskedasticity robust standard error clustered at tag-episode level in parentheses
Including tag-episode level fixed effect, weekly dummies and weekday dummies

Table A7 Low reputation users benefit more from the chat rooms, alternative threshold for low reputation.

	AcceptHour (1)	Accept2Hour (2)	Accept4Hour (3)	Accept8Hour (4)
TurnOnFeed*repaboveBottom20	-0.0006	0.0185	0.0091	0.0193
	(0.0129)	(0.0158)	(0.0132)	(0.0160)
TurnOnFeed*repBottom20	0.0474***	0.0560**	0.0612*	0.0627**
	(0.0180)	(0.0250)	(0.0322)	(0.0304)
repBottom20	-0.1286***	-0.1416***	-0.1505***	-0.1579***
	(0.0090)	(0.0093)	(0.0099)	(0.0102)
R^2	0.0120	0.0142	0.0158	0.0174
N	153,929	153,929	153,929	153,929
N. of tag-episode	93	93	93	93
TurnOnFeed	0.0132	0.0310**	0.0247**	0.0335**
	(0.0117)	(0.0150)	(0.0120)	(0.0141)
Diff: repBottom20 - repaboveBottom20	0.0480**	0.0374	0.0521	0.0434
	(0.0186)	(0.0275)	(0.0382)	(0.0384)
Mean	0.4147	0.4567	0.4868	0.5081

* p<0.1; ** p<0.05; *** p<0.01
This table examines the robustness of the results of Table 9 to an alternative definition of low reputation, defining users with reputation less than or equal to 20 as low reputation.
Heteroskedasticity robust standard error clustered at tag-episode level in parentheses
Including tag-episode level fixed effect, weekly dummies and weekday dummies

Table A8 Low reputation users benefit more from the chat rooms (controlling for low question quality)

	AcceptHour (1)	Accept2Hour (2)	Accept4Hour (3)	Accept8Hour (4)
TurnOnFeed*LowRep	0.0666***	0.0801***	0.0863***	0.0738**
	(0.0219)	(0.0263)	(0.0321)	(0.0310)
TurnOnFeed*HighRep	0.0104	0.0336**	0.0248**	0.0243
	(0.0128)	(0.0156)	(0.0120)	(0.0152)
HighRep	0.1325***	0.1452***	0.1540***	0.1613***
	(0.0092)	(0.0094)	(0.0099)	(0.0102)
TurnOnFeed*LowQuality	-0.0517*	-0.0693**	-0.0726**	-0.0257
	(0.0281)	(0.0266)	(0.0291)	(0.0274)
LowQuality	-0.0122**	-0.0142***	-0.0140***	-0.0151***
	(0.0053)	(0.0048)	(0.0048)	(0.0047)
R^2	0.0127	0.0148	0.0164	0.0180
N	153,929	153,929	153,929	153,929
N. of tag-episode	93	93	93	93
Diff (marginal): HighRep - LowRep	-0.0562***	-0.0466*	-0.0616	-0.0494
	(0.0195)	(0.0278)	(0.0381)	(0.0386)
Diff (total): HighRep - LowRep	0.0981***	0.0986***	0.0997***	0.1004***
	(0.0260)	(0.0259)	(0.0259)	(0.0260)
Mean	0.4147	0.4567	0.4868	0.5081

* p<0.1; ** p<0.05; *** p<0.01
Heteroskedasticity robust standard error clustered at tag-episode level in parentheses
Including tag-episode level fixed effect, weekly dummies and weekday dummies

Table A9 Experience: 3 way interaction, compared to noPSM tags, linear combination

	AcceptHour (1)	Accept2Hour (2)	Accept4Hour (3)	Accept8Hour (4)
Panel A				
TurnOnFeed*LowQuality*LowRep	0.1106***	0.1076**	0.0806*	0.1134***
	(0.0335)	(0.0440)	(0.0484)	(0.0326)
TurnOnFeed*LowQuality*HighRep	-0.0724**	-0.0671*	-0.0695*	-0.0226
	(0.0291)	(0.0342)	(0.0359)	(0.0372)
TurnOnFeed*HighQuality*LowRep	0.0254	0.0385	0.0576*	0.0458
	(0.0220)	(0.0241)	(0.0328)	(0.0344)
TurnOnFeed*HighQuality*HighRep	0.0193	0.0426**	0.0309**	0.0302*
	(0.0137)	(0.0165)	(0.0122)	(0.0162)
Panel B				
QAsite*LowQuality*LowRep	-0.1597***	-0.1759***	-0.1850***	-0.1933***
	(0.0107)	(0.0113)	(0.0116)	(0.0120)
QAsite*LowQuality*HighRep	-0.0058	-0.0070	-0.0067	-0.0078
	(0.0050)	(0.0050)	(0.0050)	(0.0049)
QAsite*HighQuality*LowRep	-0.1243***	-0.1360***	-0.1447***	-0.1520***
	(0.0095)	(0.0098)	(0.0105)	(0.0110)
QAsite*HighQuality*HighRep	0	0	0	0
	(-)	(-)	(-)	(-)
Panel C				
LowQuality*LowRep	-0.0688	-0.0683	-0.0672	-0.0673
	(0.0430)	(0.0429)	(0.0427)	(0.0428)
LowQuality*HighRep	-0.0746**	-0.0741**	-0.0731**	-0.0732**
	(0.0343)	(0.0342)	(0.0342)	(0.0342)
HighQuality*LowRep	-0.0980***	-0.0975***	-0.0964***	-0.0972***
	(0.0220)	(0.0219)	(0.0219)	(0.0220)
HighQuality*HighRep	0.0414**	0.0426**	0.0452***	0.0451***
	(0.0167)	(0.0165)	(0.0167)	(0.0167)
Mean	0.4147	0.4567	0.4868	0.5081

* p<0.1; ** p<0.05; *** p<0.01

The table analyzes the effect of the feed on the likelihood of receiving an answer for questions of different combinations of reputation and quality. It is the analog to Table 12 in the text for durations other than 2 hours. Panel A shows the marginal effects of the feed and how it differs based on reputation and quality. Panel B shows the likelihood a question will be answered based on quality and reputation if it is not treated by the feed. Panel C summarizes panels A and B by combining their effects to show the total relative magnitudes of the likelihood that a question will receive an answer within 2 hours when pushed into a chat room.