Analyzing Payment Based Question and Answering Service

Steve Jan*, Chun Wang*, Qing Zhang**, and Gang Wang*

*Department of Computer Science, College of Engineering, Virginia Tech
**Instructional Design and Technology, School of Education, Virginia Tech

Abstract

Community based question answering (CQA) services receive a large volume of questions today. It is increasingly challenging to motivate domain experts to give timely answers. Recently, payment-based CQA services explore new incentive models to engage real-world experts and celebrities by allowing them to set a price on their answers. In this paper, we perform a data-driven analysis on Fenda, a payment-based CQA service that has gained initial success with this incentive model. Using a large dataset of 220K paid questions (worth 1 million USD) over two months, we examine how monetary incentives affect different players in the system and their over-time engagement. Our study reveals several key findings: while monetary incentive enables quick answers from experts, it also drives certain users to aggressively game the systems for profits. In addition, this incentive model turns CQA service into a supplier-driven marketplace where users need to proactively adjust their price as needed. We find famous people are unwilling to lower their price, which in turn hurts their income and engagement-level over time. Based on our results, we discuss the implications to future payment-based CQA design.

Introduction

The success of community based question answering (CQA) services depends on high-quality content from users, particularly from domain experts. With highly engaging experts, services like Quora and StackOverflow attract over a hundred million monthly visitors worldwide (Yeung 2016). However, for most CQA systems, domain experts are answering questions *voluntarily* for free. As the question volume keeps growing, it becomes more and more difficult to draw experts' attention to a particular question, let alone getting answers on-demand (Srba and Bielikova 2016a).

To motivate experts for on-demand question answering, one possible direction is to introduce monetary incentives (Hsieh and Counts 2009). Recently, Quora started a beta test on "knowledge prize", which allows users to put cash reward on their questions[1]. While Quora is slowly accumulating interested users (less than 10 paid answers per month), another payment-based service called *Fenda*[2] is rising quickly in China. Fenda is a mobile social network app that connects users to well-known domain experts and celebrities to ask questions with payments. Launched in May 2016, Fenda quickly gained 10 million registered users, 500K paid answers and 2 million US dollar transactions in just two months (Xuanmin 2016).

Fenda and Quora lead a new wave of payment-based CQA services that socially engage users with real-world domain experts. Similar services are emerging in China (Zhihu) and US (Whale Q&A). The involvement of verified experts differs them from earlier payment-based CQA services (most defunct now) that were built on an anonymous crowd such as Mahalo Answers and ChaCha (Chen, HO, and KIM 2010; Hsieh, Kraut, and Hudson 2010; Lee et al. 2013).

So, is monetary incentive the solution to strong expert engagement in CQA systems? How does monetary incentive affect the behaviors of different players in the system and their over-time engagement? These questions are critical for payment-based CQA design, and Fenda provides a unique platform to study them. First, Fenda is the first supplier-driven CQA market, where answerers (experts) set their own price. Users ask questions to a specific person instead of an anonymous crowd using payments. In addition, Fenda's incentive model not only rewards answerers, but also those who ask good questions. After a question is answered, other users need to pay a small amount ($0.14) to listen to the answer. This money is split evenly between the asker and answerer (Figure 1). A good question may attract enough listeners to compensate the initial question cost.

In this paper, we describe our experience and findings in an effort to understand the impact of monetary incentives on CQA systems, via a detailed measurement of Fenda.[3] We collected a dataset of 88,540 users and 212,082 question-answers over 2 months in 2016 (involving 1.1 million USD financial transactions). Given the drastic differences between Fenda and mainstream CQA systems such as Quora and StackOverflow, our study can have significant implications on the direction of future CQA design. We focus on key issues faced by payment-based CQA systems, including incentivizing users without encouraging cheating/abuse, and long term sustainability of the community.

[1] https://www.quora.com/answer/prizes

[2] http://fd.zaih.com/fenda

[3] Our study has received approval from our local IRB under protocol # 16-1143.

Our study reveals a number of key findings.

- *First*, we seek to understand the effectiveness of monetary incentives to engage domain experts. Our result shows this attempt is successful. Fenda attracts a small group of high-profile experts and celebrities from 44 domains. These experts count for less than 0.5% of the user population, but have contributed a quarter of all answers and drove nearly half of the financial revenue.
- *Second*, we analyze the impact of the incentive model to both question askers and answers. We find a mixed effect. On the positive side, monetary incentive enables quick answers (average delay 8 hours); It also motivates users to ask good questions: 40% of the questions successfully drew enough interested audience to cover the askers' cost. However, we did find a small number of "bounty hunters" who have identified simple rules to make money and aggressively apply them to accumulate profits.
- *Third*, we study the dynamics between money and user engagement over time. In a supplier-driven CQA marketplace, users need to set the price of their answers. We find different pricing strategies of users have distinct impacts on their own engagement level. Users who proactively adjust their price are more likely to increase income and engagement level. Certain highly famous people, however, are unwilling to lower their price, which in turn hurts their income and social engagement.

To the best of our knowledge, this is the first empirical study on a supplier-driven CQA marketplace. In the end of the paper, we discuss the key implications of our results to CQA system design, which can benefit other arising payment-based CQA services (Quora knowledge prize, Whale Q&A, Zhihu).

Related Work

Community Based Question Answering (CQA). In recent years, researchers have studied CQA services from various aspects (Srba and Bielikova 2016b). A key direction focuses on identifying experts from the crowd (Pal, Chang, and Konstan 2012; Pal et al. 2013), and routing questions to the right experts (Li and King 2010). Other researchers have worked on assessing and predicting the quality of questions and answers (Ravi et al. 2014; Yao et al. 2014; Shah and Pomerantz 2010; Tian, Zhang, and Li 2013; Adamic et al. 2008; Harper et al. 2008), and response delay (Rechavi and Rafaeli 2011; Bhat et al. 2014). Related to quality control, researchers also developed models to characterize and detect abusive users in CQA services (Kayes et al. 2015; Srba and Bielikova 2016a).

User Incentives in CQA. Prior works have summarized the key user motivations to answer questions, which involve intrinsic, extrinsic, and social factors (Jin et al. 2013). Intrinsic factors refer to the psychic reward (*e.g.*, enjoyment) that users gain through helping others (Yu, Jiang, and Chan 2007; Nam, Ackerman, and Adamic 2009). Social factors refer to the benefits of social interaction, *e.g.*, gaining respect from other users. Intrinsic and social motivations are considered the most vital motivations for knowledge sharing in non-payment based CQA services (Jin et al. 2013). Finally, extrinsic factors include money, virtual awards (*e.g.*, badges and points), and reputation enhancement. For example, users in Naver Knowledge-iN (Nam, Ackerman, and Adamic 2009) and StackOverflow (Grant and Betts 2013) can answer questions to earn points and badges to elevate their ranking in the community.

Monetary incentive is an extrinsic factor implemented in earlier payment-based CQA services such as Google Answers, Mahalo, ChaCha and Jisiklog (Chen, HO, and KIM 2010; Hsieh, Kraut, and Hudson 2010; Lee et al. 2013). Most of them are defunct. Compared to Fenda and Quora, these systems focus on building a CQA market on an anonymous crowd, instead of a social network to engage with real-world experts. Users are primarily driven by financial incentives without a strong sense of community (Lee et al. 2013; Hsieh and Counts 2009). This is concerning since research shows monetary incentive plays an important role in getting users started in CQA, but it is the social factors that contribute to the persistent participation (Raban 2008).

Researchers have studied the impact of monetary incentives on the quality of content, but the conclusions vary. Some researchers find paying more improves answer quality (Harper et al. 2008). Other studies suggest payment has no significant impact on answer quality but merely helps to collect more answers more quickly (Chen, HO, and KIM 2010; Katmada, Satsiou, and Kompatsiaris 2016; Jeon, Kim, and Chen 2010; Mason and Watts 2010; Hsieh, Kraut, and Hudson 2010). Prior works also show payment-based CQA helps to reduce low-quality questions as users are more selective on what to ask (Hsieh and Counts 2009; Hsieh, Kraut, and Hudson 2010).

Mobile CQA. Mobile CQA systems such as ChaCha and Jisiklog allow users to ask questions to an online crowd via instant messages. Prior research shows the purposes of using mobile CQA apps often involve quick information seeking and asking for suggestions and opinions (Lee et al. 2012; 2013). Fenda is a mobile-only CQA service and is the first to implement audio-based question answering. In Fenda, users ask questions in text and answer questions in audio. In the context of education and communication, research shows the audio channel is preferred over text (Lunt and Curran 2010; King, McGugan, and Bunyan 2008), and audio provides a stronger social bounding effect (Sherman, Michikyan, and Greenfield 2013).

Research Questions and Method

Systems like Fenda and Quora are leading the way to socially engage real-world experts for question answering. The introduction of monetary incentives makes user interactions in these systems even more complex. If not carefully designed, monetary incentives can lead the systems down to the wrong path with users chasing financial profits and losing engagement in the long run. In this paper, we use Fenda as the platform to investigate how monetary incentives impact the user behavior and engagement-level in CQA systems. We choose Fenda over Quora for two main reasons: 1) Fenda has received early success with a significant vol-

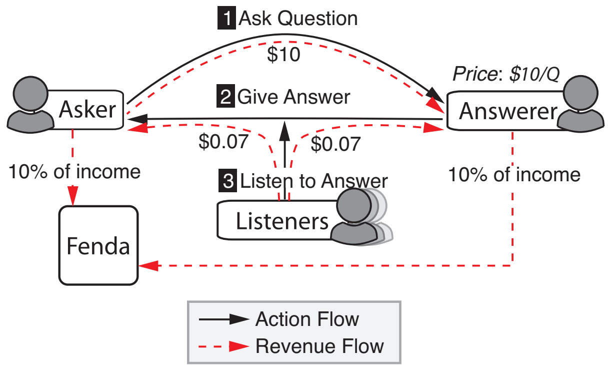


Figure 1: Fenda QA system: a user can ask the other user a question by paying her (price set by the answerer). Other users need to pay a small amount to listen to the answer ($0.14), which will be split evenly between the asker and answerer. Fenda charges 10% commission fee.

ume of data, while Quora hasn't accumulated sufficient paid content (*e.g.*, less than 10 rewarded questions in December 2016). 2) Fenda is the first supplier-driven market with a unique incentive model to motivate both question askers and respondents. We aim to understand the reasons behind its initial success and potential problems, which will benefit future CQA design.

Background of Fenda. Fenda is a mobile payment-based Q&A app in China. Fenda connects users to their friends, domain experts, and celebrities in a Twitter-like social network (*i.e.*, users following each other). Launched in May 2016, Fenda quickly gained 10 million registered users and over 2 million US dollars worth of question answers in the first two months (Xuanmin 2016; Kwong 2016).

Fenda has a unique *monetary incentive model* to reward both question askers and answerers. As shown in Figure 1, a user (asker) can ask another user (answerer) a question by paying the price set by the answerer. The answerer then responds over the phone by recording a 1-minute audio message. If the answerer doesn't respond within 48 hours, the payment will be refunded. Any other user on Fenda can listen to the answer by paying a fixed amount of 1 Chinese Yuan (US$0.14), and it will be split evenly between the asker and answerer. A good question may attract enough listeners to compensate the initial cost for the asker. Users get to decide how much they charge per question and can change the price anytime. Fenda charges 10% of the money made by any user. All financial transactions are made through a popular mobile payment service WeChat Pay[4].

There are two types of users on Fenda: verified real-world domain experts (*e.g.*, doctors, entrepreneurs, movie stars) and normal users. There is an *expert list* that contains all the experts divided based on categories. Users can browse questions from the social news feed, and there is a public stream of answers (a small sample of popular ones). To promote user engagement, Fenda selects 2-4 answers daily on the public stream for free-listening for a limited time.

[4] https://pay.weixin.qq.com/

Date (2016)	# of Questions	# of Users	# of Askers	# of Answerers
5/12 – 7/27	212,082	88,540	85,510	15,529

Table 1: Summary of Fenda dataset.

Our Questions. In the following, we use Fenda as the platform to analyze how monetary incentives impact user behavior and their engagement-level. We take a data-driven approach to answer the following key questions.

- First, as an expert-driven CQA system, to what extent does Fenda rely on experts to generate content and drive revenue? How do experts shape the interaction patterns among users?
- Second, how does the monetary incentive affect the question answering process? Does money truly enable on-demand question answering from experts? Can users make money by asking (good) questions? Will monetary incentives encourage users to game the system for profits?
- Third, in this supplier-driven market, how do users set and dynamically adjust the price of their answers? How does the pricing strategy affect their income and engagement-level over time?

Next, we introduce our data collection method, and explore answers to these questions from the data.

Data Collection

We collected a large dataset from Fenda through their public APIs. Our data collection focused on user profiles, which contained a full list of historical questions answered by the users. To obtain a large set of active users, we explored different options, some of which did not work. First, there is no central list of all registered users for us to scan the user space. Second, a user's follower list is not public (only the total number is visible), and thus social graph crawling is not feasible. To these ends, we started our crawling from the expert list. For each expert, we collected their answered questions and the askers of those questions. Then we collected the askers' profiles to get their answered question list and extract new askers. We repeated this process until no new users appeared. In this way, we collected a large set of active users who asked or answered at least one question.

We collected data in July 2016. The dataset contains 88,540 user profiles and 212,082 question-answer pairs ranging from May 12 to July 27, 2016. Each question was characterized by the asker's userID, question text, a timestamp, question price, and the number of listeners. Each answer is characterized by the answerer's userID, a length of the audio and a timestamp. UserIDs in our dataset have been fully anonymized. A data summary is shown in Table 1.

We briefly estimate the coverage of the dataset. Fenda announced that they had 500,000 answers as of June 27, 2016 (Xuanmin 2016). Up to the same date, our dataset covers 155,716 answers (31%). We believe this dataset can serve as a representative sample for our analysis.

Engaging with Domain Experts

As a CQA system driven by real-world experts, we first explore the roles and impact of domain experts in the system.

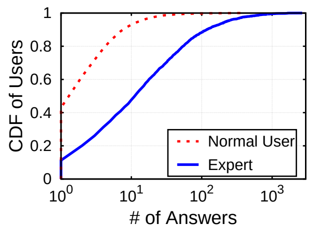


Figure 2: # of Answers per answerer.

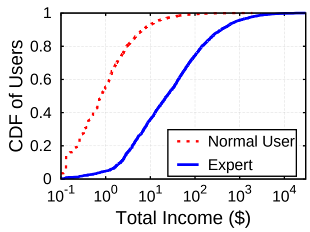


Figure 3: Total income per answerer.

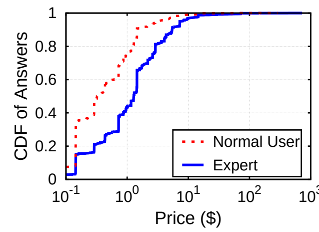


Figure 4: Price of each answer.

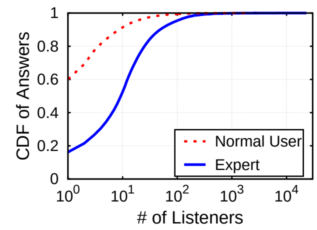


Figure 5: # of Listeners per answer.

This helps to set up the context for later sections. We analyze domain experts from two levels: individual level and macro-graph level. At the *individual level*, we assess the contributions of domain experts to the community in terms of generating content and driving financial revenue. We seek to understand how much Fenda motivates experts to contribute. Then at the *macro level*, we examine how the gravity of domain experts shapes the interaction patterns among users through a graph analysis.

Significant Expert Contribution

Fenda maintains a list of verified experts and celebrities, who are typically already well-known in the real-world in their respective domains. As of the time of data collection, there were 4370 verified experts classified into 44 categories ranging from"law", to "business", "health" and "entertainment". We refer these 4,370 users as *experts* and the rest 84,170 users as *normal users*.

Question Answering. We find that this small group of experts has contributed to a significant portion of answers. Out of the 212K answers in our dataset, 171K (81%) are from experts. Using this dataset, we can estimate the experts' contribution in the context of the entire network. On June 27 2016, Fenda officially announced total 500K answers and 10 million users (Xuanmin 2016). Up to the same date, our dataset shows the 4370 experts (0.44% of the population) have contributed 122K answers (24.4% of total answers). Individually, experts also answered significantly more questions than normal users as shown in Figure 2.

Money. Experts' answers play an important role in driving revenue to Fenda. In total, the questions in our dataset were worth $1,169,994 including payments from askers and listeners.[5] Expert's answers generated $1,106,561, counting for a dominating 95% of total revenue in our dataset. To gauge experts' contribution in the context of the entire network, we again performed estimation: Fenda reached 2 million revenue as of June 27 (Xuanmin 2016). Up to this same date (June 27), expert answers in our dataset have attracted $909,876, counting for a significant 45% of the 2 million revenue. As shown in Figure 4 and Figure 5, experts' answers usually charge higher ($2.9 vs. $1.0 on average) and draw more listeners (27 vs. 5 on average) than those of normal users.

[5] We convert Chinese Yuan to US dollar based on $1 = 6.9 Yuan, which is the exchange rate as of December 2016.

Rank	Category	Total Income $ (%)	Experts # (%)
1	Health	$123,667 (12.4%)	204 (6.6%)
2	Career	$81,950 (8.2%)	222 (7.2%)
3	Business	$81,375 (8.2%)	108 (3.5%)
4	Relationship	$73,895 (7.4%)	90 (2.9%)
5	Movies	$52,768 (5.3%)	84 (2.7%)
6	Entertainment	$52,476 (5.3%)	51 (1.7%)
7	Academia	$49,366 (5.0%)	64 (2.1%)
8	Media	$45,182 (4.5%)	138 (4.5%)
9	Real Estate	$43,842 (4.4%)	28 (0.9%)
10	Education	$39,962 (4.0%)	174 (5.7%)
11	Music	$32,037 (3.2%)	71 (2.3%)
12	Sports	$31,386 (3.2%)	67 (2.2%)
13	Science	$27,686 (2.8%)	104 (3.4%)
14	Writers	$25,253 (2.5%)	93 (3.0%)
15	Self-media	$25,026 (2.5%)	109 (3.6%)

Table 2: Top 15 expert categories with highest earnings.

Individually, experts make more money than normal users. Figure 3 shows the total income for users who answered at least one question (with Fenda's commission fee deducted). We find that 50% of experts made more than $23, while a small group of experts (5%) made more than $1000. The highest earning is $33,130 by Sicong Wang, a businessman and the son of a Chinese billionaire. He answered 31 questions related to gossip and investment. He charged $500 for each of his answers, which drew 9484 listeners ($664 extra earning) per answer on average.

Expert Categories. We find that experts of different categories have distinct earning patterns. Table 2 lists the top 15 categories, ranked by the total earnings per category. We find that professional consulting types of experts are most popular. The top category is "health", followed by "career", "business" and "relationship". Other popular categories such as "movies" and "entertainments" contains questions about insider experience, gossip and opinions on trending events.

In Figure 6, we illustrate the distinct earning patterns of experts. For each category, we compute the average price and a number of answers from each expert. The red lines represent the average values across all experts, which divide them into 4 sections. Experts in "health", "entertainment", "relationship" and "real estate" often charge high and answer more questions. These experts are among the top-earning groups (Table 2); Experts in "business" set the price high but don't answer many questions; Less-serious categories such as "funny" and "comics" have fewer and less expensive questions. "Digital" then represents the other extreme where experts answer lots of questions with a low

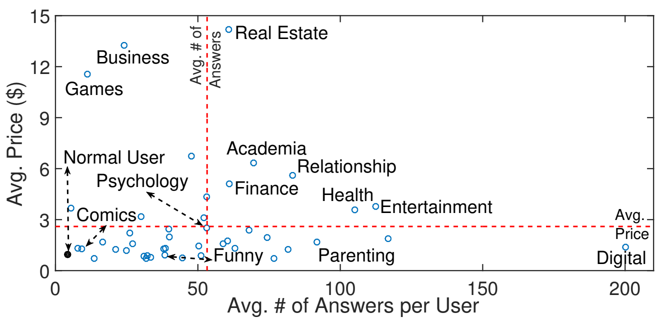


Figure 6: Scatter plot of the average number of answers per expert and average price in each expert category. The red lines represent the corresponding average values across all experts.

Graph	# of Nodes	# of Edges	Avg. Degree	Cluster Coef.	Assort. Coef.	Recip.
Fenda	87K	154K	3.54	0.02	-0.16	0.04
Quora	159K	833K	10.46	0.05	-0.03	0.01
FB	707K	1.13M	1.78	0.059	0.116	0.126
Twitter	4.32M	17.0M	3.93	0.048	-0.025	0.025

Table 3: High level statistics of different interaction graphs.

price. This reflects the different perceived values of knowledge in different domains.

Asymmetric Interaction Graph

Next, we move to the macro-level to analyze the impact of experts on the overall user interaction patterns. To do so, we build an interaction graph where each node represents a user, and a directed edge means a user has asked the other user a question. We compare Fenda's interaction graph with Quora (Wang et al. 2013), Facebook (Wilson et al. 2009) and Twitter (Xu et al. 2011). Facebook's interaction graph is based on Wall post and Twitter's interaction graph is based on Retweet. Key graph statistics are shown in Table 3.

We make three key observations. First, Fenda graph has the most negative assortativity. This metric measures how likely a user connects to other users of similar degrees. Negative assortativity means users tend to connect to others with dissimilar degrees. Fenda's assortativity value (-0.16) shows the impact of experts: normal users tend to ask experts questions while rarely ask questions among each other. This leads to a highly asymmetrical graph structure. Second, Fenda has a slightly lower clustering coefficient (0.02) compared to other graphs (0.048–0.059), indicating a sparser local connectivity. Intuitively, users who consult the same expert may not ask questions among each other. Finally, Fenda's reciprocity value is higher than Quora and Twitter, but not as high as Facebook. A closer analysis shows that only 12% of the bidirectional edges are related to experts. Connections between experts and normal users are still highly asymmetrical.

In summary, our results show that a small group of verified domain experts play an important role in Fenda who make significant contributions to generating content and financial revenues. In addition, the dominance of domain experts in Fenda also leads to asymmetric interaction patterns, differentiating Fenda from typical online social networks.

Behavior Metric	Correlation with Avg. Price
# Followers	0.53*
Avg. # Listeners	0.65*
# Questions Answered	0.04*
Avg. Response Time	0.01

Table 4: Pearson correlation coefficient between a user's answer price and key behavior metrics including # of answered questions, # of followers, # of listeners, and response time. * indicates significant correlation with $p < 0.05$.

Impact of Monetary Incentives

So far we show that Fenda is highly dependent on domain experts' contribution. Then the question is how to motivate experts to deliver timely and high-quality answers. In this section, we perform extensive analysis on the monetary incentive model in Fenda to understand its impact on user behavior. Noticeably, Fenda is the first system to use money to reward both question answerers and askers. To this end, we first analyze *question answerers* to understand how they price their answers, and whether payments lead to on-demand responses. Second, we focus on *askers* analyzing whether and how users make money by asking questions. We also seek to identify potential "bounty hunters" who aggressively or strategically game the system for profits.

Question Answerers

To motivate users (particularly domain experts) to answer questions, Fenda lets users set their own price for answering a question. In the following, we investigate how money affects the way users answer questions. Particularly, we examine if monetary incentives truly enable on-demand quick answers.

Setting the Answer Price. Fenda is a supplier-driven market where users set their own answer prices. The price can affect or be affected by various factors. In Table 4, we calculate the Pearson correlation (Sheskin 2007) between a user's price and different behavior metrics. We find that price has positive and significant correlations with the number of followers and listeners, and total questions answered. A possible explanation is that users with many followers and listeners are real-world celebrities who have the confidence to set the price high. A higher price may also motivate them to answer more questions. Note that these are correlation results, which do not reflect causality.

Surprisingly, there is no significant correlation between price and response time. This is different from existing results on customer-driven CQA markets, where an asker can set a high cash prize on her questions to collect answers more quickly (Katmada, Satsiou, and Kompatsiaris 2016; Mason and Watts 2010; Hsieh, Kraut, and Hudson 2010).

Answering On-demand? We further examine the response time to see if monetary incentives truly enable on-demand question answering. As shown in Figure 7, answers arrive fast on Fenda: 33% of answers arrived within an hour

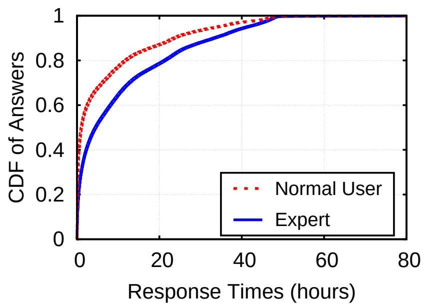


Figure 7: Response time of answers.

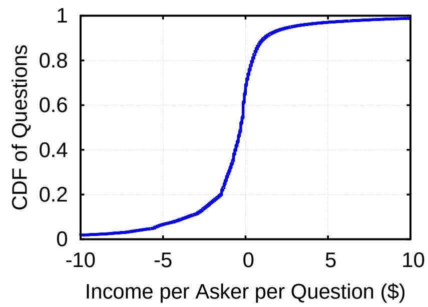


Figure 8: Income of askers per question.

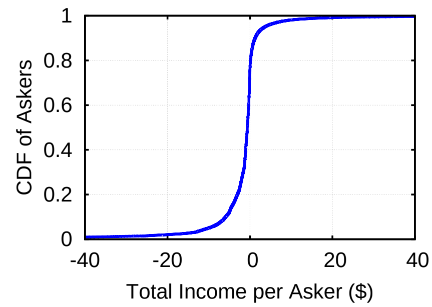


Figure 9: Total income of askers.

CQA System	Avg. Resp. Time (hr)	Payment Based?	Crowdsourcing or Targeted?
Yahoo Answers	8.25	N	Crowdsourcing
Fenda	10.4	Y	Targeted
Google Answers	36.9	Y	Crowdsourcing
Stack Overflow	58.0	N	Crowdsourcing

Table 5: Average response time of first answers (in hours). We compare Fenda with different CQA sites including Yahoo Answers (Wu and He 2014), Google Answers (Edelman 2011) and StackOverflow (Mamykina et al. 2011).

Behavior Metric	Askers w/ Income>0	Askers w/ Income $\leq$0	t-Test p value
Average # Followers	2155.5	3758.5	$< 0.001*$
Average # Listeners	55.2	16.9	$< 0.001*$
Average Price	1.58	4.58	$< 0.001*$
Average # Questions	3.99	1.86	$< 0.001*$

Table 6: Comparing the behavior metrics for askers with positive income and those with negative income. Two-sample t-test shows the differences are significant.

and 85% arrived within a day. Note that there is a clear cut-off at 48 hours. This is the time threshold when un-answered questions will be refunded. This motivates users to answer questions quickly. After 48 hours, users can still answer those questions for free. We find that only 0.7% of the answers arrived after the deadline, but we cannot estimate how many questions remain unanswered due to the lack of related data. Finally, despite the high price charged by experts, experts respond slower than normal users.

We compare the response time of Fenda with other CQA systems in Table 5. Fenda is faster than Google Answers and StackOverflow, but not as fast as Yahoo Answers. We cannot provide a related number for Quora since Quora hides the question posting time. As a payment-based system, Fenda beats Google Answers probably because Fenda only asks for a short audio message, while Google Answers require lengthy text. Compared to Yahoo Answers, we believe it is the crowdsourcing factor that plays the role. Systems like Yahoo Answers crowdsource questions to a whole community where anyone could deliver the answer. Instead, Fenda's question is targeted to a specific user, which may lead to a longer delay even with payments. For general questions, it is possible for Fenda to further speed up the response by allowing multiple answers.

Question Askers

Fenda designs the first monetary incentive model to reward users for asking good questions. Specifically, after a user's question gets answered, this user (as well as the answerer) earns $ 0.07 whenever another user listens to this answer. This model, if executed as designed, should motivate users to contribute high-quality questions for the community. In the meantime, the financial reward may also attract users who simply want to game the system for profit (*i.e.*, "bounty hunters"). In the following, we analyze whether people can make money by asking questions, and what strategies were used to bring in profit. In addition, we explore "bounty hunters" type behavior in askers.

How Do Askers Make Money? For each question, the question asker's income is half of listeners' payments, with Fenda's commission fee and initial question fee deducted. Our result shows that askers are motivated to ask good questions that attract broad interests. As shown in Figure 8, out of all questions, 40% have successfully attracted enough listeners to return a positive profit to the asker. For individual askers, Figure 9 shows 40% of them have a positive total income.

To understand why certain users make money (and others don't), we compare askers who have positive income with those with negative income in Table 6. Specifically, we examine to whom they ask questions (*i.e.*, the number followers and listeners of the answerer), average question price, and total questions asked. A two-sample t-Test (Sheskin 2007) shows the differences are significant between the two groups of askers. Not surprisingly, users of positive income are more likely send questions to people who have more listeners and charge less.

The counter-intuitive result is the *number of followers*: asking people with more followers is more likely to lose money. A possible explanation is an inherent correlation between a user's number of followers and her answer price — famous people would charge higher and the profit from listeners cannot cover the initial cost.

We also observe that askers with a higher income often asked more questions. Again, correlation does not reflect causality: it is possible the positive income motivates users to ask more questions; It is also possible that users who asked more questions get more experienced to earn money.

Finally, we examine if the types of the target experts

Rank	Askers w/ Income > 0	Askers w/ Income ≤ 0
1	Non-expert (10%)	Non-expert (25%)
2	Career (10%)	Health (12%)
3	Health (8%)	Career (7%)
4	Education (7%)	Sports (4%)
5	Others (6%)	Relationship (4%)
6	Science (5%)	Education (4%)
7	Marketing (3%)	Science (4%)
8	Writers (3%)	Media (3%)
9	Fashion (3%)	Entertainment (3%)
10	Internet (3%)	Psychology (3%)

Table 7: Comparing askers with positive and negative income on the types of experts they asked. We list the top 10 categories and % of questions in each category.

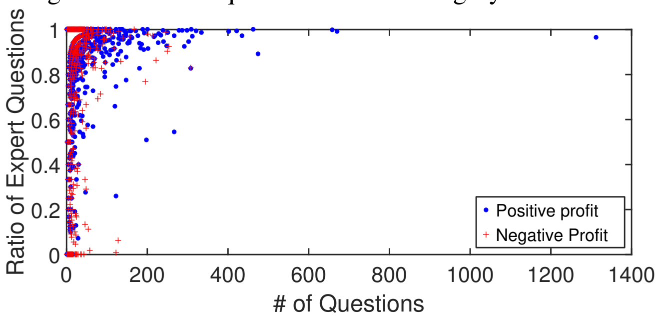


Figure 10: Total # of questions of each asker vs. the ratio of questions to experts. Blue dots (red crosses) represent askers with positive (negative) total income. The figure is better viewed in color.

would affect the askers' income. As shown in Table 7, askers with negative income are more likely to engage with non-experts (25%) compared to positive-income askers (10%). There is no clear difference in terms of expert categories. This shows that engaging with experts helps to make a profit from their listeners.

Abnormal Askers. Next, we look for potential "bounty hunters" who aggressively ask questions for profit. Our intuition is that those users would ask a lot of questions, particular to experts. To identify potential bounty hunters, we examine outliers in Figure 10, which is a scatter plot for the number of questions a user asked versus the ratio of questions to experts. We find clear outliers at the right side (*e.g.*, users with $>$100 questions). Most of these users end up with positive income. They asked way more questions than average (which is 2.27 per user), and exclusively interact with experts (ratio of expert questions is close to 1). The most extreme example is a user who asked more than 1300 questions in two months, with 95% of questions to experts. This user earned \$194.20, which is much higher than the average income of askers (-\$1.95).

To further examine these outliers, we select askers who asked more than 100 questions. This gives us 111 users who count for 0.13% of askers in our dataset but aggressively asked 11% of the questions. In addition, they seem to carefully target experts who charge a lower price (\$0.80 per answer on average) but still draw significant listeners (15.5 per answer on average). As a comparison, the rest of the experts charge \$2.49 with 23.0 listeners per answer on average.

id	Feature Name	Feature Description
1	Price Change Freq.	# of price change / # answers
2	Price Up Freq.	# price up / # answers
3	Price Down Freq.	# price down / # answers
4	Price Up - Down	(# price up - # price down) / # answers
5	Price Up Magnitude	Average percentage of price increase
6	Pirce Down Magnitude	Average percentage of price decrease
7	Consecut. Same Price	Max # consecutive same price / # answers
8	Consecut, Price Up	Max # consecutive price up / # answers
9	Consecut. Price Down	Max # consecutive price down / # answers

Table 8: A list of features for price change dynamics.

The result suggests that monetary incentives did foster undesired behavior. On the positive side, there are not many of them, and these users actually work hard to come up with interesting questions. On the negative side, beyond the answered questions we observed, these users may have asked a lot more questions that remained unanswered. These large volume of questions can act as spam to experts, which blocks other users' chance to get the experts' attention.

Dynamic Pricing and User Engagement

As a supplier-driven CQA market, Fenda lets users set the price for their answers. Our analysis has shown that price is a key factor that affects users' interactions in the community and their income. In this section, we turn to the *dynamic* aspect to analyze how users adjust their answer prices over time and how different pricing strategies affect their engagement level. Understanding this question is critical since keeping users (particularly experts) engaged in the community is the key to build a suitable CQA service.

In the following, we first identify common pricing strategies by applying unsupervised clustering on users. We seek to group users with similar patterns of price change. Then we analyze identified clusters to understand who they represent, and how their engagement-level changes over time.

Identifying Distinct Pricing Strategies

To characterize users' dynamic price change, we first construct a list of features. Then we use these features to group users with similar price change patterns to capture common pricing strategies.

Key Features. We have 9 features as shown in Table 8. For each user, we model their price change as a sequence of events. For user i, our dataset contains the complete list of her answers and the price for each answer. We use P_i to denote user i's price sequence $P_i = [p_{i,1}, p_{i,2}, ..., p_{i,N_i}]$ where N_i is the total number of answers of user i. A price change event happens when $p_{i,j-1} \neq p_{i,j}$ for any $j \in [2, N_i]$. We denote the price change sequence as $C_i = [c_{i,1}, c_{i,2}, ...c_{i,M_i}]$ where M_i is a number of times for price change and $c_{i,j}$ is a price change event (either price-up or price-down).

As shown in Table 8, our features include the overall frequency of price change (*i.e.*, $\frac{M_i}{N_i}$), a frequency for price-up and price-down, and the frequency difference between price-up and down. In addition, we consider the average price change magnitude for price-up and price-down events.

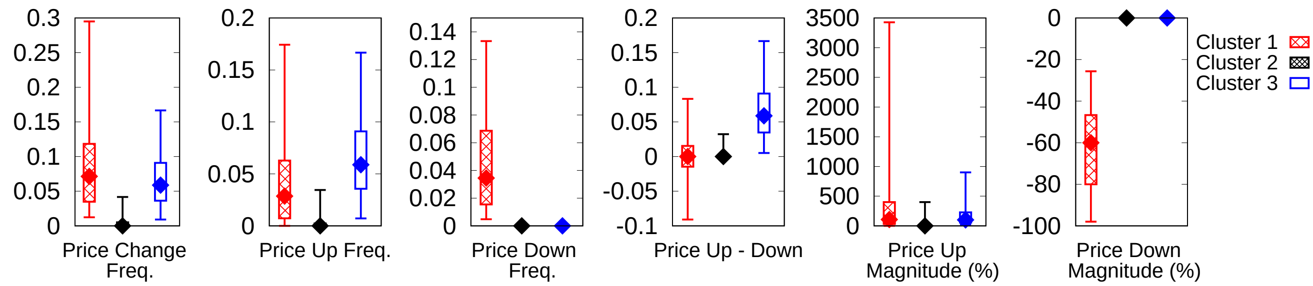


Figure 11: The distribution of top 6 features for the 3 clusters. We depict each distribution with box plot quantiles (5%, 25%, 50%, 75%, 95%). The detailed description of each feature is in Table 8.

Finally, we consider the maximum number of consecutive same price, price up and price down in the answer sequence.

Clustering Method. Based on these features, we then cluster users of similar patterns into groups. First, we use the feature vector to compute the pair-wise Euclidean distance for users. This produces a fully connected similarity graph (Wang et al. 2016) where each node is a user and edges are weighted by distance. Then, we apply hierarchical clustering algorithm (Fortunato 2010) to detect groups of users with similar price change patterns. We choose hierarchical clustering for two reasons: 1) It does not require predefining the number of clusters. 2) It is deterministic and thus the clustering result does not depend on initial seeding.

To determine the number of clusters, we use *modularity*, a well-known metric to measure clustering quality (Fortunato 2010). High modularity means groups of users are more densely connected within each group than to the rest of the users. We choose the number clusters that yields the highest modularity.

For this analysis, we only consider users who have answered enough questions (more than 10). Otherwise, discussing their price change and engagement would be less meaningful. This gives us 2094 users who have contributed 171,322 answers (85% of all answers in our dataset).

Clustering Results. Our method produces 3 clusters (highest modularity). To interpret the pricing strategy of each cluster, we plot their feature value distributions in Figure 11. Due to space limitation, we plot 6 (out of 9) most distinguishing features that have the largest variance among the 3 clusters based on chi-squared statistic (Sheskin 2007). The three common pricing strategies are:

- **Cluster 1 (32.81%):** *Frequent Price Up and Down.* This cluster has 687 users and 76% of them are experts. Price change frequency and magnitude are both high. Price up and down are almost equally frequent.
- **Cluster 2 (43.36%):** *Rarely Changing Price.* This cluster contains 908 users and 76% of them are experts. These users rarely change their price.
- **Cluster 3 (23.8%):** *Frequent Price Up.* This cluster contains 499 users and 74% of them are experts. These users increase price frequently but rarely lower their price.

We find that the 3 types of pricing patterns correspond to users of different popularity. As shown in Table 9, cluster 1 represents the least popular answerers, who have the least followers and listeners but answered more questions per day. These users constantly adjust their price, possibly to test the market. Cluster 3 represents the most popular experts and celebrities. They charge higher than others and keep increasing the price. Cluster 2 stands between cluster 1 and 3 in terms of popularity, and its users rarely change the price.

Metric	Cluster1	Cluster2	Cluster3
Average #Followers	627.60	749.46	951.39
Average #Listeners	16.62	26.96	25.88
Average Price ($)	1.70	2.42	2.58
Average #Questions/Day	3.11	2.56	2.33

Table 9: Key statistics for users in three identified clusters.

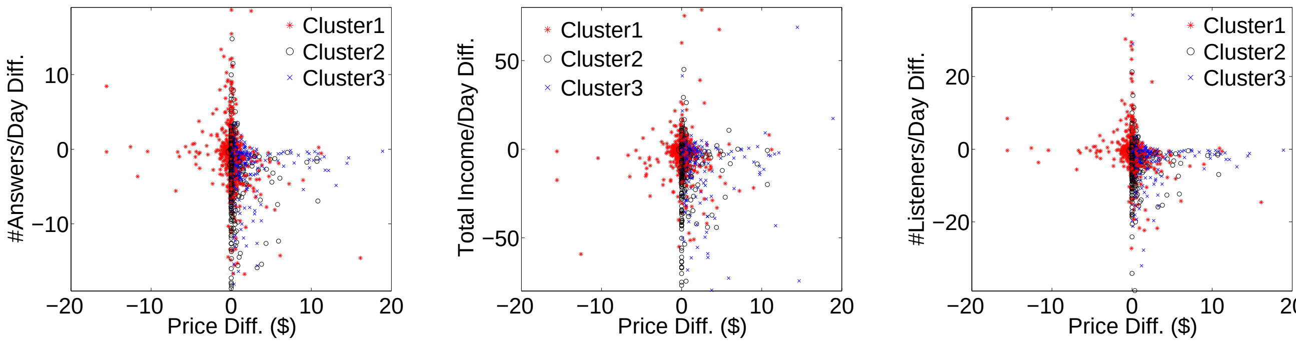
Impact on User Engagement

Next, we analyze how price adjustments affect a user's engagement-level over time. Price is a key parameter within users' control, and adjusting price is a way to test their answers' value in the market. Intuitively, this price can affect a user's incoming questions, earnings and social interactions, which are key incentive factors to keep users engaged.

Figure 12(a) shows the interplay between price change and engagement level over time for 3 identified clusters. We quantify engagement-level using number of answers per day. To measure changes over time, we divide a user's lifespan (time between her first and last answer in our dataset) into two even parts. Then we compute the differences for average price and engagement-level between the later half and first half. In a similar way, we also measure the changes in income (Figure 12(b)) and listeners (Figure 12(c)), which represent the strength of monetary and social incentives.

We observe different patterns: For cluster 2 and 3, more users are located in the lower right corner than upper right, indicating a decrease of engagement, income and number of listeners. A possible explanation is that there is a mismatch between the answer's price and its value, but users did not make the right adjustments. In contrast, we find a significant number of users in cluster 1 located in the upper left corner. By lowering their price, these users get to answer more questions, and receive more money and listeners over time.

In a supplier-driven CQA market, users need to set their price carefully to match their market value. This requires proactive price adjustments and lowering their price when necessary. Our result indicates highly popular users (*e.g.*, users in cluster 3) are less motivated or unwilling to lower



(a) Engagement. (b) Income. (c) Listeners.

Figure 12: The impact of pricing strategy to user engagement (answers per day), income, and listeners over time. We divide a user's lifespan in our dataset into two even parts, and compute the difference between the later half and the first half. A positive value indicates an upward trend. The figure is better viewed in color.

their price, which in turn hurts their income and engagement level over time.

Discussion

Key Implications. Our result shows both positive impacts of Fenda's incentive model as well as some concerning issues in the long run. Below, we discuss the key implications of our results to Fenda and other similar CQA systems.

First, *On-demand Answers.* To get quick answers from experts, Fenda adopts a supplier-driven model where experts set their price and wait for incoming questions. This model is suitable for targeted questions (users know who to ask), but can have longer delays compared to crowdsourcing (anyone can be a potential answerer). Our result shows that Fenda achieves faster response than most of CQA services, but still not as fast as the crowdsourcing based Yahoo answers. Recently in late 2016, Fenda added a new crowdsourcing channel for legal/medical questions. This channel is customer-driven: users post their questions with a cash reward, and all experts can give answers to compete for the money. We did a quick crawling on January 12, 2017 on the crowdsourcing channel and obtained 1344 questions. The average response time is 4.38 hours, which is even faster than the 8.28 hours of Yahoo Answers (Table 5). Future research is needed to understand the efficiency of the hybrid CQA marketplace with both supplier- and customer-driven channels.

Second, *Rewarding Good Questions.* Fenda is the first system rewarding question askers by using money. This leads to a mixed effect. On the positive side, users are motivated to ask good questions that attract broad interests. 40% of the questions received enough listeners to cover the asker's cost. On the negative side, this did motivate a small number of users to game the system for profits. We find these "bounty hunters" primarily target experts who charge a relatively lower price (instead of highly popular experts) to ask questions. This behavior is different from the cheating behavior in crowdsourcing where cheaters often produce low-quality work (Gadiraju et al. 2015). In this model, bounty hunters still need to come up with good questions to attract listeners. The side effect is that the large volume of questions can block other users' chance to get these experts' attention.

Finally, *Unfairness in Supplier-driven CQA Market.* In a supplier-driven CQA market, well-known experts and celebrities have a key advantage to receive questions. As a result, the financial income among answerers is highly uneven: top 5% answerers get about 90% of the total profit. To attract questions, we find that less popular users need to carefully adjust their price (including dropping the price), while more popular users tend to increase their price. To help users to bootstrap popularity, Fenda recently introduced a system update, which allows users to set their answers "free-for-listening" for 30 minutes after posting.

Study Limitations. Our study still has limitations. First, our study primarily focuses on one service. We believe a more thorough comparison with other similar systems can help to generalize our results to better understand the design space. We expect systems such as Whale Q&A and Quora knowledge prize will accumulate sufficient data over time for future studies. Second, our dataset is not perfect. The crawler produces a dataset with a complete list of experts but an incomplete of normal users. As shown in our analysis, we have used officially reported numbers from Fenda to justify our results. Also, there is no available data for us to estimate the ratio of unanswered questions in Fenda, which we consider as a limitation. Finally, Fenda is still exploring its way to building a sustainable CQA marketplace. It made a few changes right before our paper submission as discussed above. We plan to continue to collect data from Fenda and report related analysis results in future work.

Conclusion

In this paper, we discuss lessons learned from the first supplier-driven payment-based CQA system. By analyzing a large empirical dataset, we reveal the benefits of applying monetary incentives to CQA systems (fast response, high-quality questions) as well as potential concerns (bounty hunters and over time engagement). As more payment-based CQA systems arise (Whale Q&A, Quora Knowledge Prize, Zhihu), our research results can help future CQA system designers to make more informed design choices.

References

Adamic, L. A.; Zhang, J.; Bakshy, E.; and Ackerman, M. S. 2008. Knowledge sharing and yahoo answers: everyone knows something. In *Proc. of WWW*.

Bhat, V.; Gokhale, A.; Jadhav, R.; Pudipeddi, J.; and Akogl, L. 2014. Min(e)d your tags: Analysis of question response time in stackoverflow. In *Proc. of ASONAM*.

Chen, Y.; HO, T.-H.; and KIM, Y.-M. 2010. Knowledge market design: A field experiment at google answers. *Journal of Public Economic Theory* 12(4):641–664.

Edelman, B. 2011. Earnings and ratings at google answers. *Economic Inquiry* 50(2):309320.

Fortunato, S. 2010. Community detection in graphs. *Physics Reports* 486:75 – 174.

Gadiraju, U.; Kawase, R.; Dietze, S.; and Demartini, G. 2015. Understanding malicious behavior in crowdsourcing platforms: The case of online surveys. In *Proc. of CHI*.

Grant, S., and Betts, B. 2013. Encouraging user behaviour with achievements: an empirical study. In *Mining Software Repositories (MSR), 2013 10th IEEE Working Conference on*, 65–68. IEEE.

Harper, F. M.; Raban, D.; Rafaeli, S.; and Konstan, J. A. 2008. Predictors of answer quality in online q&a sites. In *Proc. of CHI*.

Hsieh, G., and Counts, S. 2009. mimir: A market-based real-time question and answer service. In *Proc. of CHI*.

Hsieh, G.; Kraut, R. E.; and Hudson, S. E. 2010. Why pay?: exploring how financial incentives are used for question & answer. In *Proc. of the CHI*.

Jeon, G. Y.; Kim, Y.-M.; and Chen, Y. 2010. Re-examining price as a predictor of answer quality in an online q&a site. In *Proc. of CHI*.

Jin, X.-L.; Zhou, Z.; Lee, M. K.; and Cheung, C. M. 2013. Why users keep answering questions in online question answering communities: A theoretical and empirical investigation. *International Journal of Information Management* 33(1):93–104.

Katmada, A.; Satsiou, A.; and Kompatsiaris, I. 2016. Incentive mechanisms for crowdsourcing platforms. In *Proc. of INSCI*.

Kayes, I.; Kourtellis, N.; Quercia, D.; Iamnitchi, A.; and Bonchi, F. 2015. The social world of content abusers in community question answering. In *Proc. of WWW*.

King, D.; McGugan, S.; and Bunyan, N. 2008. Does it make a difference? replacing text with audio feedback. *Practice and Evidence of Scholarship of Teaching and Learning in Higher Education* 3(2):145–163.

Kwong, P. 2016. China's online question-and-answer platform fenda raises us$25 million from investors. `http://www.scmp.com/business/companies/article/1982524/chinas-online-question-and-answer-platform-fenda-raises-us25`.

Lee, U.; Kang, H.; Yi, E.; Yi, M.; and Kantola, J. 2012. Understanding mobile q&a usage: an exploratory study. In *Proc. of the CHI*.

Lee, U.; Kim, J.; Yi, E.; Sung, J.; and Gerla, M. 2013. Analyzing crowd workers in mobile pay-for-answer q&a. In *Proc. of CHI*.

Li, B., and King, I. 2010. Routing questions to appropriate answerers in community question answering services. In *Proc. of CIKM*.

Lunt, T., and Curran, J. 2010. are you listening please?the advantages of electronic audio feedback compared to written feedback. *Assessment & Evaluation in Higher Education* 35(7):759–769.

Mamykina, L.; Manoim, B.; Mittal, M.; Hripcsak, G.; and Hartmann, B. 2011. Design lessons from the fastest Q&A site in the west. In *Proc. of CHI*.

Mason, W., and Watts, D. J. 2010. Financial incentives and the performance of crowds. *ACM SigKDD Explorations Newsletter* 11(2):100–108.

Nam, K. K.; Ackerman, M. S.; and Adamic, L. A. 2009. Questions in, knowledge in?: a study of naver's question answering community. In *Proc. of CHI*.

Pal, A.; Wang, F.; Zhou, M. X.; Nichols, J.; and Smith, B. A. 2013. Question routing to user communities. In *Proc. of CIKM*.

Pal, A.; Chang, S.; and Konstan, J. A. 2012. Evolution of experts in question answering communities. In *Proc. of ICWSM*.

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

Ravi, S.; Pang, B.; Rastogi, V.; and Kumar, R. 2014. Great question! question quality in community q&a. In *Proc. of ICWSM*.

Rechavi, A., and Rafaeli, S. 2011. Not all is gold that glitters: Response time and satisfaction rates in yahoo! answers. In *Proc. of PASSAT/SocialCom*.

Shah, C., and Pomerantz, J. 2010. Evaluating and predicting answer quality in community qa. In *Proc. of SIGIR*.

Sherman, L. E.; Michikyan, M.; and Greenfield, P. M. 2013. The effects of text, audio, video, and in-person communication on bonding between friends. *Cyberpsychology* 7(2).

Sheskin, D. J. 2007. *Handbook of Parametric and Nonparametric Statistical Procedures*.

Srba, I., and Bielikova, M. 2016a. Why is stack overflow failing? preserving sustainability in community question answering. *IEEE Software* 33(4):80–89.

Srba, I., and Bielikova, M. 2016b. A comprehensive survey and classification of approaches for community question answering. *ACM Trans. Web* 10(3):18:1–18:63.

Tian, Q.; Zhang, P.; and Li, B. 2013. Towards predicting the best answers in community-based question-answering services. In *Proc. of ICWSM*.

Wang, G.; Gill, K.; Mohanlal, M.; Zheng, H.; and Zhao, B. Y. 2013. Wisdom in the social crowd: an analysis of quora. In *Proc. of WWW*.

Wang, G.; Zhang, X.; Tang, S.; Zheng, H.; and Zhao, B. Y. 2016. Unsupervised clickstream clustering for user behavior analysis. In *Proc. of CHI*.

Wilson, C.; Boe, B.; Sala, A.; Puttaswamy, K.; and Zhao, B. 2009. User Interactions in Social Networks and Their Implications. In *Proc. of EuroSys*.

Wu, D., and He, D. 2014. Comparing ipl2 and yahoo! answers: A case study of digital reference and community based question answering. In *Proc. of IConf.*

Xu, T.; Chen, Y.; Jiao, L.; Zhao, B. Y.; Hui, P.; and Fu, X. 2011. Scaling microblogging services with divergent traffic demands. In *Proc. of Middleware*.

Xuanmin, L. 2016. Putting a price on knowledge. `http://www.globaltimes.cn/content/997510.shtml`.

Yao, Y.; Tong, H.; Xu, F.; and Lu, J. 2014. Predicting long-term impact of cqa posts: A comprehensive viewpoint. In *Proc. of KDD*.

Yeung, K. 2016. Quora now has 100 million monthly visitors, up from 80 million in january. VentureBeat.

Yu, J.; Jiang, Z.; and Chan, H. C. 2007. Knowledge contribution in problem solving virtual communities: The mediating role of individual motivations. In *Proc. of SIGMIS CPR*.