Cui, Lixin; Liang, Yibao; Li, Yiling

Article

The study of customer involved service innovation under the crowdsourcing: A case study of MyStarbucksIdea.com

Journal of Industry-University Collaboration (JIUC)

Provided in Cooperation with:
Research Center for China Industry-University-Research Institute Collaboration, Wuhan University

Suggested Citation: Cui, Lixin; Liang, Yibao; Li, Yiling (2020) : The study of customer involved service innovation under the crowdsourcing: A case study of MyStarbucksIdea.com, Journal of Industry-University Collaboration (JIUC), ISSN 2631-357X, Emerald, Leeds, Vol. 2, Iss. 1, pp. 22-33, https://doi.org/10.1108/JIUC-12-2019-0018

Received 20 December 2019
Revised 18 January 2020
Accepted 18 January 2020

The study of customer involved service innovation under the crowdsourcing

A case study of MyStarbucksIdea.com

Lixin Cui, Yibao Liang and Yiling Li
Beijing Institute of Technology, Beijing, China

Abstract

Purpose – Service innovation is a key source of competence for service enterprises. Along with the emergence of crowdsourcing platforms, consumers are frequently involved in the process of service innovation. In this paper, the authors describe the crowdsourcing ideation website—MyStarbucksIdea.com—and find the motivations of customer-involved service innovation.

Design/methodology/approach – Using a rich data set obtained from the website MyStarbucksIdea.com, a dynamic structural model is proposed to illuminate the learning process of consumers.

Findings – The results indicate that initially individuals tend to underestimate the costs of the firm for implementing their ideas but overestimate the value of their ideas. By observing peer votes and feedbacks, individuals gradually learn about the true value of ideas, as well as the cost structure of the firm. Overall, the authors find that the cumulative feedback rate and the average potential of ideas will first increase and then decline.

Originality/value – First, the previous researches concerning the crowdsourcing show that the creative implementation rate is low and the number of creative ideas decreases, and few scholars have studied the causes behind the problems. Second, the data used in this paper are true and valid, and it is difficult to obtain now. These data can provide strong empirical support for the model proposed in this paper. Third, it is relatively novel to combine the customer learning mechanism and heterogeneity theory to explain the phenomenon of reduced creativity and low implementation rate in crowdsourcing platform, and the research results can provide a reasonable reference for the construction of this industry.

Keywords Crowdsourcing, Structural modeling, Dynamic learning, Service innovation, Customer involved
Paper type Research paper

1. Introduction

Service innovation is the source of business competitiveness. Recently, along with the development of information technology, crowdsourcing is beginning to gain popularity in various fields. Howe (2006) defined "crowdsourcing" as "the new pool of cheap labor: everyday people using their space cycles to create content, solve problems, even do corporate R and D. Such initiatives of crowdsourcing provide a platform for everyone to post their own ideas, and these ideas are usually generated from direct or indirect service experience. Therefore, the customer group is a rich source of preference information. A typical crowdsourcing platform allows customers to support or oppose others' ideas, so that the preliminary assessment of existing ideas can be obtained. Through such initiatives, the firm can acquire a great number of ideas that are innovative and beneficial. However, arguments about the real utility of crowdsourcing have never been settled. In fact, many crowdsourcing

Journal of Industry-University Collaboration
Vol. 2 No. 1, 2020
pp. 22-33
Emerald Publishing Limited
2631-357X
DOI 10.1108/JIUC-12-2019-0018

platforms are experiencing a decrease in the number of new-posted ideas, and the feedback rate (the percentage of ideas with official feedbacks in all existing ideas) remains low. Nevertheless, we have not found enough systemic and in-depth researches on these problems.

The majority of researches on crowdsourcing are aimed at crowdsourcing contests, in which customers post ideas to compete to win an award (Archak and Sundararajan, 2009; DiPalantino and Vojnovic, 2009; Mo *et al.*, 2011; Terwiesch and Xu, 2008). Unlike crowdsourcing contests, in a permanent and open idea solicitation such as MyStarbucksIdea.com, there is no competition among idea contributors; instead, they help each other evaluate ideas. Unfortunately, only a few studies are conducted on this type of crowdsourcing initiatives. Through a reduced form approach, Bayus (2013) finds that individual creativity is positively related to current efforts, while negatively related to previous success. Di Gangi *et al.* (2010) find that there are two factors demonstrating whether an individual's idea will be adopted — the firm's ability to understand the technical requirements and to give feedback to concerns for ideas in the community. Lu *et al.* (2011) find that in crowdsourcing ideation initiatives, complementarities, and customer support provide people with chances of learning. This mechanism allows people to know about the problems that other customers have encountered, and help them come up with more ideas that are worth implementing. Yan *et al.* (2014) study the crowdsourcing platform IdeaStorm.com, which is affiliated to Dell. Their work has pioneered a research direction of structurally investigating new product ideas, as well as their development process, on the basis of real crowdsourcing data.

2. Data collection and analysis

Our data are collected from the crowdsourcing website MyStarbucksIdea.com, which is affiliated to Starbucks. This online community was established in March 2008, and it is dedicated to sharing and discussing ideas and allowing people to see how Starbucks is putting top ideas into action.

The structure of MyStarbucksIdea.com is simple, but quite efficient. Anyone (not only the customer of Starbucks) can register at the website and become a member of the online community. Afterward, anyone who owns the membership can post ideas on the website. Starbucks classifies all ideas into three categories—product ideas, experience ideas, and involvement ideas. Before an individual posts an idea, he or she shall select the category that the idea belongs to. As long as an idea is posted, other members can vote for it. If one supports an idea, he or she can submit a positive vote, which will add 10 points to the idea. And if one opposes an idea, he or she can submit a negative vote, which results in a deduction of 10 points. On the website, however, only the cumulative scores are available, while the specific number of positive and negative votes is not open to the public. Moreover, registered members can write their comments under an idea to explain detailed thoughts.

Typically, the change of an idea's status contains the following stages. Once an idea is posted, the voting and commenting function is then available to the public. The review team will select all ideas according to the scores, and deliver the good ones to the decision-makers. At this moment, the status of these ideas changes to "under review." For those ideas that are already reviewed, the status becomes "reviewed." Next, ideas that are worth implementing are selected, and their status evolves into "coming soon." Finally, when an idea is completely implemented, its status changes to "launched." In our paper, customers' learning process is gradually advancing based on two information sources—the scores and status of ideas.

There is a rich data set on MyStarbucksIdea.com, which contains detailed information about both ideas and members. We collected the public data on the website, and divided it into two groups—idea profile data and member profile data. We acquired 96,793 records of member profile data, and selected 21,305 individuals who posted more than two or more ideas

(these individuals are called "selected members"). The time ranged from January 2009 to December 2015. We found a similar distribution between ideas contributed by selected members and those by the whole member groups, so the selected member group is representative.

In Figure 1, we present the relationship between the cumulative feedback rate of the three categories and time.

3. Model construction

Taking Yan *et al.* (2014) as reference, we modify the structural model to describe the decision-making process of costumers, and further explain the data generation process. Through the explicit modeling of individuals' utility function, we can use the data to empirically recover the parameters in the analytical model.

In each time period of our model, every member will make the decision about whether to post an idea in a certain category, but whether to put the decision into practice is determined by the corresponding utility expectation. Thus, we first explain the utility function.

Suppose that individuals are indexed i, j is the index of categories ($j = 1, 2, 3$), and t denotes time. The utility function consists of four factors. The first and second factors are related to benefit, which means that if an idea is implemented, the contributor will obtain better service experience, higher online reputation, or even job opportunities. We use the parameter r_i to represent the reputation gain. Then, the third factor is the cost for posting an idea, including thinking, articling, and posting the idea. The whole cost is denoted as c_i. Finally, the fourth factor involves the discontent; for example, if an idea is not accepted or responded, the contributor will gain discontent with such situations. In the model, whether individual i is discontent in period t is denoted as D_{it} (a binary variable, $D_{it} = 1$ means that the person is discontent, while $D_{it} = 0$ means content). In addition, the degree of such discontent is measured by the parameter d_i.

Hence, the utility function is given by the following equation:

$$U_{ijt} = \begin{cases} \theta_{i0} + d_i D_{it} + \theta_{ij} + \varepsilon_{ijt} \\ \text{if the idea is implemented} \\ \theta_{i0} + d_i D_{it} + \varepsilon_{ijt} \\ \text{if the idea is not implemented} \end{cases}$$

where U_{ijt} represents the utility function when individual i posts a category j idea in period t. The parameter θ_{ij} measures the utility gain for individual i posting a category j idea, and the error term ε_{ijt} captures the random shock of decision in period t. Since c_i is linearly correlated with r_i, we cannot obtain the specific value of both of them. Thus, we combine them in a new parameter θ_{i0}, and we have $\theta_{i0} = c_i + r_i$.

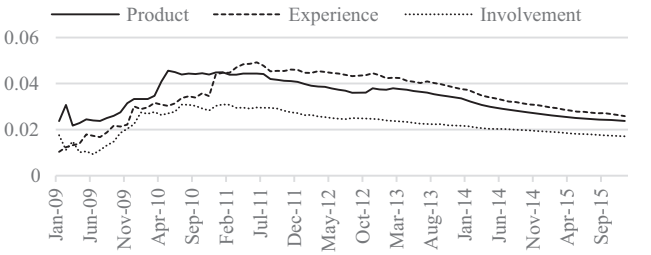


Figure 1.
Cumulative feedback rate by category

When an individual posts an idea, he or she will hold a belief that the idea will be accepted based on the existing information, that is, the expectation of the utility function, which is represented as $E(U_{ijt} \mid \text{Info}(t))$.

$$E(U_{ijt} \mid \text{Info}(t)) = \tilde{U}_{ijt} + \varepsilon_{ijt} = \theta_{i0} + d_i D_{it} + \theta_{ij} P_{ijt} \mid \text{Info}(i,\ t) + \varepsilon_{ijt} (\divideontimes)$$

where $P_{ijt} \mid \text{Info}\ (i,\ t)$ denotes the probability of acceptance based on the online information.

Every member holds an expectation of the cost and value of their idea, and they update the expectation through the information from the website. Meanwhile, they will learn about the firm's cost structure and the true value of their idea.

Suppose that the cost for implementing an idea in category j follows a normal distribution $N(C_j,\ \sigma_{yj}^2)$. Furthermore, we assume that the firm exactly knows the actual cost, while members do not. The prior belief of individuals is that the average cost for implementing an idea in category j is C_{j0}, and it follows a normal distribution $C_{j0} \sim N(C_0,\ \sigma_{C_0}^2)$. If an idea is implemented, all individuals will receive the same implementation signal, and the voting function of the implemented idea will be closed.

We allow C_{kjt} to represent the cost signal that all individuals receive, and let μ_{kjt} to represent the mean value of all cost signals in the same category, which follows the distribution $N(0,\ \sigma_\mu^2)$. The variance σ_μ^2 indicates the difference among specific cost signals. This means that the cost signals are unbiased, but noisy

$$C_{kjt} = C_j + \mu_{kjt}, \quad \mu_{kjt} \sim N\left(0,\ \sigma_\mu^2\right).$$

In a certain period of time, there could be more than one idea implemented. If there are $k_{C_{jt}}$ ideas in category j implemented in period t, the accumulated cost signal that an individual will receive can be measured by C_{sjt}, which is the average of $(C_{1jt},\ C_{2jt},\ \cdots,\ C_{k_{C_{jt}}jt})$, and it follows the following distribution

$$C_{sjt} \sim N\left(C_j,\ \frac{\sigma_\mu^2}{k_{C_{jt}}}\right).$$

We allow C_{jt-1}^e to denote one's belief of the cost for implementing an idea in category j at the beginning of period t. Based on cumulative information, the updating process of C_{jt}^e performs as follows (DeGroot, 1970).

$$C_{jt}^e = C_{jt-1}^e + \left(C_{sjt} - C_{jt-1}^e\right) \frac{\sigma_{C_{jt-1}}^2}{\sigma_{C_{jt-1}}^2 + \frac{\sigma_\mu^2}{k_{C_{jt}}}}, \tag{1}$$

$$\sigma_{C_{jt}}^2 = \frac{1}{\frac{1}{\sigma_{C_{jt-1}}^2} + \frac{k_{C_{jt}}}{\sigma_\mu^2}}. \tag{2}$$

The prior in period $t = 0$ is $C_{j0}^e = C_0$.

We calculate the voting score that one's ideas receive in different categories, and we find no significant difference. Additionally, we have verified that one's abilities of coming up with new ideas are not influenced by the learning curve effect. Therefore, we can assume that the value of one's idea remains similar and will not change over time. As soon as individual i enters the website, her prior belief of the value can be written as

$$Q_{i0} \sim N\left(Q_0, \sigma^2_{Q_0}\right).$$

Moreover, the voting score is an excellent measurement, suppose that the natural logarithm of the voting score (V) of an idea is linearly correlated with and the value of the idea.

$$V = \text{cons} + \varphi Q. \tag{3}$$

For individual i, the prior belief of the natural logarithm of the voting score can be expressed as

$$V_{i0} \sim N\left(\text{cons} + \varphi Q_0,\ \varphi^2 \sigma^2_{Q_0}\right).$$

We use the parameter Q_i to represent the average value of all ideas submitted by individual i, and we use Q_{sit} to denote the value of a specific idea raised by individual i in period t. Then, we have the following expressions

$$Q_{sit} = Q_i + \delta_{sit}, \quad \delta_{sit} \sim N\left(0,\ \sigma^2_{\delta_i}\right).$$

where δ_{sit} denotes the deviations from the mean value, while the value of δ_{sit} varies from individual to individual and changes over time.

Note that members will update their perception of their ideas' value through voting scores; suppose that the natural logarithm of the voting score that a specific idea receives is

$$V_{sit} = V_i + \xi_{sit}, \tag{4}$$

where

$$V_I = \text{cons} + \varphi Q_I,$$

$$\xi_{sit} = \varphi \delta_{sit}, \quad \xi_{sit} \sim N\left(0,\ \sigma^2_{\xi_i}\right), \quad \sigma^2_{\xi_i} = \varphi^2 \sigma^2_{\delta_i}.$$

Here, V_i is the mean value of V_{sit}, and ξ_{sit} is the deviation from V_i.

Similarly, we define Q^e_{it-1} and V^e_{it-1} to represent one's belief of the value of an idea in category j, as well as its voting score, at the beginning of period t. Thus, the updating process of Q^e_{it} and V^e_{it} follows the same rule as C^e_{it} (Erdem *et al.*, 2008):

$$V^e_{it} = V^e_{it-1} + \left(V_{sit} - V^e_{it-1}\right) \frac{\sigma^2_{V_{it-1}}}{\sigma^2_{V_{it-1}} + \sigma^2_{\xi_i}}, \tag{5}$$

$$Q^e_{it} = Q^e_{it-1} + \left(V_{sit} - V^e_{it-1}\right) \frac{\varphi \sigma^2_{Q_{it-1}}}{\varphi \sigma^2_{Q_{it-1}} + \sigma^2_{\xi_i}}. \tag{6}$$

Respectively, we have

$$\sigma^2_{V_{it}} = \frac{1}{\frac{1}{\sigma^2_{V_{it-1}}} + \frac{1}{\sigma^2_{\xi_i}}}, \quad \sigma^2_{Q_{it}} = \frac{1}{\frac{1}{\sigma^2_{Q_{it-1}}} + \frac{\varphi^2}{\sigma^2_{\xi_i}}}.$$

In addition, the prior values in period $t = 0$ are

$$Q^e_{i0} = Q_0, \quad \sigma^2_{Q_{i0}} = \sigma^2_{Q_0}, \quad V^e_{i0} = \text{cons} + \varphi Q_0, \quad \sigma^2_{V_{i0}} = \varphi^2 \sigma^2_{Q_0}.$$

The firm will take a comprehensive account of the idea's cost and value when it decides which idea to put into implementation. Suppose that the firm only accepts ideas that can bring positive net profit, and we allow the parameter π_{mjt} to represent the net profit of implementing the m th idea in category j during period t, and then we have

$$\pi_{mjt} = Q_{mjt} + C_{mjt},$$

where Q_{mjt} denotes the true value, and C_{mjt} denotes the actual cost.

As a result, the probability of implementing a specific idea can be written as

$$P_{mjt} = Pr(\pi_{mjt} > 0)$$

During the decision-making process, only the firm knows exactly about C_{mjt}, while Q_{mjt} is available to both the firm and customers. From the perspective of the firm, C_{mjt} and Q_{mjt} are known. However, for customers, C_{mjt} is a random variable satisfying the following expressions

$$C_{mjt} = C_j + \gamma_{mjt}, \quad \gamma_{mjt} \sim N\left(C_j, \sigma^2_{\gamma_j}\right).$$

Thus, in terms of the community members, the probability of implementing an idea with the value Q_{mjt} can be expressed as

$$P_{mjt} = Pr(Q_{mjt} + C_{mjt} > 0 \mid Q_{mjt}) = \Phi\left(\frac{Q_{mjt} + C_j}{\sigma^2_{\gamma_j}}\right) \tag{7}$$

Let I_{mjt} represent the decision of the firm, with $I_{mjt} = 1$, meaning that the idea is implemented, and $I_{mjt} = 0$ otherwise. Given Q_{mjt}, C_j, and $\sigma_{\gamma j}$, the likelihood of implementation is

$$L(I_{mjt}) = \Phi\left(\frac{Q_{mjt} + C_j}{\sigma^2_{\gamma_j}}\right)^{I_{mjt}} \left(1 - \Phi\left(\frac{Q_{mjt} + C_j}{\sigma^2_{\gamma_j}}\right)\right)^{(1-I_{mjt})} \tag{8}$$

As is mentioned above, members will refer to their utility function while making decisions about whether to post an idea or not. Suppose that they make their decisions independently and without the influence of idea category. Besides, they know that the firm will take into consideration both the cost and the value. Then, the $\tilde{U}_{ijt}$ in Equation (※) can be written as

$$\tilde{U}_{ijt} = \theta_{i0} + d_i D_{it} + \theta_{ij} Pr(\pi_{ijt} \mid \text{Info}(i, t) > 0), \tag{9}$$

where

$$\pi_{ijt} \mid \text{Info}(i, t) > 0 \sim N\left(Q^e_{it} + C^e_{jt}, \sigma^2_{Q_{it}} + \sigma^2_{C_{jt}} + \sigma^2_{\delta_i} + \sigma^2_{\gamma_j}\right).$$

Note that $\text{Info}(i, t)$ is the information that individuals learn about the cost and value through the two learning processes mentioned previously. This information contains the value of Q^e_{it}, C^e_{jt}, $\sigma^2_{Q_{it}}$, and $\sigma^2_{C_{jt}}$, and evolves as members update their belief of Q^e_{it}, C^e_{jt}, $\sigma^2_{Q_{it}}$, and $\sigma^2_{C_{jt}}$, over time.

We assume that the parameter ε_{ijt} in Equation (※) follows a type 1 extreme value distribution, indicating that the probability for individual i posting an idea in category j during period t meets the standard logit form. Furthermore, let A_{ijt} represent the event whether the idea is posted or not (if the idea is posted, then $A_{ijt} = 1$; otherwise, $A_{ijt} = 0$). Finally, the likelihood of posting the idea can be expressed as

$$L(A_{ijt}) = \left(\frac{\exp\left(\tilde{U}_{ijt}\right)}{1+\exp\left(\tilde{U}_{ijt}\right)}\right)^{A_{ijt}} \left(\frac{1}{1+\exp\left(\tilde{U}_{ijt}\right)}\right)^{\left(1-A_{ijt}\right)} \tag{10}$$

4. Model estimation and result analysis

4.1 Parameter estimates and analysis of cost and value

Taking into consideration both amount and validity, we select a group of individuals who proposed three or more ideas during January 2009 and February 2016 to serve as model samples. Compared with other members, the selected individuals tend to behave more actively and learn faster through a variety of signals; therefore, they fit the model well (details are explained in the following sections). Our samples contain 371 members, together with 2,466 posted ideas. Next, we estimate the model parameters through MATLAB, and the results are summarized in Table I and Table II. We set the value of some parameters which remain constant among individuals (pooled parameters) in Table I, while the estimates of the other pooled parameters are presented in Table II.

Comparing the estimates of C_1, C_2, and C_3, we find that the cost for the firm to implement a category 3 idea (an involvement idea) is the largest, the cost of implementing a category 2 idea (an experience idea) is lower, while the cost of implementing a category 1 idea (a product idea) is the lowest. In addition, C_1, C_2, and C_3 are all smaller than C_0 in terms of absolute value, which means members tend to underestimate the cost that the firm incurs when implementing an idea.

In Table II, we see that σ_μ^2 is smaller than that of C_1, C_2, and C_3 in terms of absolute value, indicating the cost signal customers receive when an idea gets implemented is fairly precise. Since the total number of implemented ideas in every month remains small, these

Table I.
Setting value of the pooled parameters

Notation	Setting value
C_0	−6
$\sigma_{C_0}^2$	50
$\sigma_{\gamma_1}^2$	50
$\sigma_{\gamma_2}^2$	50
$\sigma_{\gamma_3}^2$	50
Q_0	1
$\sigma_{Q_0}^2$	50
cons	10

Table II.
Estimates of the pooled parameters

Notation	Parameter estimates	Standard deviation
C_1	−7.12	8.59
C_2	−8.55	2.46
C_3	−14.32	10.13
ϕ	1.82	1.21
σ_μ^2	2.46	0.46

signals can speed up the learning process of the firm's cost structure. Then, we list the estimates of parameters that vary from individuals (individual-level parameters) in Table III.

In Figure 2, we plot histograms of the distribution of the individual-level parameters shown in Table III.

From Table III, we see that the average of Q_i, which represents the posteriors of the value of an idea posted by an individual, is larger than the initial value Q_0. Additionally, the variance of Q_i is quite large, meaning that the posteriors of the value of an idea significantly differs from each other among individuals. In fact, this is related to the personal posting behavior and the votes that an idea receives. That is, if a member submits more idea, he will learn the value of her ideas faster. Meanwhile, the more votes that an idea receives, the more actual value this idea will obtain.

We also observe that the average of Q_i is much smaller than the absolute value of C_1, C_2, and C_3, that is, the individual mean value of ideas is smaller than the average cost of implementing an idea. This is equivalent to saying that the firm will suffer a loss to carry out an idea, which is in agreement with the low feedback rate of ideas described before. Generally, the average and variance of $\sigma^2_{\delta_i}$, which denotes the variance of ideas' value for an individual, is quite large, meaning that an individual holds different cognition among her own ideas. Furthermore, we can see from the distribution of $\sigma^2_{\delta_i}$ that individuals with $\sigma^2_{\delta_i} \leq 50$ roughly account for half of all selected members. The small value of $\sigma^2_{\delta_i}$ implies a stable value of ideas submitted by individual *i*. Moreover, good idea contributors usually post ideas of high value, while marginal idea contributors tend to submit ideas of low value. Thus, $\sigma^2_{\delta_i}$ with a large average and variance means that the members' learning process of the value of ideas is not efficient enough, so that the filtering process of idea contributors proceeds slowly.

4.2 Analysis of parameters in utility gain

To explore the relationship between the individual mean value of ideas (Q_i) and other individual-level parameters, we present the scatter of $\sigma^2_{\delta_i}$ and d_i against Q_i in Figure 3 (a) ~ (b), respectively.

As is shown in Figure 3 (a), most points gathered in the region of $Q_i > 0$ and $\sigma^2_{\delta_i} < 10$, which means the average value of ideas posted by the majority of the selected members is slightly larger than zero, and the variance ($\sigma^2_{\delta_i}$) is large as well. The results indicate that good idea contributors only make up a small proportion of the selected members, and the general value of ideas stays at a low level. Interestingly, individuals' abilities to raise high-value ideas have not improved significantly through the learning process. Instead, their abilities tend to remain steady after they realized the true value of their ideas. In Figure 3 (b), most points gathered in the region of $Q_i > 0$ and $d_i > -1$ (which is smaller than the average of d_i, −1.22).

Notation	Parameter estimates	Standard deviation
Q_i	0.79	52.78
$\sigma^2_{\delta_i}$	52.00	116.93
d_i	−1.22	0.86
θ_{i0}	0.05	0.04
θ_{i1}	0.08	0.50
θ_{i2}	0.01	0.47
θ_{i3}	0.03	0.55

Note(s): For each individual, the individual-level parameters change over time, so they have average and variance. However, Table III only shows the average and variance from the sample group level

Table III. Estimates of the individual-level parameters

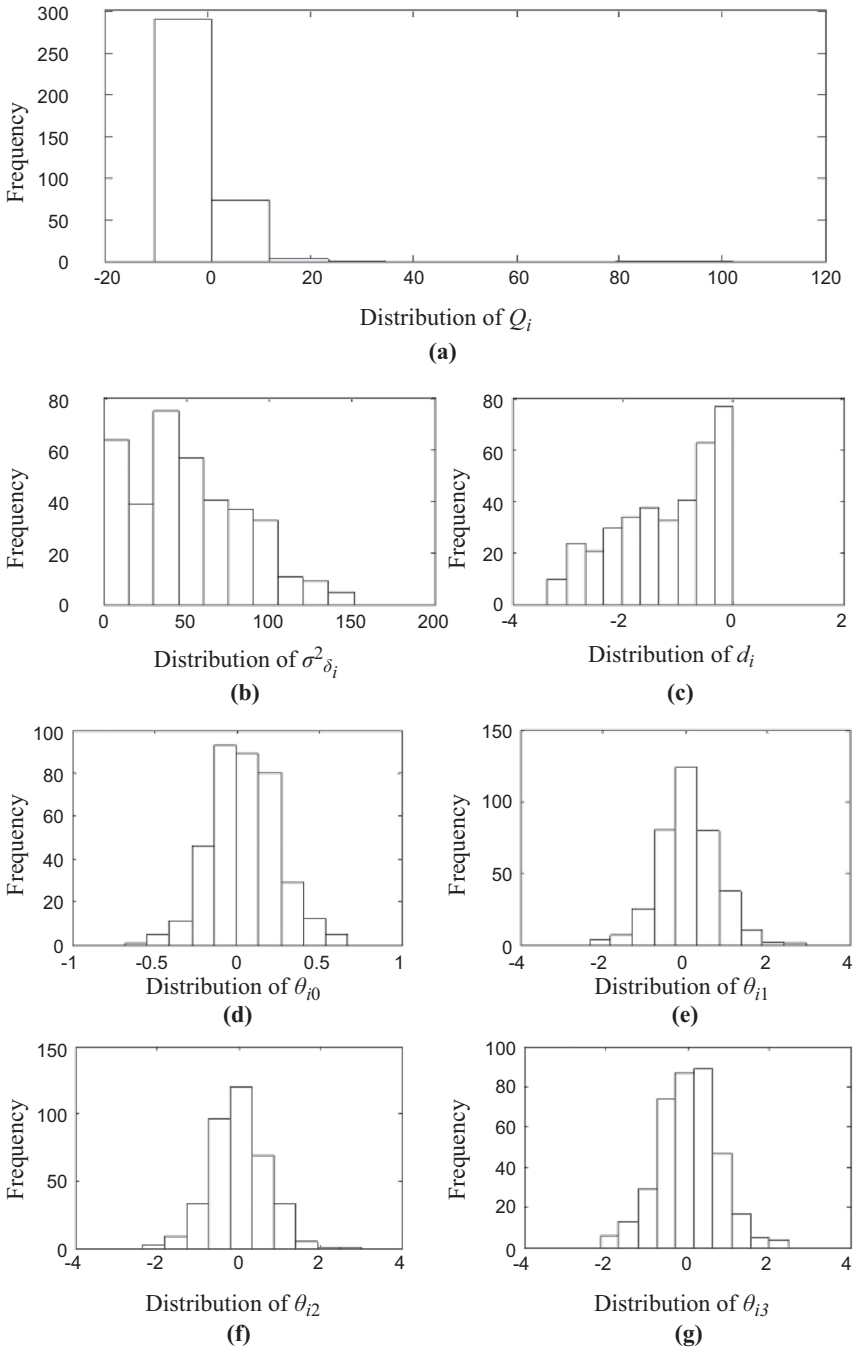
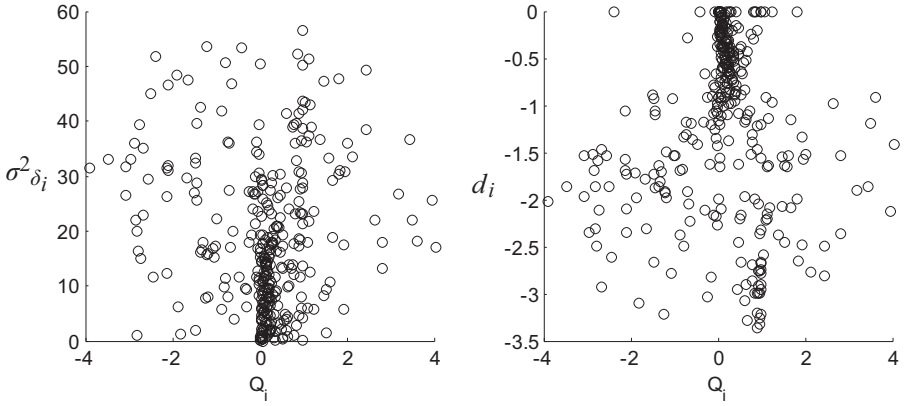


Figure 2.
Distributions of the individual-level parameters



Scatter plots of $\sigma^2_{\delta_i}$ against Q_i

(a)

Scatter plots of $\sigma^2_{\delta_i}$ against Q_i

(b)

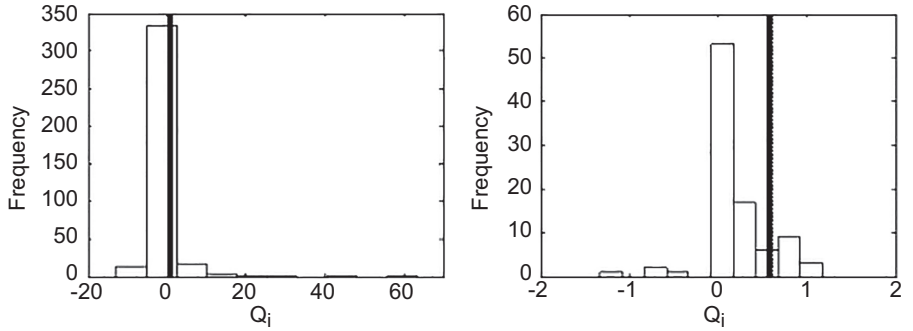
Figure 3.
Scatter plots of the selected individual-level parameters

From the analysis above, we know that although most individuals raise ideas of normal value, they are not as sensitive as those who usually post high-value or low-value ideas to the firm's feedback time. Thus, if the firm wants to collect more ideas of high value, it should promote the efficiency of responding to high-value ideas.

4.3 The filtering process of the crowdsourcing platform

Our estimates of the model parameters can explicitly represent the filtering process of idea contributors, especially marginal idea contributors, whose ideas are worse than the overall level. In the following analysis, we study the members who posted two or more ideas during the first and the last 20 months, respectively. Figure 4 (a) ~ (b) illuminate the distributions of average value in the two subgroups.

In order to elaborate the difference in abilities of posting high-value ideas between new members and early members, we present the relationship between the time when an individual first conducted their posting behavior and the average value of her ideas in Figure 5. Through the gathering state of data points, we observe that there exists a great



Individuals who posted two or more ideas in the first 20 months

(a)

Individuals who posted two or more ideas in the last 20 months

(b)

Figure 4.
Distributions of the average value of individuals in the two subgroups

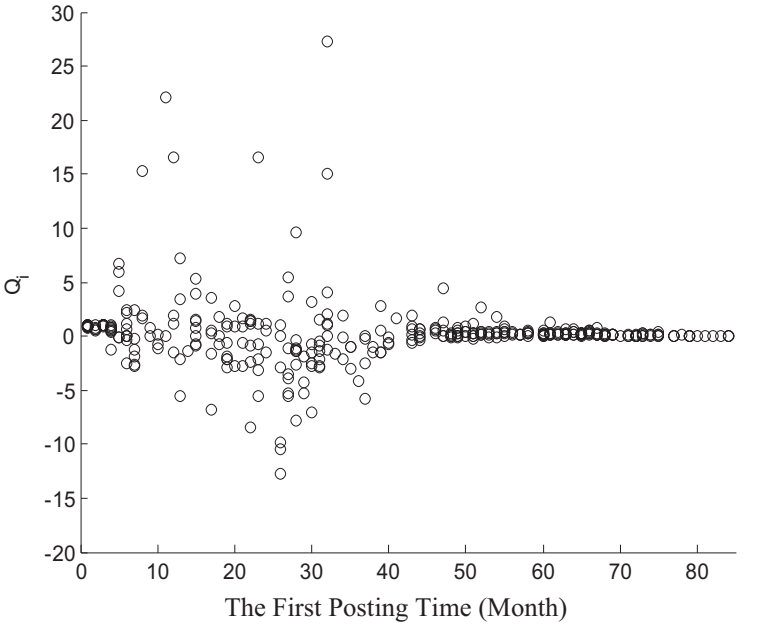


Figure 5.
Scatter plots of the selected individual-level parameters

difference in average value among individuals who posted the first idea during the first 40 months (we call them "early members"). In addition, these points scattered across the coordinate plane, and several individuals have posted ideas better than the average level. Moreover, members who posted their first idea during the last 50 months (we call them "new members") tend to submit ideas that are close to the overall mean value, and no obvious high-value proposer is found. In other words, good idea contributors who are also early members gradually become inactive, while new members cannot raise enough ideas that are as good as those of the previous contributors.

5. Conclusion

We modify an existing structural model to study customers' dynamic learning process using the actual data on MyStarbucksIdea.com. We research the efficiency of crowdsourcing initiatives in the background of customer-involved service innovation.

The results in our paper show that in the early stage of the website, members of the online community not only overestimate the value of their ideas, but also underestimate the cost for the firm to implement ideas. Therefore, members tend to be excessively optimistic and post a large number of new ideas with little value. Along with the learning process, individuals gradually realize the true value of ideas, and know about the firm's cost structure. For marginal idea contributors, the expectation of posting new ideas starts to drop. As a result, customers' learning process plays the role of self-selection, which filters out marginal idea contributors, leading to the decrease in the number of ideas after the website reaches a stable stage.

However, when the website reaches the stable stage (during the 30th and the 40th month in our model), the mean value of ideas and the cumulative feedback rate starts to fall. The reason is that only a few early members remain active; most individuals, including marginal contributors, gradually fade out. In addition, new members are not able to submit enough high-value ideas, so the vacancy for good contributors is not filled.

References

Archak, N. and Sundararajan, A. (2009), "Optimal design of crowdsourcing contests", *ICIS 2009 Proceedings Paper 200*, available at: http://aisel.aisnet.org/icis2009/200.

Bayus, B.L. (2013), "Crowdsourcing new product ideas over time: an analysis of the Dell IdeaStorm community", *Management Science*, Vol. 59 No. 1, pp. 226-244.

DiPalantino, D. and Vojnovic, M. (2009), "Crowdsourcing and all-pay auctions", *Proceedings of the 10th ACM Conference on Electronic Commerce*, *ACM*, pp. 119-128.

Di Gangi, P.M., Wasko, M.M. and Hooker, R.E. (2010), "Getting customers' ideas to work for you: learning from dell how to succeed with online user innovation communities", *MIS Quarterly Executive*, Vol. 9 No. 4, pp. 213-228.

Howe, J. (2006), "The rise of crowdsourcing", *Wired Magazine*, Vol. 14 No. 6, pp. 1-5.

Lu, Y., Singh, P.V. and Srinivasan, K. (2011), "How to retain smart customers in crowdsourcing efforts? A dynamic structural analysis of crowdsourcing customer support and ideation", *Conference of Information Systems and Technology*, Charlotte.

Mo, J., Zheng, Z. and Geng, X. (2011), "Winning crowdsourcing contests: a micro-structural analysis of multi-relational social network", *Workshop on Information Management (CSWIM)*.

Terwiesch, C. and Xu, Y. (2008), "Innovation contests, open innovation, and multiagent problem solving", *Management Science*, Vol. 54 No. 9, pp. 1529-1543.

Yan, H., Param, V.S. and Kanan, S. (2014), "Crowdsourcing new product ideas under consumer learning", *Management Science*, Vol. 60 No. 9, pp. 2138-2159.

Corresponding author
Lixin Cui can be contacted at: cuilixin@bit.edu.cn