Crowdsourcing New Product Ideas under Consumer Learning

Yan Huang

Ross School of Business

University of Michigan

yphuang@umich.edu

Param Vir Singh

Tepper School of Business

Carnegie Mellon Unversity

psidhu@cmu.edu

Kannan Srinivasan

Tepper School of Business

Carnegie Mellon University

kannans@andrew.cmu.edu

Abstract

Crowdsourcing initiatives are becoming a popular tool for new idea generation for firms. On crowdsourced ideation initiatives, individuals contribute new product ideas and vote on other's ideas which they would like the firm to implement. The firm then decides which ideas to implement. Although such initiatives are widely adopted in many different industries, they face increasing criticism as the number of ideas generated often decline over time, and the implementation rates (percentage of posted ideas that are implemented by the firm) are quite low raising concerns about their success.

We propose a dynamic structural model that illuminates the economic mechanisms shaping individual behavior and outcomes on such initiatives. Through counterfactuals, we identify the impact of several potential interventions on the success of such initiatives. We estimate the model using a rich dataset obtained from IdeaStorm.com, which is a crowdsourced ideation initiative affiliated with Dell. We find that, on Ideastorm.com, individuals tend to significantly underestimate the costs to the firm for implementing their ideas but overestimate the potential of their ideas in the initial stages of the crowdsourcing process. Therefore, the "idea market" is initially overcrowded with ideas that are less likely to be implemented. However, individuals learn about both their abilities to come up with high potential ideas as well as the cost structure of the firm from peer voting on their ideas and the firm's response to contributed ideas. We find that individuals learn rather quickly about their abilities to come up with high potential ideas, but the learning regarding the firm's cost structure is quite slow. Contributors of low potential ideas eventually become inactive, while the high potential idea contributors remain active. As a result, over time, the average potential of generated ideas increases, while the number of ideas contributed decreases. Hence, the decrease in the number of ideas generated represents market efficiency through self-selection rather than its failure. Through counterfactuals, we show that providing more precise cost signals to individuals can accelerate the filtering process. Increasing the total number of ideas to respond to and improving the response speed will lead to more idea contributions. However, failure to distinguish between high and low potential ideas and between high and low ability idea generators lead to the overall potential of the ideas generated to drop significantly.

1. Introduction

Product innovation has been an important area of business academic research. Recent advances in information technology have allowed firms to enhance their direct communication with customers, and the interaction has become an interesting source of new product ideas. Leveraging the opportunity, firms now create online idea markets where consumers can post new product ideas that are voted on by their peers. The aggregate voting score provides an indication of the potential revenue an idea can generate (hereafter, market potential). Jeff Howe (2006) named this new approach crowdsourcing, and he defined crowd as "the new pool of cheap labor: everyday people using their spare cycles to create content, solve problems, even do corporate R & D." Crowdsourcing initiatives provide individuals with a platform to express their ideas, which are typically generated from their experience with actual product usage or observing others using the product. The ideas that come from the customer crowds can reveal rich information about customers' preferences. Typical crowdsourcing platforms allow other customers to promote or demote ideas of their peers, thus providing an important early assessment of the potential of the proposed ideas. Firms can potentially obtain a large number of novel and profitable ideas at relatively low costs from such initiatives. Early adopters of this approach include some of the highly regarded business firms, such as Dell, Bestbuy, Starbucks, Nokia, Salesforce, BBC, CNN, BMW, Sears and Adobe.

Although crowdsourcing initiatives have become rapidly popular in a variety of industries, the usefulness of this new approach is still under debate. On many crowdsourced ideation platforms, the number of ideas generated decline over time, and the implementation rates (percentage of posted ideas that are implemented by the firm) are quite low.[1] Critics of such initiatives raise several concerns. First, they argue that the individuals might be too accustomed to current consumption conditions and their own specific needs and hence, are more likely to suggest ideas with little market potential (Hill 2009). Second, unlike the internal R&D teams, customers of the firm are unaware of the internal cost structure of the firm and hence, are quite likely to suggest ideas that are not viable (Schulze and Hoegl 2008). As a result, the firm typically has to invest significant effort to screen ideas, most of which have low potential and are generally infeasible. Third, individuals are often disconcerted[2] by the firm's slow or no response to their ideas and eventually stop contributing ideas. The low implementation rate of ideas and the decline in the number of ideas posted as

---

[1] Please refer to Figure 1 that shows the number of ideas contributed by consumers to Dell Ideastorm.com. Similar decrease in number of contributed ideas over time is observed for other crowdsourced ideation initiatives by Starbucks, Bestbuy, and Giffgaff. The idea implementation rates are close to 2% across these initiatives.

[2] Individuals complain that the firm ignores their ideas; thus, they are disappointed and feel that it is a waste of time to post an idea. One individual wrote in a comment, "You're also right, Tukulito (another individual's ID), that Dell has NOT responded in so MANY areas. It's been extremely frustrating." Another individual said, "Many individuals have lost interest in IdeaStorm lately because IdeaStorm, the way it stands now is, frankly, stagnant… I'm sure many individuals have lost interest in IdeaStorm in part because they're led to believe that their ideas are disregarded/ignored now… And it's not just like Dell doesn't implement any ideas now. I don't think Dell has even commented or updated many ideas lately, even the most popular or most requested ideas…"

observed in practice seem to be consistent with the arguments against crowdsourcing. If this were in fact true identifying appropriate interventions for crowdsourced ideation initiatives becomes very important. However, there is no systematic research that has investigated these issues in depth.

We provide an alternate argument that could potentially explain the decrease in number of ideas contributed over time and the low implementation rates. We argue that consumer learning and heterogeneity can explain these trends, and that such trends may in fact be a signal of market efficiency through self-selection rather than of failure. We argue that while a large number of consumers may be able to contribute only ideas of low market potential, substantial number of consumers may be able to suggest high potential ideas. Further, the consumers may not know the implementation cost for the firm or even the potential of their own ideas but they could also learn about them over time through peer feedback. For example, enthusiastic consumers may propose new product ideas, but they have no initial idea as to how good their ideas are and may simply overestimate the potential of their ideas. Peer evaluations provide a valuable and important source of real-time feedback. A strong negative vote will let the consumer know that the idea may not be that useful after all. When a string of new product ideas are turned down by peer consumers, the individual may conclude that, contrary to his/her initial belief, he/she is not a sophisticated generator of new product ideas. Thus, through learning, those customers who are 'bad' at coming up with high potential ideas (marginal idea contributors) recognize their inabilities and may reduce the number of ideas they propose over time or become inactive. In contrast, 'good' new product idea generators (good idea contributors)[3] will be encouraged to continue to provide new product ideas. Such a learning model is entirely consistent with an overall decline in number of and increase in average quality of new product ideas over a period of time observed in the dataset. Thus, a decreasing number of ideas may well reflect an efficient idea market and its resulting success rather than the ineffectiveness of the idea market.

Another important impediment to the implementation of new product idea is the cost of implementing the idea. Unfortunately, as critics argue, consumers have little or no understanding of this critical factor. However, consumers can learn or can infer the cost to implement ideas. Consumers cannot infer a firm's cost structure from unimplemented ideas because firms have not made decisions on those ideas. Nevertheless, when an idea is implemented, firms usually publicize their implementation decision and provide details about how they implement it. This information is broadcasted to all individuals in the community and by combining that information with the idea's voting score, consumers can learn how costly it is for the firm to implement similar kinds of ideas. Such sophisticated learning by consumers eventually results in the

---

[3] In our model, an individual's type is determined by the average potential of ideas generated by this person. The individual's type is continuous because average potential is a continuous variable. When we say an individual is a "good idea contributor," it means that the average potential of ideas generated by the individual falls in a higher region of the distribution. When we say an individual is a "marginal idea-contributor," it means that the average potential of ideas generated by the individual falls in the lower region of the distribution.

2

generation of ideas where cost will not be an impediment for eventual implementation. We propose and show that such a learning mechanism finds strong empirical support.

In this study, we illuminate the economic mechanisms that shape individual behavior and outcomes on crowdsourced ideation initiatives, and suggest and identify the impact of several potential interventions that could improve the efficiency and success of such initiatives. We build a structural model for crowdsourced ideation initiatives to explain contributor behavior and apply it to a rich dataset collected from IdeaStorm.com, a crowdsourced ideation initiative affiliated with Dell. We answer a number of questions: (1) Can contributors learn about the potential of their ideas and the cost for the firm to implement their ideas over time even if they do not know it initially? (2) Are individuals differ in their abilities to come up with high potential ideas? (3) How would learning about potential of ideas and cost of implementation, and individual heterogeneity shape individual behavior and affect outcomes on such initiatives? Is the downward trend in the number of ideas contributed really a sign of failure for such initiatives? (4) What policy interventions can affect the success of such initiatives? Notice that the "learning" mentioned in the first research question means "Learning about the true value (individuals realize their true ability to come up with ideas of high potential)", not "learning-by-doing (the potential of ideas generated by each individual improves as the cumulative idea he/she posts increases)". It is possible that both types of "learning" exist in our context. However, as we will discuss later in the "Data and Model Free Evidence" section, there is no evidence in our data that supports "learning-by-doing". Therefore, in our structural model, we only include the first type of learning -- "Learning about the true value". Our results show that initially contributors tend to underestimate the costs for implementing their ideas and overestimate the potential of their ideas. Therefore, marginal idea contributors initially tend to post many low potential, unviable ideas. However, as individuals learn (update their beliefs) about the firm's cost structure and the potential of their ideas, marginal idea contributors gradually become less active in generation of new ideas. A smaller fraction learns that they are good idea contributors. Consequently, although the number of ideas generated decreases over time, the average potential of ideas posted significantly increases over time. These findings show that, over time, marginal idea contributors are filtered out and that the idea market becomes more efficient. The estimation results also show that individuals learn about their own ability to come up with high potential ideas faster than they learn about the cost structure of the firm because the cost signals the firm provides are quite imprecise. We also find that individuals feel discouraged to contribute ideas if the firm does not reply to their submissions or takes an extended period of time to reply.

Our policy simulations evaluate several policy interventions and the results have important implications about how the firms can improve the performance of their crowdsourced ideation platforms. We show that Dell can accelerate the filtering out of marginal idea contributors by providing more precise cost signals. In addition, actively responding to all unimplemented ideas will adversely affect the filtering process because marginal idea-contributors who would become inactive under the current policy will stay active

3

longer under the new policy. As a result, the firm would end up with more low potential ideas. In other words, the firm is better off when it selectively responds to ideas. Providing feedback on ideas with higher votes can improve the average idea potential in the later periods; however, the improvement is insignificant. The best policy is to identify good idea contributors and respond quickly to their ideas. By doing so, good idea contributors will be less disincentivized and will be encouraged to contribute more high potential ideas. Our last set of policy simulations show that if the firm wants to provide additional incentive for consumers to contribute ideas, it should reward individuals only when their ideas are implemented, rather than reward individuals when they post ideas. By doing so, the firm can achieve the same improvement on the overall potential of ideas at a lower cost.

2. Relevant Literature

Our paper is related to the emerging literature on crowdsourcing. Although crowdsourcing has attracted enormous business and media attention, there are very few academic studies on crowdsourcing. Initiatives by established firms to encourage customers for participation in the design of new products represents the most popular form of crowdsourcing being currently used and studied (Terwiesch and Xu 2008). Such crowdsourcing initiatives soliciting new product design ideas can be classified into three types. In the first type, the creation of a vaguely specified product depends wholly on customer input. Threadless.com is an example of such an initiative where customers develop t-shirt designs on their own and submit the finished designs to Threadless. The second type of crowdsourcing is related to the first type in that the final product depends wholly on the customer input but differs from the first type in that the customers have to solve a specifically defined task or problem (Boudreau et al 2011, Jeppesen et al 2010). Crowdsourcing efforts at Topcoder or Innocentive correspond to this type. The first two types are also similar to each other in that in both of them the contributors typically compete with each other for a fixed monetary reward. Hence they are also classified as crowdsourcing contests. The third type of crowdsourcing corresponds to a permanent open call for contribution that is not directed towards any particular task or problem (Bayus 2010, Di Gangi et al. 2010). Dell Ideastorm corresponds to this type. In this type of crowdsourcing consumers typically only contribute and evaluate variety of ideas and it is up to the firm to develop and implement those ideas.

Most of the studies on crowdsourcing have analyzed crowdsourcing contests where contributors compete with each other to win a prize (Archak and Sundararajan 2009, DiPalantino and Vojnovic 2009, Mo et al. 2011, Terwiesch and Xu 2008). In contrast to crowdsourcing contests, in permanent open call crowdsourced ideation initiatives such as IdeaStorm, contributors do not compete with each other but help evaluate each other's contributed ideas. There are only a few studies on this type of crowdsoured ideation initiatives. Using a reduced form approach, Bayus (2010) finds that individual creativity is positively correlated to current effort but negatively related to past success. Di Gangi et al. (2010) find that the decision to adopt a user contributed idea is affected by the ability of the firm to understand the technical requirements and

4

respond to community concerns regarding the idea. Lu et al. (2011) find important complementarities in crowdsourced ideation and customer support initiatives. They find that customer support platforms provide opportunities for customers to learn about the problems other customers are facing and that helps them in suggesting better ideas for firm to implement. To our knowledge, we are the first to structurally examine the new product idea and development process based on actual crowdsourcing data.

Our paper is also related to the literature on consumer Bayesian learning. Bayesian learning models are widely applied to analyze consumers' choices under uncertainty[4]. Erdem and Keane (1996) and Erdem et al. (2008) investigate customer learning of brand qualities from multiple resources, such as past experience, advertisement, and price. While Mehta et al. (2003) study the formation of consideration sets, Crawford and Shum (2005) and Narayanan and Manchanda (2009) examine the physicians' learning of drug prescription. Zhang (2010) develops a dynamic model of observational learning and analyzes the kidney adoption in the U.S. kidney market. In our paper, we apply the Bayesian learning model to the individual's learning of the potential of their ideas and learning of the firms' cost structure to better understand the dynamics of idea posting behavior.

3. Research Context

Our data are from a crowdsourcing website, IdeaStorm.com, which is operated by Dell. Dell launched this website in February 2007. The goal of this initiative was to hear what new products or services Dell's customers would like to see Dell develop.

The structure of Ideastrom.com is quite simple, yet effective. Any individual (not necessarily a customer) can register on the website to participate in the initiative. Once registered, an individual can then post any relevant idea. Dell assigns 500 Dell points to the contributor for each idea[5]. Once an idea is posted, all the other individuals can vote on the idea. They can either promote the idea, which yields an additional ten points for the idea contributor, or demote the idea, which results in a ten point deduction. In the data, however, we as well as the individuals only observe the aggregate score, but not the number of "promotions" or the number of "demotions". Individuals are also allowed to comment on ideas and express their opinions in greater detail. However, in this paper, we only model individuals' submission decision. Dell uses the peer voting scores to gauge the potential of contributed ideas. Dell assigns web managers to maintain the website, and their job is to pass the ideas generated by the individuals on to the corresponding groups within the company for review. The web managers communicate with the individuals through direct comments about the ideas and changes in the status of the idea. Typically, the evolution of an idea's status is as follows.

---

[4] In this study, "learning" refers to the Bayesian updating process through which individuals update their beliefs about their own type and the firm's cost structure.

[5] This policy was changed in December 2008.

Most of the posted ideas posted are "Acknowledged" within 48 hours. If the web managers find an idea is already part of their existing product or services, they will change the status to "Already offered". Among the remaining ideas, the web managers selectively pass ideas to related departments for review, and the status is changed to "Under Review". After carefully evaluating these ideas, Dell makes one of three decisions: "Implemented", "Partially Implemented" or "Not Planned". Once an idea is "Implemented", it is closed for votes and comments. Dell also provides details regarding the decision through comments or blog posts. "Partially Implemented" and "Not Planned" ideas are not closed, which means that individuals can still vote and comment on these ideas, and it is possible that at some point, Dell will re-evaluate the ideas. Ideas that do not receive any comments within a year are "Archived" and thus no longer available for individuals to view (IdeaStorm.com). All individuals can see ideas' aggregate voting scores and which ideas have been implemented by the firm. In this way, our modeling framework allows the individuals to learn from these two observations. Dell categorizes all the ideas into three categories: Product Ideas, Dell Ideas, and Topic Ideas. When an individual posts an idea on IdeaStorm, he/she selects the Category to which the idea belongs.

Another point worth mentioning is that Ideastorm.com is a "non-competitive" (as oppose to crowdsourcing contests) and "non-collaborative" (as oppose to collaborative crowdsourcing platform such as Wikipedia) platform. On Ideastorm, individuals generate ideas independently and there is little collaboration among users when generating new ideas. There is no monetary award associated with the implementation of individuals' ideas either. Individuals submit ideas to express their consumption needs and if the firm adopts their ideas, their needs are satisfied. Users on this platform tend to express their own preferences, and have little incentive to conform to the majority's preferences. The goal of using this type of crowdsourced ideation platform is to obtain diverse ideas from outside individuals, filter these ideas using the voting system, and then implement good ideas and generate profit.

4. Data and Model Free Evidence

Our data have expanded from the initiation of IdeaStorm.com in early 2007 to the end of 2010. By the end of 2010, more than 12,000 ideas had been contributed and more than 400 had been implemented. However, we only use the data from the initiation of IdeaStorm.com to September 2008. During October 2008, a large number of material changes were made to the initiative, and therefore, we restrict our attention to data prior to these changes. We also exclude data from the first two weeks because the number of ideas contributed during this period was extremely small (≤ 5), perhaps due to public's lack of awareness of the website. Furthermore, most of the initial ideas during this period were announcements made by Dell's employees. After the elimination of the initial period, we have 84 weeks of data (Week 3 to Week 86). In our dataset, most of the ideas fall into the first two categories. There are very few ideas that belong to Category 3 (less than 10% of the number of ideas in Categories 1 and 2, see Table 1), with even fewer implemented Category

3 ideas—only 3. This makes it almost impossible to make inferences about Category 3 ideas. Therefore, our analysis focuses only on the first two categories of ideas.

A majority of individuals on the website only vote but never post any new product idea. In addition, among those who posted an idea, most posted only one idea during these 84 weeks. The notion of learning is meaningful only when a respondent posts at least two ideas. The 490 individuals who posted two or more ideas constitute fewer than 5% of the consumers on the site but account for nearly 40% of all new product ideas. Table 2 shows the important statistics of these individuals who posted two or more ideas. We observe that there is significant variation among individuals in terms of mean of the log of the voting score received by an idea (log(votes)), number of ideas generated, and first time posting.

Table 1. Summary Statistics by Category

Category	1	2	3
Category Name	Product idea	Dell idea	Topic idea
# Posted	5337 (1419)*	4243(1565)	392(108)
# Implemented	100(41)	110(54)	10(3)
% Implemented	1.87 (2.89)	2.59(3.45)	2.55(2.78)
Average log (votes)	4.626(5.286)	4.580(5.600)	4.352(4.556)
SD of log (votes)	2.160 (1.875)	2.147(1.696)	2.720(2.742)

*Numbers outside the parentheses are full sample statistics; numbers inside the parentheses are statistics of the sample of 490 selected individuals

Table 2. Summary Statistics for Individuals

Variables	Mean	Std. dev	Min	Max
Mean Log(votes)	4.819	1.513	-4.000	7.667
Number of Ideas Contributed	7.269	19.411	2	164
First Time Post (Week)	22.41	21.76	3	83

The dynamics of individual participation on the crowdsourcing website are shown in Figures 1-3. In these figures, we focus on the selected 490 individuals. From Figure 1, it is evident that the number of the ideas posted early on was very high; however, the number declined quickly over time and then stabilized. If we look at the implementation rates of different categories of ideas (Figure 2), we note that the implementation rates of both Category 1 and Category 2 ideas increase over time. In Figure 3, we note that despite some random disturbance, the weekly average log votes tend to increase over time. The data patterns shown in Figures 2 and 3 suggest that although the number of ideas generated decreases over time, the quality/potential of the ideas seems to increase. This suggests that the downward trend in the number of contributed idea may not be bad for the firm, and there may be more complicated underlying process that leads to the decline in the number of submissions and the increase in the ideas' overall potential.
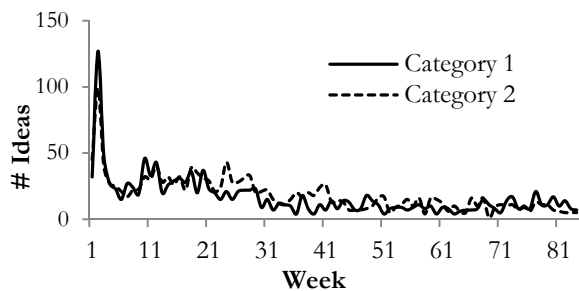
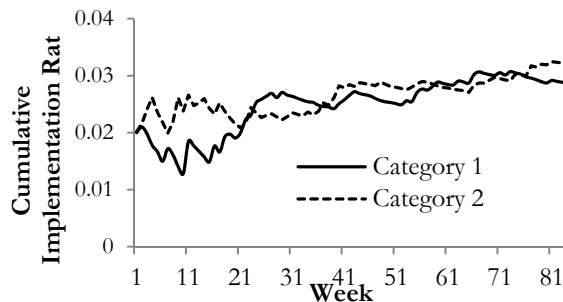**Figure 1. Numbers of Ideas Contributed in Each Week**

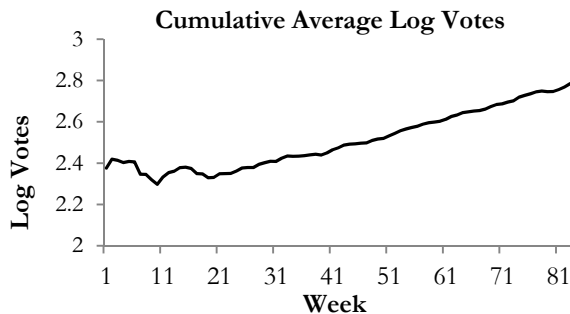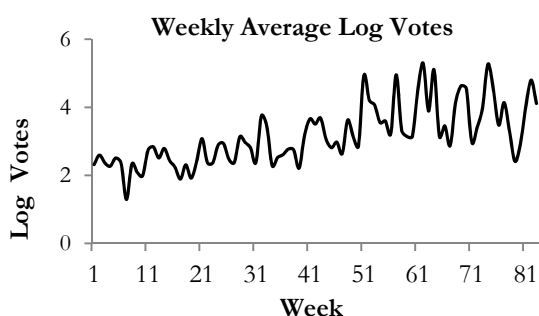**Figure 2. Cumulative Implementation Rate**

Figure 3. Weekly Average Log Votes and Cumulative Average Log Votes per Submission

One potential explanation of the increase in the overall potential of the ideas over time is a result of learning curve effect. That is, as individuals gain experience on the platform they will become more capable in generating good ideas. As voting score is used by the firm to gauge the potential of the ideas, we use average voting score as a proxy of individuals' ability to generate good ideas. We run a reduced-form regression to test whether individuals' ability to generate good ideas (as measured by voting scores) improves with past posting experience. The results are reported in Table 3. The notation for variables presented in the Table 3 is as follows. $logvote_{ijt}$ represents the log of the voting score received by the category j idea that individual i posted in period t. $\#Pastideas_{ijt}$ is the number of ideas posted by individual i till time t in category j. Note that the relationship between $logvotes_{ijt}$ and $\#Pastideas_{ijt}$ would illustrate the presence or absence of learning curve. $\#Pastideas\text{-}_{ijt}$ is the number of ideas posted by peers of individual i till time t in category j. $\#Pastideas\text{-}_{ijt}$ captures any spillover effect where individuals may learn from their peer's ideas. The test results suggest that after we control for individuals' ability (captured by the individual fixed effects in the regressions), the effect of $\#Pastideas_{ijt}$ and $\#Pastideas\text{-}_{ijt}$ is statistically insignificant on the $logvotes_{ijt}$. Therefore there is no evidence that the individuals' improve in their ability to come up with high quality ideas with experience in this setting[6]. As discussed earlier, there is no monetary award associated with the implementation of individuals' ideas either and so users on this platform have little incentive to conform to the majority's preferences. Users submit

---

[6] The results regarding non significance of classical learning curve effect in the data are robust to alternate specifications which model non-linear relationship of $\#Pastideas_{ijt}$ or $\#Pastideas\text{-}_{ijt}$ with logvotes.

ideas to express their consumption needs and will do so only when they ideas have decent chances to be implemented.

Table 3. Regression for Learning Curve Effect
D.V. *logvotes$_{ijt}$*

Variable	Coefficient (Standard error)	Coefficient (Standard error)
	Product Ideas	Service Ideas
Intercept	5.608 (1.523)***	5.081 (2.378)*
$\#Pastideas_{ijt}$	0.014 (0.014)	0.011 (0.008)
$\#Pastideas_{-ijt}/1000$	0.028 (0.022)	-0.146 (0.856)
Individual Fixed Effects	Yes	Yes

(Clustered standard errors within the same individual). Significance levels: '***'< 0.001 '**' <0.01 '*' <0.05 '.' <0.1

As shown in the above reduced-form analysis, the increase in overall quality of ideas at aggregate level overtime as shown in Figure 3 is not because individuals improve in their abilities to come up with high potential ideas. Another possible explanation of the data pattern is the dropping out of low potential idea contributor. To see whether this explanation is evident in the data, we explore the decomposition of users on IdeaStorm.com. We divide the study period (84 weeks in total) evenly into 4 chunks, each of which contains about 21 weeks, so that we can compare individuals' idea submission behavior over time. We divide all users based on their ideas' average log-vote. Average log-vote is the average of logarithm of the voting score of ideas posted by the same individual in the 84 weeks. We then divide the 490 individuals evenly using the 3rd quartile, median and 1st quartile of the average log-votes and define the four groups as "Highest average score", "Second highest average score", "Second lowest average score", and the "Lowest Average Score", which correspond to 4 levels of descending ability. Figure 4 shows that overtime, varying fractions of users in all the four groups become inactive in terms of idea contribution. More importantly, the numbers of dropouts (individuals who become inactive) are much higher in "Second lowest average score" and "Lowest Average Score" groups (marginal idea contributors) than users in the other two groups (good idea contributors). More specifically, during the last 20 weeks, only less than 1/5 of the users in "Lowest Average Score" group remain active, while about half of the top two groups of users remain active posting new ideas. Overtime, most marginal idea contributors drop out while many good idea contributors remain active, and thus users in the top two groups account for higher fraction of users who stay contributing ideas on the website. At the aggregate level, the number of submissions decreases (because marginal idea contributors post fewer ideas over time), while average voting score of the submissions increases (because most ideas posted in the later periods are contributed by good idea contributors).

9
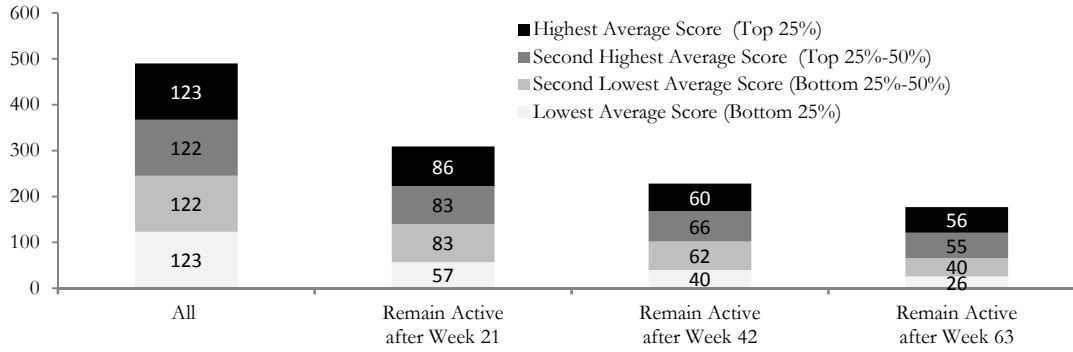
Figure 4. Composition of Users Who Remain Active

5. Model

In this section, we develop a structural model to understand the dynamics of the participation behavior of individuals. The objective in a structural model is to explain the data generation process through the explicit modeling of individual decision making process (utility function) and then uses data to empirically recover the parameters in the analytical model. As we explicitly model individuals' decision making process, we are able to run policy simulations to see how individuals' behavior will change as a policy changes (Lucas 1976). The proposed structural model incorporates learning about a firm's cost structure and the potential of own ideas. In our model, in each period an individual makes a decision whether to post an idea in a category or not. This decision is governed by the expected utility she can derive from posting the idea. Hence, we begin by first explaining the utility function of the individual.

5.1 Utility Function

There are four key components of the utility function. The first two components accounts for the benefits a user may derive from contributing ideas to the initiative. There are several reasons as to why a user may contribute an idea for the firm to implement. The first component accounts for the utility the user may derive from better performance from the improved product if her idea is implemented (Franke and von Hippel, 2003, Kuan 2001, Lakhani and von Hippel 2003). Online communities such as crowdsourced ideation initiatives provide social reputation related utility. Social reputation in online communities is considered an important extrinsic motivation because of its instrumental value in enhancing contributors' job prospects (Lakhani & Wolf, 2005). This constitutes the second component of the user utility function. Dell facilitates the social reputation mechanism by assigning 500 Dell points to a contributor for each contribution. These points are shown in the individual's profile, but they cannot be cashed in or used for discounts and thus have no monetary value. However, online reputations may translate into a number of benefits, including job offers by established companies (Kumar et al. 2011, Huang et al 2010).

In contrast to the benefits individuals derive from posting an idea, every time they post an idea they may also incur cognitive or hassle cost of coming up with, articulating, and posting the idea. This cost

10

constitutes the third component of user utility function. Our fourth component accounts for the user discontent that occurs when the firm does not respond to their posted ideas. As argued earlier, if the firm does not respond to the consumer's input the consumer may potentially get dissatisfied with the firm leading to negative effect on her utility from participation. We capture this effect through an individual-specific variable "no response" (D_{it}), which is a binary variable that equals 1 as long as there is one idea posted by an individual that has not moved to any status other than "Acknowledged" 12 weeks after it was originally posted. We chose 12 weeks as the criterion because the vast majority of the ideas that eventually moved to the next level in our dataset received the first status change within twelve weeks. If one sees that his/her ideas remain at "Acknowledged" status for more than 12 weeks, he/she may assume that this idea has little chance to be seriously reviewed, and the firm will likely not provide any feedback on the idea.[7] The effect of "no response" is denoted as d_i. We allow d_i to be different across individuals, i.e., some individuals could feel extremely disincentivized under situations where $D_{it} = 1$, while others may not share that feeling.

Besides these four components, an individual's utility function may also include factors unobserved to us. Hence, our utility function incorporates these four components as well as a random unobserved component to account for factors unobserved to us. Specifically, the utility individual i derives from posting an idea is given by the following equation

$$U_{ijt} = \begin{cases} c_i + r_i + d_i D_{it} + \theta_{ij} + \varepsilon_{ijt} & \textit{if the idea is implemented} \\ c_i + r_i + d_i D_{it} + \varepsilon_{ijt} & \textit{if the idea is not implemented} \end{cases}$$

where j represents the idea category. We adopt the classification on the website and set idea categories as Product ideas (Category 1) and Dell ideas (Category 2). The parameter, c_i, represents the cost incurred by individual i when he/she posts an idea, and r_i is the reputation gain the individual derives from the 500 IdeaStorm points. The parameter, θ_{ij}, measures individual i's utility gain from the implementation of his/her Category j idea. D_{it} represents the firm's lack of response to consumer i's ideas, and d_i denotes the extent to which such lack of response adds to individual i's cost to post an idea or how it harms the utility the individual receives from posting an idea. The error term, ε_{ijt}, captures the individual choice specific random shock in period t.

It is obvious that we cannot identify c_i and r_i simultaneously because they enter linearly in the utility function. Therefore, we combine these two terms and define $\theta_{i0} = c_i + r_i$. Thus, the individual's utility function reduces to

$$U_{ijt} = \begin{cases} \theta_{i0} + d_i D_{it} + \theta_{ij} + \varepsilon_{ijt} & \textit{if the idea is implemented} \\ \theta_{i0} + d_i D_{it} + \varepsilon_{ijt} & \textit{if the idea is not implemented} \end{cases} \tag{1}$$

---

[7] The selection of the cutoff point is subjective. We also use other time points as cutoff points, the nature of the estimation results remains unchanged with different cutoff points, and only the magnitude of the estimates slightly changes. We defer these details to the section where we discuss the robustness of our findings.

where θ_{i0} is individual specific. In each period, individuals make decisions on whether or not to post ideas in a category. Before they post their ideas, they do not know whether their idea will be implemented. However, they form an expectation on the probability of their idea being implemented. Let $E(U_{ijt}|Info(t))$ denote the expected utility individual i can obtain from posting Category j idea in period t, conditional on the information individual i has up to period t. $E(U_{ijt}|Info(t))$ then can be expressed as

$$E\left(U_{ijt}\middle|Info(i,t)\right) = \tilde{U}_{ijt} + \varepsilon_{ijt} = \theta_{i0} + d_i D_{it} + \theta_{ij} P_{ijt}|Info(i,t) + \varepsilon_{ijt} \tag{2}$$

where $P_{ijt}|Info(i,t)$ represents the perceived conditional probability of implementation.

5.2 Individual's Learning Process

Idea contribution decisions of individuals are based on their beliefs of the probability of implementation $(P_{ijt}|Info(t))$. The probability of implementation of an idea is a function of its potential and cost of implementation. The firm's decision rule for implementing ideas is explained in detail later. An individual has beliefs about the implementation cost as well as the potential of her own ideas. At the time of posting, the user uses her beliefs about potential and cost of implementation to calculate the probability of her idea's implementation which she uses to guide her posting decision. Over time, new information comes into the system. This information provides signals regarding the implementation costs or the potential of an idea. The individuals use this information to update their beliefs about the implementation cost and potential of their ideas and use these updated beliefs to guide their future contribution decisions. We model the belief update to happen in a Bayesian manner (DeGroot 1970). We explain the learning process in detail below. The first type of learning is learning by individuals about the firm's cost structure, and the second type of learning is learning the potential of one's own ideas.

Learning about the Firm's Cost Structure[8]

Suppose that implementation cost of ideas in Category j follows a normal distribution with mean C_j and variance $\sigma_{\gamma_j}^2$[9]. Note that the firm exactly knows its cost of implementation for each idea. However, the consumers may be uncertain about the cost of implementation of their ideas. At the moment when the website is launched, an individual's prior belief of the firm's average cost of implementing a Category j idea, denoted as C_{j0}, is

$$C_{j0} \sim N(C_0, \sigma_{C_0}^2) \tag{3}$$

---

[8] By "cost structure" or "cost of implementation" we imply "implementation feasibility".

[9] We assume that individuals only update the mean of the cost distribution, but not the variance. This is a standard assumption in the Bayesian learning literature. The logic behind this assumption is that as the users see the implementation rates of ideas at different voting scores, they will be able to gauge the variance of the actual implementation cost.

In Equation (3), C_0 is the prior mean and $\sigma_{C_0}^2$ is the prior variance. If individuals underestimate (overestimate) the implementation cost the data would reveal that $C_0 < C_j$ ($C_0 > C_j$). The prior variance, $\sigma_{C_0}^2$, captures the uncertainty that the individual has about the mean cost of implementation.

The event that brings in new information into the system regarding the implementation cost is the implementation of an idea. Whenever one idea (either posted by the 490 individuals in our sample or individuals outside the sample) is implemented, all individuals receive a common signal about the cost the firm incurs. This learning process is common across individuals as all of them are exposed to the same information. This is because when an idea is implemented, the firm posts an article on its official blog site describing how the firm is implementing the idea, which contains information about the firm's implementation cost. Everyone receives this information. Further, when an idea is implemented, it is closed for further voting. And so the final voting score of this idea can provide consumers with the lower bound of the cost of implementing this idea.

C_{kjt} in Equation (4) denotes the cost signal all individuals receive when one Category j idea is implemented in period t. The difference between each specific cost signal and the mean implementation cost of ideas in the same Category is captured through the parameter μ_{kjt}, which is a zero mean normally distributed random variable, and its variance, σ_μ^2, measures the variance of the implementation cost signals of ideas within the same category. This implies that while the signal is unbiased, it could be noisy. The parameter, σ_μ^2, captures the extent of noise in the signal.

$$C_{kjt} = C_j + \mu_{kjt} \tag{4}$$
$$\mu_{kjt} \sim N(0, \sigma_\mu^2).$$

In each period there could be more than one idea implemented leading to more than one signal. If there are k_{Cjt} Category j ideas implemented in period t, then the aggregate signal that individuals receive from these multiple implementations is C_{sjt}. C_{sjt} is simply the average of the k_{Cjt} signals ($C_{1jt,\dots,}C_{k_{Cjt}jt}$), and it has the following distribution

$$C_{sjt} \sim N(C_j, \frac{\sigma_\mu^2}{k_{Cjt}}) \tag{5}$$

Let C_{jt-1}^e denote individual's belief of mean of Category j idea's implementation cost in the beginning of period t. By definition, conditional on the cumulative information he/she has received by the beginning of period t, individuals update C_{jt}^e using the following Bayesian rule (DeGroot, 1970)

$$C_{jt}^e = C_{jt-1}^e + (C_{sjt} - C_{jt-1}^e)\frac{\sigma_{C_{jt-1}}^2}{\sigma_{C_{jt-1}}^2 + \frac{\sigma_\mu^2}{k_{cjt}}} \tag{6}$$

$$\sigma_{C_{jt}}^2 = \frac{1}{\frac{1}{\sigma_{C_{jt-1}}^2} + \frac{k_{Cjt}}{\sigma_\mu^2}} \tag{7}$$

13

The prior in period t=0 is $C_{j0}^e = C_0$, $\sigma_{C_{j0}}^2 = \sigma_{C_0}^2$.

5.3 Learning about the Potential of One's Own Ideas

We model individuals as heterogeneous with respect to their ability to generate ideas of high potential (good or marginal idea contributors). Further, we model individual potential to be idiosyncratic across different categories of ideas[10] and invariant over time[11]. When an individual joins IdeaStorm, her prior belief of the mean potential of her ideas is normally distributed with mean Q_0 and variance $\sigma_{Q_0}^2$

$$Q_{i0} \sim N(Q_0, \sigma_{Q_0}^2) \tag{8}$$

The information that provides a signal about the potential of one's ideas is the voting score the idea receives from peers who are also potential consumers. IdeaStorm.com allows individuals to vote on their peers' ideas, and the voting score is used as a measure of the potential of the ideas. On IdeaStorm, Dell says that IdeaStorm allows Dell "to gauge which ideas are most important and most relevant to" the public and "the Point Count (voting score) is the best way to gauge overall popularity"[12]. A high voting score means that many customers would like to see this idea implemented, while a low voting score means the idea is probably a limited idea that is favored by few. In fact, literature as well as practice in new product development tells us that firms have been long gauging potential of new product ideas or improvements on products by asking potential customers (Hauser and Urban 1977, Lilien and Kotler 1983). We assume that the natural logarithm of the votes (V) that an idea receives is linearly correlated with the potential of the idea[13]

$$V = cons + \varphi Q \tag{9}$$

The individual's prior belief about the log of voting score their ideas may receive can be written as

$$V_{i0} \sim N(cons + \varphi Q_0, \varphi^2 \sigma_{Q_0}^2) \tag{10}$$

Let Q_i denote the mean potential of ideas posted by individual i then Q_{sit} the potential of an idea posted by individual i in period t is

$$Q_{sit} = Q_i + \delta_{sit} \tag{11}$$

$$\delta_{sit} \sim N(0, \sigma_{\delta\,i}^2)$$

---

[10] Data reveals that the votes received by the two categories of ideas posted by the same individual are not statistically significantly different (p-value = 0.926).

[11] Since the data shows that individual's do not improve in their ability to come up with high potential ideas with experience, we model the individual's true type to be constant over time. In other words, we do not allow the "type" of an individual to change over time.

[12] http://www.dell.com/content/topics/global.aspx/ideastorm/moderator?c=us&l=en&s=gen

[13] In specification (9) voters are assumed to be a representative sample of Dell's consumers. Given that Dell is a technology company and Ideastorm is an online platform, the voters are likely to be a representative sample of Dell's consumers. However, in case the voters were a biased sample of Dell's consumers, our specification (9) would still be able to account for simple linear bias in voter's preferences.

where δ_{sit} is the deviation of the potential of a specific idea posted by individual i in period t from the average potential of his/her ideas. The variance of δ_{sit} is individual specific, which means that not only individuals have different potential but also they learn about the potential of their ideas at different rates over time. Note that individuals learn their potential by observing the voting scores their ideas receive.

The voting score an idea receives can be written as

$$V_{sit} = V_i + \xi_{sit} \tag{12}$$

where

$$
\begin{aligned}
V_i &= cons + \varphi Q_i \\
\xi_{sit} &= \varphi \delta_{sit} \\
\xi_{sit} &\sim N(0, \sigma^2_{\xi_i}) \\
\sigma^2_{\xi_i} &= \varphi^2 \sigma^2_{\delta_i}
\end{aligned}
\tag{13}
$$

Here, V_i is the mean value of the logarithm of votes that individual i's idea receives and ξ_{kit} is its random shock.

Let Q^e_{jt-1} and V^e_{jt-1} denote individual's belief of means of potential of her ideas and the log votes her ideas may receive at the beginning of period t, respectively. Individuals update their beliefs about V^e_{it} and Q^e_{it} together when they observe the voting scores their ideas receive. The updating rules for V^e_{it} and Q^e_{it} are (Erdem, Keane and Sun, 2008)

$$V^e_{it} = V^e_{it-1} + (V_{sit} - V^e_{it-1}) \frac{\sigma^2_{V_{it-1}}}{\sigma^2_{V_{it-1}} + \sigma^2_{\xi_i}} \tag{14}$$

$$Q^e_{it} = Q^e_{it-1} + (V_{sit} - V^e_{it-1}) \frac{\varphi \sigma^2_{Q_{it-1}}}{\varphi^2 \sigma^2_{Q_{it-1}} + \sigma^2_{\xi_i}} \tag{15}$$

where

$$\sigma^2_{V_{it}} = \frac{1}{\frac{1}{\sigma^2_{V_{it-1}}} + \frac{1}{\sigma^2_{\xi_i}}} \tag{16}$$

$$\sigma^2_{Q_{it}} = \frac{1}{\frac{1}{\sigma^2_{Q_{it-1}}} + \frac{1}{\sigma^2_{\xi_i}}} \tag{17}$$

In addition, we denote the priors for potential and for log-votes at the moment that when the individual joins IdeaStorm to be $Q^e_{i0} = Q_0$, $\sigma^2_{Qi0} = \sigma^2_{Q0}$, and $V^e_{i0} = cons + \varphi Q_0$, $\sigma^2_{Vi0} = \varphi^2 \sigma^2_{Q0}$.

5.4 Firm's Decision Rule to Implement Ideas

The firm selectively implements ideas generated by individuals. In general, the firm will consider the potential (market demand) of the ideas as well as the costs of implementing the ideas. Assume that a firm only

implements ideas that provide a positive net profit. The net profit the firm generated from implementing the mth Category j idea posted in period t can be expressed as[14]

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

where Q_{mjt} represents the true potential of the idea and C_{mjt} represents the firm's true cost associated with implementing the idea. Then, the probability that an idea will be implemented is

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

At the point that the firm makes implementation decisions, C_{mjt} is observed only by the firm, and not by consumers or researchers. However, Q_{mjt} is observed by everyone once peers have voted on the idea, given *cons* and φ. Hence, from the firm's perspective, there is no uncertainty in the decision process. For us, C_{mjt} is a random variable with mean C_j and variance σ_γ^2.

$$C_{mjt} = C_j + \gamma_{mjt}$$

where $\gamma_{mjt} \sim N(0, \sigma_\gamma^2)$.[15] This implies that we can only infer that the expected net payoff of implementing an idea is normally distributed with mean $Q_{mjt} + C_j$ and variance $\sigma_{\gamma_j}^2$. Therefore, for us the likelihood that an idea with observed potential Q_{mjt} is eventually implemented is

$$P_{mjt} = Pr\big(Q_{mjt} + C_{mjt} > 0 \,\big|\, Q_{mjt}\big) = 1 - \Phi\big(\frac{Q_{mjt} + C_j}{\sigma_{\gamma_j}}\big) \tag{18}$$

Let I_{mjt} denote the decision the firm makes on the mth Category j idea posted in period t, with value *1* indicating that the idea is implemented and *0* otherwise. The likelihood of the observed implementation decision (I_{mjt}) given Q_{mjt}, C_j and σ_{γ_j} is

$$L\big(I_{mjt}\big) = \Phi\big(\frac{Q_{mjt} + C_j}{\sigma_{\gamma_j}}\big)^{(1 - I_{mjt})}\big(1 - \Phi\big(\frac{Q_{mjt} + C_j}{\sigma_{\gamma_j}}\big)\big)^{I_{mjt}}$$

---

[14] One concern could be that the firm uses not only the voting score, but also some other factors to measure ideas' potential. In other words, the firm does not use $Q_{mjt} + C_{mjt} > 0$ as the decision rule but $Q_{mjt} + C_{mjt} > e_{mjt}$ as the decision rule. Here e_{mjt} is a random error which captures the unobserved factors that the firm uses in judging ideas' potential. Assume e_{mjt} is normally distributed with mean zero, i.e. $e_{mjt} \sim N(0, \sigma_e^2)$. Then the likelihood that an idea with observed potential Q_{mjt} is eventually implemented is

$$P_{mjt} = Pr\big(Q_{mjt} + C_{mjt} + e_{mjt} > 0 \,\big|\, Q_{mjt}\big) = 1 - \Phi\big(\frac{Q_{mjt} + C_j}{\sqrt{\sigma_\gamma^2 + \sigma_e^2}}\big).$$

We can see that if there are indeed other factors that affect firm's implementation decision and it does not have systematic bias towards a certain direction, the only modification we need to make is the interpretation of σ_γ^2 : σ_γ^2 is a combination of the true variance of the cost distribution plus the variance of e_{mjt}. If there are any other considerations that have systematic bias towards a certain direction, that is, the mean of e_{mjt} is a non-zero number *b*, *b* would be absorbed in the estimated cost. Thus, the cost we recovered from the data is actually a combination of cost and other unobserved considerations *b*.

[15] Another assumption is the random error term γ_{mjt} is i.i.d. This assumption ensures model tractability. The assumption could potentially break down under some scenarios. For example, implementation of one idea could affect the cost of implementation of other ideas.

5.5 Individual's Decision Making Problem

As previously mentioned, individuals make decisions on whether or not to post an idea in a category based on their expectation of the utility they can possibly derive from each choice. We model individuals' decisions on idea posting to be independent of each other across categories and that the individuals are aware that the firm makes implementation decisions by comparing the potential of the ideas and the implementation costs to the firm. Then, the $\widetilde{U}_{ijt}$ in Equation (2) can be expressed as

$$\widetilde{U}_{ijt} = \theta_{i0} + d_i D_{it} + \theta_{ij} Pr\left(\pi_{ijt}|Info(i,t) > 0\right) \tag{19}$$

where $\pi_{ijt}|Info(i,t) \sim N(Q_{it}^e + C_{jt}^e, \sigma_{Q_{it}}^2 + \sigma_{C_{jt}}^2 + \sigma_{\delta_i}^2 + \sigma_{\gamma_j}^2)$. $Info(i,t)$ captures individuals' perceptions about potential of their own ideas and the firm's implementation cost formed through the two types of learning processes, which contains $Q_{it}^e, C_{jt}^e, \sigma_{Q_{it}}^2$ and $\sigma_{C_{jt}}^2$. In other words, the information set $Info(i,t)$ evolves as people update $Q_{it}^e, C_{jt}^e, \sigma_{Q_{it}}^2$ and $\sigma_{C_{jt}}^2$. We assume that ε_{ijt} follows a Type 1 extreme value distribution. Hence, the probability that individual i will post a Category j idea in period t takes the standard logit form. In this case, the likelihood of observing posting outcome, A_{ijt}, can be expressed as

$$L(A_{ijt}) = \left(\frac{exp(\widetilde{U}_{ijt})}{1+exp(\widetilde{U}_{ijt})}\right)^{A_{ijt}} \left(\frac{1}{1+exp(\widetilde{U}_{ijt})}\right)^{(1-A_{ijt})} \tag{20}$$

Here, $A_{ijt} = 1$ if an individual i posts a Category j idea in period t and 0 otherwise.

6. Estimation

In the literature, most of the Bayesian learning models are estimated by (simulated) maximum likelihood estimation methods. However, in our case, due to the individual-specific parameters, Q_i, $log(\sigma_{\delta_i}^2)$, d_i, θ_{i0}, θ_{ij}, the frequentist estimation methods are inconvenient. Following Narayanan and Manchanda (2009), we apply Markov Chain Monte Carlo (MCMC) methods to estimate the individual-specific parameters. We use the Gibbs sampler to recursively make draws from the following conditional distribution of the model parameters. We briefly explain the model hierarchy here and for complete details of the estimation procedure please see the online Appendix 1.

6.1 Model Hierarchy

The parameters in our model are shown in Table 4. For identification purposes, C_1, $\sigma_{C_0}^2$, and $\sigma_{Q_0}^2$ are fixed (model identification will be briefly discussed in later sections and elaborated in Appendix 2). Among the remaining parameters, parameter vector $\boldsymbol{\alpha} = [\ C_0, C_2, \sigma_{\gamma_1}^2, \sigma_{\gamma_2}^2, \sigma_\mu^2, Q_0, cons, \varphi]$ is common across individuals, while parameter vector $\beta_i = [Q_i, log(\sigma_{\delta_i}^2), d_i, \theta_{i0}, \theta_{i1}, \theta_{i2}]$ is heterogeneous across individuals. We further assume that β_i follows the following distribution

$$\beta_i = \begin{pmatrix} Q_i \\ log(\sigma^2_{\delta\,i}) \\ d_i \\ \theta_{i0} \\ \theta_{i1} \\ \theta_{i2} \end{pmatrix} \sim MVN(\bar{\beta}, \Sigma)$$

where $\bar{\beta}$ denotes the mean of β and Σ denotes the variance and covariance matrix of β.

Table 4. Summary of the Parameters in the Model

Notation	Explanation
θ_{i0}	Cost for individual i to post an idea.
θ_{ij}	Payoffs individual i receives when his/her Category j ideas are implemented
d_i	Level of disincentive individual i receives when the status of one or more of i's ideas stays as "Acknowledged" for more than 12 weeks
D_{it}	Indicator for "no response". Binary variable that takes a value of 1 when there is at least one idea posted by individual i that has not moved to any status other than "Acknowledged" for more than 12 weeks after it is originally posted in period t.
C_0	Individual's initial prior mean costs for implementing each Category of ideas
$\sigma^2_{C_0}$	Individual's initial prior variance of the costs for implementing each Category of ideas (set to 50, assume the prior is uninformative)
C_j	The firm's mean cost for implementing Category j ideas (the mean cost for Category 1 is fixed at -6)
$\sigma^2_{\gamma_j}$	The variance of true distribution of the costs for the firm to implement ideas in Category j
σ^2_μ	Variance of cost signal
Q_0	Individuals' initial prior mean of the potential of their ideas
$\sigma^2_{Q_0}$	Individuals' initial prior variance of the potential their ideas (set to 50, assume prior is uninformative)
Q_i	Mean potential of ideas generated by individual i
$\sigma^2_{\delta\,i}$	Variability of potential of ideas generated by individual i
$cons$	Intercept of linear function between log votes and the potential
φ	Slope coefficient between log votes and potential

Conditional on *cons,* φ and $\sigma^2_{\delta\,i}$, the updating process of the potentials of individuals' ideas is deterministic because we explicitly observe the potential signal (votes). The updating process of the variance of mean cost belief is also deterministic, given σ^2_μ. Only the updating process of C^e_{jt} is stochastic. Following Narayanan and Manchanda (2009), the distribution of C^e_{jt+1}, conditional on C^e_{jt}, can be expressed as

$$C^e_{jt+1}|C^e_{jt} \sim N(\bar{C}^e_{jt+1}, v^2_{jt+1})$$

where

$$\bar{C}^e_{jt+1} = \frac{\sigma^2_{C_{jt+1}}}{\sigma^2_{C_{jt}}} \bar{C}^e_{jt} + k_{Cjt} \frac{\sigma^2_{C_{jt+1}}}{\sigma^2_\mu} C_j \tag{21}$$

$$v^2_{jt+1} = k_{Cjt} \frac{\sigma^4_{C_{jt+1}}}{\sigma^2_\mu} \tag{22}$$

Therefore, the unobserved cost belief can be drawn from the following natural hierarchy

$$C_{jt}^e | C_{jt-1}^e \sim N(\bar{C}_{jt}^e, v_{jt}^2)$$

$$C_{jt-1}^e | C_{jt-2}^e \sim N(\bar{C}_{jt-1}^e, v_{jt-1}^2)$$

$$\dots$$

$$C_{j1}^e | C_{j0}^e \sim N(\bar{C}_{j1}^e, v_{jt}^2)$$

The full hierarchical model can be specified as

$$A_{ijt} | C_{jt}^e, \sigma_{C_{jt}}^2, \sigma_{\gamma j}^2, cons, \varphi, Q_0, \ D_{it}, \ V_{si}^t, \beta_i; \ I_{mjt} | C_j, \sigma_{\gamma j}^2, \ V_{mjt}, cons, \varphi,$$

$$C_{jt}^e | C_{jt-1}^e, C_j, \sigma_\mu^2, k_{Cjt}$$

$$\beta_i | \bar{\beta}, \Sigma$$

where the additional notation V_{si}^t denotes a vector of the log voting scores that all ideas generated by i receive up to period t.

6.2 Identification

In our model the consumers make posting decisions based on their (perceived) utility. The variances of the two signals, σ_μ^2 and $\sigma_{\delta\,i}^2$, are identified from the dynamics of the posting behaviors of individuals over time. Every time an individual gets a signal, she would update her belief. This would affect her posting behavior. The variation in the posting behavior immediately after receiving a signal helps us identify the variance of the signal. If the posting behavior changes a lot then the signal is precise and the variance of the signal is small. If the posting behavior does not change much then the signal is very noisy and hence the variance of the signal is very large. The variance of the cost signal is identified from the change in posting behavior of individuals upon observing an idea getting implemented. The variance of the idea potential signal is identified upon observing the change in behavior of an individual as she receives votes on her posted idea.

The individual's prior beliefs about mean potential of her idea (Q_0) and mean cost of implementation (C_0) are identified from the direction of change in posting behavior as they observe the signals. If after observing a signal for her idea potential her probability of posting increases then it implies that individual's prior belief about mean potential of her ideas was lower than her updated belief. Similarly, if after observing a cost signal an individual's posting decreases, we can infer that in her updated belief the mean cost of implementation is higher than in her prior belief. The posting behavior would eventually stabilize after the individuals have observed numerous signals. The direction of the change and the extent of the change from the initial posting behavior identify the prior belief of an individual about mean potential of her ideas and the cost of implementation.

The relation between C_j and Q_i, as well as $\sigma_{r_j}^2$, the variance of the true cost distribution, are identified from the likelihood of an idea with certain voting score being implemented. We fix mean implementation cost (C_I) for Category 1. Further, the potential of an idea has one to one mapping with the log votes it receives. So for Category 1, the variation in implementation of ideas that receive same voting

score helps identify the variance of the implementation cost $(\sigma_{r_1}^2)$ for Category 1. Once we know $\sigma_{r_1}^2$ and C_l, we can easily identify the potential of an idea (Q_{mjt}) from variation in implementation rates across ideas. The identified potential of an idea, Q_{mjt}, and the votes it received, V_{mjt}, can be directly used to identify $cons$ and φ due to the linear relationship. The $cons$, φ, and the votes that a Category 2 idea receives can be used to calculate its potential. We can then exploit the variation in the implementation of ideas in Category 2 in the same way as we did for Category 1 to figure out C_2 and $\sigma_{r_2}^2$. The potential of a particular idea, Q_{mjt}, posted by individual i follows normal distribution with mean Q_i and variance $\sigma_{\delta\,i}^2$, which can be identified using several identified Q_{mjt} for an individual.

The overall frequency with which an individual i posts Category j ideas is jointly determined by θ_{i0} and θ_{ij}. Given everything else equal the consistent difference in probability of posting among individuals help us identify θ_{i0}. Given everything else equal, the differences in the change in probability of posting every time a cost signal is received by individuals helps us identify θ_{ij}. Individuals with higher level of θ_{ij} would have a higher change in probability of posting compared to others on receiving same cost signal given everything else equal. Everything else equal, the difference in individual$i$'s posting behavior between cases where $D_{it} = 0$ and $D_{it} = 1$ identifies d_i.

7. Estimation Results

The estimates of the parameters that do not vary across individuals (pooled parameters) are presented in Table 5. Comparing the estimate of C_2 with C_1 (fixed to -6), we see that C_2 is slightly smaller in terms of absolute value. Thus, the cost that the firm incurs when implementing Category 2 ideas is lower than the cost of implementing Category 1 ideas, which is consistent with the higher implementation rate of Category 2 ideas as compared to Category 1 ideas. The estimate for C_0 is statistically significantly higher than both C_1 and C_2, indicating that individuals initially tend to underestimate the idea implementation costs.

Table 5. Pooled Parameter Estimates			Table 6. Individual-level Parameter Estimates		
Notation	Parameter Estimates	Standard Deviation	Notation	Mean Among Individuals*	Standard Deviation Among Individuals *
C_0	-1.129	0.232			
$\sigma_{C_0}^2$	50	-- (Fixed)	Q_i	2.274	0.159
C_1	-6	-- (Fixed)	$log(\sigma_{\delta\,i}^2)$	-1.492	1.524
C_2	-5.882	0.095	d_i	-1.711	1.148
$log\,(\sigma_{\mu 1}^2)$	6.502	0.514	θ_{i0}	-4.996	0.497
$log\,(\sigma_{r1}^2)$	1.268	0.085	θ_{i1}	3.363	0.370
$log\,(\sigma_{r2}^2)$	1.443	0.103	θ_{i2}	2.938	0.587
Q_0	3.411	0.375			
$\sigma_{Q_0}^2$	50	-- (Fixed)			
$cons$	-0.514	0.106			
φ	2.352	0.033			

* For each individual, the posterior distribution of each parameter has a mean and standard deviation. The mean and standard deviation reported here are the mean and standard deviation of the individual-level parameter means.

The estimate of $log\ (\sigma_\mu^2)$ is 6.502, which is equivalent to saying that σ_μ^2 is exp(6.502)=666. This variance is quite large compared to the absolute values of C_1 and C_2, indicating that the implementation cost signals the firm provides to individuals are imprecise and consequently, that individuals cannot learn quickly about the implementation costs of the firm. Remember that $\exp(\sigma_\mu^2)$ is the variance of one signal and that there are cases where several ideas are implemented within a week. In those weeks, the variance of the cumulative signal individuals receive will be $\exp(\sigma_\mu^2)$ divided by the number of ideas implemented in each week; thus, the learning regarding the implementation would be significant in such cases. The estimates of $log\ (\sigma_{r1}^2)$ and $log\ (\sigma_{r2}^2)$ are (1.268 and 1.443 respectively); that is, $\sigma_{r1}^2 = 3.55$ and $\sigma_{r2}^2 = 4.23$. This implies there is reasonable variance in the implementation cost of ideas within Category 1 as well as Category 2.

Q_0 is also higher than the average level of Q_i, indicating that most of the individuals overestimated the potential of their ideas before their ideas were voted on[16]. The parameters $cons$ and φ determine the linear relationship between log votes and potential. The slope coefficient is 2.352, meaning that when the potential of the idea increases by 1, the log of an idea's vote increases by 2.352.

The estimation results of the mean and standard deviation of individual-level parameters are summarized in Table 6. Additionally, histograms of the distribution of the 6 individual-level parameters are shown in Figure 5. We can see that the population average of the potential of the ideas (2.274) is significantly lower than the mean cost of implementing both categories of ideas. This is consistent with the low implementation rate we observe in the data. We also observe significant differences among individuals with respect to the ability to generate ideas with high potential. The population average of variance of the potentials of ideas by one individual is small $\exp(\overline{log(\sigma_\delta^2)})= 0.225$. This result suggests that in general the potentials for the ideas posted by the same person are relatively consistent. Good idea contributors consistently generate ideas with high potential, while marginal idea contributors rarely come up with high potential ideas. This variance also implies the learning speeds of individuals with respect to the potential of their ideas. The small average variance also indicates that, on average, individuals learn quickly about the potential of their ideas. When the website was launched, many individuals, i.e., idea providers, entered the market. As they learn about the potential of their ideas and the cost for the firm to implement their ideas, marginal idea contributors dropped out, and the "idea market" became more efficient in a short time. In other words, the crowdsourcing mechanism is quite effective in filtering idea providers, and the "idea market" reaches efficiency quickly. The standard deviation of $\sigma_{\delta\ i}^2$ is relatively large (1.524), indicating that some individuals have better consistency in terms of the potential of their ideas, while others have a lower consistency.

---

[16] In the main model, we assume individuals have common prior. We show the estimation results under heterogeneous prior assumption and discuss the difference in the results under these two assumptions in the Robustness Check section.

The average level of the lack of response effect is -1.711, meaning that when individuals' ideas are not responded to in a timely manner, individuals tend to be less likely to post ideas, and the average level of this effect is equivalent to increasing the cost of posting an idea by a factor of around 0.34. Given the low overall probability of posting, the impact of such discouragement is quite large. The mean payoff individuals receive when their Category 1 ideas are implemented (3.363) is slightly higher than when their Category 2 ideas are implemented (2.938). This is consistent with the numbers of ideas posted in these two categories during the first few weeks. This finding is also intuitive because ideas in Category 1 are about product improvement, while ideas in Category 2 are related to customer services and marketing strategies. It is not surprising that individuals receive greater payoffs when the firm improves the product design according to an individual's suggestion than when the firm improves services and communications with their customers. The average cost of posing an idea is -4.996, with standard deviation 0.497.

To explore the correlation between individual mean potential (Q_i) and other individual-level parameters, we present the scatter plots of $log(\sigma_{\delta_i}^2)$ and d_i against individual mean potentials, respectively (Figure 6). Interestingly, the correlation of Q_i and $\sigma_{\delta_i}^2$ is negative, indicating that the potential of ideas generated by good idea contributors are more consistent, and thus, these individuals tend to learn more quickly about their ability. Another interesting finding is the correlation between Q_i and d_i. In other words, good idea contributors would not be as disappointed as marginal idea contributors from no response by firm to their ideas. We explore the policy implications of this finding later.
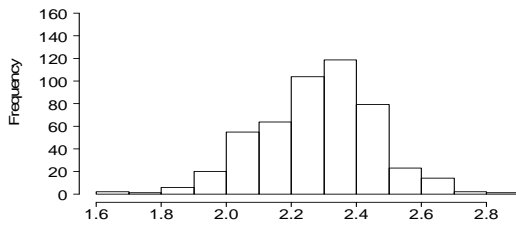

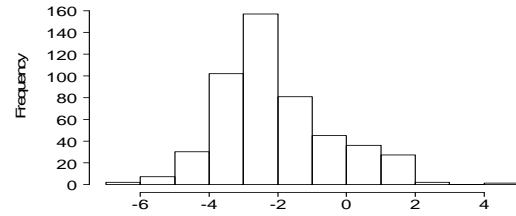**Figure 5a Distribution of Q_i**
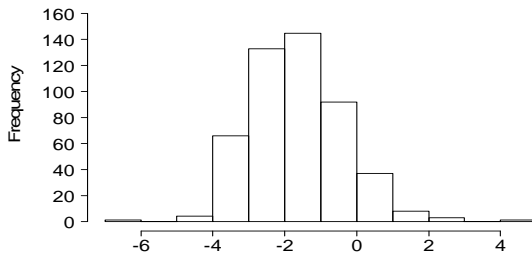
Figure 5b Distribution of $log(\sigma_{\delta_i}^2)$
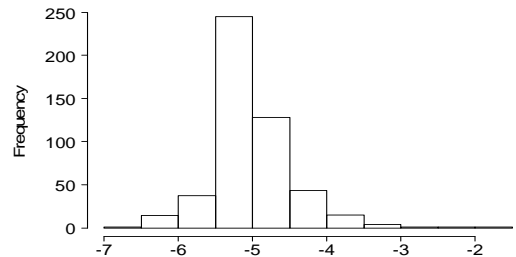
Figure 5c Distribution of d_i

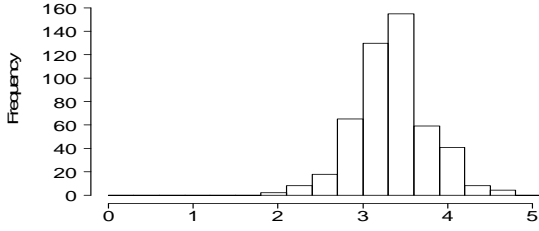**Figure 5d Distribution of θ_{i0}**
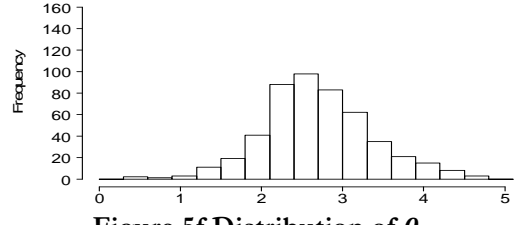
Figure 5e Distribution of θ_{i1}

Figure 5f Distribution of θ_{i2}

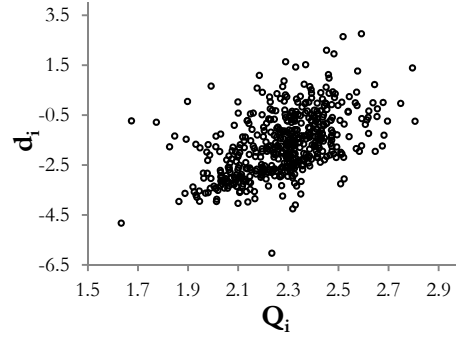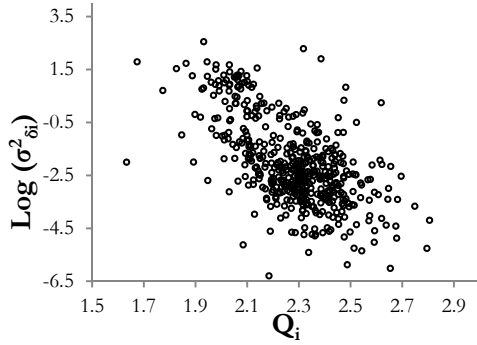**Figure 5. Distributions of Individual-level Parameters**

Figure 6. Scatter Plots of Individual-level Parameters
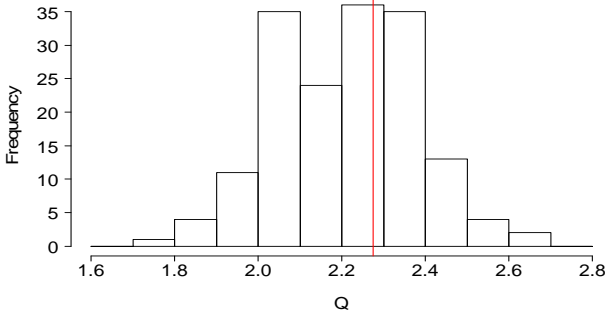
Individuals Who Posted ≥ 2 Ideas in the First 20 Weeks
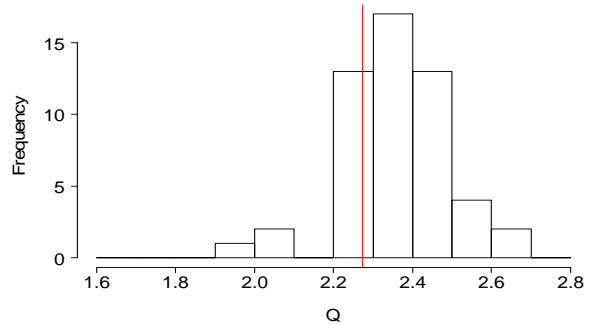
Individuals Who Posted ≥ 2 Ideas in the Last 20 Weeks

Figure 7. Comparison of the Mean Potential of Individuals Who Posted ≥ 2 Ideas in the First and Last 20 Weeks

Figure 8. Idea Generation Dynamics (Good vs. Marginal Idea-Contributors)

Our estimation process produces the posterior mean of an individual's ability to generate ideas with good potential. This allows us to explicitly examine the filtering process of idea providers in the market.

23

Figure 7 visualizes the comparison between the mean potential of individuals who post ≥2 ideas in the first 20 weeks and that of individuals who post ≥2 ideas in the last 20 weeks. The vertical line in both plots 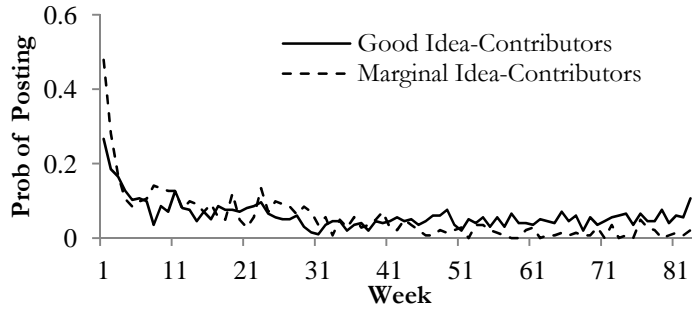is the average mean potential of the 490 individuals in our sample. From the two plots, it is evident that the distribution shifts toward the right. The majority of the individuals who post ≥2 ideas in the last 20 periods are those who have been identified as good idea contributors. From Figure 8, we can also see that in the first few weeks, marginal idea contributors post many ideas; but after sufficient learning, good idea-contributors (above average ability) are more likely to contribute ideas, while marginal idea contributors rarely post ideas in the later part of our observation period.

7.1 Model Fit and Model Comparison

In order to see whether including the learning processes can explain better the pattern we see in the data, we compare our model with three alternate models, which include the random coefficients model, the cost learning only model, and the potential learning only model. The first alternate model does not allow any learning. The second alternate model allows only cost learning. The third alternate model allows only learning about the idea potential. From Table 7, we note that our full model outperforms all other alternative models with respect to marginal likelihood and deviance information criterion (DIC). We also find that only the cost learning model slightly improves marginal likelihood when compared to the no learning model. It appears that only including cost learning does not significantly improve model fit. This is because, on the one hand, cost learning is relatively slow and therefore has limited contribution to model fit, and on the other hand, in the cost learning only model, the learning dynamics of individuals are assumed away, which deviates from reality. Not surprisingly, we find that when we include learning about idea potential, both marginal likelihood and DIC improve significantly. This suggests that learning about idea potential explains a significant amount of dynamics in the idea posting. By comparing the full model and the potential learning only model, we find that after we control for individual learning about idea potential, adding cost learning will improve the performance of the model. This suggests that although the firm only provides imprecise cost signals and individuals learn slowly about the firm's cost structure, the effect of cost learning still explains a significant degree of the remaining dynamics.

Table 7. Model Comparison

Model	Random Coefficient	Cost Learning Only	Potential Learning Only	Full Model
Log Marginal Likelihood	-8131.6	-8096.5	-7863.8	-7820.4
Difference in DIC (wrt. Full Model)*	319.4	317.3	22.7	0.0

* Difference in DIC=DIC of the model-DIC of the full model. Smaller DIC is preferred.

Table 8.1. Pooled Parameter Estimates*

Parameter	Main Model	4 Cost Signals
C_0	-1.129 (0.232)	-0.629 (0.254)

Table 8.2. Individual-level Parameter Estimates*

Parameter	Main Model	4 Cost Signals
Q_i	2.274 (0.159)	2.517 (0.185)

24

C_2	-5.882 (0.095)	-5.914(0.084)		$log(\sigma^2_{\delta\,i})$	-1.492(1.524)	-1.615(1.670)
$log\,(\sigma^2_{\mu I})$	6.502 (0.514)	6.870 (0.526)		d_i	-1.711(1.148)	-1.665 (1.149)
$log\,(\sigma^2_{\mu AO})$	--	8.597(0.461)		θ_{i0}	-4.996(0.497)	-4.615 (0.504)
$log\,(\sigma^2_{\mu NP})$	--	9.093(0.566)		θ_{i1}	3.263(0.370)	2.721 (0.333)
$log\,(\sigma^2_{\mu PI})$	--	8.770(0.593)		θ_{i2}	2.938(0.587)	2.537(0.410)
$log\,(\sigma^2_{r1})$	1.268(0.085)	1.122(0.090)				
$log\,(\sigma^2_{r2})$	1.443 (0.103)	1.295(0.126)				
Q_0	3.411 (0.375)	3.194 (0.477)				
$cons$	-0.514(0.106)	-0.610(0.114)				
φ	2.352 (0.033)	2.163 (0.099)				

7.2 Robustness Checks

In this section, we relax some of our model assumptions, test alternate explanations, and demonstrate the robustness of our results.

7.2.1 Additional Events that Could Provide Cost Signals

In the main model, we assume that individuals receive cost signals only when the firm implements ideas and we have also reasoned why the implementation of ideas would be the most important cost signal. One may argue that the three other status changes for an idea, "Already Offered", "Not Planned" and "Partially Implemented", could also provide cost signals. To see whether the estimation results are robust to models with more cost signals, we estimate the following model. We consider the four status changes, "Implemented (I)", "Already Offered (AO)", "Not Planned (NO)" and "Partially Implemented (PI)" to contain information about the firm's cost structure. Each of these status changes produces a signal which is normally distributed with mean C_j and variance $(\sigma^2_{\mu status})$. The signals from these four status changes differ from each other only in terms of variance of the signal distribution $(\sigma^2_{\mu status})$. This specification implies that the signals have different noise levels. The estimation results of the new model and the main model are summarized in Table 8.1 and 8.2. We can see that compared with the signal consumers receive from the implementation of the ideas, the variances of other signals are much higher, indicating that other signals provide very little information and so consumers cannot learn much from these signals. This is not surprising because implemented ideas have higher visibility and thus individuals receive more information from this type of cost signal. In addition, Table 8.1 and 8.2 also show that after including the extra cost signals, there is no significant change in the estimates of other parameters.

We also perform a DIC based model comparison and the results are summarized in Table 8.3. From the comparison, we can see that the log-marginal likelihood these two models produce is comparable. However, the model with four cost signals has a higher DIC, indicating that the main model still fits the data better.

Table 8.3. Model Comparison

Model	Main Model	4 Cost Signals
Log Marginal Likelihood	-7820.4	-7823.6
Difference in DIC (wrt. Full Model)*	0.0	6.3

* Difference in DIC=DIC of the model-DIC of the full model. Smaller DIC is preferred.

7.2.2 Alternative Explanations

There could be several alternative explanations of the patterns we observe in the data. We have discussed the learning curve effect earlier. Here, we discuss two more plausible alternative explanations.

Backlog of Ideas. Before IdeaStorm was launched there was no channel for customers to propose ideas to Dell for improving its products and services. When IdeaStorm was launched customers got a channel to propose ideas to Dell and the initial huge number of contributed ideas could represent the backlog of accumulated ideas that the customers had "in stock" initially due to lack of such a channel. Further, as time goes by people may find it harder to come up with more ideas. One way this could manifest in our model is through the cost of idea posting. If backlog of ideas were a significant driver of the decline in the number of ideas posted over time, then the cost of idea posting should increase over time because it would be much harder to come up with ideas over time. We estimated another model where we allow the cost of idea posting during weeks 1-42 to be different from the cost of idea posting during weeks 43-84. We do not find any statistically significant different between the costs of idea posting between these two time periods. Hence, the backlog of ideas is not a major driver of the decline in number of ideas posted over time. However it could a driver for individuals who posted less than 2 ideas (who are not in our sample). These would be the individuals who posted one idea and vanished from Ideastorm. Further, while exhaustion of ideas could potential explain the decrease in the number of ideas posted over time, it cannot explain why the average voting score increases over time.

High Potential Idea Generators Join Later. Another alternative explanation of the increase in the potential of the ideas is that high potential idea generators join the platform later. This would lead to increase in average potential of contributed ideas over time. To address this concern, we present the arrival process of the contributors as a function of their ability to come up with high potential ideas. In Figure 9, the x axis is the time and the y axis is the average potential of the idea providers who provided their first idea on the corresponding time on the x axis. It is clear from Figure 9 that there is no significant trend that may indicate that high quality idea providers come later on the platform.
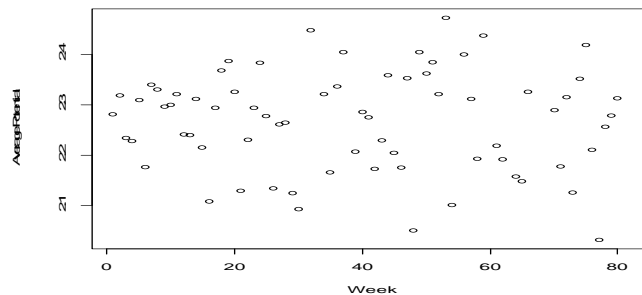

Figure 9. Arrival Process of Idea Contributors

7.2.3 Individuals Specific Heterogeneous Priors

26

One potential concern with our model is that while individuals are heterogeneous we assume individual priors about own ability to be homogenous. To test whether, our results are robust to this assumption we follow the procedure suggested by Metha et al. (2008) to allow for individual specific heterogeneous priors. The procedure followed is as follows. We use the data in the first few weeks (week 1 to 20) for initializing the priors. At week 1 all individuals have same priors. However, they can learn from their actions from week 1 to 20. They update their priors using this information and the posteriors at Week 20 become the priors at Week 21. These priors are individual specific. We use the data from week 21-84 onwards to estimate the model parameters. The results for this test are provided in Table 9.1 and 9.2. Results indicate that even when the priors are individual specific (i.e. account for heterogeneity) the main results do not change statistically significantly. From this analysis, by comparing the priors of the individuals at period 21 with their posteriors at period 84, we can see that several individuals have overestimated their true potential. In contrast only a small fraction underestimated their true potential.

7.2.4 Alternative Constructions of No Response

We employ different cutoff points for "No response" to investigate how the estimation results will change. Column 4 and 5 in both Table 9.1 and Table 9.2 show that the parameter estimates, both the pooled estimates and the mean of the individual-level parameter estimates, are quite stable. Thus, the estimation results are not sensitive to the selection of the cutoff points.

Table 9.1 Pooled Parameter Estimates*

Parameter	Main Model (Cutoff=12 Weeks)	Heterogeneous Priors	Cutoff=8 weeks	Cutoff=16 weeks
C_0	-1.129 (0.232)	-1.612 (0.326)	-1.121 (0.246)	-2.362 (0.217)
C_2	-5.882 (0.095)	-5.861(0.125)	-5.868(0.088)	-5.947(0.066)
$log\,(\sigma_\mu^2)$	6.502 (0.514)	6.751 (0.447)	6.437 (0.322)	6.756 (0.203)
$log\,(\sigma_{r1}^2)$	1.268(0.085)	1.235 (0.097)	1.119(0.097)	1.203 (0.097)
$log\,(\sigma_{r2}^2)$	1.443 (0.103)	1.410 (0.090)	1.123(0.118)	1.403 (0.101)
Q_0	3.411 (0.375)	2.854 (0.289)	3.033 (0.366)	3.629 (0.347)
$cons$	-0.514(0.106)	-0.673 (0.183)	-0.145 (0.160)	-0.876 (0.119)
φ	2.352 (0.033)	2.520 (0.087)	2.482 (0.086)	2.266 (0.067)

*Standard deviations are in parentheses.

Table 9.2 Individual-level Parameter Estimates*

Parameter	Main Model (Cutoff=12 Weeks)	Heterogeneous Priors	Cutoff=8 Weeks	Cutoff=16 Weeks
Q_i	2.274 (0.159)	2.284 (0.176)	2.006 (0.184)	2.503 (0.193)
$log(\sigma_{\delta\,i}^2)$	-1.492(1.524)	-1.680 (1.504)	-1.622 (1.184)	-1.245 (1.438)
d_i	-1.711(1.148)	-1.708 (0.861)	-1.812 (1.312)	-1.857 (1.221)
θ_{i0}	-4.996(0.497)	-4.596 (0.641)	-4.493 (0.362)	-5.445(0.454)
θ_{i1}	3.263(0.370)	3.998 (0.349)	2.802 (0.477)	3.902 (0.532)

θ_{i2}	2.938(0.587)	2.488 (0.312)	2.651(0.371)	3.865 (1.423)

*The mean and standard deviation (in parentheses) reported here are the mean and standard deviation of the individual-level parameter means.

8. Policy Simulations

We conduct three set of policy simulations to determine how a firm can improve the overall performance of its crowdsourcing ideation initiative by accounting for the heterogeneous learning by individuals about the firm's cost structure and their ability to generate ideas with good potential. The simulation results are the average across 1000 simulation iterations. The average potential and the number of Category 1 ideas generated in each period are reported. The number of Category 2 ideas generated in each period has a similar pattern as those in Category 1.

8.1 Should the Firm Provide More Precise Information on Its Cost Structure?

If the firm were to provide more detailed information about the cost of implementation then the variance of cost signal would become smaller. Hence, we implement this policy intervention by reducing the variance of the cost signal. We simulate the evolution of the average potential of individual's posts over time and the numbers of ideas in the two categories contributed each week under different standard deviations of cost signal. As shown in Figure 10, if the firm can provide more precise cost information (i.e. cost signals with smaller standard deviations), the average potential of ideas will be significantly improved after 30 weeks. We also observe a significant decrease in the numbers of ideas in each category posted each week, which can further help reduce the idea screening costs that the firm incurs. When the firm provides individuals with more precise cost signals, individuals learn more quickly about the implementation costs. Initially individuals underestimate the implementation cost. However, the quicker they learn about the true implementation cost, the sooner they update their beliefs about the probability their idea would get implemented which they had initially overestimated. Thus, individuals with lower individual mean potential will become inactive sooner. In other words, by providing more detailed feedback about their implementation costs, the firm can improve the efficiency of the idea market. We visually show this impact in the graphs for the reduction in variance of the cost of implementation. Please note that in this analysis we ignore the firm's incentive to be imprecise in signaling its implementation costs due to competitive reasons.

8.2 Should the Firm Respond to Individuals' Ideas More Quickly to Reduce Disincentive?

Our estimation results show that the firm's no or untimely response to ideas negatively affects an individual's participation in the initiative. To deal with this type of disincentive, the firm may attempt to increase the number of ideas to which it responds and to reduce the time between the posting of the idea and the firm's valuable feedback to the contributor. Although we do not know how Dell selects the ideas to which it replies to, the extremely low response rate makes the effect of their selection strategy immaterial. Almost every individual is disincentivized in the latter part of the observation period. In this set of simulations, we examine the various policies that aim to reduce an individual's feeling of disillusionment due to no response from firm.

We experimented with three potential policies. In Figure 11, "All" represents the policy under which the firm responds to all posted ideas within 11 weeks. "Differentiate Ideas" represents the policy under which the firm only replies to ideas that have a potential above the average. "Differentiate Individuals" represents the policy under which the firm identifies good idea contributors and only replies to ideas generated by them irrespective of the specific idea's potential. Not surprisingly, we find that under the three policies, the number of ideas generated increases. This is intuitive because all the policies can reduce the individual's disincentive and thus encourage them to post ideas. On the contrary, the effects that the three policies have on the average potential of ideas posted over time are very different. Interestingly, we find that in the left plot, the curve labeled "All" is below the curve representing current policy everywhere, indicating that if the firm improves the response time and response rate, it completely removes the disincentive, and the firm is worse off because it receives more ideas with significantly lower potential. Therefore, the firm should strategically select the ideas to which it responds.

When comparing the average potential in the "under the current policy" and the "Differentiate Ideas" strategy, we note that, in the beginning, the latter performs no better than the current policy. However, at approximately week 30, the "Different Ideas" strategy starts outperforming the current policy. The "Differentiate Individuals" strategy is obviously the best policy in terms of the potential of the ideas contributed by individuals, standing out immediately after 12 weeks. It also leads to more idea generation, especially in later periods.

The intuition as to why the three policies generate different results is that marginal idea-generators are more disincentivized than good idea-generators when their ideas are ignored (Figure 6). Therefore, when the firm responds to all ideas quickly, regardless of who posted the idea and idea's quality, it encourages more marginal idea-generators to contribute ideas. As a result though higher number of ideas is contributed, the average quality of contributed ideas decreases. In contrast when the firm responds to only high quality ideas, very few of such ideas are contributed by marginal idea-generators and hence majority of them still do not get any feedback from the firm and stay dissatisfied. However, a lot of good idea provider's get feedback on their ideas by the firm encouraging them to contribute more. Similarly when only high potential idea contributors receive feedback only they are the ones who are not dissatisfied with the firm and they post more, whereas low potential idea generators get more dissatisfied with firm even further as none of their ideas are responded to by the firm. In both these cases (where firm responds to high quality ideas or when firm responds to high quality idea contributors) high quality idea contributors are more likely to contribute ideas whereas low quality idea contributors become less likely to contribute ideas leading to on average a high quality of contributed ideas. "Differentiate Individuals" policy performs even better than "Differentiate ideas" policy because "Differentiate ideas" can still encourage some marginal idea contributors when they occasionally come up with good ideas; while "Differentiate Individuals" policy only encourages good idea contributors. It is easier to implement the "Differentiate Ideas" strategy because all the firm needs to do is to look at the votes and

respond to the ideas for which the log of votes is above average. Furthermore, this strategy leads to the submission of only slightly more ideas, and it outperforms the "All" strategy in terms of the potential of the ideas. If we do not consider the screening cost, this is undoubtedly the best policy because the firm will have more high potential ideas from which to choose, and furthermore.
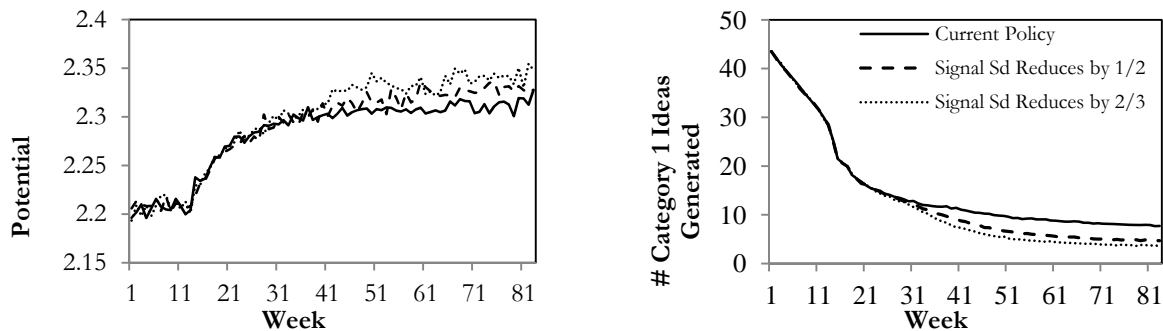
Figure 10. Simulation Results When the Firm Provides More Precise Cost Signals
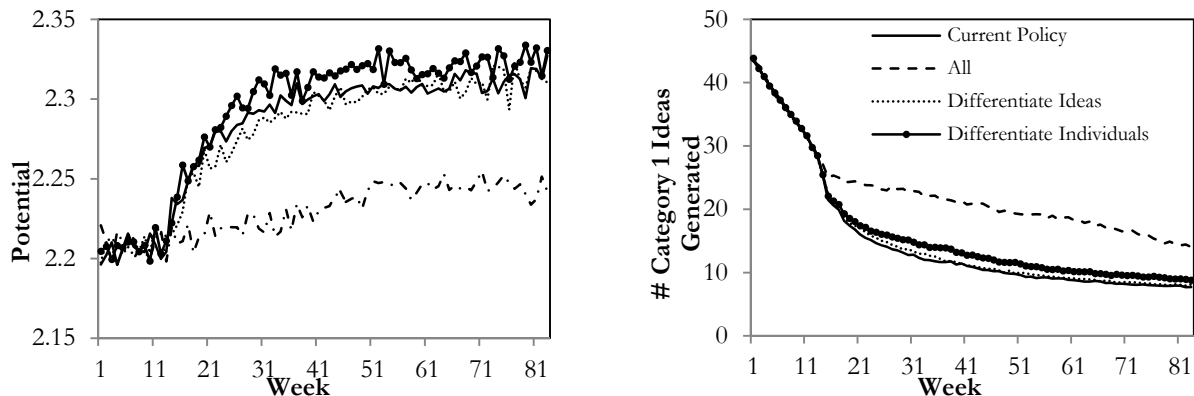
Figure 11. Simulation Results When the Firm Replies to More Ideas in a More Timely Manner

8.3 Should Individuals be Rewarded for Just Posting or for Posting Ideas that are Implemented?

Two commonly observed reward structures used on this type of crowdsourcing website include giving a reward as long as an individual posts an idea (the 500 IdeaStorm points in our case) and giving a reward to contributors only when an idea is implemented (IdeaStorm is currently applying this reward structure). In this set of policy simulations, we aim to investigate which reward structure performs better.
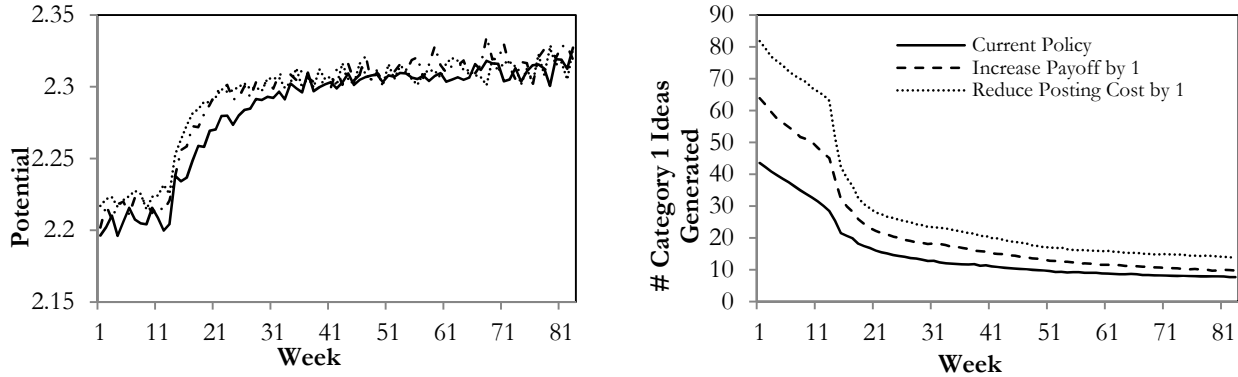
Figure 12. Simulation Results for Different Reward Structures

In Figure 12, "Reducing Posting Cost by 1" represents the policy under which individuals are rewarded as long as they post an idea. This policy will add a positive constant to the utility function of individuals, thus reducing the cost of posting by the same constant. A 1 unit increase in the utility corresponds to on average 20% decrease in the cost of posting. "Increase Payoff by 1" (equivalent to θ_{ij} is raised by 1) represents the policy under which individuals are rewarded only when their ideas are implemented. From the figures, it is evident that the effects of these two policies on the evolution of average potential are similar. Although both policies will increase postings, the "Reducing Posting Cost" policy will lead to a greater number of ideas. To determine which policy is better from the firm's perspective, we consider the cost of screening and the cost of the reward. It is obvious that the "Reducing Posting Cost" policy will cost the firm much more than the "Increase payoff" policy if the firm offers a monetary award. The idea screening cost will also be higher under the "Reducing Posting Cost" policy.

In the discussions above, we attributed the firm's objective were to maximize the average potential of contributed ideas while avoiding high screening cost. An alternative objective function is that the firm may want to maximize is the likelihood of a "killer idea" (e.g., Girotra et al. 2010). Such an objective function would favor a large number of ideas with the hope of a really good idea emerging at some point. Note that different objective functions do not change the results of the policy simulations. A firm with this objective function would choose a policy that shows that it can increase the number of contributed ideas.

9. Conclusion

Our analysis of crowdsourcing data yields several important insights.

Why Does the Number of Contributed Ideas Decrease over Time?

Our results show that, initially, individuals not only overestimate the potential of their ideas, but also underestimate the cost of implementing their ideas. Hence, individuals tend to overestimate the probability that their idea will be implemented, and therefore, they initially post many ideas. As individuals learn about the true cost structure of the firm and the potential of their own ideas the expected utility of idea posting for marginal idea contributors decreases. These learning processes cause the low potential idea contributors to

stop posting ideas. Hence, the two learning processes perform a self-selection function leading to filtering out of marginal idea contributors.

As we explained earlier, an individual's ability to come up with high potential idea does not vary over time. The average potential of contributed ideas increases over time because over time the fraction of high-potential idea contributors increases as the low potential idea providers stop contributing.

Why Does the Fraction of Ideas that are Implemented Increases over Time?

Individuals overestimate the potential of their ideas and underestimate the cost the firm incurs to implement their ideas. Once the website is launched, many individuals enter the "idea market", and thus, the market is crowded by both high and low potential ideas. As individuals learn the potential of their ideas from their experiences, marginal idea contributors tend to post fewer ideas. Consequently, at the aggregate level, the overall potential of ideas generated improves over time. From the firm's point of view, the cost associated with implementing ideas is not changed and so the implementation rate should increase over time.

The learning story we propose has basis in the literature of behavioral biases in the self-perception of individual characteristics. This stream of literature provides evidence indicating that individuals often tend to overestimate their abilities in various domains of every-day life including innovative behavior (e.g., Svenson, 1981; Dunning et al., 1989; Alicke et al., 1995). The general findings in this stream of literature are that individuals are overoptimistic about the returns of potential innovations, or the success probabilities of implementing their innovation and thus will create excessive low potential innovation (e.g. Camerer and Lovallo, 1999; Bernardo and Welch, 2001; Lowe and Ziedonis, 2006; Galasso and Simcoe, 2011, Herz et al. 2012). Peer feedback is one of the key factors that help crowdsourcing alleviate this concern. In the context of crowdsourced ideation, peer evaluation acts as a lens that provides a strong signal that individuals can use to help identify the potential of their innovations.

Facilitated by technology, crowdsourcing has become an intriguing platform for direct idea generation and implementation. The attraction of the business model lies in the real-time assessment of ideas by peers (consumers). As the business headlines on this potentially powerful new product idea source shift from hype, a sobering reality has set in as a declining number of ideas are posted and few ideas are implemented. The observed empirical trend is seen as undermining the potential of crowdsourcing. On the contrary, our analysis suggests that this can be fully justified as a natural outcome of improving the efficiency of these markets. The findings bode well for these emerging new product idea generation methods. Based on these understandings, we propose several policies that may potentially improve the performance of these crowdsourcing initiatives and simulate the overall potential of the ideas and the number of ideas generated under these proposed policies. Our policy simulations indicate that providing more precise cost signals and rewards can help a firm procure higher potential ideas. Furthermore, associating rewards with implementation is more cost-effective than offering rewards just for posting an idea. Another interesting finding in our policy

simulation is that purely increasing the numbers of ideas to respond to and shortening the time to respond without differentiating the ideas negatively affects the overall potential of ideas. In fact, a better strategy is to respond only to the ideas of good idea contributors.

Our paper also has certain limitations. First, our dataset has limited information. From the data, we know little about how the voting score of a particular idea is obtained, as we only observe the final score each idea receives. We have no information on how many people promote an idea and how many people demote the idea, which is information that may allow us to interpret the voting scores more precisely. Second, due to identification reasons, we only include individuals who posted more than 2 ideas in the study period. An understanding of the behavior of individuals who are not in our sample may also have managerial implications. Third, we treat voters as exogenous and do not consider and learning dynamics on their part. It may be interesting to consider how voters may also learn about the cost structure of the firm and use it to guide their voting behavior. Despite the limitations, our paper is the first to provide a dynamic structural framework that analyzes consumers' participation in the crowdsourcing ideation websites, helping both practitioners and researchers to understand this popular web application. We hope that our work can pave the way for future research in this important area.

References

Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., Vredenburg, D. S. 1995. "Personal contact, individuation, and the better-than-average effect" *Journal of Personality and Social Psychology* (68) 804–825.

Archak, N., Sundararajan, A. 2009. "Optimal Design of Crowdsourcing Contests," *ICIS 2009 Proceedings.*

Bayus, B. L. 2010. "Crowdsourcing and Individual Creativity Over Time: The Detrimental Effects of Past Success," Working Paper.

Bernardo, A. E., Welch, I. 2001. "On the evolution of overconfidence and entrepreneurs," *Journal of Economics and Management Strategy* (10) 301–331.

Boudreau, K., Nicola L., Lakhani, K. 2011. "Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis," *Management Science* (57:5) 843-863.

Camerer, C., Lovallo, D. 1999. "Overconfidence and Excess Entry: An Experimental Approach," *The American Economic Review* (89:1), 306-318.Crawford, G., Shum, M. 2005. "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica* (73:4) 1137–1173.

DeGroot, M. H. 1970.*Optimal Statistical Decisions*, New York, McGraw-Hill.

Di Gangi, P. M., Wasko, M., Hooker, R. 2010. "Getting customers' ideas to work for you: Learning from Dell how to succeed with online user innovation communities," *MIS Quarterly Executive* (9) 213-228.

DiPalantino, D., Vojnovic, M. 2009. "Crowdsourcing and All-Pay Auctions," Proceedings of the 10th *ACM International Conference on Electronic Commerce* 119-128.

Dunning, D., Meyerowitz, J. A., Holzberg, A. 1989. "Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability," *Journal of Personality and Social Psychology* (57) 1082-1090.

Erdem, T., Keane, M. P. 1996. "Decision-Making under Uncertainty: Capturing Dynamic Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science* (15:1) 1-20.

Erdem, T., Keane, M. P., Sun, B. 2008. "A Dynamic Model of Brand Choice When Price and Advertising Signal Product Quality," *Marketing Science* (27:6) 1111-1125.

Franke, N., von Hippel, E. 2003. "Satisfying Heterogeneous User Needs via Innovation Toolkits: the Case of Apache Security Software," *Research Policy* (32:7) 1199-1215.

Galasso, A., Simcoe, T. 2011. "CEO overconfidence and innovation," Management Science (57:8) 1469-1484.

Girotra, K., Terwiesch, C., Ulrich, K, T. 2010. "Idea Generation and the Quality of the Best Idea," *Management Science* (56: 4) 591-605.

Hauser J. R. and Urban, G. L. 1977. "A Normative Methodology for Modeling Consumer Response to Innovation", *Operations Research* (25: 4), pp. 579-619.

Herz, H., Schunk, H., Zehnder, C. 2012. "How Do Judgmental Overcondence and Overoptimism Shape Innovative Activity?" working paper.

Hill, G. 2009. "How Understanding Customer Jobs turns Crowdsourcing into Smartsourcing," *Customer Think*, http://www.customerthink.com/article/how_understanding_customer_jobs_turns_crowdsourcing_into_smartsourcing.

Howe, J. 2006. "The Rise of Crowdsourcing." *Wired Magazine* 14, 1-5.

Huang, Y., Singh, P., Ghose, A., 2010 "Show Me the Incentives: A Dynamic Structural Model of Employee Blogging Behavior" *ICIS 2010 Proceedings*.

Jeppesen, L., Lakhani, K. 2010. "Marginality and Problem-Solving Effectiveness in Broadcast Search," *Organization Science* (21) 1016-1033.

Kuan, J. 2001. "Open Source Software as Consumer Integration into Production," *Working Paper*.

Kumar, V., Gordon, B., Srinivasan, K. 2011 "Competitive Strategy for Open Source Software," forthcoming at *Marketing Science*.

Lakhani, K. R., von Hippel, E. 2003. "How Open Source Software Works: "Free" User-to-user Assistance," *Research Policy* (32:6) 923-943.

Lilien, G.L. and P. Kotler, 1983, "Marketing decision making: A model-building approach", New York: Harper and Row.

Lowe, R. A., Ziedonis, A. A. 2006. "Over-optimism and the performance of entrepreneurial firms," *Management Science* (52:2) 173–186.

Mehta, N., Rajiv, S., Srinivasan, K. 2003. "Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation," *Marketing Science* (22:1) 58-84.

Mo, J., Zheng, Z, Geng, X. 2011. "Winning Crowdsourcing Contests: A Micro-Structural Analysis of Multi-Relational Social Network," *Workshop on information Management (CSWIM)*, Harbin, China.

Narayanan, S, Manchanda. P. 2009. "Heterogeneous Learning and the Targeting of Marketing Communication for New Products," *Marketing Science* (28:3) 424-441.

Netzer, O., Lattin, J. M., Srinivasan, V. 2008. "A Hidden Markov Model of Customer Relationship Dynamics," *Marketing Science* (27: 2) 185-204.

Schulze, A., Hoegl M. 2008. "Organizational Knowledge Creation and the Generation of New Product Ideas: A Behavioral Approach, " *Research Policy* (37:10) 1742-1750.

Svenson, O. 1981. "Are we all less risky and more skillful than our fellow drivers?" Acta Psychologica (47:2) 143–148.

Terwiesch, C., Xu, Y. 2008. "Innovation Contests, Open Innovation, and Multi-agent Problem Solving," *Management Science* (54:9) 1529-1543.

Zhang, J. 2010. "The Sound of Silence: Observational Learning in the U.S. Kidney Market," *Marketing Science* (29:2) 315-335.

Online Appendix to

Crowdsourcing New Product Ideas under Consumer Learning

Appendix 1: Hierarchical Bayesian Estimation

As mentioned in the estimation strategy section, we use MCMC methods to estimate parameters in our model. To be more specific, the Gibbs sampler is applied to recursively make draws from the following conditional distribution of the model parameters:

$$\boldsymbol{\beta}_i | A_i, C_j^e, \overline{\boldsymbol{\beta}}, \boldsymbol{\alpha}$$

$$\overline{\boldsymbol{\beta}} | \boldsymbol{\beta}_i, \Sigma$$

$$\Sigma | \boldsymbol{\beta}_i, \overline{\boldsymbol{\beta}}$$

$$\boldsymbol{\alpha} | A, I, C_j^e, \boldsymbol{\beta}_i$$

$$C_{jt}^e | A, C_{jt-1}^e, C_{jt+1}^e, \boldsymbol{\alpha}, \boldsymbol{\beta}_i$$

The additional notation A_i denotes the vector of actions individual i takes in all periods, A denotes the decisions all individuals make in all periods, I denotes the decision the firm makes on all ideas posted within the observation period, $\boldsymbol{\beta}$ denotes $\boldsymbol{\beta}_i$ for all individuals, and C_j^e denotes the vector of the mean implementation cost beliefs in all periods. Further, the posterior distributions of $\boldsymbol{\beta}_i$, $\boldsymbol{\alpha}$ and C_j^e do not belong to any conjugate family, and therefore, we use the Metropolis-Hasting method to generate new draws. Each iteration involves five steps.

Step 1: Generate $\boldsymbol{\beta}_i$

The conditional distribution of $\boldsymbol{\beta}_i$ is

$$f(\boldsymbol{\beta}_i | A_i, C_j^e, \overline{\boldsymbol{\beta}}, \boldsymbol{\alpha}) \propto |\Sigma|^{-1/2} exp\left[-1/2(\boldsymbol{\beta}_i - \overline{\boldsymbol{\beta}})^{'} \Sigma^{-1}(\boldsymbol{\beta}_i - \overline{\boldsymbol{\beta}})\right] L(A_i | C_j^e, \boldsymbol{\beta}_i, \boldsymbol{\alpha})$$

Clearly, this posterior distribution does not have a closed form; therefore, we use the Metropolis-Hasting method to generate new draws with a random walk proposal density. The increment random variable is multivariate normally distributed with its variances adapted to obtain an acceptance rate of approximately 20% (Atchade, 2006). The probability that proposed $\boldsymbol{\beta}_i$ will be accepted is calculated using the following formula (the superscript $Prop$ represents the proposed new $\boldsymbol{\beta}_i$ in this current iteration, i.e., iteration r. When accept=1, $\boldsymbol{\beta}_i^{r+1} = \boldsymbol{\beta}_i^{Prop}$; otherwise, $\boldsymbol{\beta}_i^{r+1} = \boldsymbol{\beta}_i^{r}$.)

$$Pr(accept) \propto \frac{f(\boldsymbol{\beta}_i^{Prop}|A_i, C_j^e, \overline{\boldsymbol{\beta}})}{f(\boldsymbol{\beta}_i^r|A_i, C_j^e, \overline{\boldsymbol{\beta}})}$$

$$= \frac{|\Sigma|^{-1/2} exp\left[-1/2(\boldsymbol{\beta}_i^{Prop} - \overline{\boldsymbol{\beta}})^{'} \Sigma^{-1}(\boldsymbol{\beta}_i^{Prop} - \overline{\boldsymbol{\beta}})\right]L(A_i|C_j^e, \boldsymbol{\beta}_i^{Prop}, \boldsymbol{\alpha})}{|\Sigma|^{-1/2} exp\left[-1/2(\boldsymbol{\beta}_i^r - \overline{\boldsymbol{\beta}})^{'} \Sigma^{-1}(\boldsymbol{\beta}_i^r - \overline{\boldsymbol{\beta}})\right]L(A_i|C_j^e, \boldsymbol{\beta}_i^r, \boldsymbol{\alpha})}$$

Step 2: Generate $\overline{\boldsymbol{\beta}}$

$$\overline{\boldsymbol{\beta}}|\boldsymbol{\beta}_i, \Sigma \sim MVN(u, W)$$

where

$$W = (Z^{'}Z \otimes \Sigma^{-1} + W_0^{-1})^{-1}$$

$$u = W[(Z^{'} \otimes \Sigma^{-1})vec(\boldsymbol{\beta}) + W_0^{-1}u_0]$$

$$Z = vector\ of\ (1^{'}s)with\ length = N$$

$$vec(\boldsymbol{\beta}) = (\boldsymbol{\beta}_1^{'}, \boldsymbol{\beta}_2^{'}, \dots, \boldsymbol{\beta}_N^{'})$$

The priors are specified as:

$$u_0 = vector\ of\ (0^{'}s)\ with\ length = 6$$

$$W_0 = 100I_6$$

Step 3: Generate Σ

$$\Sigma|\boldsymbol{\beta}_i, \overline{\boldsymbol{\beta}} \sim IW(f_0 + N, G_0^{-1} + \Sigma_{i=1}^{N}(\boldsymbol{\beta}_i - \overline{\boldsymbol{\beta}})^{'}(\boldsymbol{\beta}_i - \overline{\boldsymbol{\beta}}))$$

where the prior hyper-parameter f_0 is set to 11, and G_0^{-1} is set to I_6.

Step 4: Generate $\boldsymbol{\alpha}$

The conditional distribution of $\boldsymbol{\alpha}$ is

$$f(\boldsymbol{\alpha}|A, I, C_j^e, \boldsymbol{\beta}) \propto |\Sigma_\alpha|^{-\frac{1}{2}} exp\left[-1/2(\boldsymbol{\alpha} - \boldsymbol{\alpha_0})^{'}\Sigma_\alpha^{-1}(\boldsymbol{\alpha} - \boldsymbol{\alpha_0})\right]L(A|C_j^e, \boldsymbol{\beta}, \boldsymbol{\alpha})L(I|\boldsymbol{\alpha})L(C_j^e)$$

where

$$L(C_j^e) = \prod_{t=1}^{T} L(C_{jt}^e|C_{jt-1}^e, \boldsymbol{\alpha})$$

Similar to what we have done for $\boldsymbol{\beta}_i$, we use the Metropolis-Hasting methods to make draws for $\boldsymbol{\alpha}$. The probability of acceptance is

$$\Pr(accept) = \frac{f(\boldsymbol{\alpha}^{prop}|A, I, C_j^e, \boldsymbol{\beta})}{f(\boldsymbol{\alpha}^r|A, I, C_j^e, \boldsymbol{\beta})}$$

$$= \frac{|\Sigma_\alpha|^{-\frac{1}{2}} exp\left[-1/2(\boldsymbol{\alpha}^{prop} - \boldsymbol{\alpha_0})' \Sigma_{\alpha_0}^{-1}(\boldsymbol{\alpha}^{prop} - \boldsymbol{\alpha_0})\right] L(A|C_j^e, \boldsymbol{\beta}, \boldsymbol{\alpha}^{prop}) L(I|\boldsymbol{\alpha}^{prop}) L(C_j^e)}{|\Sigma_\alpha|^{-\frac{1}{2}} exp\left[-1/2(\boldsymbol{\alpha}^r - \boldsymbol{\alpha_0})' \Sigma_{\alpha_0}^{-1}(\boldsymbol{\alpha}^r - \boldsymbol{\alpha_0})\right] L(A|C_j^e, \boldsymbol{\beta}, \boldsymbol{\alpha}^r) L(I|\boldsymbol{\alpha}^r) L(C_j^e)}$$

where $\boldsymbol{\alpha_0} = (0,0,\ldots 0)$ and $\Sigma_{\alpha_0}^{-1} = 100 I_8$ are diffused priors.

Step 5: Generate C_j^e

Finally, we sequentially draw C_{jt}^e for t=1 to T. The conditional distribution of C_{jt}^e is

$$f\left(C_{jt}^e | A_{jt}, C_{jt-1}^e, \boldsymbol{\alpha}, \boldsymbol{\beta}\right) \propto$$

$$|v_{jt}^2|^{-\frac{1}{2}} exp\left[-1/2\left(C_{jt}^e - \bar{C}_{jt}^e\right)'\left(v_{jt}^2\right)^{-1}\left(C_{jt}^e - \bar{C}_{jt}^e\right)\right] L(A_{jt}|C_{jt}^e, \boldsymbol{\alpha}, \boldsymbol{\beta}) L(C_{jt+1}^e | C_{jt}^e, \boldsymbol{\alpha})$$

where A_{jt} denotes the decisions all individuals make on Category j idea in period t. $\bar{C}_{jt}^e$ and v_{jt}^2 in the equation above are calculated using Equation (21) and (22). Again, because the posterior distribution does not have a close form, we have to use the Metropolis-Hasting methods to draw new C_{jt}^e.

The probability of acceptance is

$$\Pr(accept) = \frac{f(C_{jt}^{e\ Prop}|A, C_{jt-1}^e, \boldsymbol{\alpha})}{f(C_{jt}^{e\ r}|A, C_{jt-1}^e, \boldsymbol{\alpha})}$$

$$= \frac{|v_{jt}^2|^{-\frac{1}{2}} exp\left[-1/2\left(C_{jt}^{e\ Prop} - \bar{C}_{jt}^e\right)'\left(v_{jt}^2\right)^{-1}\left(C_{jt}^{e\ Prop} - \bar{C}_{jt}^e\right)\right] L(A_{jt}|C_{jt}^{e\ Prop}, \boldsymbol{\alpha}, \boldsymbol{\beta}) L(C_{jt+1}^e | C_{jt}^{e\ Prop}, \boldsymbol{\alpha})}{|v_{jt}^2|^{-\frac{1}{2}} exp\left[-1/2\left(C_{jt}^{e\ r} - \bar{C}_{jt}^e\right)'\left(v_{jt}^2\right)^{-1}\left(C_{jt}^{e\ r} - \bar{C}_{jt}^e\right)\right] L(A_{jt}|C_{jt}^{e\ r}, \boldsymbol{\alpha}, \boldsymbol{\beta}) L(C_{jt+1}^e | C_{jt}^{e\ r}, \boldsymbol{\alpha})}$$

Appendix 2: Model Identification

We now briefly discuss some intuition as to how the parameters in our model are identified. In our model the consumers make posting decisions based on their (perceived) utility. With this assumption, we can infer individual's utility derived from posting different categories of ideas from their posting decisions. The basic logic behind the identification strategy that the "true" parameters in the utility function, as well as the "true" learning parameters, will lead to a utility function that can best predict the data we observe in the reality.

In the estimation, we fix the mean cost of one category (product ideas) and the variance of individuals' initial belief about the cost distribution and potential distribution. We have to fix the mean cost of one category because if we add a constant to all Q_is and then add the same constant to all C_js, we will obtain exactly the same utility value. When we fix C_1, we will be able to identify Q_is and C_2. As a result, the estimated values of C_2 and Q_i should be interpreted as relative to C_1. We set the initial variance of individuals' initial belief about the cost distribution and potential distribution to a large value to reflect the fact that individuals' prior believe is non-informative.

The variance parameters σ_μ^2 and $\sigma_{\delta_i}^2$ are both identified from the dynamics of the posting behaviors of individuals over time. We are able to identify σ_μ^2 and $\sigma_{\delta_i}^2$ simultaneously because the signals of the implementation costs and the potentials are generated from different events. σ_μ^2 is identified through the dynamics of the choice probabilities at the population level. For example, if one idea is implemented in period t, the perceived cost of implementation for all individuals will be updated. For those who do not post in this period, their perception about the potential of their ideas has not changed before or after the period, and the changes in the probability of posting ideas after they receive the cost signal help us to identify σ_μ^2. If σ_μ^2 is very small, which means that the cost signals individuals receive are precise, then individuals can learn faster, their perceptions converge to the true value quickly, and vice versa. Similarly, the average learning speed (how much adjustment individuals make to their perceptions) of the potential of the ideas is affected by both $\sigma_{\delta_i}^2$ and the slope parameter φ. In addition, from Equation (10), we know the relationship between $\sigma_{\xi_i}^2$, the variance of the voting scores individual i's ideas receive, which can be directly estimated from the voting score data, and the variance of potential of the individuals i's ideas, $\sigma_{\delta_i}^2$, is $\sigma_{\xi_i}^2 = \varphi^2 \sigma_{\delta_i}^2$. Therefore, individuals' learning speed (how much their behavior change after receiving a potential signal) observed in the data can help us identify φ. Once φ is identified, $\sigma_{\delta_i}^2$ is also identified. Note that φ is a population level parameter. It is possible that there still remain variations in individuals' learning speed of the potential of the ideas, after controlling for $\sigma_{\delta_i}^2$. Thses remaining variations will be captured by θ_{ij}, which we will explain in detail later.

The overall frequency that an individual i posts Category j ideas is jointly determined by θ_{i0} and θ_{ij}. However, we are able to separately identify θ_{i0} and θ_{ij} because they enter the utility function in different ways. If we observe an individual who posts frequently, it could be because 1) he/she incurs low cost to post an idea; or 2) he/she receives higher payoffs when his/her Category j ideas are implemented. θ_{i0} is the constant term in the utility function, which does not change as individuals receive signals over time; while θ_{ij} is multiplied by the perceived probability of individuals' ideas being implemented. For example, when the firm implements a Category j idea and so all individuals' perceive costs of implementing Category j idea are updated. Individuals whose θ_{ij} is larger will be affected more significantly. In addition, the magnitude of θ_{ij} is also reflected in the changes in individuals posting behavior after they receive a potential signal. For example, consider two hypothetical individuals, namely A and B. From the voting score data, we find the mean and variance of their ideas' voting score are very similar. This implies that A and B updates their perception of the potential of their ideas in a similar way. However, individual A's probability of posting a Category j idea changes dramatically after she receives a new potential signal, while individual B's probability of posting Category j idea does not change a lot. The only cause of this difference is different θ_{ij}. Therefore, such variation help identify θ_{ij}. Similar logic can be applied for the identification of θ_{ij} for the same

individual. Assume that after receiving a potential signal, individual A's probability of posting a Category 1 idea changes significantly, while her probability of posting a Category 2 idea only changes slightly, we can conclude that $\theta_{A1} > \theta_{A2}$. Once θ_{ij} is controlled, θ_{i0} can be identified from the overall frequency that individual i posts ideas (after controlling for θ_{ij}). d_i can be easily identified because D_{it} is observed for every i in every period. The difference in individual i's posting behavior between cases where $D_{it} = 0$ and $D_{it} = 1$ identifies d_i. The binary construction of D_{it} can help disentangle the effects of learning and dissatisfaction.

The identification of Q_i and σ_γ^2 relies on two sets of observations. The behavior of "well-informed" individuals, whose perception about the firm's cost structure and potential of their ideas is very close to the true value, is an important source of the identification of Q_i and σ_γ^2. Note that we observe the voting score an idea receives is a linear function of the idea's potential, or $V_i = cons + \varphi Q_i$. V_i can be easily estimated by averaging all individual i's ideas' voting scores; and the identification of φ has been discussed previously. Given V_i and φ, identifying Q_i is equivalent to identifying $cons$. Consider a hypothetical "well-informed" individual's probability of posting a Category 1 idea is 0.1, i.e. $\exp(\widetilde{U}_{ijt})/(1 + \exp(\widetilde{U}_{ijt})) = 0.1$. Solving for $\widetilde{U}_{ijt}$, we get $\widetilde{U}_{ijt}$=-2.303. Given $\theta_{i0}, \theta_{i1}, d_i$ and C_1, as well as the variance parameter $\sigma_{\delta_i}^2$, Equation (19) is an equation of two unknown parameters is Q_i and σ_γ^2, or equivalently $cons$ and σ_γ^2. Another source of identifying Q_i and σ_γ^2 is the likelihood of observed implementation decisions on all Category 1 ideas. From our dataset, we observe the decisions the firm makes on each idea, given its voting score. In Equation 18, Q_{mjt} can be calculated by $Q_{mjt} = (V_{mjt} - cons)/\varphi$. Assume that a Category 1 idea with log-voting score equally 2 has 0.01 chance to be implemented, then $[(V_{mjt} - cons)/\varphi + C_1]/\sigma_\gamma^2 = $ -2.326. Given V_{mjt} and φ, it is also an equation with two unknowns parameters $cons$ and σ_γ^2. Combining these two constraints, Q_i (or equivalently $cons$) and σ_γ^2 can be identified. Once Q_i is identified, C_2 can be identified through the probabilities that individuals post Category 2 ideas and the firm's decisions on Category 2 ideas, given the votes each idea receives. C_0 can be identified through the probability of posting in the first seven weeks as no idea was implemented before the seventh week. In these seven weeks, individuals have not received any cost signals, and their beliefs about the cost structure stay at C_0, but they receive signals about the potential of their ideas when they post. Given Q_i, C_0 can be easily identified. Given C_j and C_0, Q_0 can be identified through the probability of posting for the latecomers throughout the whole observation period. Before an individual posts any ideas on the website for the first time, his/her beliefs about his/her idea's potential is always Q_0, while his/her beliefs about the implementation cost is updated. Given the different C_{jt}^e for different t's, Q_0 can then be identified.

Appendix 3: Derivation of the Updating Rules

We begin with the Bayes rule. The Bayes rule is

$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \propto P(B|A)P(A)$. Now let us explain how we use the Bayes rule in coming up with our updating functions. In the Bayesian updating process, A represents people's belief about a certain parameter, B represents signal. Let us begin with the learning process of the implementation cost. Assume that individuals' prior belief about the mean of implementation cost in period t follows a normal distribution N $(C_{jt-1}^e, \sigma_{C_{jt-1}}^2)$ and the cost signal individuals receive in period t is $C_{kjt} \sim N(C_j, \sigma_\mu^2)$. The updated (posterior) distribution of the cost distribution is $N(C_{jt}^e, \sigma_{C_{jt}}^2)$.

The prior of the mean implementation cost in period t follows a normal distribution N $(C_{jt-1}^e, \sigma_{C_{jt-1}}^2)$, and this is similar to the term $P(A)$ in the equation above. As we are dealing with a continuous distribution, instead of probability mass P, we use the probability density function of the normal distribution N $(C_{jt-1}^e, \sigma_{C_{jt-1}}^2)$ below:

$$p\left(C_j \middle| C_{jt-1}^e, \sigma_{C_{jt-1}}^2\right) = \left(2\pi\sigma_{C_{jt-1}}^2\right)^{-\frac{1}{2}} exp\left[-\frac{1}{2\sigma_{C_{jt-1}}^2}\left(C_j - C_{jt-1}^e\right)^2\right].$$

As we assume that the cost signal C_{kjt} follows a normal distribution $C_{kjt} \sim N(C_j, \sigma_\mu^2)$, the probability density of observing a cost signal of a value C_{kjt} is:

$$p(C_{kjt}|C_j) = \left(2\pi\sigma_\mu^2\right)^{-\frac{1}{2}} exp\left[-\frac{1}{2\sigma_\mu^2}\left(C_{kjt} - C_j\right)^2\right]$$

Let $D = (C_{1jt}, \dots, C_{k_{Cjt}jt})$ be the cost signals individuals receive in period t, with k_{Cjt} indicating the number of such signals. The likelihood of observing D, given C_j and σ_μ^2 is simply the product of the $p(C_{kjt}|C_j)$ over $k=1$ to k_{Cjt}.

$$p(D|C_j) = \prod_{k=1}^{k_{Cjt}} p(C_{kjt}|C_j) = \left(2\pi\sigma_\mu^2\right)^{-\frac{k_{Cjt}}{2}} exp\left[-\frac{1}{2\sigma_\mu^2}\sum_{k=1}^{k_{Cjt}}\left(C_{kjt} - C_j\right)^2\right]$$

$$\propto exp\left[-\frac{1}{2\sigma_\mu^2}\sum_{k=1}^{k_{Cjt}}\left(C_{kjt} - C_j\right)^2\right]$$

Here, $p(D|C_j)$ is similar to $P(B|A)$ in the first equation. Finally, $p(C_j|D)$ corresponds to $P(A|B)$. Following the Bayes rule $P(A|B) = \propto P(B|A)P(A)$, the posterior of mean cost of implementation is

$$p(C_j|D) \propto p(D|C_j)p\left(C_j\middle|C_{jt-1}^e, \sigma_{C_{jt-1}}^2\right)$$

$$\propto exp\left[-\frac{1}{2\sigma_\mu^2}\sum_{k=1}^{k_{Cjt}}\left(C_{kjt} - C_j\right)^2\right]exp\left[-\frac{1}{2\sigma_{C_{jt-1}}^2}\left(C_j - C_{jt-1}^e\right)^2\right]$$

$$= exp\left[-\frac{1}{2\sigma_\mu^2}\sum_{k=1}^{k_{Cjt}}\left(C_{kjt}^2 + C_j^2 - 2C_{kjt}C_j\right)^2 - \frac{1}{2\sigma_{C_{jt-1}}^2}\left(C_j^2 + C_{jt-1}^{e}{}^2 - 2C_jC_{jt-1}^e\right)^2\right]$$

$$\propto exp\left[-\frac{C_j^2}{2}\left(\frac{1}{\sigma_{C_{jt-1}}^2} + \frac{k_{Cjt}}{\sigma_\mu^2}\right) + C_j\left(\frac{C_{jt-1}^e}{\sigma_{C_{jt-1}}^2} + \frac{\sum_{k=1}^{k_{Cjt}}C_{kjt}}{\sigma_\mu^2}\right) - \left(\frac{C_{jt-1}^{e}{}^2}{2\sigma_{C_{jt-1}}^2} + \frac{\sum_{k=1}^{k_{Cjt}}C_{kjt}^2}{2\sigma_\mu^2}\right)\right] \quad (*)$$

So far, we have derived the posterior distribution of C_j in terms of the prior and the signals. The posterior is also normally distributed and parameterized by two parameters—mean and variance. Therefore, if we can recover the mean and the variance of the posterior distribution, we can fully describe the posterior distribution. We do this in the following steps. Given C_{jt}^e and $\sigma_{C_{jt}}^2$

$$p(C_j|D) = \left(2\pi\sigma_{C_{jt}}^2\right)^{-\frac{1}{2}} exp\left[-\frac{1}{2\sigma_{C_{jt}}^2}\left(C_j - C_{jt}^e\right)^2\right] \propto exp\left[-\frac{1}{2\sigma_{C_{jt}}^2}\left(C_j{}^2 + C_{jt}^{e\,2} - 2C_jC_j^e\right)^2\right]. (**)$$

This is just the definition of the posterior distribution. That is, Equation (*)≡Equation (**). To connect prior distribution and signals with posterior variance $\sigma_{C_{jt}}^2$, we match coefficients of $C_j{}^2$ in Equation (*) and Equation (**). We then have

$$\frac{-C_j{}^2}{2\sigma_{C_{jt}}^2} = \frac{-C_j{}^2}{2}\left(\frac{1}{\sigma_{C_{jt-1}}^2} + \frac{k_{Cjt}}{\sigma_\mu^2}\right)$$

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

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

which is Equation (7) in the paper. Similarly, we match coefficients of C_j in equation (*) and equation (**), and then get

$$\frac{2C_jC_{jt}^e}{2\sigma_{C_{jt}}^2} = C_j\left(\frac{C_{jt-1}^e}{\sigma_{C_{jt-1}}^2} + \frac{\sum_{k=1}^{k_{Cjt}}C_{kjt}}{\sigma_\mu^2}\right)$$

$$\frac{C_{jt}^e}{\sigma_{C_{jt}}^2} = \left(\frac{C_{jt-1}^e}{\sigma_{C_{jt-1}}^2} + \frac{\sum_{k=1}^{k_{Cjt}}C_{kjt}}{\sigma_\mu^2}\right)$$

$$C_{jt}^e = \left(\frac{C_{jt-1}^e}{\sigma_{C_{jt-1}}^2} + \frac{\sum_{k=1}^{k_{Cjt}}C_{kjt}}{\sigma_\mu^2}\right)\sigma_{C_{jt}}^2$$

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

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

$$= C^e_{jt-1}\left(1 - \frac{k_{Cjt}\sigma^2_{C_{jt-1}}}{k_{Cjt}\sigma^2_{C_{jt-1}} + \sigma^2_\mu}\right) + C_{sjt}\left(\frac{k_{Cjt}\sigma^2_{C_{jt-1}}}{k_{Cjt}\sigma^2_{C_{jt-1}} + \sigma^2_\mu}\right)$$

$$= C^e_{jt-1} + \left(C_{sjt} - C^e_{jt-1}\right)\frac{\sigma^2_{C_{jt-1}}}{\sigma^2_{C_{jt-1}} + \frac{\sigma^2_\mu}{k_{Cjt}}}$$

which is Equation (6) in the paper.

The proof of the updating rules of V^e_{it} and $\sigma^2_{V_{it-1}}$ is almost identical to the proof we derive above. The derivation of the updating rules of Q^e_{it} and $\sigma^2_{Q_{it-1}}$ is only slightly different. Assume for a moment that people directly observe the potential signals Q_{sit}, then the updating rules for Q_i will be

$$\sigma^2_{Q_{it}} = \frac{1}{\frac{1}{\sigma^2_{Q_{it-1}}} + \frac{k_{Qit}}{\sigma^2_{\delta\, i}}}$$

$$Q^e_{it} = Q^e_{it-1} + (Q_{sit} - Q^e_{it-1})\frac{\sigma^2_{Q_{it-1}}}{\sigma^2_{Q_{it-1}} + \frac{\sigma^2_{\delta\, i}}{k_{Qit}}}$$

However, in reality, the potential signal Q_{sit} is not directly observed. Instead, individuals observe the voting score V_{kit} and then use the linear relation between V_{kit} and Q_{kit} to recover the potential signal Q_{kit}. As in the paper, we assume the relationship between V_{kit} and Q_{kit} as

$$V_{kit} = \text{cons} + \varphi Q_{kit}$$

$$V_{kit} = V_i + \xi_{kit}$$

$$\sigma^2_{\xi\, i} = \varphi^2 \sigma^2_{\delta\, i}$$

Therefore, $Q_{kit} = (V_{kit} - cons)/\varphi$, $Q^e_{it-1} = (V^e_{it-1} - cons)/\varphi$ and $\sigma^2_{\delta\, i} = \sigma^2_{\xi\, i}/\varphi^2$. Now the two updating rules discussed above can be rewritten as:

$$\sigma^2_{Q_{it}} = \frac{1}{\frac{1}{\sigma^2_{Q_{it-1}}} + \frac{k_{Qit}}{\sigma^2_{\delta\, i}}} = \frac{1}{\frac{1}{\sigma^2_{Q_{it-1}}} + \frac{\varphi^2 k_{Qit}}{\sigma^2_{\xi\, i}}}$$

$$Q^e_{it} = Q^e_{it-1} + (Q_{sit} - Q^e_{it-1})\frac{\sigma^2_{Q_{it-1}}}{\sigma^2_{Q_{it-1}} + \frac{\sigma^2_{\delta\, i}}{k_{Qit}}}$$

$$= Q^e_{it-1} + \left(\frac{V_{kit} - cons}{\varphi} - \frac{V^e_{it-1} - cons}{\varphi}\right)\frac{\sigma^2_{Q_{it-1}}}{\sigma^2_{Q_{it-1}} + \frac{\sigma^2_{\xi\, i}}{\varphi^2 k_{Qit}}}$$

$$= Q_{it-1}^e + (V_{sit} - V_{it-1}^e) \frac{\varphi \sigma_{Q_{it-1}}^2}{\varphi^2 \sigma_{Q_{it-1}}^2 + \frac{\sigma_{\xi_i}^2}{k_{Qit}}}$$

These are equation (17) and (15) in the paper.

In this learning model, we impose two assumptions. First, the implementation cost is normally distributed which is a continuous distribution. Second, we assume individuals' prior belief about the mean implementation cost is also normally distributed, which is also a continuous distribution. The normal prior assumption provides tractability benefits, as the posterior will also have a closed form representation. For this reason, in learning literature, most models use this formulation. This type of learning model has some nice features. For example, from Equation (7), $\sigma_{C_{jt}}^2 = \frac{1}{\frac{1}{\sigma_{C_{jt-1}}^2} + \frac{k_{Cjt}}{\sigma_\mu^2}}$, we can see that $\sigma_{C_{jt}}^2$ is monotonically decreasing.

This means that as individuals receive more signals, the variance of the posterior distribution keeps decreasing, and so their uncertainty is reduced. From Equation (6), $C_{jt}^e = C_{jt-1}^e + \left(C_{sjt} - C_{jt-1}^e\right) \frac{\sigma_{C_{jt-1}}^2}{\sigma_{C_{jt-1}}^2 + \frac{\sigma_\mu^2}{k_{cjt}}}$, we can see that individuals' new belief about the mean of the cost distribution is affected by their prior belief and the new signal they receive. Individuals adjust their belief by comparing the new signals they receive and their prior belief. $\frac{\sigma_{C_{jt-1}}^2}{\sigma_{C_{jt-1}}^2 + \frac{\sigma_\mu^2}{k_{cjt}}}$ tells us the weight individuals assume to the new signals. $\frac{\sigma_{C_{jt-1}}^2}{\sigma_{C_{jt-1}}^2 + \frac{\sigma_\mu^2}{k_{cjt}}}$ is always between 0 and 1. σ_μ^2 represents the variance of the signals. When σ_μ^2 is small, which means that the signal is precise, individuals assign a larger weight to the new signals and so their beliefs get updated faster. In addition, when σ_μ^2 is fixed, the weight assigned to the difference is bigger when the variance of the prior $(\sigma_{C_{jt-1}}^2)$ is large. This indicates that individuals learn very quickly in the beginning. As $\sigma_{C_{jt}}^2$ becomes smaller individuals' learning progress slows down, their belief will tend to stabilize. These features match individuals' real-world behavior well.

Appendix 4 Convergence of the Markov Chain

In our model, we have two sets of parameters and we will show the convergence of the chains for the two sets of parameters separately. Parameter vector $\alpha = [\, C_0, C_2, \sigma_{\gamma 1}^2\, \sigma_{\gamma 2}^2, \sigma_\mu^2, Q_0, cons, \varphi]$ is common across individuals, while parameter vector $\beta_i = \beta_i = [Q_i, log(\sigma_{\delta_i}^2), d_i, \theta_{i0}, \theta_{i1}, \theta_{i2}]$ is heterogeneous across individuals. We further assume that β_i follows the following distribution

$$\beta_i = \begin{pmatrix} Q_i \\ log(\sigma_{\delta_i}^2) \\ d_i \\ \theta_{i0} \\ \theta_{i1} \\ \theta_{i2} \end{pmatrix} \sim MVN(\bar{\beta}, \Sigma)$$

where $\bar{\beta}$ denotes the mean of β and Σ denotes the variance and covariance matrix of β.

We plot the series of draws of α and $\bar{\beta}$ separately. The Markov chain was run a total of 45,000 iterations, and plotted is every 30th draw of the chain. The figures indicate that chain converged after about 9,000 iterations. And the convergence of $\bar{\beta}$ is slightly faster.