Analysis of the Rank-Order Mechanism for User-Generated Content through Simulations of Game Play

Honors Thesis in Statistics at UC Berkeley

Weijian Han

November 23, 2015

I. Introduction

This honors thesis will look at the ranking mechanisms of user-generated contents on the internet from a game-theoretic perspective. The first half of this thesis will look at an article published on the Journal of Economic Theory by Arpita Ghosh and Patrick Hummel, called “A Game-Theoretic Analysis of Rank-Order Mechanisms for User-Generated Content”, and attempt to replicate its results through repeated simulations. The second half will use a simplified model to study how variations in the payoff function affect the agents’ behaviors.

In Ghosh and Hummel’s paper, the set-up of the rank-order mechanism is as follows: Contributors of contents are strategic agents who benefit from attention and have a cost to quality. Viewers of the contents can either give each content a positive review or a negative one, and the contents are ranked based on the proportion of positive reviews they receive. Each contributor (agent) chooses whether or not to contribute and the quality of the contribution to maximize their payoff. The paper theorizes that in repeated plays of this game, a symmetric mixed strategy Nash equilibrium will emerge, where all agents will participate with the same probability and choose their qualities from a common distribution. Furthermore, the paper theorizes that with a sufficiently large total attention level, every agent will tend toward contributing at the highest possible quality.

The actual implementation of the rank-order mechanism is as follows: In each round, a total of *K = 10* agents each contribute a piece of content. Each piece of content has a quality *V* associated with it, where $0 \leq V \leq 1$. This quality value is the probability of a viewer giving the content a positive review. Each contribution is rated by a given number of viewers (selected as *T = 10,000* for the rest of this paper), and the proportion of positive viewers is used to rank the content among all of the contributions. Each agent also incurs a cost *c(V)* associated with the quality of their contribution, and this cost function is convex, continuously differentiable, increasing with respect to quality, and when the quality is zero, the function is positive and has zero derivative. As quality approaches one, the cost function increases infinitely. Each agent gains a benefit *v(A)* depending on the attention *A* they receive for each contribution. The attention each contribution receives is a function of the contribution’s ranking and the total amount of attention available. The benefit function is concave, continuously differentiable, increasing with respect to *A*, and zero when *A = 0*.

In each round, each agent decides on whether or not to contribute, and if so, the quality for their contribution. Each agent incurs a net payoff equal to the benefit from attention minus the cost of contributing. The agents try to maximize payoff with a hill-climbing approach by slightly changing the quality of their contributions and probability of contributing each round in the direction of increasing payoffs.

According to Ghosh and Hummel’s paper, as long as the above conditions are met, a symmetric mixed strategy equilibrium should emerge in the probability of participation and a common distribution of contribution quality.

II. Simulation Approach

In an attempt to simulate the problem as described in Ghosh and Hummel's paper, I wrote a program in Java to display how the qualities of the contributions change over time. The following are some of the specifications I chose, as well as alterations from the original game setting:

1. Cost Function

In the original problem, the cost function has to satisfy the following constraints:

- Convex
- Continuously differentiable
- Increasing in quality *V*
- Positive and has zero derivative at *V = 0*
- Approaches infinity as *V* approaches one

This is the cost function I chose:

$$c(V) = 50 * \tan(V * \frac{\pi}{2})$$

(This cost function does not satisfy one of the assumptions in Ghosh and Hummel's paper. Namely, the derivative at *V = 0* is not zero.)

2. Benefit Function

In the original problem, the benefit function has to satisfy the following constraints:

- (Weakly) concave
- Continuously differentiable
- Increasing in attention *A*
- Zero at *A = 0*
- Derivative bounded away from zero for all values of *A*

This is the benefit function I chose:

$$v(A) = A$$

3. Contribution Quality

In the original paper, each agent decides whether or not to contribute, and if so, what quality to contribute in. In my set up, in every round, each agent decides on two values: *p* and *q*, where *p* is the probability of contributing in this round, and the quality of the contribution is derived from *q*: $V(q) = binom(1000, q)/1000$. The parameter *q* is essentially an estimate of *V*, because the distribution of $binom(1000, q)/1000$ is approximately normal, centered at *q*. This set up is chosen because in the original paper, the mixed strategy equilibrium refers to the convergence of the distribution of quality values, and simply using a value *q* as the quality values cannot achieve this convergence, since the quality value would be deterministic rather than a probabilistic distribution. The original paper by Ghosh and Hummel did not specify the type of probabilistic distribution for the quality value, and this set up would be putting the additional constraint that the distribution is binomial.

4. Attention

The total attention is distributed among all of the agents who have made a contribution in each round. The agent that had the highest ranked contribution gets 10% of all of the total attention, and the next ranked contribution allocates 10% of the remaining attention. This process continues until all agents have received some attention, and it repeats again, starting from the highest ranked contribution, until all of the attention are distributed.

III. Convergence of q for a Fixed p

The first thing I tried to do was fixing the value for *p* and seeing if the values for *q* converges over time. I updated the *q* values using a learning algorithm. There are two learning rates that change over time: α and β. These two rates represent two different ways of learning respectively: looking at how payoff changes when changing the parameter *q*, and looking at what other agents chose for *q*, and which *q* values achieved the maximum payoff. The update function is as follows:

$$q_{t+1} = q_t + \alpha * D * R + \beta * (q_t^{max} - q_t) * p$$

$$D = \begin{cases} 0 & if\ a\ contribution\ was\ not\ made \\ 1 & if\ payoff_t > payoff_{t-1}\ and\ q_t > q^* \\ & or\ payoff_t < payoff_{t-1}\ and\ q_t < q^* \\ -1 & otherwise \end{cases}$$

$(q^*\ is\ the\ previous\ q\ that\ resulted\ in\ a\ contribution)$

$R: Random\ number\ between\ 0\ and\ 1$

$q_t^{max}: The\ q\ value\ that\ achieved\ the\ highest\ payoff\ among\ all\ agents$

$$\alpha = \begin{cases} 0.02 & if\ t < 5{,}000 \\ \frac{100}{t} & if\ t \geq 5{,}000 \end{cases}$$

$$\beta = 0.005 \quad if\ t < 20{,}000$$

$$\frac{100}{t} \qquad if\ t \geq 20{,}000$$

This update function puts weight on both an agent's performance over time as well as how they perform relative to other agents. Initially, more weight is put on improving an agent's own performance, but after time 20,000, the weights for both factors become equally important. The intuition here is that initially, all agents are exploring different *q* values and going toward the direction of higher payoff while slightly being pulled toward the value that achieved the highest payoff. As time elapses and the *q* values converge toward the same general area, both learning rates asymptotically approach zero, which means the *q* values will be more stable and converge to a more precise value. The second term is also multiplied by *p* because the first term is only non-zero if a contribution was made, so for lower values of *p*, the second term becomes more dominant, so I decreased its weight by multiplying it by *p*. This approach worked quite well for the setup of 10 agents, 1000 total attention, and a fixed *p* value of 1:

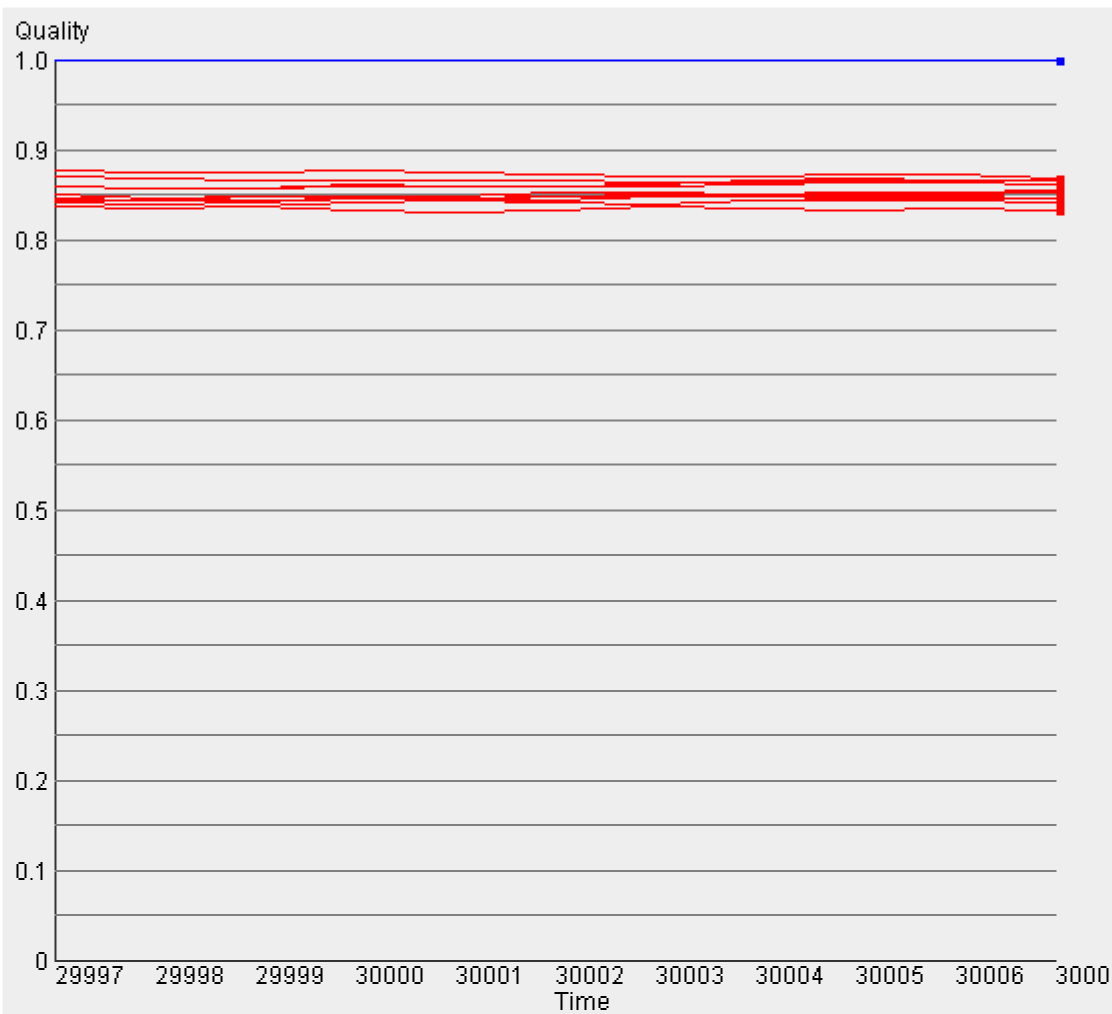

Figure 1*: Convergence of the ten agents' p and q values after 30,000 rounds*

The blue line indicates the p value of all ten agents (it is actually ten blue lines falling on the same value), and the red lines represent their q values. Three simulations of $t = 30{,}000$ yielded the following three average q values: 0.852, 0.864, and 0.863.

I went on to test the different values of p and plotted the q values that it converged to (the mean q value after 30,000 iterations). For lower values of p, the variance of the q values become much bigger, so it is hard to see what the actual relationship is between p and q, which is why I stopped running the simulations with p less than 0.6.

P	1	0.9	0.8	0.7	0.6
Q	0.852	0.836	0.844	0.823	0.824
	0.864	0.844	0.844	0.815	0.782
	0.863	0.844	0.841	0.82	0.812

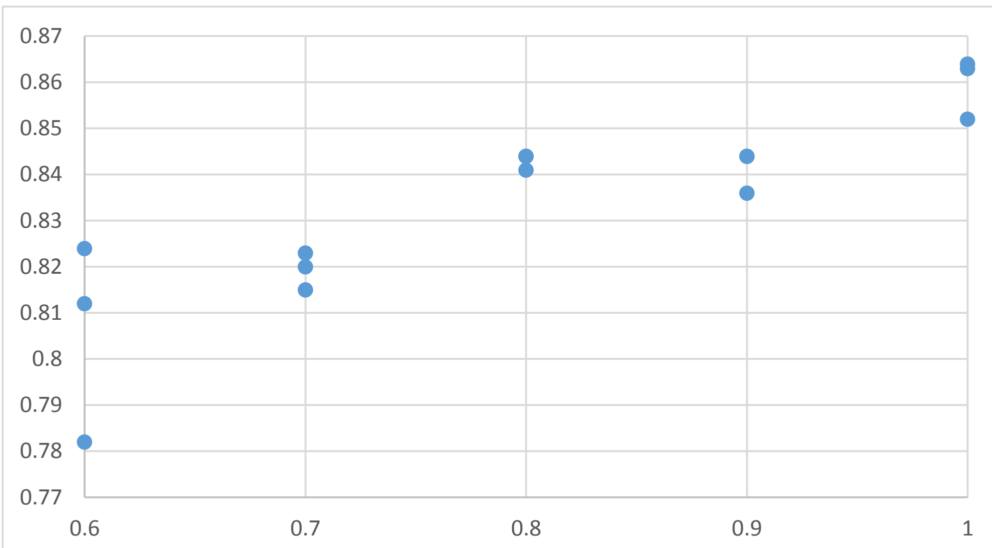

***Figure 2**: Average q values converged to for each chosen p value between 0.6 and 1*

For a fixed total attention level, when fixing the p value for all of the agents, the q values tend to converge, reaching a mixed strategy Nash equilibrium, as proposed by Ghosh and Hummel's paper. To see how this converge happens, we can look more closely at how each individual agent reacts to the strategies of the other agents.

IV. Convergence of the *k*th Agent

By the definition of Nash equilibrium, when all agents are playing with the strategy that achieves the mixed strategy Nash equilibrium, no single agent can improve his own expected payoff by deviating from this strategy. So to test whether or not a pair of p and q values is a Nash equilibrium, I fixed the first *k-1* agents to always play with the said p and q values and allowed the *k*th agent to freely choose his strategy. The goal is to see whether or not the *k*th agent will play with the same strategy as the rest of the agents in the long run.

First, I fix the *p* and *q* values of the *k-1* agents to 0.5 and 0.6 respectively. There are two observations:

1. The *k*th agent tends to continue to contribute at roughly the same *p* and *q* values, always slightly higher than everyone else, but not too high.
2. If the *k*th agent starts contributing at *p* and *q* values below that of the other *k-1* agents, it continues to contribute at the lower *p* and *q* values and even decreases them over time.

Thus, I suspect that the strategy pair (0.5, 0.6) is not a Nash equilibrium, and when everyone else is participating with this strategy, the *k*th agent can maximize its payoff by either making a slightly better contribution than everyone else or not contributing at all.

I wrote a program to calculate the expected payoffs of choosing different *q* values for the *k*th agent, while fixing the *p* values of everyone and the *q* values of the *k-1* agents at *p** and *q**. For $(p^*, q^*) = (0.5, 0.6)$, the following is the plot of the expected payoffs with respect to *q*:

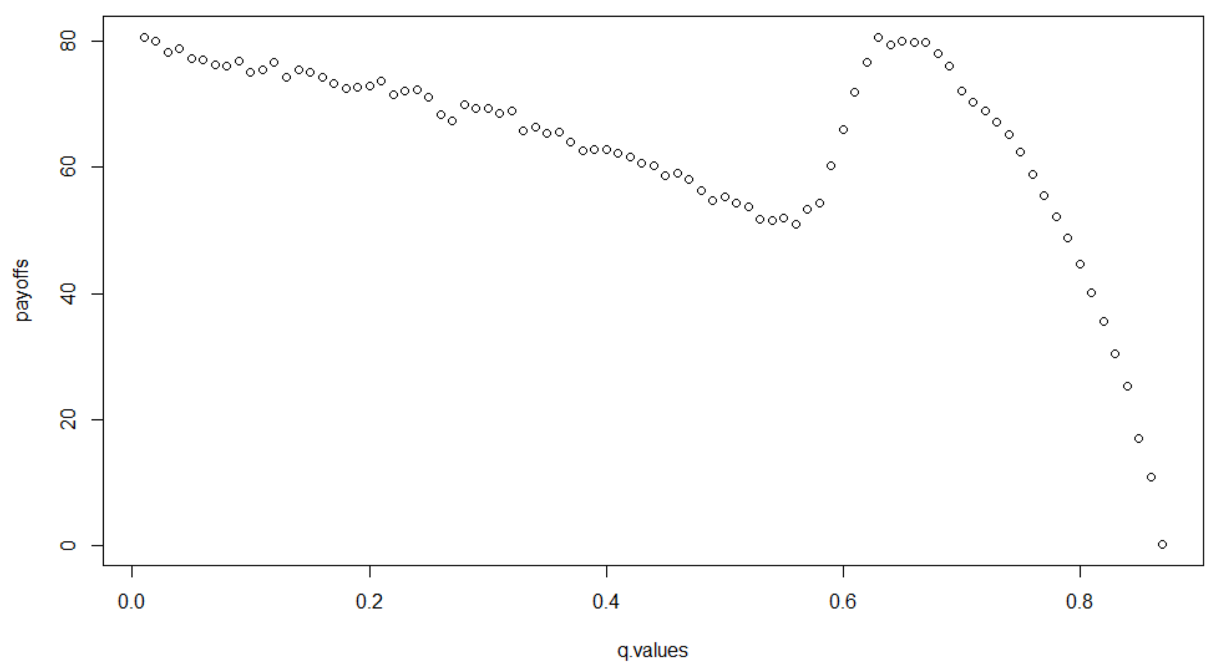

Figure 3: *Expected payoff of the kth agent for choosing different q values while the other k-1 agents are fixed at q = 0.6 and all k agents have a fixed p = 0.5*

This graph supports the previous observations that when everyone else is fixed at a *q** of 0.6, the *k*th agent can maximize its payoff by either choosing a slightly higher *q* or a very low *q*, thus the strategy pair (0.5, 0.6) is not a Nash equilibrium.

From the results of Section III, I hypothesize that the strategy pair (0.8, 0.84) could be very close to a Nash equilibrium, and using this as the parameter for the same simulations as done previously, the results are as follows:

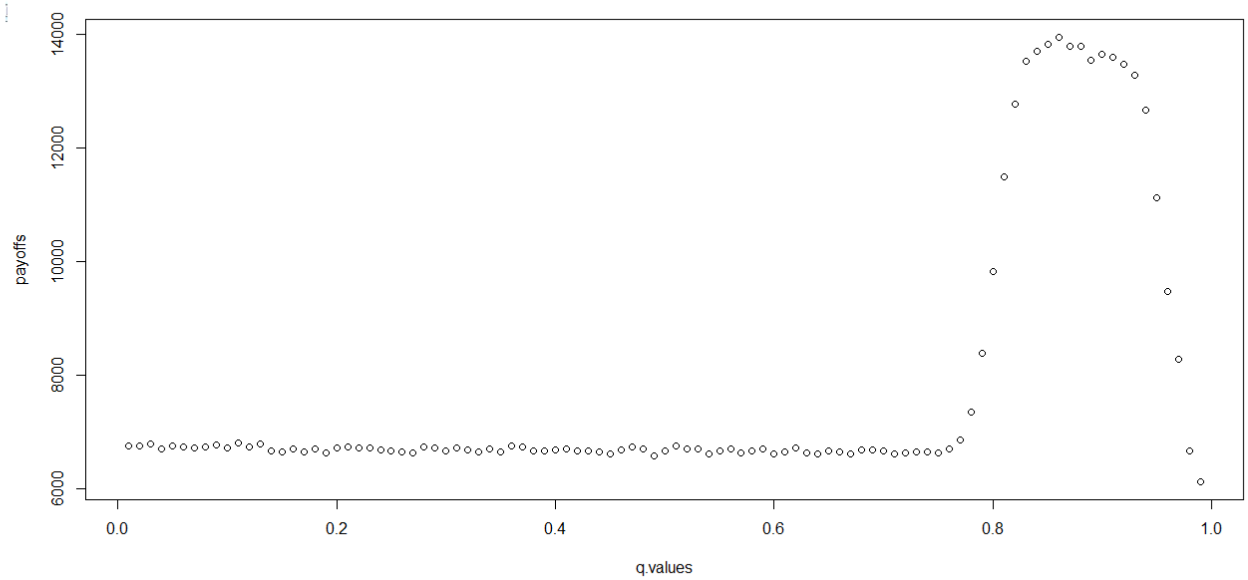

Figure 4*: Expected payoff of the kth agent for choosing different q values while the other k-1 agents are fixed at q = 0.84 and all k agents have a fixed p = 0.8*

This result verifies the hypothesis that the strategy pair (0.8, 0.84) is very close to a Nash equilibrium, because when all *k* agents are fixing their choice of *p* at the *p** value of 0.8 and *k-1* agents are contributing at *q* = *q** = 0.84, the choice of *q* that maximizes the *k*th agent's payoff is very close to *q**.

Now that the existence of a Nash equilibrium is verified for a fixed attention level, the next step is to test whether or not the agents' p and q values would converge to 1 when the total attention level increases indefinitely.

V. Test of Convergence with Increasing Total Attention

Taking a similar approach as the simulations in Section III, with the only adjustment of increasing total attention by 10 points per round, or 0.01% of the initial total attention level, we can see that the *p* and *q* values of all agents converge to 1 after 10,000 runs.

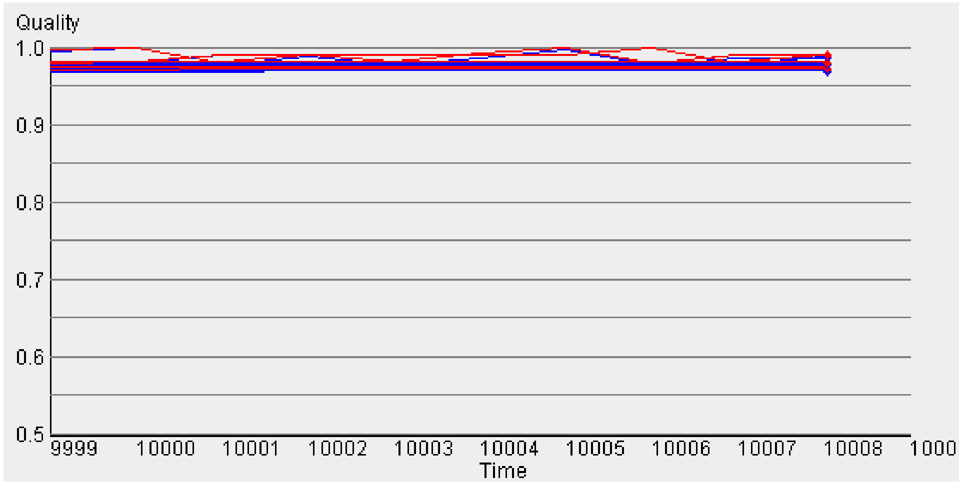

Figure 5: *Convergence of all agents' p and q values after 10,000 rounds with increasing total attention*

The same simulations from Section IV, fixing all agents' *p* values at 0.5 and the *k-1* agents' *q* values at 0.6, when done with an increasing total attention level, yields the following result:

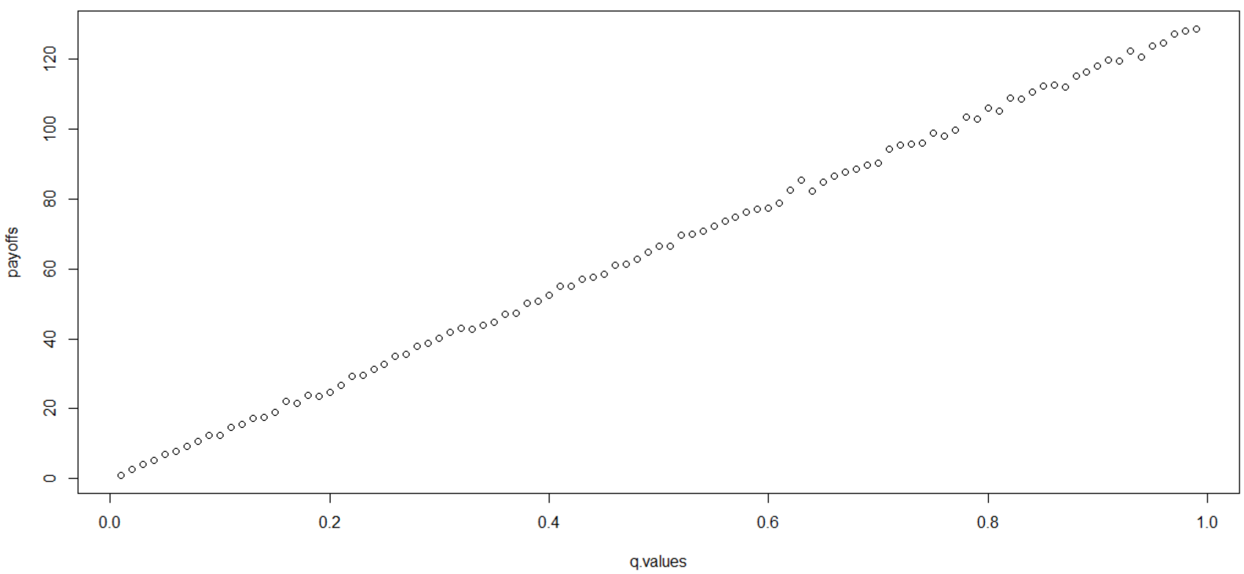

Figure 6: *Expected payoff of the kth agent for choosing different q values while all k agents fixed p at 0.5 and the other k-1 agents fixed q at 0.6*

This result supports the idea that when the total attention level increases indefinitely, the optimal strategy for any agent is to contribute with the highest quality. This result has realistic implications: as an increasing number of people visit a certain website or blog, as long as the contributors to the site has some kind of incentive to provide content to attract viewers, the quality of the contributions will keep improving to reach the best of the contributors' ability.

VI. A New Design for Modeling the Evolution of User-Generated Content

In the previous model, proposed in Ghosh and Hummel's paper "A Game-Theoretic Analysis of Rank-Order Mechanisms for User-Generated Content", the contribution quality is taken as a distribution, although the exact distribution is not specified in the paper. This creates an extra level of uncertainty, and when a contributor updates his choices for the next contribution, he can only update the probability of contributing and the distribution of the quality of the next contribution rather than directly choosing the quality as a fixed value. For the next part of the analysis, I will simplify the model such that the contributor will have full control of the quality with no chance involved, and every contributor will always produce some content every turn.

In this new model, each contributor starts by initiating the random variable *V*, which is the same as the quality value of the previous models – the probability of getting a positive review. Then, the number of positive reviews are simulated through the *binomial(T, V)* distributions, where *T* is the total number of viewers, which is constant for all of the contributors. The contributions are ranked, and the total attention is distributed among the contributors.

After the first round, the contributors will adjust their contribution qualities every round in order to achieve higher payoffs. Each contributor will consider two alternative qualities, one slightly higher than the previous quality, and one slightly lower. He will then simulate the payoffs of using each alternative by holding the qualities of the other contributors to be constant and same as the previous round. Each contributor will pick the quality that yields the higher payoff from the simulation for the next contribution. The contributions are then ranked and the attention distributed, and this process is repeated until the contribution qualities converge.

Using this methodology, I intend to test several payoff functions to study the behavior of how the contribution qualities change over time. First, I set the cost function to be a constant 0, while varying the benefit function to be different powers of the attention received:

$$v(A) = A^x$$

I observe the path of the average *V* value over repeated rounds of simulation. It is intuitive that when the cost of contributing is zero, the best action is to contribute at the highest quality possible every time. This behavior is certainly observed, as shown in the plots below. More interestingly though, the speed at which the contributors converge to the maximum quality does not seem to depend on the benefit function. Through repeated simulations, I find that the average *V* increases at each round with probability ~0.75, regardless of the benefit function. This might be because *A* is not a deterministic function of *V*, so there's a chance that choosing a higher *V* yields a lower *A*.

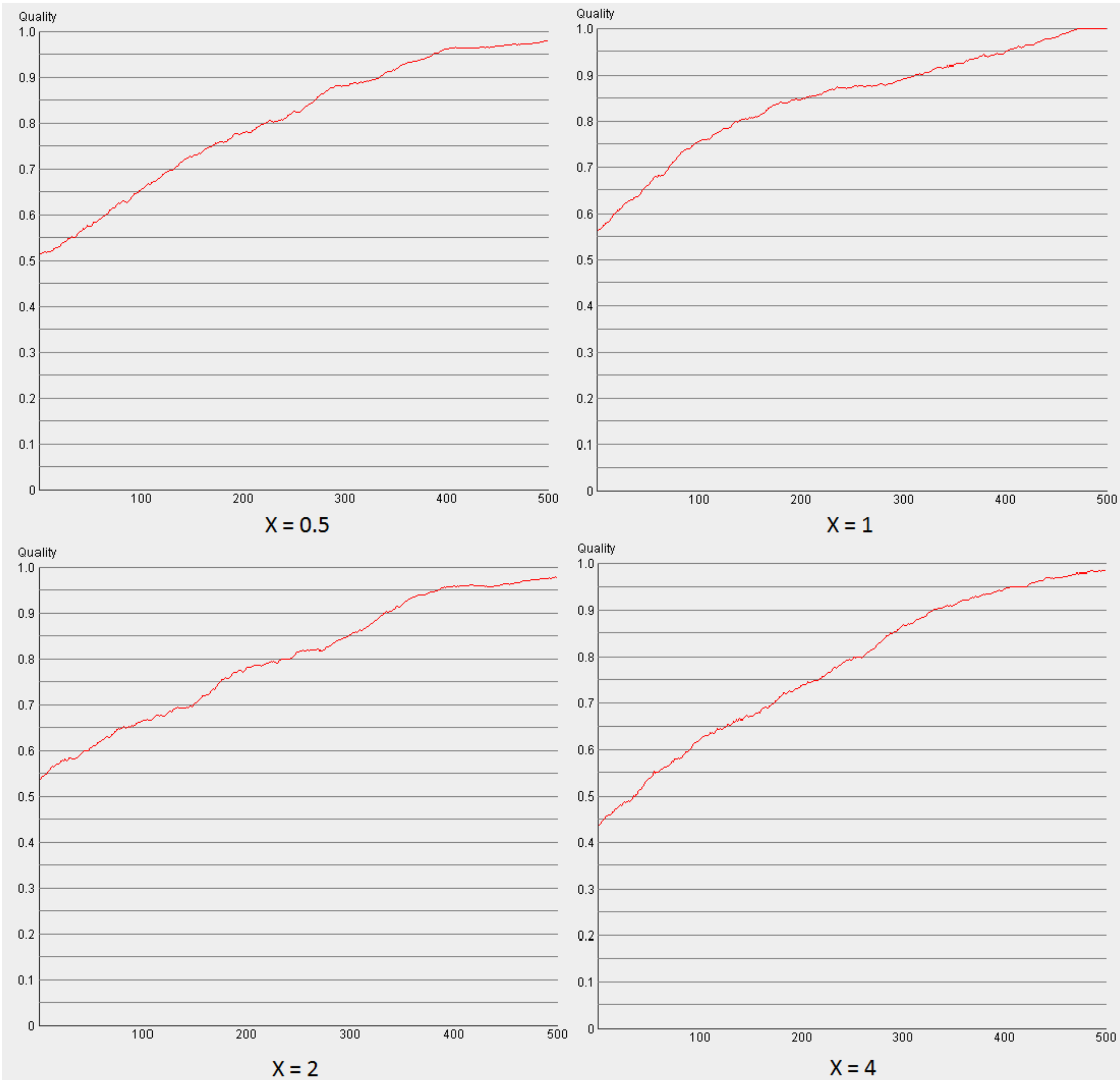

Figure 7: *Path of the average V value over time under different benefit functions*

Focusing on the cost function now, I arbitrarily set the benefit function to be linear with respect to attention received, and I proceed to test different cost functions to see how that affects the way the average *V* value changes over time.

I chose two different convex functions to model cost: tangent (the cost function used in sections III through V of this paper) and logit. I chose these two functions because they can be scaled to be 0 when the argument is 0, and approach infinity as the argument asymptotically approaches 1.

Comparing the path of the average *V* value over time between the two potential cost functions while setting the benefit function to be linear in attention:

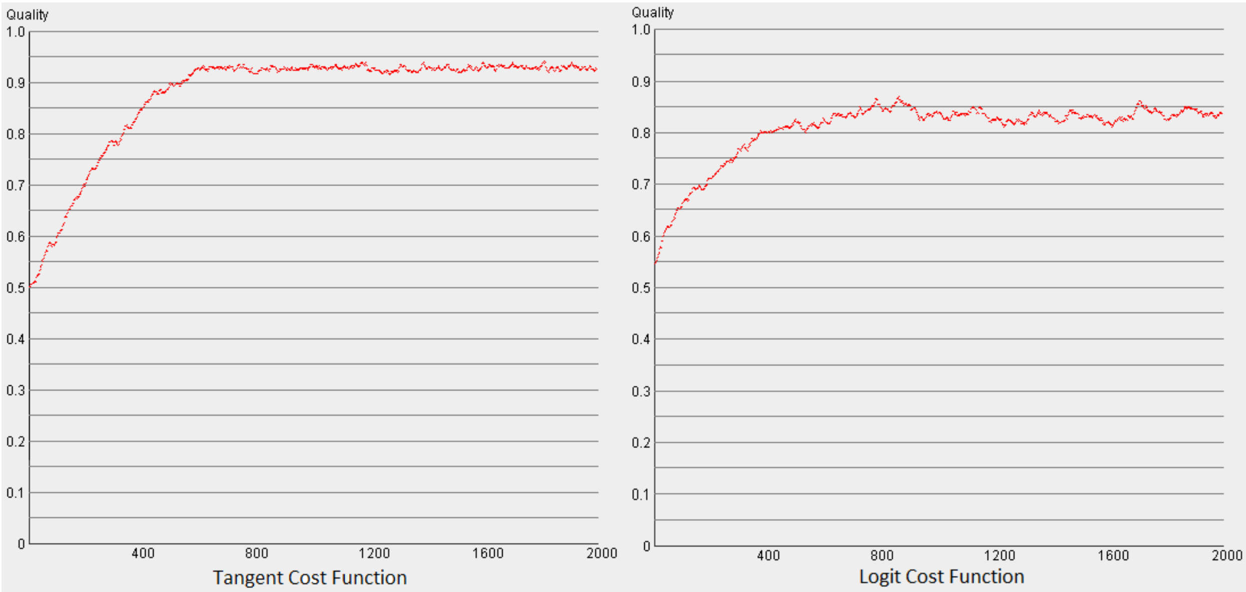

***Figure 8**: Comparison of the paths of the average V value for the tangent and logit cost functions under the linear benefit function*

For the plot on the left, the cost function is defined as:

$$c(V) = 100 * \tan(V * \frac{\pi}{2})$$

And the plot on the right corresponds to a cost function of:

$$c(V) = 1000 * \log_{10} \frac{V+1}{1-V}$$

For both cost functions, the *V* value eventually converges, but it does so at different values for each cost function. With the tangent cost function, *V* converges at approximately 0.93, and with the logit cost function, *V* converges at approximately 0.83. Plotting the two cost functions, we can see that when $c(V) = 1000$, the *V* value of each cost function is close to the value at which they converge to:

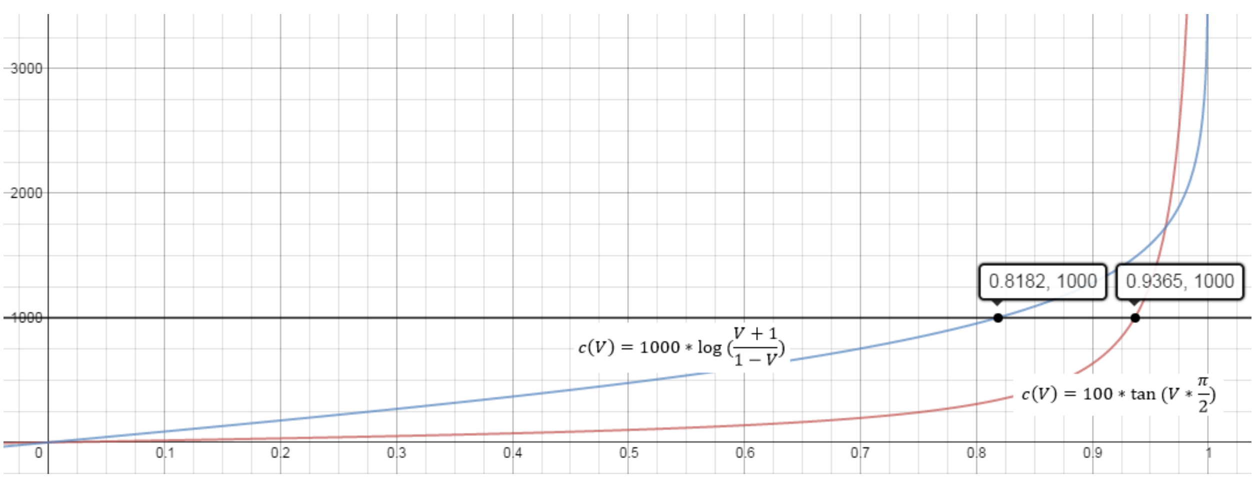

__Figure 9__: Plot of the tangent and logit cost functions. Note that c(V) = 1000 when V is close to the value at which the path of the average V converges to for both cost functions

To analyze why the optimal strategy seems to be contributing with a *V* value that yields a cost of around 1000, I will look at the distribution of the benefits received by the contributors. Since the benefits are functions of rankings, and there is a constant number of contributors at each round, then then distribution of benefits received by the contributors will remain constant at each round.

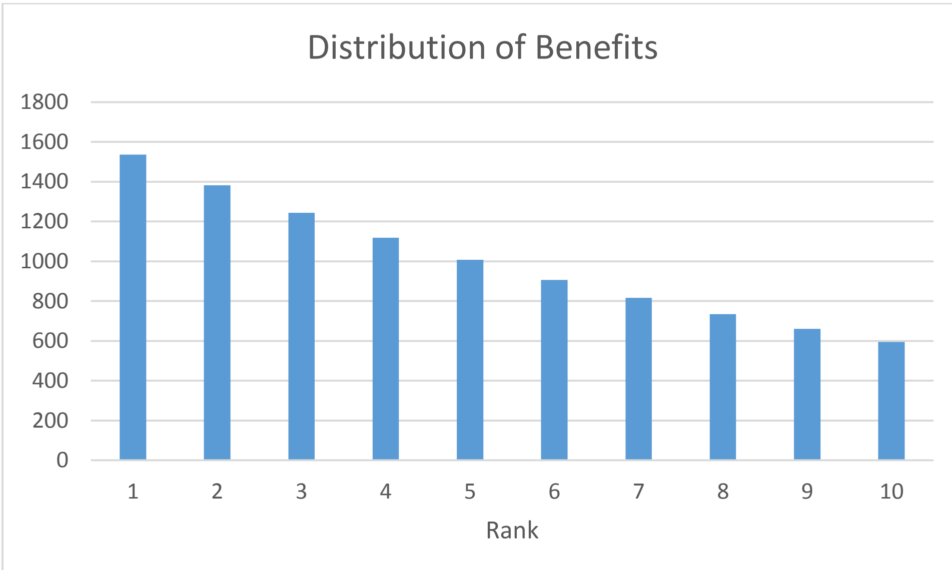

***Figure 10**: Benefit received by the contributor vs. the rank of the contribution. (Total attention = 10,000)*

As can be seen from the above graph, half of the contributors will receive a benefit of more than 1000, and the other half will receive a benefit of less than 1000, which means that if every contributor incurs a cost of 1000, half of them will end up with a positive payoff, and half with a negative payoff. Therefore, when the expected payoff is 0, a Nash equilibrium is reached.

The intuition is as follows: when all of the agents are contributing at a quality lower than the Nash equilibrium value, any agent can deviate by contributing at a slightly higher quality to beat all of the other agents and achieve a higher payoff. When the quality is lower than the Nash equilibrium, the extra cost in increasing the quality is less than the benefit gained from it. At a certain point, the extra cost becomes so high that increasing the quality is no longer optimal, but lowering the quality would mean lower benefits and therefore lower payoffs, so a Nash equilibrium is reached.

This can be seen from the plot of expected payoff vs. quality (*V*) in a game where all of the *k-1* agents are playing at the Nash equilibrium quality, and the *k*th agent is allowed to deviate. By contributing at the same Nash equilibrium quality as everyone else, the *k*th agent maximizes his expected payoff.

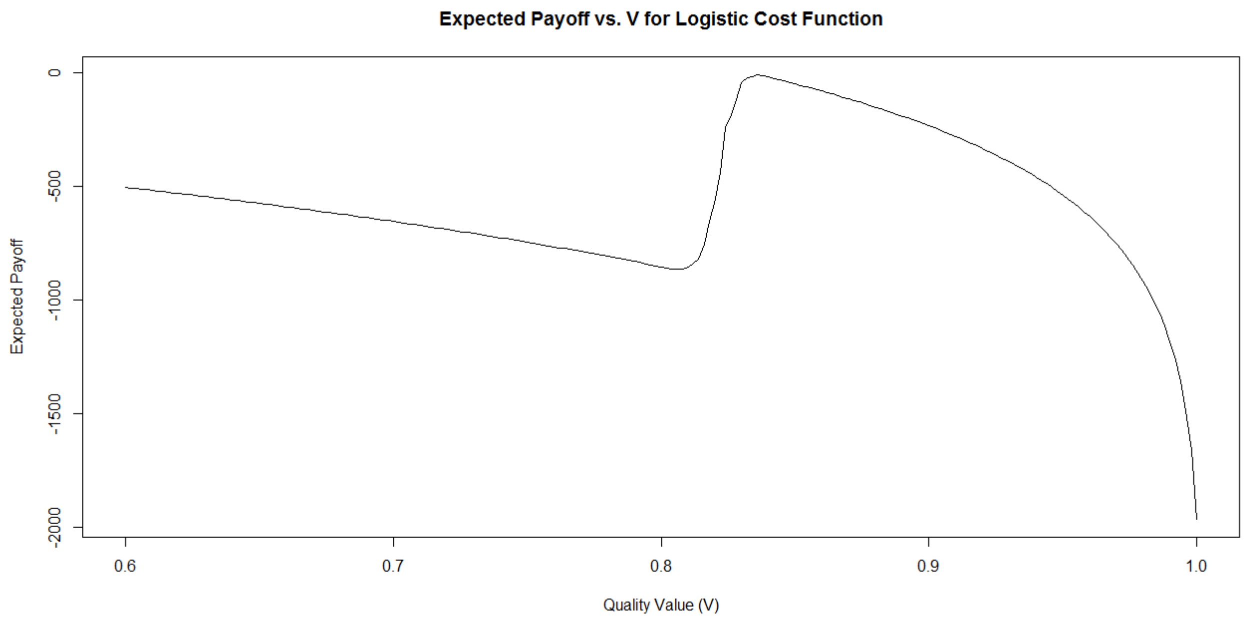

Figure 11: *Expected payoff vs. quality when using the logistic cost function (simulated from 1,000 games for a selection of quality values)*

VII. Conclusion

In this honors thesis, I replicated and confirmed the results of Arpita Ghosh and Patrick Hummel's paper "A Game-Theoretic Analysis of Rank-Order Mechanisms for User-Generated Content". It is indeed true that under certain assumptions (outlined in Section II), contributors of ranked contents will converge in distribution to a certain Nash equilibrium distribution of contribution qualities, and if the total attention is large enough (i.e. the rewards are high enough), the lower bound of this distribution will approach the highest possible quality.

I proceeded to study a simplified deviation of the model proposed in Ghosh and Hummel's paper in order to focus on the characteristics of the payoff function. I found that as long as the benefit function follows the description in Section II, the quality of the contributions will eventually converge, although the value that it converges to depends on the convexity of the benefit function. The cost function also affects the value that the qualities converge to, and more specifically, this is a value at which the expected payoff is 0.

Future studies of this topic may include adding different types of viewers to study strategies based on focusing on a specific type of viewers. The way attention is distributed could also be changed to study how strategies would differ if more attention is given to the top ranked contents or if the attention is distributed more evenly.