Comput Econ
DOI 10.1007/s10614-014-9453-8

Word-of-Mouth Communication and Demand for Products with Different Quality Levels

Bharat Bhole · Bríd G. Hanna

Accepted: 7 July 2014

Abstract We analyze a market with two product alternatives that differ in quality. Consumers choose between these products based on consumer reviews and their own experience. We examine how the market share of the superior product is affected by (i) the number of reviews obtained by consumers; and (ii) the type of information conveyed in these reviews. We find that when consumers randomly sample reviews from the entire population, an increase in the number of reviews can *decrease* the market share of the superior product. This, however, is not the case when consumers seek out reviews on each product. Further, we find that the market share of the superior product can be significantly lower when reviews convey subjective satisfaction compared to when they convey objective payoffs. This effect depends on the degree of heterogeneity in consumer expectations.

Keywords Word-of-mouth communication · Product quality · Product reviews · Bounded rationality · Computational approach · Agent-based modeling

1 Introduction

Consumers often need to choose between alternative products without knowing which product is superior. For example, given a choice between two cold medicines, a consumer may not know which of them would be more effective. Even if a consumer has had some experience with both medicines, the experience may not be a perfect indicator of the quality. This is because the final outcome is influenced by factors other

B. Bhole (✉) · B. G. Hanna
Department of Economics, Rochester Institute of Technology, Rochester, NY 14623, USA
e-mail: bharat.bhole@rit.edu

B. G. Hanna
e-mail: bxggse@rit.edu

Published online: 31 July 2014

than the inherent quality of the medicine. For instance, it is possible that consumer took the less effective medicine when she was infected by a milder strain of a virus, or had greater immunity, and hence she was cured quickly. Similarly, it is possible that she took the more effective medicine when she was infected by a more severe strain of virus, or had a lower immunity, and hence took longer to recover.[1]

In such situations where knowledge is imperfect, it is possible that consumers repeatedly choose an inferior alternative without giving the choice further thought. At the same time, however, it seems reasonable that once in a while they reexamine their choice more carefully. This reexamination often involves obtaining opinions or *reviews* from other consumers, who also have imperfect knowledge such as their own, and then making a decision based on their own experience and the information obtained from others. When consumers make decisions in this manner, it is important to know the extent to which the market learns to use the superior product without any external intervention. The objective of the paper is to explore this question. If it is the case that the lower quality product maintains a substantial market share in the absence of any regulatory intervention, this has implications for both the need for regulation as well as the choice of product quality by the firms.

To address this question we consider a variety of settings with the following common features. There are two products, H and L. On average product H yields a higher net payoff than product L. These average payoffs are unknown to the consumers. In each period some consumers, called *potential switchers*, consider a switch in their product choice. Each potential switcher obtains a sample of reviews and chooses the product with the higher average payoff in her sample. If the potential switcher's current product is the only product represented in the sample, then she chooses that product (in other words, she does not switch).

In such a context we study how the market share of product H is affected by (i) the number of reviews obtained by each potential switcher (studied in Sects. 2, 3, 4, and 5); and (ii) the type of information conveyed in these reviews (studied in Sect. 6). For the latter, we compare the market share of H when reviewers communicate precise payoffs experienced by them (referred to as *objective payoff communication*) with the market share when they communicate subjective satisfaction on a 5-level scale (referred to as the *stars rating system*). For the relationship between the number of reviews and the market share of H, we examine how it is affected by:

1. The sampling procedure of the potential switchers. Do they obtain a random sample of reviews from the entire population (referred to as *simple random sampling*), or do they obtain an equal number of reviews on each product (referred to as *equal-reviews sampling*)?
2. Who considers a switch in the product? Are those who consider a switch randomly determined (referred to as *exogenous switching*), or are only the dissatisfied consumers likely to consider a switch (referred to as *endogenous switching*)?
3. The type of information conveyed in the reviews.

Our main findings are the following:

[1] Some other examples of markets that may have this feature are the services of medical doctors, car mechanics, and lawyers.

1. The use of simple random sampling results in a counterintuitive U-shaped relationship between the number of reviews obtained by each potential switcher and the market share of the superior product.
2. The use of equal-reviews sampling results in a positive monotonic relationship between the number of reviews obtained by each potential switcher and the market share of the superior product.
3. In the long-run, simple random sampling with as few as 2 reviews yields at least as large a market share of the superior product as equal-reviews sampling with many more reviews.
4. Findings 1–3 hold in both the exogenous and endogenous switching scenarios, and whether objective payoff communication or a stars rating system is assumed.
5. If consumer expectations (about product performance) vary substantially, then a change from objective payoff communication to the stars rating system significantly reduces the market share of the superior product. In contrast, if consumer expectations are similar, then a change in informational content of reviews has a negligible effect on the market share of the superior product.

There is a vast literature on the question of whether the population will learn to consume a higher quality alternative when quality is uncertain. This literature can be classified along the following important dimensions. First, how do consumers use the information they receive? Are they Bayesian (Alcalá et al. 2006; Bala and Goyal 1998; Kondor and Ujhelyi 2005; Bikhchandani et al. 1998; Banerjee and Fudenberg 2004) or do they use rules-of-thumb (Izquierdo and Izquierdo 2007; Ellison and Fudenberg 1995; Smallwood and Conlisk 1979; Rogerson 1983; Zhao and Duan 2014; Lamberson and Page 2008; Chatterjee and Xu 2004)? Papers in the former category assume that consumers start with certain priors, use Bayesian updating as they receive new information, and then adopt the alternative that maximizes their expected payoff. In contrast, papers in the latter category assume that consumers are boundedly rational and use simple rules to decide amongst the alternatives. Some examples of such rules are: choose the alternative with the highest sample average payoff; or choose the alternative that has the highest market share.

Second, what information is gathered by the consumers? Do they rely only on their own past experience (Rogerson 1983; Zhao and Duan 2014), or market share (Smallwood and Conlisk 1979; Bikhchandani et al. 1998), or word-of-mouth communication (Bala and Goyal 1998; Kondor and Ujhelyi 2005; Bikhchandani et al. 1998; Alcalá et al. 2006; Izquierdo and Izquierdo 2007; Ellison and Fudenberg 1995) or on both market share and on word-of-mouth communication (Banerjee and Fudenberg 2004; Lamberson and Page 2008)?

Third, within the word-of-mouth category, papers differ in terms of the information communicated amongst consumers. Bala and Goyal (1998), Alcalá et al. (2006), Izquierdo and Izquierdo (2007), Ellison and Fudenberg (1995) assume that consumers convey payoff information. Kondor and Ujhelyi (2005) assume that consumers provide information only on whether they are satisfied or dissatisfied. In Bikhchandani et al. (1998) consumers convey either their ex ante signals about quality or their choices.

Our paper assumes that consumers are boundedly rational and use rules-of-thumb to make decisions based on their own experience and on word-of-mouth communication.

We consider two cases, one in which reviewers convey payoff information and another where they use a stars rating system (specifically, reviewers convey their satisfaction level on a 5-level scale). Our model is very similar to that of Ellison and Fudenberg (1995). In fact, we start with a version of their model to facilitate a comparison with their results. However, we then extend the model in various ways to address additional questions that they do not consider.[2] Also in contrast to Ellison and Fudenberg, we use a computational approach to analyze the model. There are several reasons for using the computational approach. First, it allows us to examine the market shares both in the short-run and in the long-run. Second, the computational approach makes it feasible to consider the various extensions of the Ellison and Fudenberg (1995) model. Finally, this approach makes it possible to get a sense of the quantitative importance of different features of word-of-mouth communication on the market shares of the two alternatives.[3]

Izquierdo and Izquierdo (2007) and Lamberson and Page (2008) are also similar to our paper in the sense that they assume rule-of-thumb decision making and word-of-mouth communication of payoffs. Also, both use a computational approach. However, there are a few important differences between these papers and our paper. Izquierdo and Izquierdo (2007) focus on showing how quality uncertainty can result in market failure when consumers rely only on their own experience.[4] While they discuss how the use of other consumers' reviews can mitigate this effect of quality uncertainty on market failure, they do not study the effect of sample size on the market share of the superior product under the different sampling procedures and switching rules. Further, they do not study the effect of subjective communication on the market share of the superior product.

Like Izquierdo and Izquierdo (2007), Lamberson and Page (2008) do not study how subjective communication affects the market share of the superior product. Further, Lamberson and Page (2008) do not study the effect of the number of reviews obtained by consumers; in their paper consumers act sequentially and take into account the experience of all past consumers. Also, their paper differs from ours in the rules-of-thumb used by consumers. In their model, a consumer's decision is based on a comparison of the *aggregate* payoffs across the two alternatives. In contrast, in our model a consumer chooses an alternative based on a comparison of the *mean* payoffs of the two alternatives.[5]

[2] Ellison and Fudenberg (1995) do not consider equal-reviews sampling, endogenous switching, nor do they consider the use of a stars rating system.

[3] Ellison and Fudenberg (1995) say the following about their study, "We have not been able to completely determine the long-run dynamics of our model. Rather than simplify the model further, we have chosen to provide a partial characterization."(p. 101) The computational approach enables us to provide a more complete characterization of the outcomes with a more realistic model. See Judd and Page (2004) who advocate a computational approach for this and other reasons.

[4] It may appear that Izquierdo and Izquierdo (2007) consider a single product. However, their setup can be interpreted as one with the following two characteristics: (1) there are two products, one that has uncertain quality and another that has a known and certain quality that is normalized to zero; and (2) the product with uncertain quality has higher average quality, but consumers are not aware of this fact. With this interpretation, market failure in their model occurs when consumers consume the lower quality product.

[5] Lamberson and Page (2008) do not explicitly talk about payoffs. However, they have a variable referred to as "feedback", which can be interpreted as the consumer's payoff.

The rest of our paper is organized as follows: In Sects. 2–5 we study the relationship between the number of reviews obtained by each potential switcher and the market share of product H. Section 2 presents the benchmark model that incorporates the simple random sampling procedure, objective payoff communication, and exogenous determination of potential switchers. Section 3 modifies the benchmark model to consider equal-reviews sampling. Section 4 introduces endogenous determination of potential switchers; the case of simple random sampling is studied in Sect. 4.1 and equal-reviews sampling is studied in Sect. 4.2. Section 5 modifies the model in Sect. 4 to consider reviewers' use of the stars rating system. Section 6 changes the focus from the earlier sections to investigate whether (compared to objective payoff communication) the stars rating system significantly affects the market share of product H.

2 The Benchmark Model and the Results

2.1 The Model

As noted in the introduction, our benchmark model is similar to that of Ellison and Fudenberg (1995). The society consists of P identical consumers.[6] Each consumer chooses between the two alternatives, H and L, in each time period, $t = 1, 2, 3, \ldots 5{,}000$. The supply of both alternatives, H and L, is perfectly elastic. The payoff experienced by a consumer in any time period depends on the alternative consumed in that period and on an idiosyncratic shock that is independent across consumers and across time. Specifically, the net payoff (net of price) of consumer i who consumes alternative g ($g \in \{H, L\}$) in time period t is given by:

$$U_{igt} = \theta_g + \varepsilon_{it} \tag{1}$$

where $\varepsilon_{it} \sim N(0, \sigma^2)$ is an idiosyncratic shock that is identically and independently distributed across time and consumers. Parameter θ_H (respectively, θ_L) denotes the average net payoff from alternative H (respectively, L). Alternative H is superior to alternative L in the sense that it yields a higher average payoff, i.e., $\theta_H > \theta_L$. This fact is unknown to consumers.[7]

Following Ellison and Fudenberg (1995), it is assumed that in any time period a constant fraction of the population, denoted by α, considers a switch in their product. Each consumer in the remaining population (those not considering a switch) consumes the same product she consumed in the previous period. For convenience, we refer to any consumer considering a switch as a *potential switcher*. Further, if a potential switcher's most recent consumption is H (resp. L) we refer to her as an H *potential*

[6] Ellison and Fudenberg (1995) have a continuum of agents, whereas we have a finite population.

[7] In our model θ_H and θ_L are constant over time; that is, there are no product-specific, population-wide shocks. In contrast, in Ellison and Fudenberg (1995) θ_H and θ_L are random; $\theta_H - \theta_L$ equals $\theta > 0$ with probability p and $-\theta$ with probability $(1 - p)$. In their model, product H is superior in the sense that $p > 1/2$.

switcher (resp. *an L potential switcher*). The potential switchers in any time period are randomly chosen from the entire population.[8]

Any potential switcher in period t uses the following decision process:

(i) She samples N consumers from the entire population using simple random sampling.[9] Each sampled consumer communicates her review, which contains information about the product she consumed and the payoff she experienced in period $t-1$. We refer to these sampled consumers as reviewers.[10] A reviewer who consumed product H (respectively, L) in period $t-1$ is referred to as an H (respectively, L) reviewer. Let k denote the number of L reviewers in the sample; it follows that $N-k$ is the number of H reviewers in the sample.

(ii) The potential switcher adds her own most recent (period $t-1$) experience to the information obtained from the N other reviewers. Hence, if she is an H potential switcher she will have $N-k+1$ sample observations on H and k observations on L. Similarly, an L potential switcher will have $N-k$ sample observations on H and $k+1$ observations on L.

(iii) If all the reviews (including her own experience) relate to the same product then the potential switcher does not switch. Following Ellison and Fudenberg (1995), we refer to this as the *must-see* condition. On the other hand, if reviews for both the alternatives are present in the sample, she separately calculates the mean payoff of L reviewers and of H reviewers, and in period t chooses the alternative with the higher mean payoff. In terms of the notation, this implies the following decision rule for an H potential switcher:

$$\begin{aligned}
&\text{if } k = 0, \text{ choose } H \text{ with certainty}\\
&\text{if } k > 0 \begin{cases} \text{choose } H \text{ if } \dfrac{\sum_{i=1}^{N-k} U_{i,H,t-1} + U_{own,H,t-1}}{N-k+1} \geq \dfrac{\sum_{i=1}^{k} U_{i,L,t-1}}{k} \\ \text{otherwise, choose } L. \end{cases}
\end{aligned} \tag{2}$$

where, $U_{own,H,t-1}$ denotes the H potential switcher's most recent payoff.[11] Similarly, an L potential switcher's decision rule is:

---

[8] This assumption is changed in Sect. 4, where it is assumed that those who are more dissatisfied are more likely to consider a switch.

[9] A potential switcher obtains a new sample in every time period that she becomes a potential switcher.

[10] One could assume that sampled reviewers report a weighted average of their past payoffs instead of only their most recent payoff. As long as the reviewers report the (weighted average) payoff for the most recently consumed product only, our results continue to hold.

[11] Note that this assumes that a potential switcher places the same weight on her own payoff as she places on other reviewers' payoffs. One could consider a situation where a potential switcher places a higher weight on her own payoff. This is an interesting alternative to consider in future work.

Table 1 Base parameter values for the Benchmark Model simulations

Parameter	Description	Value
P	Population size	500
N	Number of reviews / sample size	10
α	Fraction of P that consider switching	0.1
θ_H	Average net payoff from H	100
θ_L	Average net payoff from L	75
σ	Variance in net payoff from H and L	20
m_H^0	Initial market share of H	0.5

$$\text{if } k = N, \text{ choose } L \text{ with certainty}$$
$$\text{if } k < N \begin{cases} \text{choose } H \quad \text{if} \quad \dfrac{\sum_{i=1}^{N-k} U_{i,H,t-1}}{N-k} > \dfrac{\sum_{i=1}^{k} U_{i,L,t-1} + U_{own,L,t-1}}{k+1} \\ \text{otherwise,} \quad \text{choose } L. \end{cases} \tag{3}$$

Given these assumptions about the available alternatives, consumer knowledge and consumer decision making, we are interested in exploring how the market share of the higher quality product, H, evolves over time and how this evolution is affected by the parameter N. In order to do so we use a computational approach. The details are explained in the following subsection.

2.2 Methodology

We encoded the above set of assumptions into a computer program and then used this program to simulate the market share of product H from time periods 1 through 5,000.[12,13] The base parameter values for the benchmark model are given in Table 1. When the discussion or the figures do not explicitly specify the value taken by a parameter, it is implied that the parameter is set at its base value.

The model has several stochastic components: the payoff experienced by each consumer in each time period, the set of potential switchers in each time period, and the sample of reviews obtained by each potential switcher in each time period. For this reason we ran 100 replications, each with a different random seed, for each combination of parameter values considered. We then calculated the average market share of product H across these 100 replications at every time period. We compare

[12] We used Python and C programming languages for this purpose. The code is available upon request from the authors.

[13] We believe that 5,000 time periods is sufficient because: One, for the parameter combinations considered, the market share of H converges by time period 5,000. Two, even if one model time period corresponds to as short as one calendar day, which means that the product is consumed once everyday, 5,000 time periods corresponds to about 14 calendar years. In practical terms, we believe that what happens beyond a 14 year time horizon is unlikely to be relevant for most policy planning purposes. For example, for questions such as whether regulatory intervention is necessary for promotion of the higher quality product, or what quality level a firm should provide, it may not be very helpful to know what happens beyond a 14 year horizon.

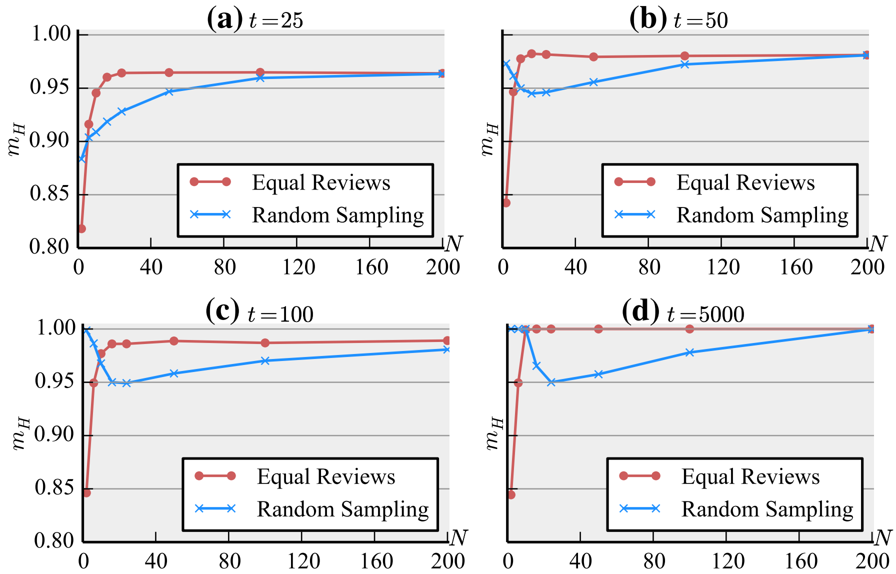

Fig. 1 Market share of H (m_H) for exogenous switching; $N = 2, 6, 10, 16, 24, 50, 100, 200$; $t = 25, 50, 100, 5{,}000$; $\sigma = 20$

these average market shares for the different parameter combinations. Henceforth, the average market share across replications is denoted by m_H.

To facilitate discussion of the results, we present results only for time periods 25, 50, 100, and 5,000. For convenience, and for the reasons noted in footnote 13, we refer to time period 5,000 as the long-run when discussing the results.

2.3 The Results

In this subsection we examine the relationship between the number of reviews (N) obtained by each potential switcher and the market share of product H (m_H). To do so we simulate m_H for $N = 2, 6, 10, 16, 24, 50, 100, 200$. The results are presented as "Random Sampling" plots in Fig. 1.

We see in Fig. 1 that except in time period 25, the relationship between N and m_H is U-shaped. That is, in the time periods 50 and later, and in the lower range of N values ($N = 2, 6, 10$), the more reviews each potential switcher obtains before making a decision, the smaller is the market share of H. This is a counter-intuitive result. One would normally expect that the more reviews any consumer obtains before making a decision, the more likely she will make the correct choice. In our model the correct choice is H, and therefore a greater likelihood of making the correct choice should cause the market share of H to be higher in any time period. But, in fact, we see the opposite.

To understand why the relationship between N and m_H is U-shaped we need to understand how N affects a potential switcher's likelihood of choosing H. We explain this in the following discussion.

First, consider a scenario where a potential switcher has sampled X reviewers (including herself), and at least one of these X reviewers is an L reviewer. Let us refer to this as *scenario A*. In this scenario, the probability that a potential switcher chooses H is either zero (applicable in the case of an L potential switcher with no H reviewers in the sample), or it is given by one of the following: the probability that the inequality in expression (2) is satisfied (this is applicable when the potential switcher is of type H); or the probability that the inequality in expression (3) is satisfied (this is applicable when the potential switcher is of type L and there is at least one H reviewer in his sample). In the case where the potential switcher is of type L and has no H reviewers in his sample, it is straightforward that obtaining an additional review increases the probability that H is chosen. In the other two cases also it can be shown that both the inequalities referred to above are more likely to be satisfied when an additional review is obtained.[14] Hence, sampling an additional review in scenario A unambiguously increases the probability that a potential switcher chooses product H.

Now consider the alternative scenario where the X reviewers in the potential switcher's sample (including herself) do not contain any L reviewers. Let us refer to this as *scenario B*. This scenario can arise only when the potential switcher is of type H. In this scenario, the must-see condition implies that the H potential switcher will choose H with certainty. In this case, obtaining an additional review increases the likelihood that there will be an L reviewer in the sample, which unambiguously *decreases* the likelihood of choosing H.[15] It is worth emphasizing that, in this scenario, it is the *must-see* condition that causes the likelihood of choosing H to fall when the sample size increases.

Hence, whether an increase in N will increase or decrease m_H depends on which of the two scenarios, A or B, is more prevalent. When there are relatively more H potential switchers and they mostly face scenario B, a higher N would result in a lower m_H. This is likely to be the case when (i) m_H is relatively large and, (ii) when N is small. Therefore, in the later time periods when m_H has reached a sufficiently high level and in the range of smaller N values, we see a negative relationship between N and m_H. In all other situations (such as, in the early time periods when m_H is relatively small; or in the later time periods and in the range of larger N values), an increase in N results in a higher m_H.

Ellison and Fudenberg (1995) find that (in the long-run) for smaller sample sizes there is convergence to the superior product, but for large sample sizes diversity obtains, i.e., both H and L have positive market shares. However, Ellison and Fudenberg identify the effect of sample size on only the long-run market shares, whereas we

[14] The probability that these inequalities are satisfied is given by the probability that $\bar{U}_H - \bar{U}_L > 0$, where $\bar{U}_g$ is the sample average payoff from product $g \in \{H, L\}$. Given the model assumptions, $\bar{U}_H - \bar{U}_L \sim N\left(\theta_H - \theta_L, \frac{\sigma^2}{k^2+(X-k)^2}\right)$ when there are X reviewers and k of these X reviewers are L reviewers. When one additional reviewer is sampled, the variance of this distribution decreases: it either becomes $\left(\frac{\sigma^2}{k^2+(X+1-k)^2}\right)$ when the additional reviewer is an H reviewer, or $\left(\frac{\sigma^2}{(k+1)^2+(X-k)^2}\right)$ when the additional review is an L reviewer. This decrease in variance implies an increase in the probability that $\bar{U}_H - \bar{U}_L > 0$.

[15] As can be seen from Eq. 2, an H potential switcher chooses H with certainty when $k = 0$. However, the probability she chooses H is less than 1 when $k > 0$.

identify both short-run and long-run effects. In addition, the approach in Ellison and Fudenberg (1995) does not allow for a comparison of market shares corresponding to different sample sizes when there is diversity. Our approach allows us to do so.

We tested the sensitivity of the U-shaped relationship between N and m_H to changes in the other parameters. Different values of P, α, σ, θ_H and θ_L resulted in a similar pattern: in early time periods a higher N implies a greater m_H but in later periods we get a U-shaped relationship between N and m_H.[16]

Another interesting feature of the random sampling plots in Fig. 1 is that in time periods 100 and 5,000, a sample size of 2 yields a (weakly) higher m_H than any other sample size considered. Even a sample size of 200 does not result in a higher m_H than a sample size of 2.[17]

This suggests that if consumers adopt the decision rule described in this Sect. 2.1 then in the long-run (see results for periods 100 and 5,000) it is socially optimal that they sample no more than 2 other reviewers.[18] Further, if learning others' payoffs can be likened to reading their detailed reviews or personally communicating with them (as opposed to, for example, just looking at the average number of stars a product has), it seems unlikely that most consumers will obtain samples larger than 50 before making their choice. This is because the cost of obtaining samples larger than this magnitude is likely to be prohibitive. In that case, a sample size of 2 yields the socially optimal outcome even in earlier time periods (see, for example, random sampling plots for $t = 50$ in Fig. 1).

3 Equal Reviews Sampling

One might argue that in reality potential switchers do not obtain a random sample of reviews, but instead seek out reviews for each product that is in their choice set. That is, in terms of the setting here, a potential switcher will seek out both L and H reviewers before making a decision about which product to consume. In this section we modify the sampling procedure to allow for this. Specifically, we assume that any potential switcher obtains an equal number of reviews on each alternative; if their goal is to obtain a total of N reviews, then they obtain $N/2$ reviews on each of the two products H and L. We refer to this as *equal-reviews sampling*.[19]

The $N/2$ reviews of any given product are obtained using simple random sampling from the population of current consumers of that product. If there are less than $N/2$ current consumers of a given product, then potential switchers obtain reviews from all

[16] We tried $P = 500, 1{,}000$; $\alpha = 0.01, 0.1, 0.5, 1$; $\sigma = 10, 30$; and (θ_H, θ_L) pairs (100, 75), (250, 225), (375, 350), (1,000, 975). Not surprisingly, the larger the α the sooner the U-shaped pattern appears. Also, we found that for $\sigma = 10$ an increase in N has little effect on m_H, while for $\sigma = 30$ the U-shaped relationship between N and m_H is more pronounced than for the base case of $\sigma = 20$. In the interest of space, these robustness checks are not presented in the paper but are available from the authors upon request.

[17] This feature also holds for the other parameter combinations we tried but that are not presented here.

[18] This assumes that it is socially optimal to produce and consume only H, i.e., $U_H - c_H > U_L - c_L$, where c_i is the constant marginal cost of alternative i, where $i \in \{H, L\}$.

[19] This sampling procedure is equivalent to the $S(K)$ sampling procedure (with $K = N/2$) introduced by Osborne and Rubinstein (1998) and used by Spiegler (2006) and Szech (2011).

the consumers of that product. In this case they have less than a total of N reviews. The other assumptions are the same as in Sect. 2. As in the previous section, we simulate m_H for $N = 2$, 6, 10, 16, 24, 50, 100 and 200. The results (obtained with all other parameters at their base values) are presented as "Equal Reviews" plots in Fig. 1.

It is clear from Fig. 1 that with equal-reviews sampling the relationship between N and m_H is positive across all values of N. To understand why, note that with equal-reviews sampling, every potential switcher has at least one L reviewer in her sample (except in the extreme case where no one consumes L, in which case the sampling procedure is irrelevant). This situation is the same as scenario A described in the explanation of the U-shaped relationship in Sect. 2. In this scenario an increase in the sample size increases a potential switcher's likelihood of choosing H, which in turn causes m_H to be higher.

An interesting question is whether one sampling procedure is better than the other; better in the sense of yielding as high or a higher market share of H at the same or a smaller sample size. To ascertain this, we compare m_H across the two sampling procedures for the base parameter values (see Fig. 1).[20,21] The following are the results of these comparisons.

In time periods 100 and 5,000, the largest m_H for random sampling is as high or higher than the largest m_H with equal-reviews sampling. Further, this largest m_H in random sampling occurs at the same or a smaller N than that required for the largest m_H in equal-reviews sampling. Hence, in these later periods simple random sampling clearly yields the socially superior outcome.

In time period 50, neither sampling procedure is clearly superior to the other. Generally, the largest m_H in equal-reviews sampling is as large or larger than its counterpart in random sampling. However, this largest m_H in equal-reviews sampling requires a larger N. That is, the equal-reviews procedures yields a higher m_H at the cost of requiring a larger sample size. Since we have not explicitly modeled the cost of obtaining reviews, we cannot further analyze this trade-off.

In time period 25, we find that the equal-reviews sampling procedure is generally superior. The largest m_H in equal-reviews sampling is as high or higher than that in the case of random sampling and it does not require a larger N to achieve this.

To summarize, in initial time periods equal-reviews sampling is superior, in intermediate time periods neither is clearly superior, but eventually random sampling becomes superior. The specific length of these time ranges will vary with parameter values. For example, the larger is α the shorter the range of time for which equal-reviews sampling is superior.

---

20 To check for the robustness of the findings we also compared m_H across the two sampling procedures for $\sigma = 10, 30$; $\alpha = 0.01, 0.5, 1$; and $(\theta_H, \theta_L) = (250, 225)$, (375, 350) and (1,000, 975). We changed one parameter at a time with other parameters left at their base value. For example, when σ was set to 10, all the other parameters were at their base level. In the interest of space, we have not presented these results for the different σ, α and (θ_H, θ_L) values, but they can be requested from the authors.

21 The reader may find it helpful to know that a difference of even 1 percentage point in m_H across the random sampling and equal-reviews sampling procedures is sufficient for statistical significance at the 1 % level (see Tables 3, 4 and 5 in the Appendix).

4 Endogenous Switching

In Sects. 2 and 3 we assumed that potential switchers are randomly chosen from the population at large. In reality, people who are dissatisfied with their choice are more likely to consider a switch than those who are satisfied. In this section we endogenize the determination of potential switchers to reflect this, and investigate whether this modification affects the results obtained in the previous sections.

To endogenize the determination of potential switchers we assume that consumers have preconceived expectations about the payoff. If the payoff that a consumer experiences in any period meets or exceeds her expectations in that period then she is satisfied and does not consider a switch in the next period. If her payoff falls short of her expectations then she is deemed dissatisfied and there is a chance that she considers a switch. The probability that she considers a switch is an increasing function of the extent by which her actual payoff falls short of the expected payoff.[22]

In terms of notation, let consumer i's expectation in time period t be denoted by A_{it}. If in time period t, $U_{igt} \geq A_{it}$, then consumer i is deemed to be *satisfied* with her choice and does not consider a switch. If $U_{igt} < A_{it}$ then the consumer is deemed *dissatisfied*. It is assumed that a dissatisfied consumer considers a switch in period $(t+1)$ with probability p given by:

$$p(U_{igt}, A_{it}) = \left(1 - \frac{U_{igt}}{A_{it}}\right). \tag{4}$$

We assume that all consumers start with the same expected payoff in period 1, i.e., $A_{i1} = A_{j1} = A$ for all consumers i, j, where A is some constant. We refer to A as the *initial expected payoff*. Consumers update their expected payoff over time based on their experience. Specifically, we assume that a consumer's expected payoff in period $(t+1)$ is given by:

$$A_{i,t+1} = \gamma U_{igt} + (1-\gamma)A_{it}. \tag{5}$$

Note that even though initial expected payoff is identical across consumers, the process of updating introduces heterogeneity in their expected payoffs. The degree of heterogeneity is positively related to the size of γ. We now examine whether the above described endogenous switching process affects the results we obtained in Sects. 2 and 3. Subsection 4.1 considers the case of simple random sampling and Subsect. 4.2 considers the case of equal reviews.

4.1 Endogenous Switching with Random Sampling

All assumptions, other than the endogenous determination of potential switchers, are the same as in the benchmark model. Further, the base values of the relevant parameters

[22] The view that satisfaction/dissatisfaction results from a comparison of expectations with perceived performance is supported by the "Disconfirmation of Expectations" theory in the marketing literature (see Oliver 1980 and Spreng et al. 1996).

Table 2 Base parameter values of A and γ for the endogenous switching simulations

All other parameter values are as in Table 1

Parameter	Description	Value
A	Initial expected payoff	90
γ	Weight on current payoff in expected payoff calculation	0.5

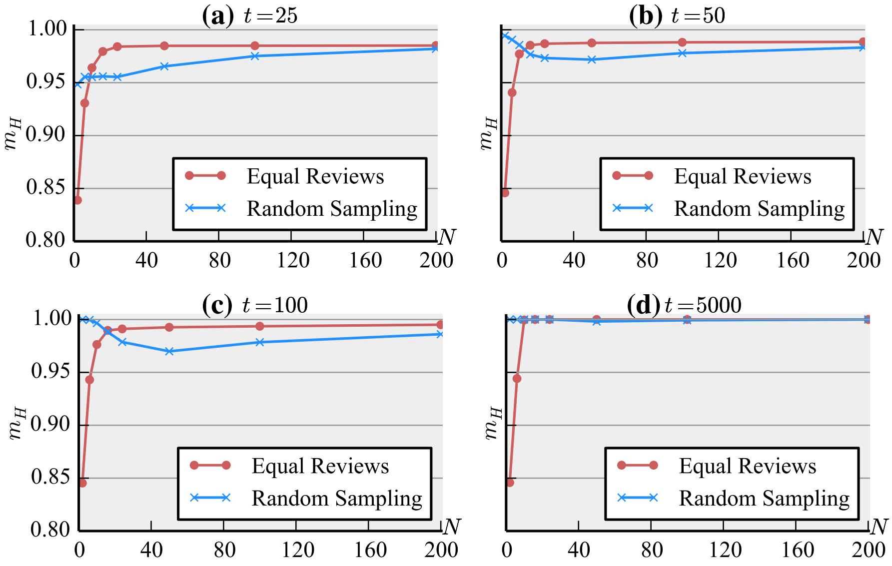

Fig. 2 Market share of H (m_H) for endogenous switching; $N = 2, 6, 10, 16, 24, 50, 100, 200$; $t = 25, 50, 100, 5{,}000$; $\sigma = 20$

are also the same as in the benchmark model. For the new parameters introduced in this section, A and γ, we assume the base values of 90 and 0.5 respectively (Table 2).

The ‘Random Sampling’ plots in Fig. 2 show the market share of H for $N = 2$, 6, 10, 16, 24, 50, 100, 200 in time periods 25, 50, 100 and 5,000. For the base parameter values we observe similar qualitative features as in the case of the benchmark model. In the very initial time periods (such as time period 25), there is a positive relationship between N and m_H. But in the later time periods the relationship takes on a U-shape. Further, as in the benchmark model, we find that in time periods 50, 100 and 5,000, m_H is largest when $N = 2$.

These two results (the U-shaped relationship between N and m_H and that in periods 50 and later m_H is largest when $N = 2$) are robust to changes in σ (we ran simulations for $\sigma = 10, 30$). The results are also mostly robust to changes in γ and A. We simulated m_H for all possible combinations of $\gamma = 0.01, 0.05, 0.1, 0.5, 0.9$ and $A = 25, 90, 150$ for each value of N, and found that the results hold in all but three cases: ($\gamma = 0.01$, $A = 25$, $t \in \{50, 100\}$) and ($\gamma = 0.05$, $A = 25$, $t = 50$). For these three cases we found a positive relationship between N and m_H.[23]

[23] In the interest of space, these robustness checks are not presented in the paper but are available from the authors upon request.

To understand why we have an exception in these three cases we need to recall from Sect. 2.3 that for m_H to fall as N increases, three conditions need to be satisfied: first, a sufficiently high proportion of the potential switchers have to be H potential switchers; second, a sufficiently high proportion of consumers have to be H consumers; and third, N has to be sufficiently small. With exogenous determination of potential switchers, the first condition follows from the second. As a result, when m_H is sufficiently large in time periods 50 and later, we see a negative relationship between N and m_H in the range of small N values. In contrast, with endogenous determination of potential switchers it is not obvious that most potential switchers will be H potential switchers even when the population consists of predominantly H consumers.

As shown in Eq. 4, whether a consumer becomes a potential switcher depends both on her expected payoff and her actual payoff; the smaller the ratio of actual payoff to expected payoff, the more likely it is that a consumer will become a potential switcher. On average, the L potential switchers experience lower actual payoffs compared to H potential switchers. Hence, if expected payoffs are similar across L and H consumers, then potential switchers will be predominantly of type L even when consumers are predominantly of type H. Given identical initial expected payoffs and how the expected payoffs adjust over time in our model (see Eq. 5), when γ is small, the expected payoffs will be similar across H and L consumers even in the later time periods. In these cases, therefore, the potential switchers will be predominantly L consumers and we will either not see a U-shaped relationship or it will be less prominent. In the three cases identified above as being exceptions, we can see that γ is small and this is why we do not see the U-shaped relationship.

In summary, the introduction of endogenous switching does not change the qualitative results found in the benchmark model: we see a similar U-shaped pattern and in time periods beyond 50 the sample size of 2 generally yields as high or greater m_H than any other sample size.

4.2 Endogenous Switching with Equal-Reviews Sampling

Results for the case where potential switchers use equal-reviews sampling (and all parameters are at their base levels) are presented as "Equal-Reviews" plots in Fig. 2. We find similar results as those found in the case of exogenous switching (see Sect. 3): (i) the relationship between N and m_H is positive across all values of N; and (ii) in initial time periods equal-reviews sampling is superior to random sampling, in intermediate time periods neither sampling procedure is clearly superior, but eventually random sampling becomes superior. The specific length of these time ranges will vary with parameter values. Further, we found these conclusions to be robust to changes in A and γ.[24]

[24] As in the previous subsection, we checked whether these conclusions hold for all possible parameter combinations formed with $\gamma = 0.01, 0.05, 0.1, 0.5, 0.9$ and $A = 25, 90, 150$.

5 Endogenous Switching with Subjective Communication

In Sects. 2–4 we assumed that consumer reviews communicate objective payoffs. However, in reality, consumer reviews generally convey subjective satisfaction on a discrete scale, and not the precise objective payoff information. For example, on the Amazon website people rate products on a 5-star scale.[25] We refer to this communication of subjective satisfaction on a discrete scale as the *stars rating system*.

Relevant information is lost when reviewers use the stars rating system. First, the discrete nature of the response results in a lumping of different payoff levels, which means that potential switchers can identify only a range of possible payoffs from each review. For example, a consumer may give a product 5 stars when she experiences a payoff of 100, but she may also give it the same 5 stars when she experiences a payoff of 120. Second, how satisfied a consumer is depends on how her payoff or experience compares with her expectations. For example, consumer A may expect a payoff of 80 and is extremely satisfied when she experiences a payoff of 100 and gives the product 5 stars. Whereas consumer B may expect a payoff of 125 and is dissatisfied when he experiences a payoff of 110 and gives the product only 3 stars. Due to this relative nature of the satisfaction level and the fact that the payoff and expectations are private information, it is not possible for the potential switchers to identify even the range of possible payoffs from each review, and therefore they cannot rank the payoffs of the reviewers.

In this section we investigate whether the use of the stars rating system (instead of objective payoff communication) affects our findings about the relationship between N and m_H. Whether the stars rating system has a significant quantitative effect on m_H is explored in Sect. 6.

We assume that in any period t a consumer's rating on this 5-star scale is determined by the ratio of her payoff U_{igt} to her expectations A_{it}. The higher the ratio U_{igt}/A_{it}, the more stars the consumer awards the product in her review. Specifically, we assume that there are thresholds $T_k,\ k \in \{1, 2, 3, 4, 5\}$ and the consumer awards the product k stars when the ratio U_{igt}/A_{it} exceeds T_k but is less than T_{k+1}.[26] In the base case we assume that $T_1 = 0$, $T_2 = 0.7$, $T_3 = 0.9$, $T_4 = 1$, and $T_5 = 1.1$. This is expressed more succinctly as $T = [0, 0.7, 0.9, 1, 1.1]$.

Expectations evolve over time in the same manner as discussed in Sect. 4. Further, in the base case the other parameters take the same values as in the base case of endogenous switching (see Sect. 4.1). We first address the case of simple random sampling and then consider equal-reviews sampling.

The random sampling plots in Fig. 3 show the market share of H for N = 2, 6, 10, 16, 24, 50, 100, 200 and time periods 25, 50, 100 and 5,000 when potential switchers

[25] In addition to rating on a discrete scale, many consumers also provide written descriptions of their experience with the product. These descriptions can provide payoff-relevant information in addition to that provided by a rating on a 5-star scale, and hence can affect choices made by potential switchers. In this paper, however, we do not take account of these verbal descriptions. We think it is important to first examine the difference between the stars rating system and the objective payoff communication because if this difference is not large, a study of how verbal descriptions additionally affects m_H is not warranted.

[26] This rule, along with restriction to maximum of 5 stars implicitly means that there is a threshold level $T_6 = \infty$.

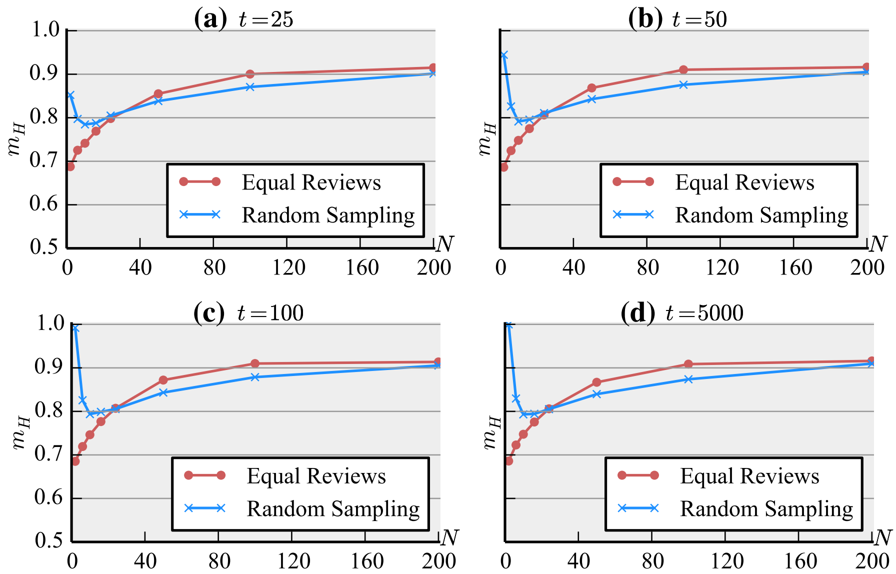

Fig. 3 Market share of H (m_H) for subjective communication; $N = 2, 6, 10, 16, 24, 50, 100, 200$; $t = 25, 50, 100, 5{,}000$; $\sigma = 20$

use the simple random sampling procedure. We observe the same qualitative results as in Sect. 4.1. The relationship between m_H and N is U-shaped, and in time period 50 and after m_H is largest when $N = 2$. Further, as in Sect. 4.1, we find that the results are generally robust to changes in parameters σ, A and γ, with the same exceptions as noted in that section.[27]

To model subjective communication we introduced a new parameter: the thresholds used by consumers for determining the number of stars to assign to a product. We also investigated the robustness of the results to changes in these thresholds. In addition to the base parameter value $T = [0, 0.7, 0.9, 1.0, 1.1]$, we ran simulations for: $T = [0, 0.75, 0.95, 1.1, 1.25]$, $[0, 0.6, 0.75, 0.9, 1]$, $[0, 1, 1.2, 1.4, 1.6]$, $[0, 0.2, 0.4, 0.6, 0.8$ and $[0, 0.65, 0.75, 1.1, 1.2]$.[28] The results identified in the base case were robust to these changes in the thresholds.[29]

To summarize, when potential switchers use the simple random sampling procedure, replacing objective payoff communication with the stars rating system does not change the qualitative results pertaining to the relationship between N and m_H.

[27] We considered the same set of A and γ combinations as considered in Sect. 4.1; these are all possible combinations formed using values $\gamma = 0.01, 0.05, 0.1, 0.5, 0.9$ and $A = 25, 90, 150$. For σ we considered values $\sigma = 10, 30$. In the interest of space, these robustness checks are not presented in the paper but are available from the authors upon request.

[28] Recall that a threshold value of $T = [T_1, T_2, T_3, T_4, T_5]$ means that a consumer awards k stars to product g consumed by her in time period t if $T_k \leq U_{igt}/A_{it} < T_{k+1}$.

[29] In the interest of space, these robustness checks are not presented in the paper but are available from the authors upon request.

The results for the equal-reviews sampling procedure are shown as equal-reviews plots in Fig. 3. We observe the same qualitative results as we did in Sect. 4.2. Further, these results are robust to the different values of A, γ and T considered.[30]

6 Quantitative Effect of Subjective Communication on Market Shares

So far in this paper we have focused on the qualitative relationship between N and m_H. We now discuss whether the loss of information that results from the use of the stars rating system *significantly* affects the absolute level of m_H. We mean *significant* in a practical sense and consider the effect on market share to be significant if it is 10 percentage points or larger. This difference is also highly statistically significant.[31]

For simplicity we focus our discussion on time period 5,000. However, we do present the figures for the other time periods ($t = 25, 50$ and 100). We discuss the case of random sampling and then the case of equal-reviews sampling.

6.1 The Case of Random Sampling

We find that whether or not subjective communication significantly affects m_H is determined mainly by two parameters, N and γ. Recall that γ is the weight that consumers place on the current payoff when they update their expected payoff (see Eq. 5), and the larger is γ, the larger is the variation in expected payoff across consumers. Figure 4 compares m_H in the case of objective payoff communication with that in the stars rating system when all parameters are at their base values. As can be seen in this figure, in time period 5,000 when $N = 2$ there is no difference between m_H across the two cases. However, when N equals 10 and 16, the difference between m_H across the two cases is more than 20 percentage points. For the other sample sizes considered, this difference is between 10 and 20 percentage points.

Recall that $\gamma = 0.5$ in the base case. As long as γ is held at 0.5, the foregoing observations do not change either with initial expected payoff (A), nor with the thresholds.[32] However, a change in γ is an important determinant of the difference in m_H across the two cases. As Fig. 5 shows, when $\gamma = 0.01$ (and other parameter values are at the base level) there is no difference in m_H across the two cases in the long-run for any sample size. This result is robust to changes in A and T.

---

[30] We ran simulations for the same set of A, γ and thresholds values as in the previous sections. These results are available upon request from the authors.

[31] We investigated the statistical significance of the difference in m_H across the objective payoff communication and the stars rating system (for both random sampling and equal-reviews sampling) for the following cases: (i) the base case, (ii) $\sigma = 10$, (iii) $\sigma = 30$ and (iv) $\gamma = 0.01$. We found that differences in m_H of 0.3 to 1.4 percentage points are sufficient for statistical significance at the 1 % level.

[32] We ran simulations for $A = 25$ and 150 in addition to the base level of 90. For thresholds, in addition to the base level of $T = [0, 0.7, 0.9, 1.0, 1.1]$, we simulated m_H for $T = [0, 0.75, 0.95, 1.1, 1.25]$, [0, 0.6, 0.75, 0.9, 1], [0, 1, 1.2, 1.4, 1.6], [0, 0.2, 0.4, 0.6, 0.8] and [0, 0.65, 0.75, 1.1, 1.2]. The results for these cases are available upon request from the authors.

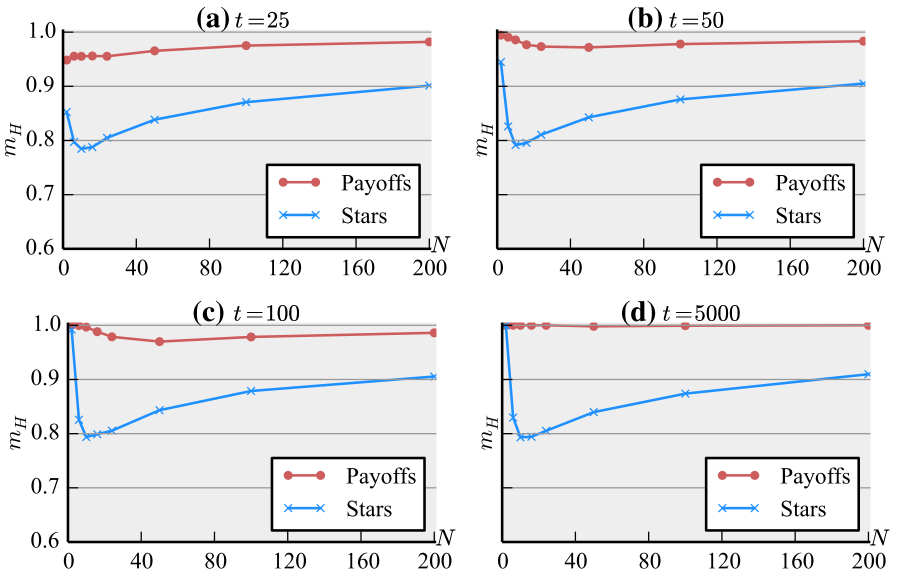

Fig. 4 Market share of H (m_H) for random sampling; $N = 2, 6, 10, 16, 24, 50, 100, 200$; $t = 25, 50, 100, 5{,}000$; $\gamma = 0.5$

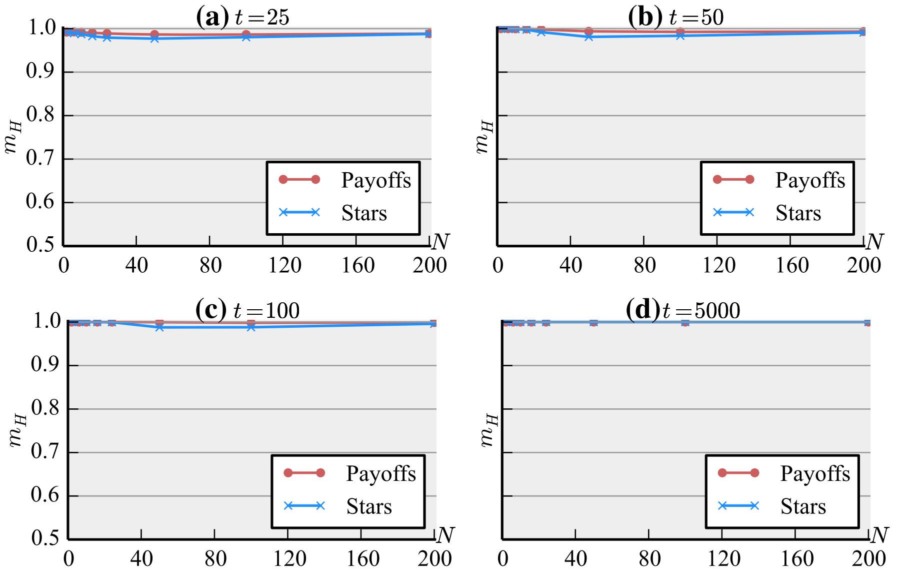

Fig. 5 Market share of H (m_H) for random sampling; $N = 2, 6, 10, 16, 24, 50, 100, 200$; $t = 25, 50, 100, 5{,}000$; $\gamma = 0.01$

The above results suggest that the loss of information due to lumping continuous payoffs into discrete categories does not play any role in affecting the market share of product H. However, the potential switcher's inability to rank the experience of differ-

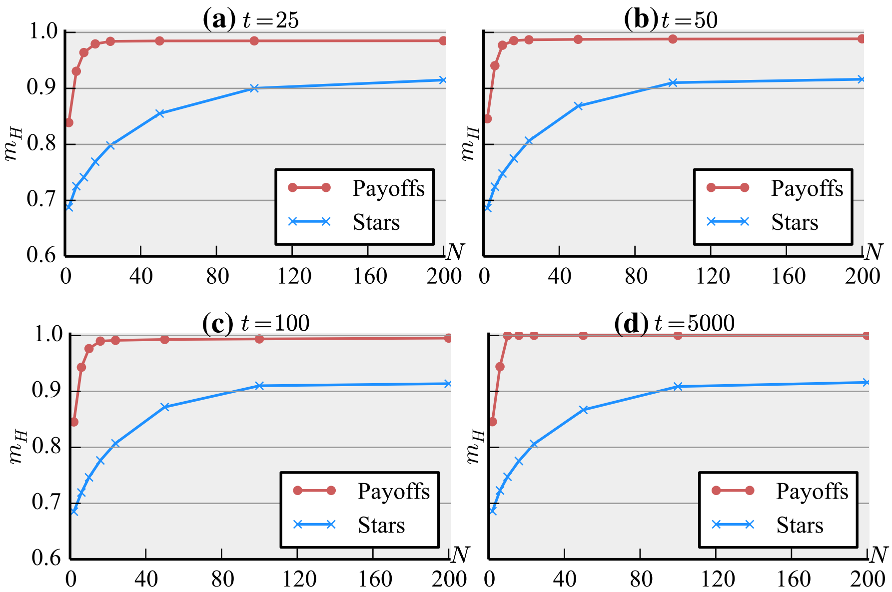

Fig. 6 Market share of H (m_H) for equal-reviews sampling; $N = 2, 6, 10, 16, 24, 50, 100, 200$; $t = 25, 50, 100, 5{,}000$; $\gamma = 0.5$

ent consumers based on their ratings (because of differences in reviewer expectations) does have a significant effect on m_H.[33] When γ is small (less than 0.1), expectations are not very different across reviewers and hence there is little subjectivity in the reviews. In this case, we see that m_H is similar irrespective of whether people communicate precise payoff information or lump it in categories.[34]

However, when γ is large (for example, 0.5 or 0.9) then expectations vary significantly across reviewers and subjective ratings do not reflect relative payoffs across consumers. In this case we see that m_H is significantly affected, it is lower by about 10 to 20 percentage points in the long-run, except when $N = 2$.

6.2 The Case of Equal-Reviews Sampling

Figure 6 compares objective payoff communication with the stars rating system when potential switchers use equal-reviews sampling procedure, and all parameters are at their base values. We find that the stars rating system has a similar effect on m_H as in the case of random sampling: the difference in m_H is between 10 to 20 percentage points in time period 5,000. Further, the same two parameters, N and γ, determine

[33] See the second paragraph of Sect. 5, where we discuss two ways in which subjective communication results in loss of relevant information.

[34] The fact that there is little subjectivity in reviews when γ is small is due to the assumption that the initial expected payoff, A, is the same across all consumers. If we were to relax this assumption then a smaller γ would not necessarily imply less subjectivity. In this case, the difference in m_H across objective payoff communication and the stars rating system would likely be large even for small values of γ.

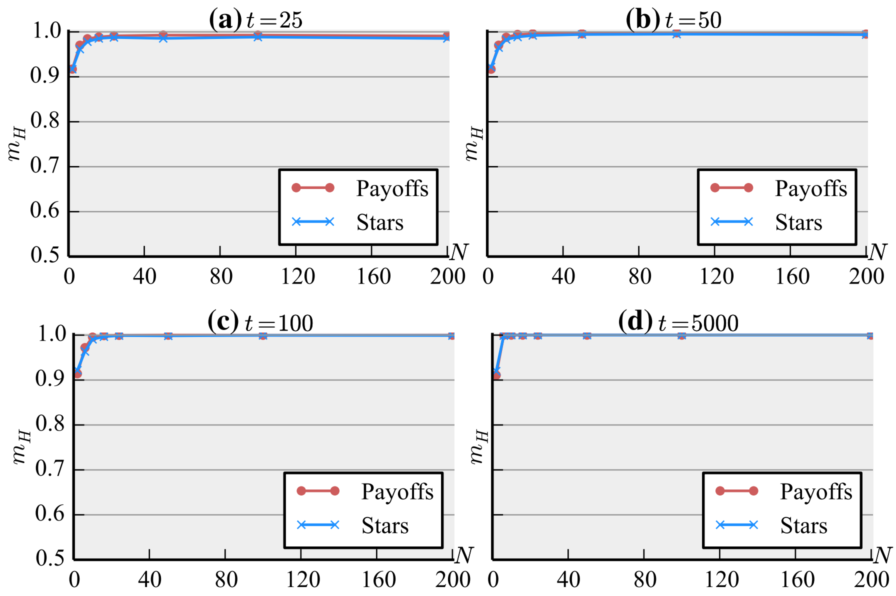

Fig. 7 Market share of H (m_H) for equal-reviews sampling; $N = 2, 6, 10, 16, 24, 50, 100, 200$; $t = 25, 50, 100, 5{,}000$; $\gamma = 0.01$

the extent to which m_H is affected. In particular, we find that when γ is small, m_H is practically the same regardless of whether reviewers communicate objective payoff information or subjective satisfaction on a discrete scale (see Fig. 7).

7 Concluding Remarks

In this paper we analyzed a market with two product alternatives where one is superior to the other in the sense of yielding a larger payoff on average. Consumers choose between these alternatives based on reviews from other consumers and their own experience. In this context, we examined how the market share of the superior product is affected by the number of reviews obtained by consumers.

We find the surprising result that, in some circumstances, an increase in the number reviews obtained by consumers results in a lower market share of the superior product. Specifically, this counter-intuitive result is realized when consumers obtain reviews using simple random sampling from the entire population, and do not switch unless they have at least one observation on the other product. This result holds whether consumers who consider a switch are exogenously chosen from the population, or whether only those who are dissatisfied with their current choice consider a switch. Further, whether reviews convey objective payoffs or whether they convey subjective satisfaction on a 5-level scale, this result continues to hold. This result suggests that, in some cases, less information can result in better market outcomes.

If consumers seek reviews on both products (equal-reviews sampling) then we do not see the negative relationship referred to above. A comparison of simple random sampling and equal-reviews sampling reveals that, in the long-run, the former procedure with 2 product reviews results in as high or a higher market share of the superior product than the latter with any number of reviews up to 200. This is an interesting result to the extent that the natural tendency of most consumers is to seek out reviews on all products under consideration. For example, most online stores don't allow for a simple random sampling of reviews. Reviews are organized on a per product basis, which only allows for a sampling procedure similar to the equal-reviews approach discussed in this paper. Allowing for simple random sampling would involve a procedure where consumers identify the set of alternatives they are considering and then the online store provides a random sample of reviews from amongst all the reviews for all these alternatives. Our results suggest that under some circumstances such a procedure may result in a higher market share of the superior product.

We also studied whether a change in the informational content of reviews from objective payoffs to subjective satisfaction (expressed on a 5-level scale) affects the market share of the superior product. We find that such a change can significantly reduce the market share of the superior product, regardless of whether consumers use simple random sampling or equal-reviews sampling. Given that in reality people convey subjective satisfaction through their reviews, as opposed to precise objective payoffs, this result suggests that the lower quality product can maintain a non-negligible market share even in the long-run.

Acknowledgments We appreciate the helpful comments from our fellow 2013 Southern Economics Association session participants and 2014 Eastern Economics Association participants

Appendix

Table 3 Mean and standard deviation of the market share of H across 100 replications: base parameter values

Exogenous switching					
N	t	Random sampling		Equal reviews sampling	
		Mean	SD	Mean	SD
2	25	0.88	0.0175	0.82	0.0207
2	50	0.97	0.0104	0.84	0.0171
2	100	1.00	0.0025	0.85	0.0155
2	5,000	1.00	0.0000	0.84	0.0168
6	25	0.90	0.0136	0.92	0.0146
6	50	0.96	0.0127	0.95	0.0128
6	100	0.99	0.0116	0.95	0.0122
6	5,000	1.00	0.0000	0.95	0.0139
10	25	0.91	0.0146	0.95	0.0104
10	50	0.95	0.0157	0.98	0.0122
10	100	0.97	0.0167	0.98	0.0119
10	5, 000	1.00	0.0000	1.00	0.0047
16	25	0.92	0.0126	0.96	0.0086
16	50	0.95	0.0150	0.98	0.0170
16	100	0.95	0.0127	0.99	0.0133
16	5,000	0.97	0.0259	1.00	0.0000
24	25	0.93	0.0110	0.96	0.0077
24	50	0.95	0.0125	0.98	0.0185
24	100	0.95	0.0116	0.99	0.0182
24	5,000	0.95	0.0158	1.00	0.0000
50	25	0.95	0.0085	0.96	0.0084
50	50	0.96	0.0093	0.98	0.0240
50	100	0.96	0.0091	0.99	0.0179
50	5,000	0.96	0.0100	1.00	0.0000
100	25	0.96	0.0068	0.96	0.0092
100	50	0.97	0.0080	0.98	0.0250
100	100	0.97	0.0069	0.99	0.0220
100	5,000	0.98	0.0134	1.00	0.0000
200	25	0.96	0.0068	0.96	0.0070
200	50	0.98	0.0068	0.98	0.0263
200	100	0.98	0.0094	0.99	0.0197
200	5,000	1.00	0.0022	1.00	0.0000

Table 4 Mean and standard deviation of the market share of H across 100 replications: base parameter values

Endogenous switching

N	t	Random sampling		Equal reviews sampling	
		Mean	SD	Mean	SD
2	25	0.95	0.0139	0.84	0.0166
2	50	0.99	0.0044	0.85	0.0177
2	100	1.00	0.0005	0.85	0.0190
2	5,000	1.00	0.0000	0.85	0.0183
6	25	0.96	0.0121	0.93	0.0157
6	50	0.99	0.0067	0.94	0.0120
6	100	1.00	0.0016	0.94	0.0150
6	5,000	1.00	0.0000	0.94	0.0137
10	25	0.96	0.0117	0.96	0.0103
10	50	0.99	0.0102	0.98	0.0118
10	100	1.00	0.0057	0.98	0.0149
10	5,000	1.00	0.0000	1.00	0.0038
16	25	0.96	0.0105	0.98	0.0090
16	50	0.98	0.0126	0.99	0.0131
16	100	0.99	0.0129	0.99	0.0140
16	5,000	1.00	0.0000	1.00	0.0000
24	25	0.96	0.0108	0.98	0.0051
24	50	0.97	0.0114	0.99	0.0128
24	100	0.98	0.0128	0.99	0.0160
24	5,000	1.00	0.0000	1.00	0.0000
50	25	0.97	0.0091	0.98	0.0055
50	50	0.97	0.0109	0.99	0.0178
50	100	0.97	0.0095	0.99	0.0146
50	5,000	1.00	0.0079	1.00	0.0000
100	25	0.98	0.0076	0.98	0.0054
100	50	0.98	0.0089	0.99	0.0180
100	100	0.98	0.0099	0.99	0.0157
100	5,000	1.00	0.0043	1.00	0.0000
200	25	0.98	0.0052	0.99	0.0050
200	50	0.98	0.0109	0.99	0.0172
200	100	0.99	0.0108	1.00	0.0144
200	5,000	1.00	0.0000	1.00	0.0000

Table 5 Mean and standard deviation of the market share of H across 100 replications: base parameter values

Endogenous switching with the stars rating system

N	t	Random sampling		Equal reviews sampling	
		Mean	SD	Mean	SD
2	25	0.85	0.0293	0.69	0.0231
2	50	0.94	0.0194	0.69	0.0220
2	100	0.99	0.0092	0.69	0.0206
2	5,000	1.00	0.0000	0.69	0.0236
6	25	0.80	0.0268	0.73	0.0225
6	50	0.83	0.0275	0.72	0.0249
6	100	0.83	0.0274	0.72	0.0242
6	5,000	0.83	0.0267	0.72	0.0235
10	25	0.78	0.0223	0.74	0.0240
10	50	0.79	0.0233	0.75	0.0214
10	100	0.79	0.0214	0.75	0.0216
10	5,000	0.79	0.0206	0.75	0.0226
16	25	0.79	0.0238	0.77	0.0259
16	50	0.80	0.0214	0.77	0.0261
16	100	0.80	0.0218	0.78	0.0229
16	5,000	0.79	0.0222	0.78	0.0256
24	25	0.80	0.0228	0.80	0.0270
24	50	0.81	0.0219	0.81	0.0267
24	100	0.81	0.0232	0.81	0.0261
24	5,000	0.81	0.0212	0.81	0.0259
50	25	0.84	0.0212	0.86	0.0280
50	50	0.84	0.0196	0.87	0.0265
50	100	0.84	0.0223	0.87	0.0262
50	5,000	0.84	0.0248	0.87	0.0270
100	25	0.87	0.0227	0.90	0.0274
100	50	0.88	0.0219	0.91	0.0299
100	100	0.88	0.0195	0.91	0.0305
100	5, 000	0.87	0.0218	0.91	0.0271
200	25	0.90	0.0239	0.92	0.0297
200	50	0.91	0.0229	0.92	0.0341
200	100	0.91	0.0214	0.91	0.0299
200	5,000	0.91	0.0223	0.92	0.0325

References

Alcalá, F., González-Maestre, M., & Martınez-Pardina, I. (2006). Competition versus information under the word-of-mouth effect. In *Tech. rep. Working paper*

Bala, V., & Goyal, S. (1998). Learning from neighbours. *The Review of Economic Studies*, *65*(3), 595–621.

Banerjee, A., & Fudenberg, D. (2004). Word-of-mouth learning. *Games and Economic Behavior*, *46*(1), 1–22.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *The Journal of Economic Perspectives*, *12*(3), 151–170.

Chatterjee, K., & Xu, SH. (2004). Technology diffusion by learning from neighbours. *Advances in Applied Probability*, *36*(2), 355–376.

Ellison, G., & Fudenberg, D. (1995). Word-of-mouth communication and social learning. *The Quarterly Journal of Economics*, *110*(1), 93–125.

Izquierdo, S. S., & Izquierdo, L. R. (2007). The impact of quality uncertainty without asymmetric information on market efficiency. *Journal of Business Research*, *60*(8), 858–867.

Judd, K., & Page, S. E. (2004). Computational public economics. *Journal of Public Economic Theory*, *6*(2), 195–202.

Kondor, P., & Ujhelyi, G. (2005). *Social learning with subjective communication and self-selection*. Manuscript: Univ Houston

Lamberson, PJ., & Page, SE. (2008). Long term efficiency and short term lock-in: Market predictability with positive feedbacks. In *Working Paper*

Oliver, RL. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, *17*(4), 460–469

Osborne, MJ., & Rubinstein, A. (1998). Games with procedurally rational players. *American Economic Review*, *88*(4), 834–847

Rogerson, WP. (1983). Reputation and product quality. The Bell Journal of Economics pp 508–516

Smallwood, D. E., & Conlisk, J. (1979). Product quality in markets where consumers are imperfectly informed. *The Quarterly Journal of Economics*, *93*(1), 1–23.

Spiegler, R. (2006). The market for quacks. *The Review of Economic Studies*, *73*(4), 1113–1131.

Spreng, RA., MacKenzie, SB., & Olshavsky, RW. (1996). A reexamination of the determinants of consumer satisfaction. *The Journal of Marketing*, *60*(3), 15–32

Szech, N. (2011). Becoming a bad doctor. *Journal of Economic Behavior & Organization*, *80*(1), 244–257.

Zhao, L., & Duan, W. (2014). Simulating the evolution of market shares: The effects of customer learning and local network externalities. *Computational Economics*, *43*(1), 53–70