Online Demand under Limited Consumer Search and Limited Review Reading

Marton Varga and Paulo Albuquerque

PRELIMINARY AND INCOMPLETE! TYPOS POSSIBLE.

1 Introduction

Enabling online shoppers to have access to consumer generated reviews has become the standard in retail and services industries, and a large fraction of consumers read reviews before they purchase a product (Chevalier and Mayzlin 2006, choose where to dine (Wu et al. 2015), or decide about their holiday destination (Vermeulen and Seegers 2009). The information available to consumers that is related to reviews often includes the mean rating of all reviews and also a set of individual reviews, with both a rating and text providing additional input about the product or service.

Even though reading more reviews could be useful to make better decisions, consumers are unlikely to read all available reviews (Kim et al. 2011). Reading reviews is costly, as consumers need to spend time to go over the large number of available comments. In addition, after consulting a sample of reviews, consumers may conclude that they have enough information, believing that further reading will not lead to substantial new input to influence product choice. Hence, as some but not all reviews influence search and purchase decisions, it is critical to managers to better understand the role that limited viewing of reviews and respective ratings has on consumer behavior.

The goal of this paper is two-fold. First, methodologically, we develop a sequential search model that provides an economic rationale for consumers' online choices, including (i) search for products, (ii) reading reviews, and (iii) product choice. Related models in the literature (Kim et al. 2010, Ursu 2018) incorporate the search and purchase components, but ignore the review reading decision. We fill this gap in the literature by a novel search framework that, we believe, is better suited to study online search behavior as a whole. Second, substantively, we explain variation in decisions, over products and consumers, related to search, review usage, and purchase, such as (1) the decision whether to read reviews, (2) the amount of reviews read, and (3) the influence of review readership on search and purchase decisions.

Our approach assumes that before browsing for products, consumer forms a belief, with some uncertainty, about the product quality while viewing the product characteristics on the category page of the retailer. The consumer is motivated to read reviews to reduce uncertainty about product quality, updating the initial belief as a function of the discovered ratings. In our model, there is an optimal strategy for the amount of reviews being read: consumers decide to stop reading reviews about a product when the gain coming from reduced uncertainty is lower than the cost of further reading. The information from the ratings of the discovered reviews influences further search and purchase, creating an economic link between the different decisions made during consumer journey.

We estimate the model on a rich click-stream data from a large UK retailer for the digital TV recorders category. The data set includes information about which reviews consumers read, their browsing decisions, and purchases.

We use the model to evaluate several counterfactual situations. First, we study the impact of a price discount, following the arrival of a negative review to a popular product. In the control scenario, the retailer does not react to the negative review, while in the the counterfactual scenario, the retailer monitors the changes in the set of consumer generated reviews and adjusts the product price accordingly. We show that using our algorithm, the retailer can gain ...

2 Literature Review

Primarily, this paper builds on and contributes to two streams of literature: consumer search and online word-of-mouth.

In terms of consumer search, our approach is most related to methods that assume that consumers search for information sequentially (Weitzman 1979). In terms of applications of the Weitzman model (Weitzman 1979) for the online setting, related applications include Kim et al. (2010), Honka and Chintagunta (2016), Kim et al. (2016), and Ursu (2018). These papers however do not model consumer decisions to read consumer-generated reviews. In this paper, we contribute to this literature by developing a model that uses the sequential search approaches as foundation and adds the decision to read reviews.

This study is also related to the literature on the relationship between reviews and consumer purchase decisions (Babić Rosario et al. 2016, Floyd et al. 2014, You et al. 2015). Wu et al. (2015) study how consumers update their beliefs about restaurant quality from available reviews but without modeling the decision to read reviews. Varga and Albuquerque (2019) show that reviews can modify the size of the consideration set and influence product choice, but they do not model search or reviewer reading decisions. Moe and Schweidel (2011) focus on why people post reviews. There are a variety of papers behind the motivation of reading reviews (Hennig-Thurau et al. 2003, Burton and Khammash 2010), but, to the best of our knowledge, there has not been work on modeling the decision to read reviews and the length of that search behavior. Consequently, we believe that we can contribute to the literature with our proposed sequential search model that provides an economic reasoning behind review reading behavior of online shoppers.

3 Model

3.1 Illustrative Example

Before we formally describe the model, we provide an illustrative example of a decision-making process that we address with our approach. Assume that a risk-averse person is considering purchasing a product, and she has initial beliefs about the mean and variance of the probability that, if she decides to purchase the item, she has a good or bad experience with the product. For example, she may believe that the mean probability of getting a good experience is 50% (and 50% of getting a bad experience), with a variance of 20%.

Assume then that she can read reviews from previous buyers of the product, describing their experiences, and that these product owners are similar in terms of their preferences. If the consumer reads a positive review, this additional input leads to an update on both the mean and variance of the beliefs. For example, it could be that the new mean probability of getting a good experience with the product goes up to 60%, while the uncertainty about that probability reduces from 20% to 10%.

At this point she might decide to read another review. If she does so, her point estimates will be updated and the respective variance will drop again. As she keeps on reading reviews, the probability distribution will converge to a point estimate. This is because if she reads many reviews (approaching infinite), there will be no more ambiguity: the probability of getting a good experience with the product will be equal to the fraction of good reviews out of all reviews read.

3.2 Framework

In this section, we describe the set up behind the model in an online setting where consumers shop at a retailer's website. We start by assuming that consumers have knowledge about the main attributes of each product before visiting any product page. This is similar to a situation where a consumer starts their search at the category page or at a search engine page where she observes a certain set of product attributes for all alternatives in the category, such as brand, average rating, price, and number of reviews. Consumers can then decide to visit a product page. At a product page, consumers can scroll down to a review section, where reviews are exogenously sorted. If consumers decide to scroll down to the review area, they have the possibility of making an additional decision to paginate to other pages related to the product.

Besides the attributes known before search begins, each consumer values additional functional product attributes that are unknown before search, as well as experiential quality, which is independent of product characteristics. We define the term *experiential quality* as the subjective consumer-experience with the product, or the individual-specific experiential value of using the product. At the search stage, this match value or quality is uncertain to consumers and it is only realized after consumers purchase and use the product. Reading reviews can provide input into this experiential quality, as explained below. In contrast, consumer valuation for the functional product attributes can be realized during search, upon browsing the product page.

When visiting the review part of the product page, consumers find out more about the population distribution (mean and variance) of the experiential quality. To match our empirical setting, we assume that this distribution is discrete with the number of support points being equal to the number of ratings that can be given in reviews. As consumers read reviews, they update their knowledge about the distribution of the quality, changing the mean accordingly and reducing uncertainty about that distribution with more information. Hence, search includes two components in our approach: (1) browsing the product page to find out about the functional attributes that are unknown pre-search; and (2) reading reviews to update consumer knowledge about the distribution of experiential product quality. In a world without search and reading costs, consumers would visit all product pages, read all the reviews, and would always know the population distribution both pre-search unknown terms.

We note that we do not model the creation of reviews. We also assume that there are no dynamics in review creation, i.e., the posting of a rating is not influenced by recently posted ratings. As the number of posted reviews increases, the distribution of the posted ratings converges to the true probability distribution of evaluations. Finally, we assume that older reviews are just as informative as more recent ones.

3.3 Model Development

In this section, we present the different components of the proposed approach: (1) Utility function, (2) consumer beliefs about experiential quality, (3) the uncertainty surrounding experiential quality, (4) the resulting marginal benefits and reservation utilities, and (5) the optimal number of reviews to be read and search path by consumers.

3.3.1 Utility Function

Assume that consumer i obtains utility from purchasing product j as a function of the number of reviews she has read about it m_{ij}

$$\mathbb{E}(U_{ij}|m_{ij}) = x_j\beta + \mathbb{E}(Q_{ij}|m_{ij}) - \text{Var}\left(Q_{ij}|m_{ij}\right) + \epsilon_{ij},$$

where x_j includes the log price and log number reviews (plus one) of product j; Q_{ijt} is the quality of the product about which $\mathbb{E}_t(Q_{ij}|m_{ij})$ expectation is formed given the set of reviews sampled m_{ij}; $\text{Var}(Q_{ij}|m_{ij})$ indicates uncertainty about the quality; ϵ_{ij} is the impact of additional attributes not known ahead of visiting a product page that can be resolved by viewing the product-page. We assume it follows a normal distribution with parameters $(0, \sigma^{\epsilon})$.

By introducing the notations

$$\mathbb{E}(U_{ij}|m_{ij}) = U_{ij}^m,$$

$$Q_{ij}^m = \mathbb{E}(Q_{ij}|m_{ij}),$$

and

$$V_{ij}^m = \text{Var}\left(Q_{ij}|m_{ij}\right),$$

we can simply write

$$U_{ij}^m = x_j\beta + Q_{ij}^m - V_{ij}^m + \epsilon_{ij}.$$

By visiting a product page, the consumers immediately solves the uncertainty about ϵ_{ij}. By scrolling down to a review space and noting its ratings and content, consumers update their belief about the distribution of the product's quality, Q_{ij}. The outside utility is denoted as $U_{i0} = \epsilon_{i0}$ and it follows a normal distribution of parameters $(\mu^{\epsilon_0}, \sigma^{\epsilon})$.

3.3.2 Beliefs about Quality

We model the consumer i's subjective experience upon purchase - henceforth (subjective) quality - as a weighted average of the probabilities S_{ij}^k denoting that (after experiencing the item) consumer i would evaluate product j with rating $k \in K$ (regardless whether this evaluation actually materializes in form of writing a review). As an example with $K = 5$, if $S_{ij} = [0, 0, 0, 0, 1]$ the consumer obtains a product which she would certainly evaluate with a rating of 5 (or 5-star if ratings are represented by stars). The true quality - i.e. the experience level upon purchase - of product j for consumer i then equals to:

$$Q_{ij} = \sum_{k=1}^{K} \gamma^k S_{ij}^k,$$

where $\|\boldsymbol{S}\|_1 = 1$. Negative γ values decrease, zero γ values does not change, while positive γ values increase product quality.

Consumer i, however, does not now the true quality Q_{ij} since she does not know the vector S_{ij}. However, she forms expectations about the latter. In other words, for each potential rating level, the consumer forms a belief about the probability that after purchase she would rate the item as k. Therefore, we write the expected product quality as a conditional expression that depends on the reviews read about product j:

$$\mathbb{E}(Q_{ij}|m_{ij}) = \sum_{k=1}^{K} \gamma^k \mathbb{E}(S_{ij}^k|m_{ij}).$$

Given the set of reviews read m_{ij}, the expected probability vector is a combination of (i) the consumer's prior belief before reading any review about product j (but already knowing the mean rating and the total number of reviews as these are observed on the category-page) and (ii) the reviews discovered so far during search. Each time the consumer read a set of reviews, these beliefs are update. Thus we can write

$$\mathbb{E}_i(S_{ij}^k|m_{ij}) = \frac{\psi_{ij}^k(m_{ij}, \psi_{ij}^{0k})}{\sum_{k=1}^{K} \psi_{ij}^k(m_{ij}, \psi_{ij}^{0k})},$$

where ψ_{ij}^k is a vector of length K, governing the evolution of beliefs. Here, ψ_{ij}^{0k} indicates the prior of the consumer formed after seeing the category-page but before seeing any further pages.

The vector of priors ψ_{ij}^{0k} at the beginning of search (so after seeing the category-page but nothing more) is a function of a belief parameter that is shared by all consumer $\overline{\Psi}_j$. Consumers initial belief is, however, heterogeneous. $\overline{\psi}_{ij}^k$ follow a Dirichlet distribution with parameter vector $\overline{\Psi}_j$, thereby reflecting individual deviation from the mean belief.[1]

The intuition of our approach to generate initial beliefs ψ_{ij}^0 is the following. We assume that the consumer knows the empirical distribution of the reviews. However, when she sees the average rating and number of submitted reviews about a product she does not yet know a single rating:

[1]The Dirichlet distribution is convenient, as elements of ψ_{ij}^0 sum to one so these can directly indicate perceived shares.

but she infers those. In other words, the consumer forms a best guess about the distribution of the ratings accessible at the website.

One can think about ψ_{ij}^0 as a standardized deviation from the average belief about the probability that product j is of quality k before seeing any reviews (but knowing the average rating and the number of submitted reviews). Each ψ_{ij}^0 is informative about the direction of this deviation and measures the relative distance of individuals from the common belief.

The consumer updates her ψ_j after each new sample of reviews the following way:

$$\psi_{ij}'' = \psi_{ij}' + \tau r_{ij},$$

where τ is the speed of update and r_{ij} is a vector of length K indicating the number of reviews with rating k within a recently read set of reviews. As $0 \leq \tau$, the updated belief is a combination of the prior belief and the newly discovered ratings.

3.3.3 Uncertainty about Quality

Uncertainty about product quality might be reduced by reading available reviews. So far we have discussed that by sampling more reviews $\mathbb{E}(Q)$, the expected probability that product j is of quality k, gets updated trough ψ and $\mathbb{E}(S)$. However, we have not yet discussed how sampling reviews impacts uncertainty about this expectation.

We propose that as consumer i samples reviews of product j, her uncertainty about the probability that j is of quality k is also updated. Due to the property of the Dirichlet distribution, the variance of this probability is:

$$\mathrm{Var}[S_{ij}^k] = \frac{\psi_{ij}^k(\phi_{ij} - \psi_{ij}^k)}{\phi_{ij}^2(\phi_{ij}+1)},$$

where $\phi_{ij} = \sum_{k=1}^{K} \psi_{ij}^k$.

Since for at least one $k \in \{1, ..., K\}$, ψ_{ij}^k is an increasing function of the number of reviews read, it is easy to see that

$$\lim_{\phi_{ij} \to \infty} \mathrm{Var}[S_{ij}^k] = 0.$$

This means that in the limit, as more and more reviews are sampled, the uncertainty about the probability that j is of quality k will be eliminated. Intuitively, if an infinite number of past buyers evaluated a product and a new consumer reads all these reviews, the latter will be certain about the probability distribution of the product quality.

Note that $\mathrm{Var}[S_{ij}^k]$, the uncertainty about the probability of quality level k decreases with additional sampled reviews as long as ψ_{ij}^k is low enough, or more precisely as long as[2]

$$\psi_{ij}^k < \frac{\phi_{ij}(2\phi_{ij}+1)}{3\phi_{ij}+2}.$$

This means that in most cases new reviews lower $\mathrm{Var}[S_{ij}^k]$ for each k, except when ψ_{ij}^k is high but the new reviews do not increase ϕ_{ij} enough. This can occur when the consumer initially has a high confidence that the quality will be k^{HIGH}, but then the review just sampled indicates a low quality level. In this case, the consumer might become more uncertain about the probability that the product is of quality k^{HIGH}.

Uncertainty about subjective product quality is then to be captured by the term

$$\mathrm{Var}(Q_{ij}|m_{ij}) = \sum_{k=1}^{K} \mathrm{Var}[S_{ij}^k|m_{ij}],$$

indicating that it is a function of the m_{ij} ratings viewed by the consumer, and that variance of each evaluation level k equally contributes to quality uncertainty.

3.3.4 Marginal Benefit from Reading Reviews

Denote the number of reviews read so far as m_0, the total product reviews of item j as M and the consumer's subjective quality about the product as S (i.e. ignore the i and j subscripts for notational simplicity). At any given point during product search, the consumer believes that the probability that an additionally sampled rating is k is $\mathbb{E}(S^k|m_0)$. Intuitively, the consumer's best forecast for the probability that a randomly selected review has a rating of k equals her best forecast for the probability of that product having a level of quality k.

We propose that the consumer forms rational expectations about her expected belief parameters will look like following an update. Henceforth, we assume that the consumer expects her quality belief to be updated according to the formula

$$\mathbb{E}_{m_0}[\psi_{m_0+m}^k] = \psi_{m_0}^k + \tau\mathbb{E}_{m_0}[r^k] = \psi_{m_0}^k + \tau m \frac{\psi_{m_0}^k}{\sum_{k=1}^K \psi_{m_0}^k},$$

where the last equality comes from the expected value of the Dirichlet distribution.[3]

In the Appendix we prove that the expected benefit from reading m additional reviews is

[2] Taking the derivative of the variance we obtain

$$\frac{\mathrm{dVar}[S^k]}{\mathrm{d}\phi} = \frac{\psi^k}{\phi^2(\phi+1)} - \frac{2\psi^k(\phi-\psi^k)}{\phi^3(\phi+1)} - \frac{\psi^k(\phi-\psi^k)}{\phi^2(\phi+1)^2}.$$

With $0 < \psi^k, 0 < \phi$, we then have

$$\frac{\mathrm{dVar}[S^k]}{\mathrm{d}\phi} < 0 \iff \psi^k < \frac{\phi(2\phi+1)}{3\phi+2}.$$

[3] We further assume that when the consumer makes inference about the variance (not the mean) about an arbitrary function $f(\psi_{m_0+m})$, she considers ψ_{m_0+m} as if it was known and equal to $\mathbb{E}[\psi_{m_0+m}]$. From the perspective of the consumer, this assumption implies bounded rationality in the sense that the consumer only has to calculate her updated belief parameters. From the perspective of the researcher, this assumption is convenient for computational purposes on which we elaborate in the Appendix.

$$\mathbb{B}_{m+m_0,m_0} = \frac{1}{\phi^2}\sum_{k=1}^{K}\left[\psi^k\phi^-\left(\frac{1}{\psi^k+1} - \frac{1}{\psi^k+1+\frac{m\tau\psi^k}{\phi}}\right)\right],$$

where $\phi = \sum_{k=1}^{K}\psi^k$ and $\phi^- = \sum_{l=1,l\neq k}^{K}\psi^l$.[4]

It is easy to see that $\mathbb{B}_{m+m_0,m_0}$ is positive. Because of the term $\frac{1}{\phi^2}$,

$$\lim_{\phi\to\infty}\mathbb{B}_{m+m_0,m_0} = 0.$$

Since ϕ is unbounded whenever at least one of its components is unbounded, for any $k \in (1, .., K)$ we also have

$$\lim_{\psi^k\to\infty}\mathbb{B}_{m+m_0,m_0} = 0,$$

meanings that the marginal benefit from sampling further ratings goes to zero because as at least one element of ψ can unboundedly increase (the vector of the empirical distribution of the ratings must contain at least one positive element). In other words, in the limit if the consumer has already sampled enough ratings, sampling more provides no benefit whatsoever.

It also holds that

$$\frac{\partial\mathbb{B}_{m+m_0,m_0}}{\partial m} > 0, \text{ and } \frac{\partial\mathbb{B}_{m+m_0,m_0}}{\partial\tau} > 0,$$

i.e. the marginal benefit is an increasing function of m (the number of reviews that are going to be sampled) and of τ (the update parameter). Intuitively, one learns more from more ratings. Also, if a consumer updates her belief upon reading reviews heavily, the benefit from learning the ratings of not yet seen reviews is higher.

The following simple examples illustrate the shape of $\mathbb{B}_{1+m_0,m_0}$ (the benefit from sampling one more rating) for three selected cases ($\psi^1 = 0.1, \psi^1 = 1, \psi^1 = 10$) with $K = 2, \tau = 1, \gamma^1 = -1, \gamma^2 = 1$. Since $\gamma^1 < 0$ and $\gamma^2 > 0$, in this example quality decreases with more ratings of type $k = 1$ (low), and increases with more ratings of type $k = 2$ (high). On Figures 1-3, we show how the marginal benefit depends on ψ^2 (the parameter of high quality probability) for a fixed ψ^1 (the parameter of low quality probability). As ψ^2 goes to zero, the expected probability of a high draw also goes to zero. This implies that around $\psi^2 = 0$, there is very few benefit from sampling. As ψ^2 moves away from zero, the expected probability of a high draw increases, and the variance of the high draw probability shrinks - leading to a larger marginal benefit from sampling. Where the marginal benefit reaches its maximum, the probability of good and bad draw are close to each other (exactly how close depends on the other parameters). Further increasing ψ^1 after this peak point now decreases the marginal benefit. The more likely a high draw is, the the less informative is an additional rating about quality. As ψ^2 goes to infinity (i.e. probability of the high draw is almost 1), the benefit from sampling almost fully evaporates.

[4]We omitted the m_0 subscripts in ϕ to avoid notational clatter.

Figure 1: Marginal benefit ($\psi^1 = 0.1$)

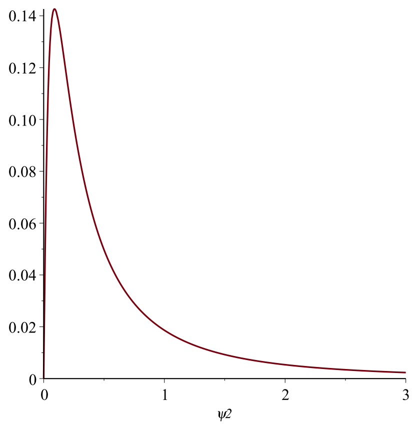


Figure 2: Marginal benefit ($\psi^1 = 1$)

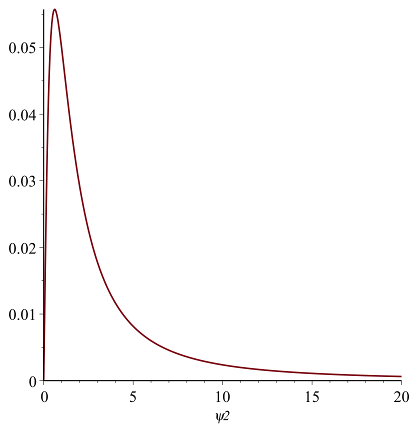
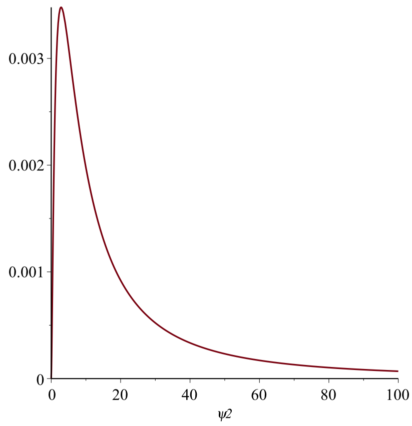


Figure 3: Marginal benefit ($\psi^1 = 10$)



Notice that for higher ψ^1 values the marginal benefit is smaller - for the same ψ^2/ψ^1 ratio. This happens because the more $\psi^1 + \psi^2$ is, the more certain is the consumer about the probability distribution of bad/good outcomes. In the limit as the sum of the elements goes to infinity, she perfectly knows the probabilities regarding the next draw, so further sampling is not informative any more.

3.3.5 Optimal number of reviews to read before page-view

In our model, the consumer calculates the optimal number of reviews to read, $m^*_{m_0}$, at each point during her search whenever she has to make a decision about whether to collect further information about the item. This also holds for the decision whether to view the item's product-page. We denote the optimal number of reviews to read before opening the product-page as $m^*_{\ominus}$, where the subscript $\ominus$ indicates that the product page has not been opened ct (no reviews about a product can be read without opening its product-page first). As search progresses and quality beliefs are consequently augmented, the optimal number of reviews to read about a product can change, corresponding to the sequential nature of the search process. Following earlier works (Kim et al. 2010, Ursu 2018), acquired information about product j leads to no information gain about any other product but j.

The net benefit from review readings before viewing the product page is the difference between the gain from reading the ex-ante optimal number of reviews about the product minus the cost of reading these reviews

$$\mathbb{B}_{m^*_{\ominus}} - c^R m^*,$$

where m^* is the solution to

$$m^*_{\ominus} = \arg\max_{m \in R} \mathbb{B}_{m_{\ominus}} - c^R m,$$

and c^R is the marginal cost of reading a review following a lognormal distribution with parameters μ^{c^R}, σ^{c^R}. (Expectations about the page-view shock ϵ are zero for any product, hence do not appear in the above expression.) The set R is the amount of reviews potentially seen that depends on the website: e.g. if the website displays reviews in batches of 5 then R contains positive multiples of 5 up to the total number of posted reviews M.

Then

$$\mathbb{B}_{m^*_\ominus} = \mathbb{E}\left[U|m^*_\ominus\right] - \left[U|m_\ominus\right] = \frac{1}{\phi^2_{m_\ominus}} \sum_{k=1}^{K} \left[\psi^k_{m_\ominus} \phi^-_{m_\ominus} \left(\frac{1}{\psi^k_{m_\ominus}+1} - \frac{1}{\psi^k_{m_\ominus}+1+\frac{m^*_\ominus \tau \psi^k_{m_\ominus}}{\phi_{m_\ominus}}} \right) \right] =$$
$$\frac{1}{\phi^2} \sum_{k=1}^{K} \left[\psi^{0^k} \phi^- \left(\frac{1}{\psi^{0^k}+1} - \frac{1}{\psi^{0^k}+1+\frac{m^*_\ominus \tau \psi^{0^k}}{\phi}} \right) \right].$$

The consumer calculates $m^*_\ominus$ then right at the beginning of search (once on the category-page) by calculating the net benefit of reading $m \in R$ reviews about product j and picking $m^*_\ominus \in R$ that provides the highest benefit. This is then done for all product as well at the beginning of search. The consumer then opens the product page with the corresponding optimal number of reviews-to-read in mind.

The realized number of reviews - which we denote as $\tilde{m}$ - can differ from this ex-ante optimal number depending on the realized page-view shock and reviews. This is because reviews that will be opened will not necessarily have the ratings the consumer expects. If that happens, it will shock the system: beliefs about the quality distribution will not be updated as forecasted, so quality uncertainty and the product's expected utility will change.

Identification is possible as the researcher observes which reviews the consumer has seen at the end and in which order.

3.3.6 Reservation Utilities

The reservation utility is defined as the hypothetical value that makes the consumer equivalent between sampling the item and seizing search. Only products have reservation utilities, reviews do not.

Sampling product j is an option that adds a new option: scrolling to reviews of j. This then adds a new option: paginating to the second review-page of j. This also adds a new option: paginating to the third review-page of j, etc. This chain provides an option value for product j that must be taken into account when calculating the reservation utility of sampling the product.

A we described, at the beginning of search the consumer can select the optimal number of reviews to be read about each product. By construction, a new realization of ψ_{ij} - i.e. the review-shock - modifies the utility of the product that is currently in view. However, the reservation utility of searching for a not yet seen products does not change with ψ_{ij}. Thus, discovered reviews only affect the benefit from reading further reviews of the current product, and any information discovered during search - including reviews - leaves the reservation utility of a not yet viewed product unchanged - as it is the standard in the literature (Kim et al. 2010; Ursu 2018). This is key, as it implies that we can use the reservation utilities of products right at the beginning of the search to infer the order in which the products are to be sampled. Furthermore, since reviews can only be read about the last searched product, we do not face the curse of dimensionality. For a proposed set of parameters, the researcher knows at start in which order the consumer is going to sample the products.

The before page-view utility of a product can be written as

$$U_{m^*_\ominus} = x\beta + \mathbb{E}(Q|m^*_\ominus) - \mathrm{Var}\left(Q|m^*_\ominus\right) + \mathbb{E}(\epsilon).$$

Notice that

$$\mathbb{E}(Q|m^*_\ominus) = \sum_{k=1}^{K} \gamma^k \mathbb{E}(S^k|m^*_\ominus),$$

$$\mathbb{E}(S^k|m^*_\ominus) = \frac{\psi^k(m^*_\ominus, \psi^{0k})}{\sum_{k=1}^{K} \psi^k(m^*_\ominus, \psi^{0k})} = \frac{\psi^k(0, \psi^{0k})}{\sum_{k=1}^{K} \psi^k(0, \psi^{0k})} = \psi^{0k},$$

and

$$\mathrm{Var}(Q|m^*_\ominus) = \sum_{k=1}^{K} \mathrm{Var}[S^k|m^*_\ominus] = \sum_{k=1}^{K} \mathrm{Var}[S^k|0] =,$$

$$\mathrm{Var}[S^k|0] = \frac{\psi^k_0(\phi_0 - \psi^k_0)}{\phi_0^2(\phi_0+1)} = \frac{\psi^{0k}(\phi_0 - \psi^{0k})}{\phi_0^2(\phi_0+1)},$$

where $\phi_0 = \sum_{k=1}^{K} \psi^k_0 = \sum_{k=1}^{K} \psi^{0k}$.

As a consequence, all terms are known in the before-pageview utility equation, which we can then write as

$$U_{m^*_\ominus} = x\beta + Q_{m^*_\ominus} - \mathrm{Var}\left(Q_{m^*_\ominus}\right).$$

Consequently, the consumer assigns reservation utilities at the very beginning of search to each item. These reservation utilities will not change during search because none of the resolved page-view shocks (ϵ), neither the uncovered reviews influence the utility of other items other than product in view.

The products are to be searched in the decreasing order of the reservation utilities z_j that are solutions to

$$c^P = \mathbb{A}_j(z_j) = \int_{z_j}^{\infty} (U_{m^*_{\ominus j}} - z_j) f(U_{m^*_{\ominus j}}) \, \mathrm{d}U_{m^*_{\ominus j}},$$

assuring that the marginal cost of searching product j equals the marginal benefit of searching it. (For the derivation of this equation see Kim et al. (2010).) We assume consumers value their time equally when it is spent on reading reviews and when it is spent on reading product-pages. I other words, for a consumer the per second effort of information collection is the same, although this per second cost is heterogeneous across consumers. Therefore the cost of reading a product page is linearly related to the cost of reading a review:

$$c_i^P = \lambda c_i^R, \ \forall i \in \{1, ..., N\}$$

3.3.7 Consumer Search Path

Search starts with observing the category-page, and calculating the reservation utilities z of all products based on information available there. Then the consumer ranks the products in decreasing order of reservation utilities and the search for products begins. (The product with the highest reservation utility is always searched.) Denote by $O(n)$ the identity of the product with the n^{th} largest reservation utility. Then formally (with consumer subscripts omitted for simplicity),

$$z_{O(n)} \geq \max_{k=n+1}^{J} z_{O(k)}, \ \forall n \in \{1, ..., J-1\},$$

where J is the number of products available to consumer i.

Just as in the previous literature (Kim et al 2010, Ursu 2018), the consumer entirely seizes her search at the retailer's website, provided that at that given point in the search process all unsearched products have a lower reservation utility than the maximum utility of the searched alternatives, including the outside option (utility in hand). On the other way round, if the consumer makes an n^{th} search, it must be that the reservation utility of that product exceeds her utility in hand. Mathematically,

$$z_{O(n)} \geq \max_{k=0}^{n-1} U_{O(k)}^{\tilde{m}}, \ \forall n \in \{1, ..., s\},$$

and

$$z_{O(n')} \leq \max_{k=0}^{s} U_{O(k)}^{\tilde{m}}, \ \forall n' \in \{s+1, ..., J\},$$

where s denotes the number of products searched.

In addition to this, we introduce that the consumer stops sampling any further reviews of a product whenever the utility of that product upon reading more reviews cannot be higher than the maximum utility of the searched alternatives, including the outside option and the last searched product about which m_0 reviews have been read. Intuitively, if the consumer, at any stage of the search process, forecasts that despite learning any possible number of additional ratings of a product it will not be better than her best alternative so far, it is not economic to spend more effort on reading any more reviews about that item. This is because, the consumer expects that regardless of the additional number of ratings to be seen, she will not buy that product. Formally, in order to read further reviews there must be some positive integer m for which the following condition holds

$$\exists_{m \in R^{m_0^+}} \max_{k=0}^{n} U_{O(k)}^{\tilde{m}} < \mathbb{E}[U_{O(n)}|m + m_0], \ \forall m_0 \in \tilde{R}^{m_0},$$

where $R^{m_0^+}$ is the set of additional reviews possible to read at the website provided m_0 reviews has been read and $\tilde{R}^{m_0^-}$ is the set of number of reviews the consumer has read about the product except the finally read number $\tilde{m}$.

Since we have shown that the marginal benefit in terms of reducing quality uncertainty by reading reviews is an increasing function of the number of reviews that are going to be sampled

$$\frac{\partial \mathbb{B}_{m+m_0,m_0}}{\partial m} > 0,$$

it must be true that if there exists such an m that satisfies the above condition, the following also holds

$$\max_{k=0}^{n} U_{O(k)}^{\tilde{m}} < \mathbb{E}_{m_0}[U_{O(n)}|M], \ \forall m_0 \in \tilde{R}^{m_0},$$

saying that it if further review reading happened following the sampling of m_0 reviews, it must be that reading all available reviews M was expected to result in an utility that is larger than the utility in hand.

Review sampling for a given product also stops whenever the marginal benefit from reducing uncertainty about quality (as described in 2.3) exceeds its marginal cost. Thus if further reviews sampling occurred, it must be that

$$\mathbb{B}_{m_1+m_0,m_0} > c^R m_1, \ \forall m_0 \in \tilde{R}^{m_0^-},$$

where $m_1 = \min R^{m_0^+}$ is the smallest number of reviews possible to read provided m_0 reviews have been read (e.g. if reviews are displayed in batches of 5, $m_1 = 5$).

The two conditions above imply that review reading for a product stops at $\tilde{m}$ reviews read whenever (i) there is no number of ratings to sample after which the expected utility is higher than the utility in hand *or* (ii) the marginal benefit of reducing quality uncertainty about the product is lower than the marginal cost of reading reviews. Formally,

$$\max_{k=0}^{n} U_{O(s)}^{\tilde{m}} \geq \mathbb{E}_{\tilde{m}}[U_{O(s)}|M],$$

or

$$\mathbb{B}_{m_1+\tilde{m},\tilde{m}} \leq c^R m_1.$$

Finally, the consumer chooses the product (or the outside good) with the highest utility provided that search for product and reviews are both over. Formally,

$$U_j^{\tilde{m}} \geq \max_{k=0}^{s} U_{O(k)}, \ \forall j \in O \cup \{0\}.$$

3.4 Parameter Estimation

The probability that the consumer searches in order O, reads exactly $\tilde{m}$ reviews about the J products, and chooses product j is

$$P_{jO\tilde{m}} = \Pr\Bigg(z_{O(n)} \geq \max_{k=0}^{n-1} U^{\tilde{m}}_{O(k)} \cap z_{O(n)} \leq \max_{k=0}^{s} U_{O(k)} \cap$$
$$\max_{k=0}^{n} U^{\tilde{m}}_{O(k)} < \mathbb{E}_{m_0}[U^{\tilde{m}}_{O(n)}|M] \cap \mathbb{B}_{m_1,m_0} > c^R m_1 \cap$$
$$\big(\max_{k=0}^{n} U^{\tilde{m}}_{O(s)} \geq \mathbb{E}_{m_0}[U^{\tilde{m}}_{O(s)}|M] \cup \mathbb{B}_{m_1,\tilde{m}} \leq c^R m_1\big) \cap U^{\tilde{m}}_j \geq \max_{k=0}^{s} U^{\tilde{m}}_{O(k)} \Bigg) =$$
$$\mathrm{I(cond)} f_\epsilon(\epsilon) f_{\psi^0}(\psi^0) f_{c^R}(c^R) \mathrm{d}\epsilon \mathrm{d}\psi^0 \mathrm{d}c^R,$$

where I(cond) is an indicator function for whether the search conditions above hold. Then the log-likelihood becomes

$$LL = \sum_i \sum_{O(i)} \sum_{\tilde{m(i)}} \sum_j d_{ijO(i)\tilde{m}(i)} \log P_{ijO(i)\tilde{m}(i)}$$

As this integral has not closed-form solution, we approximate the log-likelihood of the model with is simulated counterpart

$$SLL = \sum_i \sum_{O(i)} \sum_{\tilde{m(i)}} \sum_j d_{ijO(i)\tilde{m}(i)} \log \hat{P}_{ijO(i)\tilde{m}(i)}$$

using the following logit-smoothed AR simulator strategy.

3.4.1 Initial Beliefs

Drawing ϵ and c^R is a standard procedure, unlike drawing ψ^0. To ensure that a product with a ceteris paribus higher average rating will have a belief that indicates higher quality on average, we propose the following algorithm to generate draws for ψ^0 (j subscripts omitted for simplicity):

- Create K bins according to the number of potential evaluations. Initialize M reviewers, where M equals the number of reviews submitted for the product.
- Draw the evaluation of the first reviewer ($l = 1$) resulting in rating k_1. Each such rating draw comes from the empirical distribution of the ratings on the retailer's website in the respective product category.
- Evaluate the following condition

$$\min \sum_{n=m+1}^{M} d_n < AR * M - d_l - \sum_{n=1}^{m-1} d_n < \max \sum_{n=m+1}^{M} d_n,$$

where AR is the average rating of the product. The above condition states that one should be always able to get further draws such that it is not a-priori impossible to reach a number of draws that gives rise to the observed average rating.

- If this condition is violated by a recent draw d_m, disregard it and get a new one.
- Draw d_m for $m = 1, 2, 3, ...M$ as long as a draw is obtained such that it does not violate the above condition.
- Once M draws are obtained, collect the accepted draws to obtain the size the bin D^k for $k = 1, 2, ..., K$.
- Draw D 1000 times, and denote its mode as $\overline{D}$.
- The common prior is obtained by $\overline{\Psi} = c(\overline{D}) + 1$, where the unit vector ensures positivity required for Dirichlet parameters, and c is a constant chosen by the researcher. A higher c leads to more homogeneous initial beliefs, while a lower c allows for a wide dispersion of initial beliefs among consumers. For products without reviews, D simply corresponds to the empirical distribution of ratings, which is assumed to be known to the consumer.
- Importantly, $\overline{\Psi}$ - is drawn only once, at the beginning of the estimation.
- With a $\overline{\Psi}$ in hand, we draw $\psi_i^0 \sim Dir(\overline{\Psi})$, the individual level initial beliefs.

Notice, that by the functional assumption we used to generate the common belief parameter, the higher the number of reviews a product has the smaller will be the spread of ψ_{ij}^{0k} across consumers, for each k. This is due to the Dirichlet distribution as higher Dirichlet parameters result in lower variance of the draws. This is a key property of our model. It implies that consumers are more certain about the quality of a product if that has been reviewed by more previous customers. It also implies that if a product has few or no reviews, consumers are more uncertain about its benefits.

3.4.2 The Logit-Smoothed AR Simulator

Define the set of products as $j \in \{1, ..., J\}$, with x_j and y_j unique to each product, where y_j denotes the vector of accessible review ratings for product j in the order as they can be accessed by the consumers. Some products might not be available to all consumers visiting the site at time t, for two reasons: the retailer does not offer that product at the time t; or the product has a different y_j at time t than time $t' \neq t$. The latter occurs because whenever the set of accessible reviews for product j changes (because of the arrival of new reviews), a new product j' is created. In other words, consumers who visit the website with a new set of reviews accessible cannot view the same product with the old reviews: i.e. the product with the old reviews is not available to them. We set the reservation utility of products that are not available to consumer i to $-\inf$.

We take the following steps to obtain the logit-smoothed AR simulator for our model likelihood given the set of parameters $\Omega = \{\beta, \gamma, \sigma^\epsilon, \mu^{cost^R}, \sigma^{cost^R}, \tau, \lambda\}$.

1. For each consumer $i \in \{1, ..., N\}$ in a dataset, generate D pseudo-consumer

- who face the same product offering - i.e. have the same set of available products - as consumer i;
- whose taste (ϵ_{id}), belief (ψ^0_{id}) and cost (c^R_{id}) draws come from the distributions as of consumer i's: $(f_\epsilon|\sigma^\epsilon)$; $(f_{\psi^0}|\bar{\Psi})$; $(f_{c^R}|\mu^{cost^R}, \sigma^{cost^R})$.

2. Use the draws to form $m^*_{\ominus i_d}$ and by it $U_{m^*_{\ominus i_d}}$ for each pseudo-consumer.

3. Form reservation utilities z_{i_d} for each pseudo-consumer.

4. Define the followings for each pseudo-consumer i_d had she sampled exactly the same products and exactly the same reviews as consumer i did (i.e. pseudo-consumers do not have their own choices, they follow the path of consumer i, we just require them to reveal their respective utility values):
   - $v^1_{i_d} = z^{O_i(n)}_{i_d} - \max_{k=0}^{n-1} U^{\tilde{m}O_i(k)}_{i_d}$, i.e. the differences between the pseudo-consumer's reservation utilities and up-to-that-point in-hand utility for each product searched (except for the always searched first product);
   - $v^2_{i_d} = \max_{k=0}^{h} U^{\tilde{m}O_i(k)}_{i_d} - z^{O_i(n')}_{i_d}$ i.e. the differences between the pseudo-consumer's in-hand utility at the end of search and the reservation utilities of each product not searched;
   - $v^3_{i_d} = \mathbb{B}^{m_1+m_0,m_0}_{i_d} - c^R_{i_d} m_1$, i.e. the pseudo-consumer's differences between the benefit of reading the next set of reviews and the cost of that for each product and intermediate number of read reviews consumer i have read (i.e. at point where consumer i decided to sample more ratings);
   - $v^4_{i_d} = c^R_{i_d} m_1 - \mathbb{B}^{m_1+\tilde{m},\tilde{m}}_{i_d}$, i.e. the pseudo-consumer's differences between the cost of reading the next set of reviews and the benefit of that for each product and total reviews consumer i have read about those;
   - $v^5_{i_d} = \mathbb{E}_{m_0}[U^{O(n)}_{i_d}|M] - \max_{k=0}^{n} U^{\tilde{m}O(k)}_{i_d}$, i.e. the difference between the pseudo consumer's expected utility after reading all reviews about the product and her utility in hand at points when consumer i decided to sample further ratings;
   - $v^6_{i_d} = \max_{k=0}^{n} U^{\tilde{m}O(k)}_{i_d} - \mathbb{E}_{\tilde{m}}[U^{O(n)}_{i_d}|M]$, i.e. the difference between the pseudo consumer's utility in hand and expected utility after reading all reviews about the product and at points when consumer i decided to sample no further ratings;
   - $v^7_{i_d} = U^{\tilde{m}O_i(k)}_{i_d} - \max_{k=0}^{h} U^{\tilde{m}O_i(k)}_{i_d}$, i.e. the difference between the searched products post-review utilities and the chosen product's post-review utility.

5. Define smoothing parameters $w^O < 0$, $w^R_1 < 0$, $w^R_2 < 0$, $w^P < 0$. There is no general guidance how to choose these parameters. The larger smoothing parameters (in magnitude) are used, the lower is the bias as the simulator approaches the AR simulator. However, too large values can lead to numerical instability. For instance, Ursu (2018) performs a grid search over these values to find the best values for the Monte-Carlo simulation, while Kim

et al. (2010) simply set these to -5. We follow the grid-search strategy of Ursu (2018). We explain our strategy to pick these smoothing parameters in the respective chapter in more details.

6. Compute the followings

   - $S_{i_d}^O = \frac{1}{1+\sum_i \exp(w^O * v_{i_d}^1) + \exp(w^O * v_{i_d}^2)}$;
   - $S_{i_d}^P = \frac{1}{1+\sum_i \exp(w^P * v_{i_d}^7)}$;
   - $P_{i_d}^{m_0} = \exp(-w^{R_1} v_{m_0 i_d}^3) + \exp(-w^{R_2} v_{m_0 i_d}^4)$;
   - $P_{i_d}^{\tilde{m}} = \min\left(\exp(w^{R_1} v_{m_0 i_d}^3), \exp(w^{R_2} v_{m_0 i_d}^4)\right)$;
   - $\mathbb{P}_{i_d} = \frac{1}{s_i} \sum_{s_i} \left(P_{i_d}^{\tilde{m}} \prod_{m_0} P_{i_d}^{m_0}\right)$;
   - $\hat{P}_i^O = \frac{1}{D} \sum_d S_{i_d}^O$;
   - $\hat{P}_i^R = \frac{1}{D} \sum_d \frac{1}{1+\mathbb{P}_{i_d}}$;
   - $\hat{P}_i^P = \frac{1}{D} \sum_d S_{i_d}^P$.

7. The simulated probability for the entire search path of consumer i is

$$\hat{P}_{ijO(i)\tilde{m}(i)} = \hat{P}_i^O \hat{P}_i^R \hat{P}_i^P$$

3.5 Identification of Model Parameters

In principle, the functional form allows the joint identification of μ^{c^P} and σ^ϵ, but - as it is standard practice in the literature (Kim et al 2010, Ursu 2018) - it is advised to fixed one of these two. This is because product search intensity increases both with a lower c^P an with a higher σ^ϵ.

Similarly to the product search intensity, review search intensity is also affected by two parameters, namely τ and μ^{c^R}. Once again, in theory the functional form would allow for joint identification of these two, but the researcher would require an extraordinary rich dataset (and an extraordinary powerful computer) to do so. Therefore, we opt for fixing one of these two as well. As τ also enters indirectly to the reservation utilities, while c^R does not our choice is to fix τ. Fixing the update parameter might be justifies by an experiment in which the researcher quantifies how strong recipients update their quality beliefs. We plan to carry out such as experiment for the updated version of this paper. Alternatively, the researcher can fix τ, estimate the model and obtain a pre-defined predictive fit using the estimated parameters. Then, the model can be re-estimated with different fixed τ, and the one with the best predictive fit on data (or test set) could be chosen.

Data on each decision of the consumer whether to read more reviews is required for the identification of τ. Such data also contributes to the identification of γ and β, which are also identified by the product search path and purchase decisions.

The cost ratio between reading a review and reading a product-page, just as in the earlier literature, c^P $(\lambda * c^R)$ is identified through the product search path.

Parameters that affect utility through perceived quality are identified from search and purchase decisions (better perceived products are searched and purchased more often), but importantly also from the decision whether to read more reviews about the product. For instance, if the expected utility heavily drops after encountering a low rating (leading to worse expected quality), consumers are more prone to stop reading reviews about that item. From the γ vector, one element needs to be fixed for identification purposes. We suggest normalizing the parameter corresponding to the lowest evaluation level to zero.

Cost heterogeneity is identified by the differing length of search across consumers, just as by the different number of reviews read.

3.5.1 Monte-Carlo Simulation

In order to show that the proposed logit-smoothed AR simulator can recover the model parameters, I test its performance on simulated data. We generate 5000 consumers and 500 products with unique prices and reviews, leading to 25,000 observations (an observation is a product-consumer combination - searched or not, reviewed or not). Each consumer has exactly 5 products to search for and to choose from, not including the outside good. Products have integer ratings between 1 and 5, for each of their reviews. The empirical rating distribution across all products is uniform and known to the consumers. We draw 20 pseudo-consumers for each consumers, with pseudos facing the same product offering as their matching consumer.

In the generated dataset (42.8%,21.1%,16.2%,12.9%,6.9%) of the consumers have searched for (1,2,3,4,5) products, respectively. 47.2% choose to buy a product instead of the outside good. 70.2% in the cases where reviews were accessible consumers read no reviews, while 5 reviews were read in 19.8% of the cases and 25 reviews were read in 6.7% of the cases (other values than these were read in less than 1% of the cases).

As our retailer displays 5 reviews on the products' first review-page (and 20 additional reviews on higher order review-pages accessible through pagination), we define $cost^R = 5c^R$, and parameterize $\lambda = \frac{c^P}{cost^R}$ instead of $\lambda = \frac{c^P}{c^R}$. Thus, a $\lambda = 1$ would mean that reading the product page is just as costly to the consumer as reading the first set of reviews.

We fix the parameter c that control the magnitude of the heterogeneity of the initial beliefs to 1000, resulting in very small initial belief heterogeneity. We further fix $\tau = 0.5$, which corresponds to a belief update of moderate intensity; and $\gamma(1) = 0$ which might be interpreted as products of 1-star (subjective) quality have zero contribution to the utility. We parameterize $\gamma(k) = \sum_{k=1}^{5} |\bar{g}(k)|$ and estimate $\bar{g}$ instead of directly estimating γ. Finally, we fix the cost heterogeneity parameter $\mu^{cost^R} = 0$.

To select the best hyper-parameters we run the estimation for all possible combinations of

$$w^k \in \{-1, -3, -5, -10\}$$

for $k \in \{O, P, R_1, R_2\}$.[5]

To measure how good we can predict consumers' behavior at the website, we define a prediction error function. With the optimal hyperparameters from the grid-search, the value of this function is the lowest among all combinations. The prediction error function we use is

$$RMSE(\boldsymbol{w}, \tau, \hat{\Omega}) = \sqrt{\sum (S - \hat{S}(\boldsymbol{w}, \tau, \hat{\Omega}))^2} + \sqrt{\sum (P - \hat{P}(\boldsymbol{w}, \tau, \hat{\Omega}))^2} + \sqrt{\sum (R - \hat{R}(\boldsymbol{w}, \tau, \hat{\Omega}))^2},$$

where S is the distribution of number of products searched among consumers (in percentage points), P is the market share of outside good (in percentage points), R is the distribution of reviews read about a product (in percentage points), while $\hat{S}, \hat{P}, \hat{R}$ are the predicted values of these vectors given hyperparameters used to maximize the model likelihood. We set $S, \hat{S}$ to zero for searches longer than 5 products; and $R, \hat{R}$ to zero for number of read reviews different than $0, 5, 25$, which in most cases corresponds to the reading of $0, 1, 2$ review-pages.

We find that the hyperparameters $w^O = -3$, $w_1^R = -10$, $w_2^R = -1$, $w^P = -3$ provide the best fit. The true and estimated parameter values are shown in Table 1. We can see that even with a small number of pseudo consumers the parameters are recovered quite well.[6]

Table 1: Monte-Carlo Results

Description	Parameter	True value	Estimate
Price (log)	$\beta(1)$	1	1.06
#Reviews (log)	$\beta(2)$	1	1.02
Rating = 2	$\gamma(2)$	1	1.03
Rating = 3	$\gamma(3)$	2	2.85
Rating = 4	$\gamma(4)$	4	4.30
Rating = 5	$\gamma(5)$	7	7.44
Cost level	μ^{cost^R}	-3.6	-3.48
Cost ratio	λ	1	0.72
Outside utility	μ^{ϵ_0}	4	4.10
Log-likelihood			-11,976
Observations			25,000

[5]For the empirical estimation, had we possess a large enough sample, we would use part of our dataset as training data to calibrate the hyperparameters to avoid overfitting. In the absence of a large dataset, we have to rely on the whole data to select the best set of hyperparameters. For retailers who would like to run our model, calibrating it on both a training and a test set should mean no problem: e.g. they could use half year of historical data as training set, while another half year as test set - or a combination of these.

[6]Ursu (2018) uses 150 draws for her model requiring fewer parameters to be estimated. Number of pseudo-consumers is to be increased in later versions of the current paper, the small number here is due computational reasons.

4 Empirical Application

4.1 Search, Reading Reviews, and Purchases at an Online Retailer

We have obtained a data set that covers the click-stream activity of visitors to a large UK retailer in February and March 2015. We have access to the consumers' activity in the Digital TV Recorders category, more precisely we know (i) what products they searched for and what time, (ii) which reviews they read about the searched products, and (iii) which product they purchased. The retailer was listing 16 products over the two month period. Reviews and prices of these products, however, could have changed over time. Since both reviews and prices are crucial inputs to consumer decisions, if the same product at two different times has different reviews and/or price we define these two as distinct items. Therefore, from now on, when we mention products we refer to the 107 different product-variants consumers could have searched for.

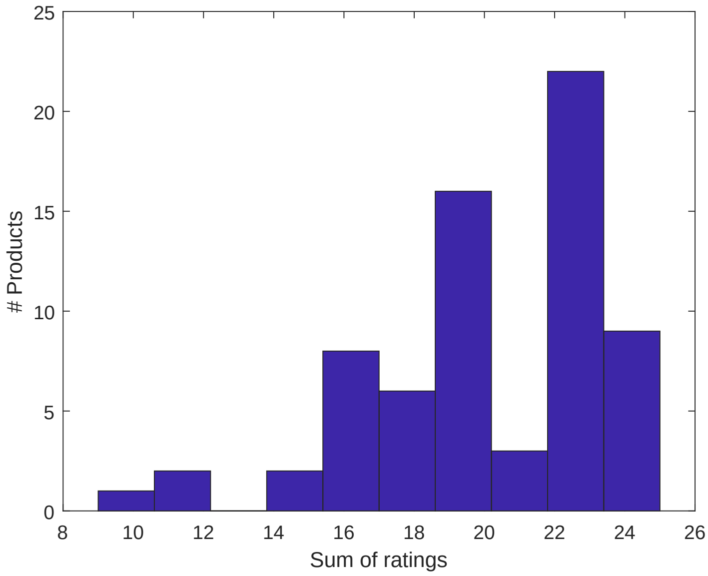
When consumers land at the retailer's website, they can see the (at time available) products listed in the category-page along with their price, number of reviews and mean rating. There, they can click on a product to receive additional textual and graphical information, defined as *page-view*. Once the product is searched for, they can scroll down to its reviews: at the bottom of the product page, the newest 5 reviews are displayed (these reviews are not visible without scrolling) which we define as *first review-page*. Should they wish to sample further reviews, consumer have the choice of paginating the review-pages; if they do so, they will view 20 more reviews. Consumers might purchase any of the previously searched item.

The mean price in the category was GBP 148.9, while the mean number of available reviews for a product was 34.2. The average of the mean product ratings was 4.2 out of the maximum 5. 9.5% of the accessible reviews had a 1-star rating, 2.7% had a 2-star rating, 4.7% had a 3-star rating, 25.5% had a 4-star rating, while 57.6% had a 5-star rating. All 107 products had at least one review available.

After dropping eight outliers consumers who carried out extra long searches, we remain with 532 consumers browsing in the category. 67.2% of them searched for one product only, 18.3% searched for two, while the longest search included 7 products. 98.3% of the consumers decided not to buy any of the products offered. 75.1% did not read any review for a search product that had accessible reviews, 20.2% read only the first review-page, 4.2% read the first two review-pages, while 0.5% the first three review-pages.

To illustrate that there is significant heterogeneity among the ratings of the first review-pages of different products, we sum the first 5 ratings of those products that have at least this amount of reviews. Figure 4 shows the distribution of this sum across products provided these products have been searched and reviewed by at least one consumer in the digital television recorder category. The graphs reveals that there is a wide range of ratings on the first review-pages, which allows us to identify the model parameters from consumers' search and purchase patterns following the discovery of these ratings.

Figure 4: Sum of first set of ratings



In order to gain insight how product characteristics influence whether consumer search it or not, we run a linear probability regression (Table 2). Results suggest that consumer tend to search for products with more reviews and higher average rating. We also run a regression to explain whether a product is reviewed, and a different one to see whether it is purchased. We find none of the product characteristics to be significant, presumably due small sample size (reviewing a product is not common, while purchasing one is very rare in our data).

Table 2: Linear Probability Estimates

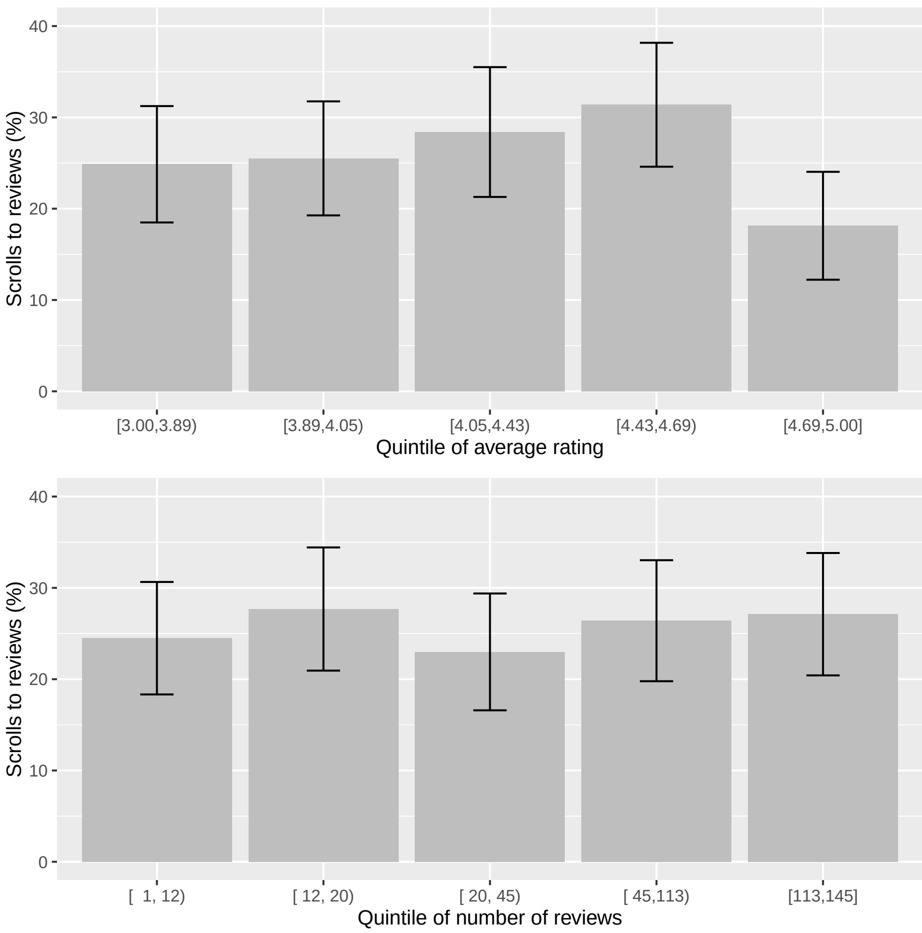
	Product searched	Product reviewed	Product purchased
Intercept	0.108 (0.118)	0.718 (0.459)	-0.034 (0.109)
Log #reviews	0.017*** (0.003)	0.014 (0.016)	-0.004 (0.003)
Mean rating	0.029** (0.011)	0.001 (0.004)	0.003 (0.011)
Log price	-0.111 (0.114)	-0.015 (0.013)	0.008 (0.027)
Product reviewed			0.011 (0.008)
Observations	8,384	839	839

***p<0.001 **p<0.01 *p<0.1.

Note: Standard errors in parentheses.

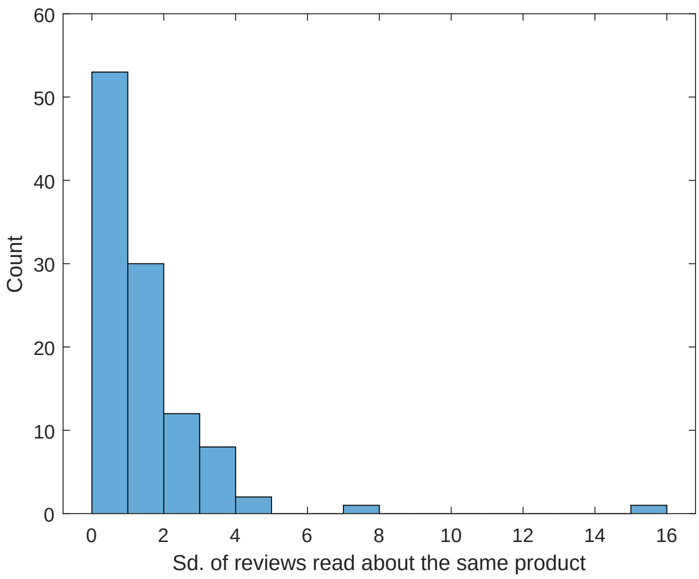
Although we see no clear evidence that mean ratings would influence whether a product is reviewed, non-linear effect might still play a part. To investigate this, we rank the product searches of consumers by the mean rating of the item and groups these observations into quintiles according to their rank. For each of these groups, Figure 5 displays the fraction of consumers who decided to read reviews after searching the product. We see that in the highest quintile the chance of reading reviews is significantly smaller than in the second quintile. Our economic argument behind this phenomenon is that in these cases consumers face low uncertainty about product quality (good quality with high probability), which lowers the benefit from reading the reviews of the product. In other words, with high certainty about product (good) quality, consumers are more likely to decide not to engage in costly review reading. The second part of the figure shows, however, that products with a high number of reviews - unlike products with high average rating - are not necessarily considered as high certainty - high quality products. Here, we have ranked the consumers according to the number search product's reviews into quintiles. Unlike before, we see no difference between the review reading propensity of consumers in different groups.

Figure 5: Fraction of consumers reading reviews



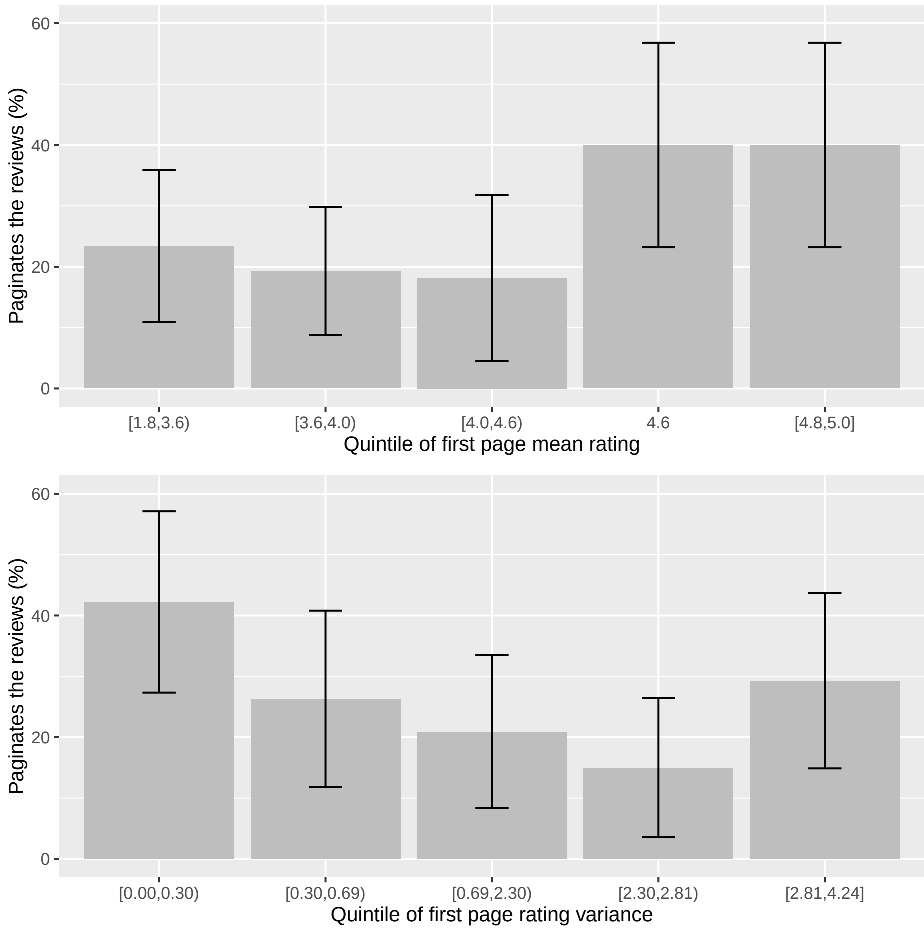
In the data, we observe that different consumers read different amount of reviews about the same product. Figure 6 displays the histogram of the standard deviation of the read reviews by products. Heterogeneity in the number of read reviews for a given product (with its reviews unchanged) can first of all come from heterogeneous cost of effort among consumers. People value time differently, so some might wish to spend more time on acquiring additional information from eWOM, others might prefer less so. Furthermore, we believe that a consumer who have already discovered an appealing item - e.g. due to her extensive product search - is less motivated to read reviews about a recently searched item than one who has not yet found such an appealing product (or one whose outside options are lower). Our modeling assumptions, then, should reflect these data patterns.

Figure 6: Heterogeneity in reading reviews



The consumer might decide not to read any more reviews about a given product once a discouraging information is discovered. If the first set of ratings are low, this can deter the consumer from the product, and then there is no further incentive to engage in costly review sampling for a not desired item. On Figure 7 we show the fraction of consumers who decided to paginate the reviews as function of the mean rating of the first 5 reviews. The bars suggest that review reading for a product is more likely to stop once low ratings are encountered. This supports our assumption that consumers update their beliefs about the product quality dynamically as their sample the reviews.

Figure 7: Fraction of consumers reading higher order review-pages



The lower part of Figure 7 groups the observations by the variance of the first 5 ratings. The data suggests a non-linear pattern here. For the lowest quintile of variance we see that around 40% of people paginated the reviews, which ratio shrinks as we move to higher quintiles up to the fourth. We interpret this as low variance maintains the desirability of the product. However, pagination propensity is high again (30%) in the fifth quintile. We believe that this is because high variance increases the uncertainty about product quality, which consumers are willing to reduce by continued reading of reviews. In either case, consumers seem to take into account the sampled ratings when making inference about product quality and further search decisions.

4.2 Estimation Results

To estimate the model on the Digital TV Recorder data, we draw 20 pseudo-consumers similarly to the Monte-Carlo simulation. Each consumer has her respective pseudo-consumers facing the same product palette just as the consumer had available at time of her search (due to review and price changes some products become "unavailable", while some become "available" to the consumers). The distribution of the ratings is skewed towards the left, which means that the chance of encountering a 2-star or a 3-star rating is very low for a consumer. Therefore, we collect the 1-star, 2-star and 3-star ratings into one group of "low" ratings. This does not affect how

$\bar{D}$ is drawn, i.e. pseudo-consumers still form an expectation about the distribution of accessible ratings such that the mean of those corresponds to the observed mean, which is displayed on the category-page. This grouping only changes $\bar{\Psi}$, and consequently the ψ^0 draws. Now $\bar{\Psi}$ is a vector of length 3, with

$$\bar{\Psi}(1) = c\left(\bar{D}(1) + \bar{D}(2) + \bar{D}(3)\right) + 1,$$

and

$$\bar{\Psi}(2) = c\left(\bar{D}(4)\right) + 1,$$

$$\bar{\Psi}(3) = c\left(\bar{D}(5)\right) + 1.$$

Analogously to the Monte-Carlo simulation, we fix c to 1000; $\gamma(1)$ to 0 (products of "low" quality have zero contribution to the utility); μ^{cost^R} to 0; and σ^ϵ to 1. As weights, we use $w^O = -3$, $w_1^R = -10$, $w_2^R = -1$, $w^P = -3$, the ones we found best in the Monte-Carlo simulation.

Table 3 shows the estimated coefficients. For this specification, the full model estimates do not match our expectations, as the coefficient for the 5-star quality is lower than for the 4-star quality. We plan to investigate how we could obtain reasonable empirical estimates and avoid potential convergence problems.

We compare our model estimates to the literature standard sequential search model (Ursu 2018, Kim et al. 2010), which does not take into account that consumers read and learn from product reviews. In this simplified version, we add the mean rating as product characteristics with a respective coefficient to $\beta(3)$ to estimate, and we use $w^O = -3$, $w^P = -3$.[7]

Results of the standard sequential search model are displayed in the last column of Table 3. For the standard model, our estimates look reasonable.

Table 3: Empirical Estimation

Description	Parameter	Estimate (Full model)	Estimate (Standard model)
Price (log)	$\beta(1)$	0.267	0.162
#Reviews (log)	$\beta(2)$	0.064	0.050
Mean rating	$\beta(3)$	-	0.132
Rating = 4	$\gamma(2)$	0.778	-
Rating = 5	$\gamma(3)$	0.269	-
Cost level	μ^{cost^R}	-1.602	-3.752
Cost ratio	λ	0.673	-
Outside utility	μ^{ϵ_0}	1.006	2.918
Log-likelihood		-1,751	-863
Observations		8,384	8,384

[7]For products without reviews, we use the average of mean ratings across products.

4.3 Counterfactual Simulation

5 Conclusion

References

Babić Rosario, A., F. Sotgiu, K. De Valck, and T. H. Bijmolt (2016). The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors. *Journal of Marketing Research 53*(3), 297–318.

Burton, J. and M. Khammash (2010). Why do people read reviews posted on consumer-opinion portals? *Journal of Marketing Management 26*(3-4), 230–255.

Chevalier, J. A. and D. Mayzlin (2006). The effects of word of mouth on sales: Online book reviews. *Journal of Marketing Research 43(3)*, 345–354.

Floyd, K., R. Freling, S. Alhoqail, H. Y. Cho, and T. Freling (2014). How online product reviews affect retail sales: A meta-analysis. *Journal of Retailing 90*(2), 217–232.

Hennig-Thurau, T., G. Walsh, and G. Walsh (2003). Electronic word-of-mouth: Motives for and consequences of reading customer articulations on the internet. *International journal of electronic commerce 8*(2), 51–74.

Honka, E. and P. Chintagunta (2016). Simultaneous or sequential? search strategies in the us auto insurance industry. *Marketing Science*.

Kim, E. E. K., A. S. Mattila, and S. Baloglu (2011). Effects of gender and expertise on consumers' motivation to read online hotel reviews. *Cornell Hospitality Quarterly 52*(4), 399–406.

Kim, J. B., P. Albequerque, and B. J. Bronnenberg (2010). Online demand under limited consumer search. *Marketing Science 29(6)*, 1001–1023.

Kim, J. B., P. Albequerque, and B. J. Bronnenberg (2016). The probit choice model under sequential search with an application to online retailing. *Management Science*.

Moe, W. W. and M. Trusov (2011). The value of social dynamics in online product ratings forums. *Journal of Marketing Research 48*(3), 444–456.

Ursu, R. M. (2018). The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions. *Marketing Science 37*(4), 530–552.

Varga, M. and P. Albuquerque (2019). Measuring the impact of a single negative consumer review on online search and purchase decisions through a quasi-natural experiment. *Marketing Science Institute Report 103*(1).

Vermeulen, I. E. and D. Seegers (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. *Tourism management 30*(1), 123–127.

Weitzman, M. L. (1979). Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society*, 641–654.

Wu, C., H. Che, T. Cha, and X. Lu (2015). The economic value of online reviews. *Marketing Science 5*(34), 739–754.

You, Y., G. G. Vadakkepatt, and A. M. Joshi (2015). A meta-analysis of electronic word-of-mouth elasticity. *Journal of Marketing 79*(2), 19–39.

Appendix

The marginal benefit from reading m additional reviews is obtained as follows.

First note that the expected values for the update are

$$\mathbb{E}[\phi_{m+m_0,m_0}] = \sum_{k=1}^{K} \left(\psi^k_{m_0} + \tau m \frac{\psi^k_{m_0}}{\sum_{k=1}^{K} \psi^k_{m_0}} \right) =$$
$$\sum_{k=1}^{K} \psi^k_{m_0} + \frac{\tau m}{\sum_{k=1}^{K} \psi^k_{m_0}} \sum_{k=1}^{K} \psi^k_{m_0} = \phi_{m_0} + \tau m.$$

We can write

$$\begin{aligned}
\mathbb{B}_{m+m_0,m_0} = \mathbb{E}_{m_0}\left[U|m_0+m\right] - \mathbb{E}_{m_0}\left[U|m_0\right] = \\
x\beta + \mathbb{E}_{m_0}[Q|m_0+m] - \mathrm{Var}_{m_0}(Q|m_0+m) + \epsilon \\
- x\beta + \mathbb{E}_{m_0}[Q|m_0] - \mathrm{Var}_{m_0}(Q|m_0) + \epsilon = \\
\mathbb{E}_{m_0}[Q|m_0+m] - \mathbb{E}_{m_0}[Q|m_0] + \mathrm{Var}_{m_0}(Q|m_0) - \mathrm{Var}_{m_0}(Q|m_0+m) = \\
\mathbb{E}_{m_0}[Q|m_0+m] - Q^{m_0} + V^{m_0} - \mathrm{Var}_{m_0}(Q|m_0+m) = V^{m_0} - \mathrm{Var}_{m_0}(Q|m_0+m),
\end{aligned}$$

where the last equality comes from the fact that expectations about quality do not change unless new reviews are actually sampled.

The above can be proven as follows. The expected quality without reading more reviews is

$$\mathbb{E}_{m_0}[Q|m_0] = \sum_{k=1}^{K} \gamma^k \frac{\psi^k_{m_0}}{\sum_{k=1}^{K} \psi^k_{m_0}}.$$

Analogously, the prior-reading expectation about past-reading expected quality is

$$\mathbb{E}_{m_0}[Q|m_0+m] = \sum_{k=1}^{K} \gamma^k \frac{\mathbb{E}[\psi^k_{m_0+m,m_0}]}{\sum_{k=1}^{K} \mathbb{E}[\psi^k_{m_0+m,m_0}]}.$$

Substituting

$$\mathbb{E}[\psi^k_{m_0+m,m_0}] = \psi^k_{m_0} + \tau m \frac{\psi^k_{m_0}}{\sum_{k=1}^{K} \psi^k_{m_0}},$$

to the expected quality after reviews sampling gives

$$\mathbb{E}_{m_0}[Q|m_0+m] = \sum_{k=1}^{K} \gamma^k \frac{\psi_{m_0}^k + \tau m \frac{\psi_{m_0}^k}{\sum_{l=1}^K \psi_{m_0}^l}}{\sum_{k=1}^K \psi_{m_0}^k + \tau m \frac{\psi_{m_0}^k}{\sum_{l=1}^K \psi_{m_0}^l}} = \sum_{k=1}^{K} \gamma^k \frac{\psi_{m_0}^k \left(1 + \tau m \frac{1}{\sum_{k=1}^K \psi_{m_0}^k}\right)}{\sum_{k=1}^K \psi_{m_0}^k \left(1 + \tau m \frac{1}{\sum_{k=1}^K \psi_{m_0}^k}\right)} =$$

$$\sum_{k=1}^{K} \gamma^k \frac{1 + \tau m \frac{1}{\sum_{l=1}^K \psi_{m_0}^l}}{1 + \tau m \frac{1}{\sum_{l=1}^K \psi_{m_0}^l}} \frac{\psi_{m_0}^k}{\sum_{k=1}^K \psi_{m_0}^k} = \sum_{k=1}^{K} \gamma^k \frac{\psi_{m_0}^k}{\sum_{k=1}^K \psi_{m_0}^k} = \mathbb{E}[Q|m_0] = Q^{m_0}.$$

The first part of the benefit from scrolling is therefore zero

$$\mathbb{E}_{m_0}[Q|m_0+m] - Q^{m_0} = 0.$$

Marginal benefit from reading m more reviews then equals:

$$\mathbb{B}_{m+m_0,m_0} = V^{m_0} - \mathrm{Var}_{m_0}(Q|m_0+m) = -(\mathrm{Var}_{m_0}(Q|m_0+m) - V^{m_0})$$
$$-\left(\mathbb{E}_{m_0}\left[\sum_{k=1}^K \mathrm{Var}[S^k|m_0+m]\right] - \mathbb{E}_{m_0}\left[\sum_{k=1}^K \mathrm{Var}[S^k|m_0]\right]\right) =$$
$$-\left(\mathbb{E}_{m_0}\left[\sum_{k=1}^K \frac{\psi_{m_0+m}^k(\phi_{m_0+m} - \psi_{m_0+m}^k)}{\phi_{m_0+m}^2(\phi_{m_0+m}+1)}\right] - \sum_{k=1}^K \frac{\psi_{m_0}^k(\phi_{m_0} - \psi_{m_0}^k)}{\phi_{m_0}^2(\phi_{m_0}+1)}\right).$$

Applying the bounded rationality assumption

$$\boldsymbol{\psi_{m_0+m}} = \mathbb{E}[\boldsymbol{\psi_{m_0+m}}],$$

omitting the m_0 subscripts, and collecting terms we obtain

$$\mathbb{B}_{m+m_0,m_0} = \frac{1}{\left(\sum_{k=1}^K \psi^k\right)^2} \sum_{k=1}^K \left[\psi^k \left(\sum_{l=1,l\neq k}^K \psi^k\right)\left(\frac{1}{\psi^k+1} - \frac{1}{\psi^k + 1 + \frac{m\tau\psi^k}{\sum_{k=1}^K \psi^k}}\right)\right] =$$
$$\frac{1}{\phi^2}\sum_{k=1}^K \left[\psi^k \phi^- \left(\frac{1}{\psi^k+1} - \frac{1}{\psi^k + 1 + \frac{m\tau\psi^k}{\phi}}\right)\right],$$

where $\phi = \sum_{k=1}^K \psi^k$ and $\phi^- = \sum_{l=1,l\neq k}^K \psi^l$.