WHEN DO CONSUMERS TALK?

By

Ishita Chakraborty, Joyee Deb and Aniko Oery

February 2023

COWLES FOUNDATION DISCUSSION PAPER NO. 2254R3

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

When do consumers talk?

Ishita Chakraborty, Joyee Deb and Aniko Öry*

February 2023

Abstract

The propensity of consumers to talk after a good versus bad experience with a product can differ based on information available from other marketing channels, for example the brand image or advertising. This can result in selection of positive/negative word-of-mouth for reasons outside of product quality. We develop a unifying framework of WOM, brand image, product advertising, and pricing with a focus on the instrumentality motive of word-of-mouth: early adopters talk to inform new buyers' purchasing decisions. The different marketing channels shape the *information sharing* behavior of the early adopter as well as the target consumer's *purchase decision*. We show that if the brand image is strong, then in equilibrium only negative WOM can arise. In contrast, with a weak brand image, positive WOM must occur. We also show that holding product quality fixed, a positive advertising signal realization leads to a more positive WOM selection. The model can be applied to both one-one informal WOM as well as online reviews. The assumptions and main predictions of our model are consistent with those that we identified from a primary survey and observational Yelp data.

Keywords: brand image, costly communication, recommendation engines, review platforms, word of mouth

*Chakraborty: University of Wisconsin-Madison, email: ishita.chakraborty@wisc.edu; Deb: Yale School of Management, email: joyee.deb@yale.edu; Öry: Yale School of Management, email: aniko.oery@yale.edu. We would like to thank Alessandro Bonatti, Judy Chevalier, Gizem Ceylan-Hopper, Preyas Desai, Kristin Diehl, T. Tony Ke, Zikun Liu, Miklos Sarvary, Stephan Seiler, Katja Seim, Jiwoong Shin, K. Sudhir, Jinhong Xie and seminar audiences at Chicago Booth, London Business School, Marketing Science conference, MIT, Stanford GSB, SICS, Warrington College of Business, Wharton and Yale SOM for helpful suggestions.

Introduction

Many consumption decisions are influenced by reviews and word of mouth. Empirical research shows that, on average, higher reviews tend to increase sales for all types of businesses (Chevalier and Mayzlin (2006); Luca (2016); Liu, Lee and Srinivasan (2019); Reimers and Waldfogel (2021)). Firms also invest in referral rewards and other marketing tactics to encourage word of mouth. In platforms like Amazon.com or Yelp, search results are often sorted by average overall rating which puts businesses with lower ratings at a significant disadvantage. Hence, it is important for both businesses and platforms to understand how valence of user-generated ratings could differ due to reasons outside of product quality or popularity.

In this paper, we study how the selection of user-generated ratings interacts with information coming from product advertising and the brand image— an under-explored relationship between important marketing elements (Lovett, Peres and Shachar (2013)). Reduced-form evidence from Yelp.com shows the following, perhaps surprising, patterns: On a 5-star scale, the mean rating for national established chain restaurants is 2.3 stars, the modal rating being 1 star (46.9% in our data), but the mean rating is 3.8 for comparable independent restaurants, the modal ratings being 4 or 5 stars (41.2%). Furthermore, stores from national established chains with a *high* overall store rating receive fewer reviews than stores with low average ratings. In contrast, comparable independent stores receive relatively more reviews if they have a high average rating. This data suggests negative selection of reviews for established chains and positive selection of reviews for individual restaurants.[1] Two surveys that we conducted in the restaurant and Amazon review context, respectively, confirm this hypothesis: the propensity to review for strong brands is higher after a negative experience, while it is higher after a good experience for weak brands.

Such differences in the selection of reviews could be driven by differences in the incentive

[1]Reviews are well-known to be skewed (see Schoenmüller, Netzer and Stahl (forthcoming)). Chevalier and Mayzlin (2006) and Fradkin, Grewal, Holtz and Pearson (2015) document positive skews in user ratings for books and home rentals, respectively.

to review or the informational content of products with a strong brand image. Indeed, respondents of our surveys report different motives to review for products of strong versus weak brands. Among chain restaurants, the second biggest motivation to review is "venting" while in independent restaurants, there is a strong motive to "promote" a good restaurant. However, a key driver of word-of-mouth for all kinds of restaurants and experiences is the desire to improve decision-making for subsequent buyers. Can the incentive to be useful alone lead to differential selection of WOM for products by established brands versus new brands?

Motivated by these observations, we develop a unifying framework for word-of-mouth (WOM) communication, brand image, advertising, and pricing. In modeling WOM, our key premise is that engaging in WOM or writing reviews is costly, and early adopters of a product share their experience only if they can instrumentally affect purchase decisions. This assumption is in line with the findings from our survey and also consistent with research in psychology and marketing that highlights two complementary functions of WOM: First, WOM helps consumers acquire information when they are uncertain about a purchase decision. Second, people engage in WOM to enhance their self-image and/or to persuade others, which is both consistent with them only sharing information with instrumental value.[2]

We define a brand image to be the distribution of consumers' private brand associations about brand quality. The strength of the brand image is the precision of this distribution. For instance, a new restaurant of an established chain is likely to have a strong brand image, i.e., consumers agree on the quality of the brand and know what to expect from it, while a new independent restaurant likely starts with a weak brand image, i.e., consumers have very different priors about its quality. Consumers care about the underlying brand quality because it is informative about the quality of any new products that is launched under the brand name.

When a new product is launched, early adopters may generate WOM that informs new

[2]See Berger (2014) for a survey.

consumers about its quality in addition to their brand associations and public product advertising. We show that strength of the brand image determines whether WOM can be instrumental or not. Our model predicts that a product launched by a brand with a strong brand image only receives negative WOM while a product launched under a weak brand must receive positive WOM. The reason is that if the brand image is strong, the optimal pricing leads to most consumers buying as a default, but for a product with a weak brand image, there are many consumers whose default action is not to buy at the optimal price but who can be convinced to buy. Further, holding quality fixed, a positive advertising realization leads to a more positive selection of WOM.

Formally, we consider a firm with a given brand image that is optimally setting a price for a new product of uncertain quality. A representative target customer makes a purchase decision based on his private brand association, public information about the new product, e.g., via advertising, and WOM. We are interested in a setting where only a few consumers have tried the product. Hence, from a new consumers perspective, a representative potential early adopter has experienced the product only with a small probability. Experience provides a private noisy signal of quality that can be shared. An early adopter with product experience can thereby influence a target customer's purchase decision. If the early adopter decides not to talk or has not experienced the product yet, the target customer receives no message. We characterize perfect Bayesian equilibria of this game.

In order to know if her information has instrumental value, the early adopter needs to take into account the target customer's purchase decision in the absence of WOM. If the early adopter thinks that a representative target consumer is likely to buy given the public advertising message, brand image and the price, then there is no reason for her to engage in positive WOM after a good experience, but she may affect the target consumer's action through negative WOM after a bad experience. Conversely, if a target consumer is likely to not buy in the absence of WOM information, an early adopter wants to share only a positive experience, since sharing a negative experience has no incremental instrumental value. The firm sets its price accordingly to maximize revenues from the target customer. In practice, the price may be a combination of the posted price, promotions, extra benefits, etc. We show that a well-entrenched brand optimally sets a price that induces buying for most customers while a weak brand cannot make sure that all customers buy absent WOM. Consequently, for a well-entrenched brand the unique equilibrium results in only negative WOM, i.e., experiences above a threshold are not being shared. For a sufficiently weak brand the unique equilibrium results in positive WOM, i.e., experiences below a threshold are being shared.

Finally, we show that more positive advertising outcomes result in higher ratings on average. The reason is that in a negative-WOM equilibrium, the threshold experience above which consumers do not talk increases in the advertising outcome, and that it is more likely that the equilibrium regime shifts to one with positive WOM. Intuitively, if there is a lot of positive public information about a product, a weakly negative signal can become instrumental for consumers with a weak brand association.

Literature Review

Our paper is substantively related to research on diffusion of information through WOM, pioneered by Bass (1969). WOM can occur via platforms, social networks or traditional networks. Most early papers in this area treat WOM as a costless mechanical process, and focus on how the social network structure affects information percolation about the existence of a product: See for instance Galeotti (2010) or Galeotti and Goyal (2009).[3] We contribute to the more recent literature that considers the strategic motive of consumers to engage in costly WOM and the literature that links branding to WOM.

[3]Similarly, Leduc, Jackson and Johari (2017) study the diffusion of a new product when consumers learn about the quality in a network and the firm can affect the diffusion through pricing and referral incentives. Campbell (2013) instead analyzes the interaction of advertising and pricing. See also Godes, Mayzlin, Chen, Das, Dellarocas, Pfeiffer, Libai, Sen, Shi and Verlegh (2005) for a survey of the literature.

Strategic Incentives for WOM

One of the first papers to study the strategic incentive to engage in WOM is by Campbell, Mayzlin and Shin (2017) who focus on how the firm should balance WOM and advertising if consumers' incentive to talk stems from a desire to signal social status. They find that advertising crowds out consumers' incentives to engage in WOM. Joshi and Musalem (2021) also study the relationship between advertising and WOM and develop insights on how the firm should adjust advertising spendings based on the valence and volume of WOM interactions. Other authors focus on WOM and referral programs. In Biyalogorsky, Gerstner and Libai (2001) a firm can encourage WOM through the price or a referral program. Unlike in our model, a reduced price induces senders to talk because it "delights" them. Kornish and Li (2010) also consider the trade-off between referral rewards and pricing in a model where the sender cares about the receiver's surplus. Kamada and Öry (2017) consider a contracting problem in which the incentive to talk is driven by externalities of using a product together. They show that offering a free contract can make WOM more attractive since receivers are more likely to start using the product. We consider WOM not about the existence of a product, but about the experience, and we highlight how the brand image affects an instrumental incentive to talk in that case. [4] Mayzlin (2006) considers a firm's strategic incentive to engage in promotional chat alongside an early adopter.

There is a growing empirical literature that studies the impact of review statistics, like volume, valence (positive or negative) and variance, on business outcomes (e.g., sales).[5] Luca (2016) finds that a one-star increase in Yelp ratings can increase revenue by 5-9 percent. Chintagunta, Gopinath and Venkataraman (2010) show that an improvement in reviews

[4]The incentive to talk in our paper is similar to the incentive to search in Mayzlin and Shin (2011): The marginal value of information must be larger than the marginal cost of information dissemination or acquisition, respectively. In empirical work, Toubia and Stephen (2013) have discussed a similar incentive (image-related) which promotes people to share content on a platform like Twitter.

[5]For example, Nosko and Tadelis (2015), Dhar and Chang (2009) and Duan, Gu and Whinston (2008) show that the volume of reviews matter (rather than the rating), and Sun (2012) show that high variance in reviews corresponds to niche products, valued highly by some buyers but not by others. Onishi and Manchanda (2012) show a positive impact of blogging on sales.

leads to an increase in sales for movies and Seiler, Yao and Wang (2017) document that micro blogging has an impact on TV viewership. More specifically, the asymmetric impact of valence on profit-relevant outcome variables has been studied in some empirical contexts. Mittal, Ross Jr and Baldasare (1998) find that negative information has larger impact on consumer purchase decisions compared to positive information. Chevalier and Mayzlin (2006) find that negative reviews have a larger effect on sales than positive reviews.[6] Chen and Xie (2008) additionally discuss under what conditions firm advertising and WOM are complements or substitutes. Their findings suggest that allowing users to post reviews is detrimental to firm's profit if reviews are not sufficiently informative and also it may not be a good strategy to allow consumer reviews at a very early stage of product introduction.

To the best of our knowledge, our paper is the first to provide an information-theoretical foundation for what determines valence of WOM and user-generated reviews. We highlight how asymmetry in the propensity to engage in WOM can be driven by the dispersion of consumer beliefs about quality and the firm's pricing decision. The only other paper that studies different propensities to review after positive versus negative experiences is by Angelis, Bonezzi, Peluso, Rucker and Costabile (2012), who argue using experimental evidence that consumers with a strong self-enhancement motive generate a lot of positive WOM and transmit more negative WOM about other people's experiences: Differences in valence simply arise from differences in the type of people who choose to be early adopters. Chakraborty, Kim and Sudhir (2019) also study selection in reviews using text analysis, but their focus is primarily on what drives content selection among different types of reviewers.

Branding and WOM

There are several empirical papers that have documented the impact of WOM like reviews on brand image and brand equity (Sundaram and Webster, 1999; Luo, 2009; Bruhn, Schoenmueller and Schäfer, 2012; Hollenbeck, 2018; Baker, Donthu and Kumar, 2016; Luca, 2016).

[6] Also, Godes (2016) studies how the type of WOM affects the incentives of firms to invest in product quality. Mayzlin, Dover and Chevalier (2014) empirically demonstrate how some branded hotels may be receiving more negative fake reviews, however, our work abstracts from this incentive.

Specifically, Luo (2009) finds that negative WOM has a medium-term and long-term effect on brand equity. In a study of hotel brands, Hollenbeck (2018) shows that the volume of reviews is not correlated with sales for chain restaurants, but increases sales for small brands. He argues that with the rise of review platforms small brands can compete more equally with larger brands. Note that chain hotels systematically solicit WOM reviews from regular repeat customers, which may effectively alleviate potential negative selection for chain hotels predicted by our analysis.[7] Lovett et al. (2013) note that though brands and WOM are cornerstones of marketing, the interconnection between the two is surprisingly understudied. In their empirical study on the impact of certain brand characteristics on WOM, they find that differentiation, complexity and excitement impact the volume of online and offline WOM. Our paper complements this literature by developing a game-theoretical framework to explain how brand image strength (defined as the precision of distribution of brand associations) can impact the valence of the WOM generated due to WOM selection, and offers some suggestive evidence for the model predictions from Yelp.com.

Model

We consider a firm (it) that is launching a new product and trying to sell to a representative target customer (he). The new product can be under an existing brand name—e.g., the retail chain Starbucks is opening a new store location—or under a completely new brand name. We normalize the marginal cost of production to zero.

New product quality and brand quality. A brand is defined by an underlying *brand quality* $b \in \mathbb{R}$ that governs the quality of all products sold under that brand name. The value of b is unknown to consumers and the firm. Consumers and the firm share a common prior

[7] Note that we are also not considering the fact that the number of consumers is likely higher for good products than for bad products. This effect can eliminate any correlation of review ratings and quality of products for chain restaurants which is consistent with the finding in Luca (2016) that sales are not affected by reviews for chain restaurants.

belief that $b \sim \mathcal{N}(b_0, \sigma_0^2)$. The quality of a new product is given by

$$\theta = b + \eta,$$

where $\eta \sim \mathcal{N}(0, \sigma_\theta)$ is a product-specific quality component that is independent of b. Since we assume that each firm is associated with a single brand, we use brand and firm interchangeably.

As an example, think of b as the underlying quality of the Starbucks brand. The quality of a specific new Starbucks location j, denoted by $\theta_j := b + \eta_j$ then depends on b and the idiosyncratic quality of the specific location denoted by η_j.[8] For a new independent coffee shop, b would be the quality component of the new store that can be conveyed through branding. For example, a new independent coffee shop that positions itself as a specialized German bakery and cafe can send this message through its branding and marketing activity, prior to opening, which conveys a public signal about its quality. We use these examples throughout this section to explain our model.

This notion of brand quality is related but conceptually different from the notion of brand positioning in Ke, Shin and Yu (2020). In their model, the brand positioning is the average position of all its products. In our model, b is not equal to the mean of all individual product/store location qualities θ_j. Instead, b is an underlying parameter that affects all θ_j equally in a stochastic sense. In other words, the mean quality of all individual products is an unbiased estimate of b. b_0 represents the mean brand quality and σ_0^2 represents the public uncertainty about brand quality. For Starbucks, σ_0^2 is likely small. For a new specialized Germany bakery and cafe σ_0^2 is large. Next, we outline how consumers form idiosyncratic posterior beliefs.

Brand associations, brand image, product advertising and word-of-mouth. The representative target customer learns about the product quality θ from several sources before

[8]Note that we will not have a subscript j in our model as we are interested in the reviewing of a single representative new product.

making a purchase decision.

1. **Private brand association.** A target customer observes a private signal $b_i = b + \epsilon_i$ where $\epsilon_i \sim \mathcal{N}(0, \sigma_1^2)$. Specifically, the target customer's posterior belief about b given his private signal b_i is given by $b|b_i \sim \mathcal{N}(\mu_b(b_i), \sigma_{b|b_i}^2)$ where

$$\mu_b(b_i) := \frac{\sigma_1^2}{\sigma_0^2 + \sigma_1^2} b_0 + \frac{\sigma_0^2}{\sigma_0^2 + \sigma_1^2} b_i, \qquad \sigma_{b|b_i}^2 := \frac{\sigma_0^2 \sigma_1^2}{\sigma_0^2 + \sigma_1^2}.$$

We call $\mu_b(b_i)$ the target customer's *brand association*. The precision $\frac{1}{\sigma_{b|b_i}^2}$ captures how confident the target customer is about the brand quality, given his private brand association.

We can think of b_i as reflecting a customer's idiosyncratic information based on his own past experiences. In the Starbucks example, b_i is based on a customer's experiences with different Starbucks locations in the past, and $\mu_b(b_i)$ is the customer's resulting brand association about the Starbucks brand.[9] In our example of a completely new independent coffee shop that is positioned as a German bakery-cafe, b_i can be based on a customer's past experiences with coffee shops with similar branding: Some consumers may associate it with high quality due to past positive experiences with European coffee houses and German bread, while others might expect low quality coffee due to an earlier bad experience with German coffee.

The distribution of private brand associations $\mu_b(b_i)$ in the population of target customers constitutes the *brand image* of the firm. Formally, the brand image is given by a distribution $\mathcal{N}(b_0, \sigma_b^2)$, where $\sigma_b^2 := \frac{\sigma_0^4}{\sigma_0^2+\sigma_1^2}$. $\frac{1}{\sigma_b^2}$ is the *strength of the brand image*. This is a common definition of brand image in the marketing literature—brand image is usually considered to be the combined effect of brand associations (e.g., Newman, 1957; Biel, 1992; Lee, James and Kim, 2014).[10] Note that consumers have similar

[9]Brand association can be defined as an individual consumer's contact with a brand (Aaker 1991).

[10]Kotler (2000) also argues that consumer beliefs form brand images which in turn influence consumer purchase decisions writes: "A belief is a descriptive thought that a person holds about something. Beliefs

brand associations when the public uncertainty about the brand σ_0 is relatively low. This is likely the case for established brand names, e.g., Starbucks. For an independent Germany bakery-cafe, the high public uncertainty and variance in experiences give rise to a weak brand image.

2. **Product advertising.** The target customer also has access to a public signal $a = \eta + \epsilon_a$ of product specific quality η, where $\epsilon_a \sim \mathcal{N}(0, \sigma_a^2)$ is independent of ϵ_i and b. We can interpret a as advertising or information from past reviews about the new Starbucks location.[11] The precision $\frac{1}{\sigma_a^2}$ is the informativeness of product advertising.

3. **Word-of-mouth communication.** Finally, the target customer can acquire information about the new product via word-of-mouth communication from an early adopter (she).[12] An early adopter is someone who had the opportunity to experience the new product recently, and therefore observed an independent quality signal about the new product $s = \eta + \epsilon_s$ where $\epsilon_s \sim \mathcal{N}(0, \sigma_s^2)$ is independent of ϵ_a, ϵ_i and b. Henceforth, we refer to this signal as the early adopter's *experience*. The early adopter can choose to share her signal through a WOM message m. WOM can be shared on a review platform, a social network platform, or simply verbally if the early adopter is a friend of the target customer. Our focus is on one such representative WOM channel.

   The target customer believes that with probability $\beta \in (0, 1)$ a user of a platform or a friend whom he has recently met is an early adopter. In the Starbucks example, β can be interpreted as the probability that there is a *potential* Yelp.com reviewer who has recently experienced the specific new location.[13] In the theoretical analysis we focus

may be based on knowledge, opinion, or faith (...) manufacturers are very interested in the beliefs that people have about their products and services. These beliefs make up product and brand images, and people act on their images." The brand strength then represents how consistent consumers' beliefs about the product are.

[11]Note that a only informs about the idiosyncratic product quality component η. If it was a signal about the full product quality $\theta = b + \eta$, our qualitative results remain unchanged but would unnecessarily complicate the analysis.

[12]https://www.brightlocal.com/research/local-consumer-review-survey/ finds that consumers care most about recent reviews

[13]Note that not every Starbucks customer is a potential Yelp.com reviewer, but only someone who is active

on small β, i.e., the probability of meeting an early adopter is small in the relevant period. For example, on Yelp.com, the mean number of reviews per day across a wide range of restaurants is 0.12.

Communication is verifiable, so an early adopter either sends a message $m = s$ (where s is her experience) or $m = \emptyset$ which represents no WOM.[14] Note that a target customer cannot distinguish between the absence of an early adopter and the decision of an early adopter to not engage in WOM, and β determines the informativeness of $m = \emptyset$. Small β means that the lack of WOM ($m = \emptyset$) is more likely due to the absence of an early adopter than due to the decision of an early adopter to not engage in WOM.

Engaging in WOM ($m = s$) costs the early adopter $c > 0$, but if the message is instrumental, she experiences a positive utility $r > 0$. Thus, she receives r if and only if

(a) she sends a message $m = s$ and the target customer buys, but would not have bought with $m = \emptyset$,

(b) or if she sends a message $m = s$ and the target customer does not buy, but would have bought with $m = \emptyset$.

Let $\xi := \frac{c}{r}$. We assume $1 - \beta > \xi$, to rule out the trivial case of early adopters never engaging in WOM because they are too unlikely to face a target customer.

The instrumental motive of WOM can be rationalized by the self-enhancement and persuasion motives as described in Berger (2014). He argues that when people care about impression management, they are "more likely to share things that make them look good rather than bad." So, the early adopter would like to provide useful information that makes her look good because it is instrumental for the target consumer's decision making. Then, r can be interpreted as the early adopter's utility of an enhanced

on the platform and sometimes writes reviews.

[14]We do not consider review manipulation as in Mayzlin et al. (2014), Luca and Zervas (2016), and He, Hollenbeck and Proserpio (2020).

self-image from providing information of instrumental value.[15] Because messages are verifiable, the utility specification above reflects also the persuasive motive, where a sender engages in word-of-mouth to influence others and change their action.[16] We also provide some evidence in Section that is consistent with an instrumental motive of reviewers.[17]

Table 1 summarizes the relevant notation on brand association, brand image, product advertising and WOM.

Table 1: Summary of Key Notation

Notation	Meaning	Notation	Meaning
$b \sim \mathcal{N}(b_0, \sigma_0^2)$	brand quality	$\mu_b(b_i) = \mathbb{E}[b \mid b_i]$	brand association
$\theta = b + \eta, \eta \sim \mathcal{N}(0, \sigma_\theta^2)$	product quality	$\mu_b(b_i) \sim \mathcal{N}\left(b_0, \sigma_b^2\right)$	brand image
$a = \eta + \epsilon_a, \epsilon_a \sim \mathcal{N}(0, \sigma_a^2)$	product advertising	$\frac{1}{\sigma_b^2}$, $\sigma_b^2 = \frac{\sigma_0^4}{\sigma_0^2 + \sigma_1^2}$	strength of brand image
$b_i = b + \epsilon_i, \epsilon_i \sim \mathcal{N}(0, \sigma_1^2)$	private brand signal	$m \in \{\emptyset, s\}$	WOM message
$s = \eta + \epsilon_s, \epsilon_s \sim \mathcal{N}(0, \sigma_s^2)$	early adopter's experience		

Timing. The game proceeds as follows:

1. The representative target customer observes his private signal b_i, shaping his brand association $\mu_b(b_i)$.

2. The brand launches a new product, all players see public information about the new product a, and the the brand publicly chooses a price p.

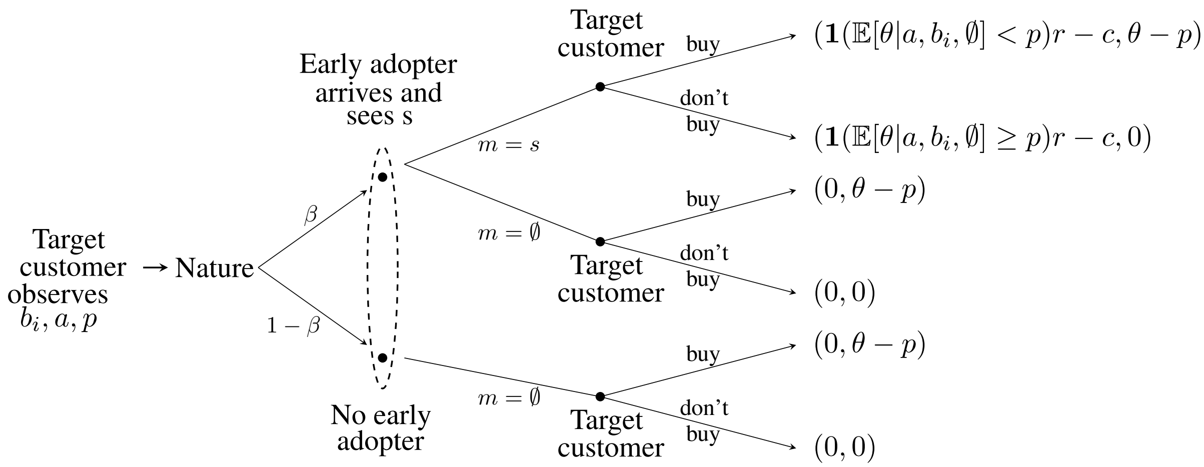
3. With probability β an early adopter has tried the product. If present, the early adopter decides whether to engage in WOM ($m = s$) or not ($m = \emptyset$). If no early adopter is present, a message $m = \emptyset$ is sent.

[15]Restaurant reviewers on Yelp.com cite simplified decision-making for first-time visitors as one of the reasons for writing a review. See Carman (2018).

[16]This is also consistent with the Gricean maxims proposed in Grice, Cole, Morgan et al. (1975) that when engaging in a conversation, people should make it relevant to the audience and provide enough information, but not more than required. We thank Kristin Diehl and Gizem Ceylan-Hopper for pointing us to this reference.

[17]There are many concurrent motives to engage in WOM, such as social bonding or simply emotional regulation. We abstract from them in this study, but argue that the instrumental motive can be one important reason for differential selection of WOM.

Figure 1: Game tree of WOM subgame given b_i, a and p



Notes: The payoffs in brackets at the end of the game tree are the early adopter's and the target customer's payoffs.

4. The representative target customer updates her belief about product quality θ based on b_i, a, and m, and decides whether to buy or not.

Figure 1 summarizes the WOM subgame after the target customer has formed $\mu_b(b_i)$, and has observed the price p and public signal a.

Strategies and equilibrium. The firm's strategy simply comprises a price p for any public advertising signal a. The early adopter's WOM strategy maps the price p, advertising signal a and experience s to a probability with which an early adopter engages in WOM. The target customer's purchasing strategy maps the price p, public information a, private brand association b_i and the WOM message m to a probability to purchase.

We consider *perfect Bayesian equilibria (PBE)*. A PBE comprises a tuple of strategies of all players and the target customer's posterior belief $\hat{\theta}(a, b_i, m)$ given a, b_i and m such that all players play mutual best-responses. For a given PBE, we denote by $\mathcal{S}_{a,p}$ the set of experiences after which the early adopter talks, given an advertising message a, equilibrium price p and equilibrium strategy of the early adopter. The complement $\mathcal{S}^c_{a,p} := \mathbb{R} \setminus \mathcal{S}_{a,p}$ are the experiences after which the early adopter sends $m = \emptyset$.

WOM Subgame

To characterize equilibria, we proceed by backwards induction. Consider the subgame that starts after the private brand association is formed, the public signal a is observed, and the price is set. We call this the "WOM subgame" and its equilibria "WOM equilibria." Figure 1 depicts this subgame. In the WOM subgame, we first characterize the target customer's purchase decision and then step backwards to the early adopter's WOM decision. Finally, in Section , we solve for the firm's optimal price-setting decision, given the behavior in the WOM subgame. All proofs are in the Appendix.

Purchase decision of a target customer

Posterior belief about θ. Given a public advertising signal a, it is optimal for a target customer with brand association b_i and WOM message m to purchase if and only if his expected utility from purchasing $\hat{\theta}(a, b_i, m) := \mathbb{E}[\theta | a, b_i, m]$ exceeds the price: $\hat{\theta}(a, b_i, m) - p \geq 0$. Prior to receiving any WOM message, the target customer forms a posterior belief about the product quality $\theta | a, b_i \sim \mathcal{N}(\mu_\theta(a, b_i), \sigma^2_{b|b_i} + \frac{\sigma_a^2 \sigma_\theta^2}{\sigma_a^2 + \sigma_\theta^2})$ based on a and $\mu_b(b_i)$, where

$$\mu_\theta(a, b_i) := \mathbb{E}[\theta | a, b_i] = \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_{\mathrm{a}}^2} \underbrace{a}_{\text{advertising}} + \underbrace{\mu_b(b_i)}_{\substack{\text{brand} \\ \text{association}}} .$$

The value of $\mu_\theta(a, b_i)$ is the sum of the two pieces of information a and $\mu_b(b_i)$, where the product specific signal a is weighted by the variance of η relative to the variance of the public signal. Finally, after receiving the WOM message $m = s \in \mathbb{R}$, the target customer's posterior belief about product quality is given by

$$\begin{aligned} \hat{\theta}(a, b_i, s) \quad &:= \quad \mathbb{E}[\theta | a, b_i, s] = \frac{\sigma_s^2 \sigma_\theta^2}{\Sigma} \underbrace{a}_{\text{advertising}} + \underbrace{\mu_b(b_i)}_{\substack{\text{brand} \\ \text{association}}} + \frac{\sigma_a^2 \sigma_\theta^2}{\Sigma} \underbrace{s}_{\text{WOM}} \\ &= \quad \mu_\theta(a, b_i) - \frac{\sigma_a^2 \sigma_\theta^2}{\Sigma} \left(\frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a - s \right). \end{aligned}$$

where $\Sigma := \sigma_s^2(\sigma_\theta^2+\sigma_a^2)+\sigma_a^2\sigma_\theta^2$. This is again a linear combination of a, $\mu_b(b_i)$ and s weighted by relative variances. After receiving no WOM message ($m = \emptyset$), a target customer's equilibrium posterior belief about quality is given by

$$
\begin{aligned}
\hat{\theta}(a, b_i, \emptyset) :=& \beta\, \mathbb{E}[\theta \big| a, b_i, s \in \mathcal{S}_{a,p}^c] + (1-\beta)\, \mu_\theta(a, b_i) \\
=& \mu_\theta(a, b_i) + \beta\, \frac{\sigma_a^2\sigma_\theta^2}{\Sigma}\, \frac{\int_{\mathcal{S}_{a,p}^c} s \frac{1}{\sqrt{2\pi\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}} \exp\left(-\frac{\left(\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2}a-s\right)^2}{\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}\right) ds}{\int_{\mathcal{S}_{a,p}^c} \frac{1}{\sqrt{2\pi\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}} \exp\left(-\frac{\left(\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2}a-s\right)^2}{2\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}\right) ds}.
\end{aligned}
$$

Cutoff strategy with respect to $\mu_\theta(a, b_i)$. Given an experience s, a consumer buys if and only if $\hat{\theta}(a, b_i, s) \geq p$ which is equivalent to

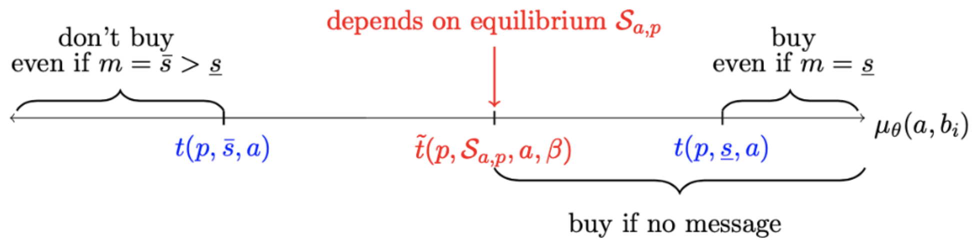
$$
\underbrace{\mu_\theta(a, b_i)}_{\substack{\text{target customer's} \\ \text{private info}}} \quad \geq \quad \underbrace{p + \frac{\sigma_a^2\sigma_\theta^2}{\Sigma}\left(\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2}a - s\right) =: t(p, s, a).}_{\text{early adopter's private info until she talks}}
$$

Note that the realized threshold $t(p, s, a)$ is not observed by the firm or the target customer prior to WOM as it depends on the early adopter's private experience s. Similarly, the realized $\mu_\theta(a, b_i)$ is private information to the target customer. Using the expression for $\hat{\theta}(a, b_i, \emptyset)$ above, we know that after receiving non WOM message ($m = \emptyset$), a target customer buys if and only if

$$
\mu_\theta(a, b_i) \geq p + \beta\, \frac{\sigma_a^2\sigma_\theta^2}{\Sigma}\, \frac{\int_{\mathcal{S}_{a,p}^c} s \frac{1}{\sqrt{2\pi\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}} \exp\left(-\frac{\left(\frac{\sigma_\theta^2}{2\sigma_\theta^2+\sigma_a^2}a-s\right)^2}{\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}\right) ds}{\int_{\mathcal{S}_{a,p}^c} \frac{1}{\sqrt{2\pi\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}} \exp\left(-\frac{\left(\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2}a-s\right)^2}{2\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}\right) ds} =: \tilde{t}(p, \mathcal{S}_{a,p}, a, \beta).
$$

Figure 2 illustrates the role of the thresholds $t(p,s,a)$ and $\tilde{t}(p,\mathcal{S}_{a,p},a,\beta)$ for two different experiences $\underline{s} < \bar{s}$. The horizontal line represents different values of the mean interim belief $\mu_\theta(a,b_i)$.

Figure 2: Target consumer's decision



WOM decision of an early adopter

Given the target customer's optimal buying strategy, we can infer the early adopter's communication decision. Since $\mu_b(b_i) = \frac{\sigma_1^2}{\sigma_0^2+\sigma_1^2}\ b_0 + \frac{\sigma_0^2}{\sigma_0^2+\sigma_1^2}\ b_i \sim \mathcal{N}\left(b_0, \sigma_b^2\right)$, the early adopter believes that

$$\mu_\theta(a,b_i) \sim \mathcal{N}\left(\bar{\mu}(a,b_0), \sigma_b^2\right), \text{ where } \bar{\mu}(a,b_0) := b_0 + \tfrac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2}\ a$$

given the observed public advertising signal realization a. We denote the cumulative distribution function (cdf) of $\mu_\theta(a,b_i)$ by Φ_{μ_θ}. Note that if $\sigma_b^2 = 0$, then $\mu_b(b_i) \equiv b_0$ and $\mu_\theta(a,b_i) = \bar{\mu}(a,b_0)$. We discuss this limiting benchmark case of well-entrenched brands in Section . Here, we assume that $\mu_b(b_i)$ admits a continuous cdf, i.e., $\sigma_b^2 > 0$.

Then, an early adopter with experience s such that $\tilde{t}(p,\mathcal{S}_{a,p},a,\beta) \geq t(p,s,a)$ weakly prefers to engage in WOM if and only if

$$\underbrace{r\ \left(\Phi_{\mu_\theta}(\tilde{t}(p,\mathcal{S}_{a,p},a,\beta)) - \Phi_{\mu_\theta}(t(p,s,a))\right)}_{\text{benefit of talking}} \geq \underbrace{\quad c \quad}_{\text{cost of talking}} . \tag{1}$$

Whenever an early adopter with experience s such that $\tilde{t}(p,\mathcal{S}_{a,p},a,\beta) \geq t(p,s,a)$ wants to talk, then an early adopter with experience $s' > s$ also wants to talk, since $t(p, s, a)$ is decreasing in s. Similarly, if s is such that $\tilde{t}(p, \mathcal{S}_{a,p}, a, \beta) \leq t(p, s, a)$, an early adopter weakly prefers to engage in WOM if and only if

$$\underbrace{r\left(\Phi_{\mu_\theta}(t(p, s, a)) - \Phi_{\mu_\theta}(\tilde{t}(p, \mathcal{S}_{a,p}, a, \beta))\right)}_{\text{benefit of talking}} \geq \underbrace{c}_{\text{cost of talking}}. \tag{2}$$

Hence, if an early adopter with experience s such that $\tilde{t}(p, \mathcal{S}_{a,p}, a, \beta) \leq t(p, s, a)$ wants to talk, then an early adopter with $s' < s$ wants to talk as well. Using this insight we can derive the following lemma:

Lemma 1 *Suppose $\sigma_b^2 > 0$. In any equilibrium, $\mathcal{S}_{a,p}$ takes one the following forms:*

1. No-WOM equilibrium*: $\mathcal{S}_{a,p} = \emptyset$.*

2. Full-WOM equilibrium*: $\mathcal{S}_{a,p} = (-\infty, \underline{s}] \cup [\bar{s}, \infty)$ with cutoffs $\underline{s}$, $\bar{s}$, such that $t(p, \bar{s}, a) \leq \tilde{t}(p, \mathcal{S}_{a,p}, a, \beta) \leq t(p, \underline{s}, a)$ and*

$$\Phi_{\mu_\theta}(\tilde{t}(p, \mathcal{S}_{a,p}, a, \beta)) - \Phi_{\mu_\theta}(t(p, \bar{s}, a)) = \xi \tag{3}$$

$$\Phi_{\mu_\theta}(t(p, \underline{s}, a)) - \Phi_{\mu_\theta}(\tilde{t}(p, \mathcal{S}_{a,p}, a, \beta)) = \xi. \tag{4}$$

   Thus, $p \in (t(p, \bar{s}, a), t(p, \underline{s}, a))$.

3. Positive-WOM equilibrium*: $\mathcal{S}_{a,p} = [\bar{s}, \infty)$ with cutoff $\bar{s}$ such that $t(p, \bar{s}, a) \leq \tilde{t}(p, \mathcal{S}_{a,p}, a, \beta)$ and* (3)*.*

4. Negative-WOM equilibrium*: $\mathcal{S}_{a,p} = (-\infty, \underline{s}]$ with cutoff $\underline{s}$ such that $t(p, \underline{s}, a) \geq \tilde{t}(p, \mathcal{S}_{a,p}, a, \beta)$ and* (4)*.*

The absence of WOM ($m = \emptyset$) can be interpreted as "good news" in a negative-WOM equilibrium, and "bad news" in a positive-WOM equilibrium. The probability that an early adopter is present β, determines the informativeness of $m = \emptyset$. As mentioned earlier, intuitively, when β is small, the absence of WOM is more likely to be driven by the absence of an early adopter altogether, rather than the decision of an early adopter not to share her experience. This is the case we focus on in our theoretical analysis.

Main Results

Given an advertising outcome a and price p, Lemma 1 characterizes the continuation equilibrium of the WOM subgame. In this section we determine what type of WOM equilibrium is observed if the firm chooses a profit-maximizing price. We start with a benchmark case when the strength of the brand image is infinite, i.e., $\frac{1}{\sigma_b^2} = \infty$. Intuitively, this is a setting in which there is no uncertainty about brand quality, i.e., the brand association is equal to b_0. We call such a brand a well-entrenched brand. We show that well-entrenched brands can only generate negative WOM equilibria for small β. Then, we show that in general, the type of WOM depends crucially on the strength of the brand image. In particular, we show that for weak brands, a positive-WOM equilibrium is the unique outcome for small β.

Negative WOM for well-entrenched brands

For a well-entrenched brand with $\sigma_b^2 = 0$, the brand image is a point distribution given by $\mu_b(b_i) \equiv b_0$. Then, conditional on the public information, the interim belief is given by

$$\mu_\theta(a, b_i) \equiv \bar{\mu}(a, b_0) = b_0 + \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2}\, a.$$

All target customers have the same interim belief $\theta | a, b_0 \sim \mathcal{N}\left(\bar{\mu}(a, b_0), \frac{\sigma_\theta^2 \sigma_a^2}{\sigma_\theta^2 + \sigma_a^2}\right)$. Finally, an experience s leads to a posterior belief about quality given by

$$\hat{\theta}(a, b_0, s) = \frac{\sigma_s^2 \sigma_\theta^2}{\Sigma}\, a + b_0 + \frac{\sigma_a^2 \sigma_\theta^2}{\Sigma}\, s.$$

Now, after message $m = \emptyset$, the target customer (weakly) prefers to buy if and only if $\tilde{t}(p, \mathcal{S}_{a,p}, \beta) \leq \bar{\mu}(a, b_0)$ and he (weakly) prefers to buy after a message $m = s$ if an only if $t(p, s) \leq \bar{\mu}(a, b_0)$. Monotonicity of $t(p, s)$ implies an analogous statement to Lemma 1, namely that $\mathcal{S}_{a,p}$ has to be either equal to $\emptyset$, or of the form $(-\infty, \underline{s}] \cup [\bar{s}, \infty)$ (full WOM), or $[\bar{s}, \infty)$

(positive WOM), or $(-\infty, \underline{s}]$ (negative WOM). Using these ingredients we can characterize WOM equilibria and derive our first main result: A well-entrenched brand generates only negative WOM for small β.

Proposition 1 *(Negative WOM for well-entrenched brands) Let $\sigma_b^2 = 0$ and $\xi < 1$.*

1. *There exists a $\bar{\beta} \in (0, 1)$ such that for $\beta < \bar{\beta}$, the unique equilibrium is a negative-WOM equilibrium for any realization of public signal a.*
2. *In the negative-WOM equilibrium, the cutoff $\underline{s}^*(a, \beta)$ is increasing in a and decreasing in β.*

The proof of the result is in the appendix. Intuitively, what is central is the way "no WOM" is interpreted in equilibrium. If the target customer expects only negative experiences to be shared, then no WOM becomes a positive signal. With few early adopters, no WOM is observed with high probability, and so an equilibrium with only negative WOM becomes optimal for the firm. If the probability β is smaller, $\underline{s}^*(a, \beta)$ is larger, which means early adopters engage in negative WOM after less extreme negative experiences.

The cutoff $\underline{s}^*(a, \beta)$ is also increasing in a, i.e, early adopters are willing to engage in negative WOM for less extreme experiences if a is high. Thus, higher a increases the average rating. Intuitively, if the advertising message is very positive, the firm charges a higher price and even less extreme bad experiences have incremental instrumental value.

Strong versus weak brands

Next, we consider brands with $\sigma_b^2 > 0$, and analyze how the brand strength $\frac{1}{\sigma_b^2}$ affects equilibrium WOM and pricing. We continue to focus on small β. It is useful to define the following notions that capture how the distribution of target consumers' beliefs about product quality prior to any WOM relates to the price.

Definition 1 *Given a price p and public signal a, we call the target customer*

- pessimistic *if* $\Phi_{\mu_\theta}(p) > \xi > 1 - \Phi_{\mu_\theta}(p)$*;*

- optimistic *if* $1 - \Phi_{\mu_\theta}(p) > \xi > \Phi_{\mu_\theta}(p)$*;*

- uninformed *if* $1 - \Phi_{\mu_\theta}(p), \Phi_{\mu_\theta}(p) > \xi$*;*

- well-informed *if* $1 - \Phi_{\mu_\theta}(p), \Phi_{\mu_\theta}(p) < \xi$*.*

Intuitively, the target customer is said to be pessimistic if in the absence of WOM, he buys with a small probability. Analogously, he is optimistic if in the absence of WOM, he buys with a high probability. A target customer is well-informed (uninformed) if WOM changes the purchase decision with a small (high) probability. Lemma 2 shows that this determines the type of WOM that arises.

Lemma 2 *Given a price p and a public signal a,*

1. *if target consumers are pessimistic, a* positive-WOM equilibrium *exists;*

2. *if target consumers are optimistic, a* negative-WOM equilibrium *exists;*

3. *if target consumers are uninformed, a* full WOM equilibrium *exists;*

4. *if target consumers are well-informed, a* no WOM equilibrium *exists.*

For sufficiently small β, these are the unique WOM equilibria.

To finally characterize what WOM will be observed in equilibrium, we need to derive the firm's profit maximizing price, and thus infer what WOM is induced by that price. To this end, note that, for any given set of parameters, the firm's revenue in the absence of WOM is given by $(1 - \Phi_{\mu_b(b_i)}(p))p$, which attains a unique maximum. This is because, $\mu_b(b_i)$ is normally distributed, and $\frac{1-\Phi_{\mu_b(b_i)}(p)}{\phi_{\mu_b(b_i)}(p)} - p$ is strictly decreasing. A unique profit maximizing price for sufficiently small β allows us to infer whether consumers are pessimistic, optimistic, uniformed or well-informed, and then we use Lemma 2 above to determine the type of WOM.

In the proposition below, we present our main result. First, we show that Proposition 1 about well-entrenched brands extends to brands with sufficiently strong brand image. In contrast, with sufficiently weak brand image, early adopters always share positive experiences whenever there is WOM. If the cost of talking ξ is low, both positive and negative experiences are shared, but the early adopter remains silent after intermediate experiences. If the cost of talking is high, only positive experiences are shared. The proposition also shows how advertising can affect what type of WOM arises: In particular, positive (negative) public information via advertising can leads to positive (negative) WOM.

Proposition 2 (Strong and weak brands) *Fix $\xi < 1$. Then, there exists a $\bar{\beta}$ so that for $\beta < \bar{\beta}$, the following hold:*

1. *If the brand image is sufficiently weak ($\frac{1}{\sigma_b}$ small), any equilibrium entails only positive WOM: For small ξ, the unique equilibrium entails full WOM, for large ξ it entails positive WOM.*

2. *If the brand image is sufficiently strong ($\frac{1}{\sigma_b}$ large), the unique equilibrium entails only negative WOM.*

3. *If the public signal a is high, the unique equilibrium entails only positive WOM. If a is sufficiently negative, then the unique equilibrium is a negative-WOM equilibrium. For intermediate values of a, either a full or no WOM equilibrium arises depending on whether ξ is small or large, respectively.*

To see why positive WOM arises if the brand image is sufficiently weak, first note that a weak brand image leads to more dispersed demand, making it impossible for a firm to price so that everyone buys absent WOM (as also discussed in Johnson and Myatt (2006)). But if target customers will not buy in the absence of WOM, then positive WOM has instrumental value, and therefore arises in equilibrium. Conversely, if the brand image is sufficiently strong, a firm can price to ensure that target customers all buy in the absence of WOM. But in this case, only negative WOM can be instrumental.

To see the intuition behind part 3, notice that the public advertising signal simply shifts the mean of demand rather than the variance. When a is positive, the firm has an incentive to shift the price up by more than a to maximize profits. This reduces the number of target customers who are willing to buy absent WOM, thus making positive WOM instrumental, leading to positive-WOM equilibria. Conversely, if the public signal about quality is very bad (i.e., sufficiently negative a) then the firm is willing to charge a very low price. This makes target customers willing to buy even in the absence of WOM, then making negative WOM instrumental. All in all, the impact on average reviews is similar to Proposition 1: Higher a results in higher average ratings.

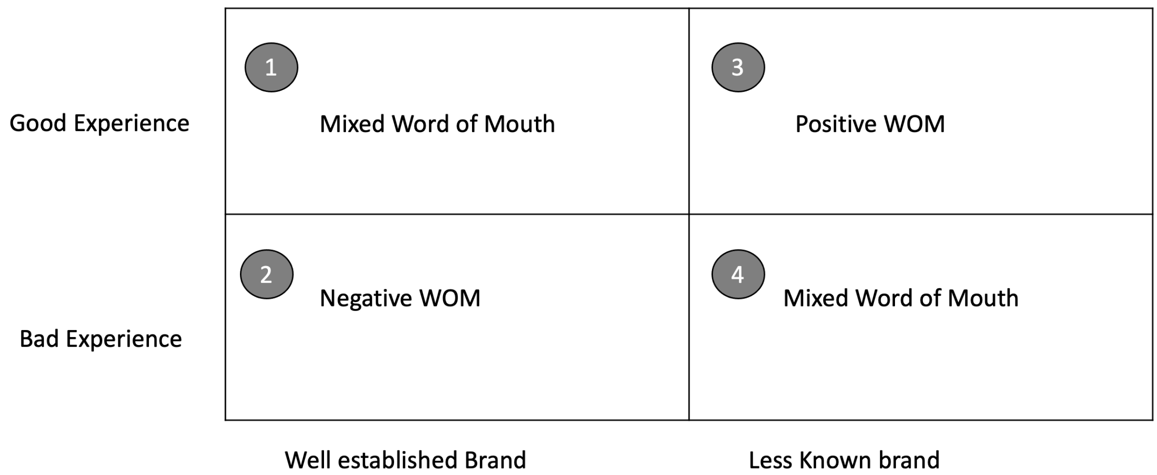
Suggestive Evidence from Data

A premise of our analysis is that early adopters engage in WOM if their information has instrumental value, and we show, under this premise, that the selection of positive versus negative WOM depends on two factors: the strength of the brand image and other public information about the product itself. A stark testable prediction is that with well-entrenched brands (strong brand image), there can only be negative WOM. On the other hand, for weak brands, we expect to see positive WOM under the same conditions. We expect to see mixed WOM for intermediate strength of brand image. In this section, we first provide some experimental evidence for selection of positive versus negative WOM depending on brand strength and for the instrumental motive of WOM. We then show some reduced-form analysis from Yelp.com that further provides suggestive evidence for our theory. Thus, our theoretical model of word of mouth and branding is relevant for a broad set of empirical settings to understand the nature of selection in online word of mouth.

Experimental Evidence for Selection and Instrumentality

To understand whether there are indeed differences in consumers' propensity to review after good/bad experiences with strong versus weak brands and the role of instrumentality, we conducted an online survey on Prolific.com with 400 participants. In this survey, partic-

Figure 3: Experiment Design (n=396)



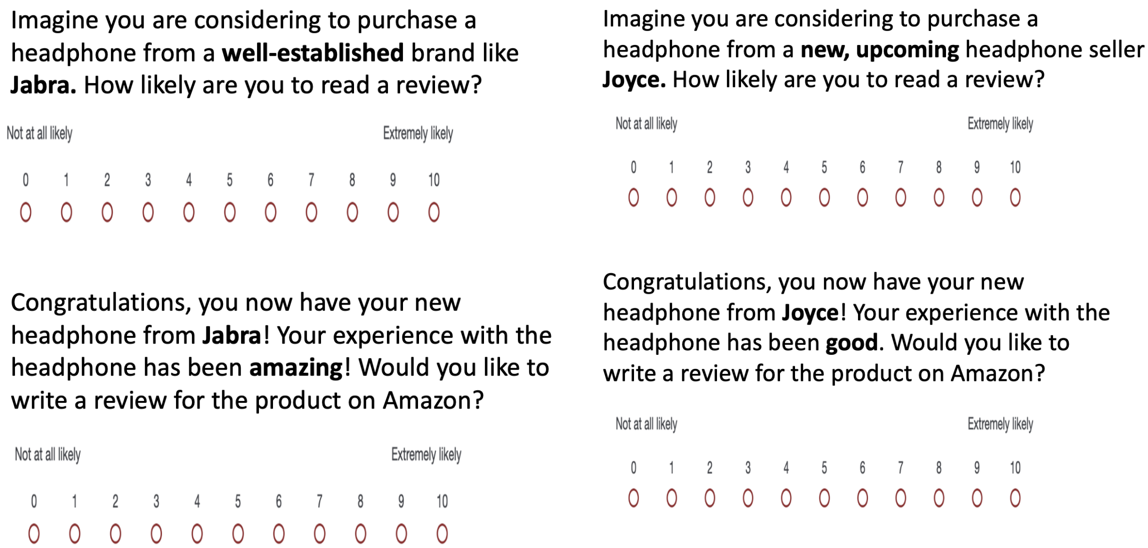
Participants are randomly assigned to one of the 4 conditions (approx 100 in each condition). The theoretical model predicts mixed word of mouth in conditions 1 and 4, positive WOM in condition 3 and negative WOM in condition 2

ipants were asked questions about their reviewing behavior on one of two types of online platforms— e-retailers (like Amazon.com) or restaurant portals (e.g., Yelp.com, Tripadvisor). Both restaurants and online retailers are characterized by ex ante consumer uncertainty about product quality, hence reviews are an important decision component. [18]

We now describe the survey design for the online retailer experiment. Fig 3 is a snapshot of the experiment design which closely maps to the scenarios in our theoretical model. For the online retailer scenario, participants are asked to imagine that they are purchasing a headphone from a large online retailer like Amazon.com. It is $2X2$ between subjects design where survey participants are randomly assigned to one of the 4 conditions — well-established brand & good experience (Condition 1), well-established brand & bad experience (Condition 2), less-known brand & good experience (Condition 3) and less known brand & bad experience (Condition 4) respectively. To give an example, participants in Condition 1 are told to imagine that they are considering to purchase a headphone from a well-established

[18]For quality control of responses, we only used workers who have a minimum acceptance rate of 80% on Prolific.com. We further added some screening questions (to test if they have prior experience with online retail and restaurant reviewing) and an attention check question.

Figure 4: Likelihood to read/write reviews (Condition 1 & 3)



brand (like Jabra, Sony) on Amazon.com. Subsequently, they are told that they had a *good* post-purchase experience with this brand. Within the larger class of good/bad experiences we also manipulate intensity (great, good, ok, not so good, bad) as this is closer to the 5-point ratings on most online review platforms. Fig 4 describes in more detail what participants in Conditions 1 and 3 see. Likewise, participants in the other conditions are asked similar questions modifying the *brand* and *experience* fields as per their assignment. To test their propensity to read/write reviews when they are in each of these conditions, they are asked the following questions - (a) *How likely are they to read a review* before purchase (b) Conditional on their post-purchase experience, *how likely are they to write a review?* on the same platform (c) For those participants who show a propensity to review, we ask two follow-up questions — What are some of your *motivations to review* (experience sharing, venting, getting a promotional coupon, helping others make the right decision, promoting the product and others)? and What is your *most important motive*?

Our survey results have two main findings : First, from Table 2, we can see that the propensity to review is much higher for a good(bad) experience for less-known(well-established) brand. For a well-established brand, the propensity to review is extremely low for a good or average experience (24-29% people will write a review following a good/average experience), somewhat higher for a great experience (41% people would review after a great experience) but the highest for a bad or horrible experience. 88% of consumers would consider writing a review when they have a horrible experience with a well-established brand. On the other hand, 53% consumers would review a less-known brand even if they have a reasonably good experience. This hints at a higher propensity to write negative reviews for well-established brands and positive reviews for less well-known brands. The second key finding from the survey is that most people who write a review (irrespective of brand or experience), do it with an intention of *helping others to make the right decision*. This is consistent with our assumption on the role of instrumentality i.e., people write reviews when they think their review can help others make a better decision. While other incentives exist (experience sharing being another important one), instrumentality does emerge as one of the important motivations.

We get qualitatively similar results in the restaurant review scenario. For chain restaurants which have well-established brand identities, there is higher propensity to write negative reviews. In fact, the second biggest motivation to review chain restaurants is *venting* while in independent restaurants, there is a strong motive to *promote* a good restaurant. However, a key driver of word-of-mouth for all kinds of restaurants and experiences is the desire to improve decision-making for subsequent buyers i.e. instrumentality. We report these results in Web Appendix Section B.1. We now present some data from Yelp.com to further support these experimental findings.

Table 2: Propensity to review (Online Retailer)

	Great	**Good**	**Ok**	**Bad**	**Horrible**
Well Est	41%	24%	29%	71%	88%
Less Est	53%	53%	6%	82%	76%

Table 3: Most Important Motivation to Review (Online Retailer)

Experience	Well Established	Less Established
Great	Help others make the right decision	Share my experience
Good	Give feedback to the company	Help others make the right decision
Ok	Help others make the right decision	Share my experience
Bad	Help others make the right decision	Help others make the right decision
Horrible	Share my experience	Help others make the right decision

Selection and Instrumentality Motives on Yelp.com

Restaurant review platforms present a good empirical setting for our theory. Restaurants are experience goods whose quality cannot be fully ascertained a priori (Nelson, 1970; Luca, 2016) and people often rely on recommendations.[19] Moreover, the industry comprises entities that vary widely in the strength of their brand image: At one extreme, national chains like Domino's Pizza or Starbucks have invested millions of dollars to create a well-entrenched brand image with a clearly communicated brand promise and product portfolio. These firms have numerous stores (products) in different parts of the world under the same brand, allowing consumers to have several contacts (associations) with the brand, thus leading to a strong brand image. At the other extreme, there are independent restaurants that are one-store entities with no strong brand identity. Note also that unlike some other product categories like hotels, cars or movies, which also have active review forums, restaurants typically do not employ strong loyalty programs where reviewers expect loyalty rewards or other external incentives in return for good reviews. This makes Yelp.com a particularly good environment to study the instrumentality motive of WOM.

Finally, it is well-known that consumers on Yelp.com are interested in mainly recent reviews. Our data indicates that the probability of a review on a given day is small – the mean number of reviews per day across a wide range of restaurants is 0.12– thus resulting in a small number of recent reviews. This is consistent with our assumption of small β. Earlier ratings can be captured in the parameter a that reflects public information.[20]

[19] 94 % of US diners are influenced by online reviews as per the Trip Advisor "Influences in Diner Decision-Making" survey 2018. BrightLocal's 2017 Local Consumer Review Survey estimated this number at 97 %

[20] The model assumes that early adopters ignore that their review might also have an impact on a for

Data Description and Summary Statistics

We construct our dataset from the Yelp Data Challenge 2017 and websites of major restaurant brands. The Yelp dataset has business, review and reviewer information for restaurants in several US cities (majorly Pittsburgh, Charlotte, Las Vegas, Cleveland and Phoenix) between the years 2004-2017. Every review in this dataset has a unique identifier, an overall rating, review text and timestamp. Reviews can be linked to a specific reviewer and business through unique business and reviewer identifiers. For every business, we know the name and exact location. For every reviewer we know when they joined the platform, how many years they have been part of the Yelp Elite program, number of friends and fans and how many compliments they have received. We augment this dataset with other business characteristics like whether the business is a chain or not, age of the brand and number of stores of the brand in US (from Statista.com and company websites).[21] We also derive the *cuisine* variable using information from corporate reports for chains and name-matching for independent restaurants.[22]

We restrict attention to cuisines for which there exist both independent restaurants and chains. We identify 72 chains and cluster them based on two dimensions, age of the chain and number of stores in the United States. Seven chains are classified as *national established chains* with a median brand age of *62* years and median spread of *15K* stores per chain. These are Burger King, Domino's Pizza, Dunkin' Donuts, KFC, McDonald's, Pizza Hut and Subway.[23] We have $30,419$ reviews from 2834 such national established chain stores. We combine two additional clusters into a category called *less established chains*. These are either old brands with limited coverage e.g., Carl's Jr and Chick-fil-A or relatively newer brands and

target customers in the far future.

[21]There could be some other ways to quantify brand strength, for example surveys like Brand Asset Valuation (BAV). We find that these survey-based measures are highly correlated with our data-driven clustering based on age and number of stores.

[22]Independent restaurants often have the cuisine in their name for e.g., Otaru Sushi or Mooyah Burgers. We ignore restaurants for which we cannot identify the cuisine.

[23]Table A.8 in the Online Appendix summarizes details about these chains such as revenue, brand value and proportion of positive, negative and neutral word of mouth.

cuisines e.g., Applebee's Neighborhood Grill & Bar, Red Lobster and Chipotle. Their median brand age is *48* years and spread is 1000 stores across US. We have 86,359 reviews from 2913 less established chains. Most of the national established chains are sandwich, pizza, burger joints and coffee shops whereas the less established chains have a wider variety of cuisines e.g., "delis", "chinese", "breakfast" and "steak". To ensure fair comparison between chain restaurants and independent restaurants, we chose independent restaurants serving the same cuisines by name-matching on "sandwich", "pizza", "burger","steak", "deli", "breakfast (or brunch)", "chinese" and "coffee" categories. This gives us a total of 307,622 reviews from 6228 independent restaurants.

Reduced-form Evidence for WOM Selection in Valence: We start with presenting summary statistics and distributions of ratings for different types of restaurants. These are consistent with the theoretical predictions that a strong (weak) brand image leads to negative (positive) selection of WOM (i.e., Propositions 1 and 2).

First, we calculate two review statistics: the average *review-level* star rating and the average store-level star rating. Review rating is simply the average of all reviews for a restaurant. Store rating is the average of the aggregate ratings at an individual store-level. Thus, the store rating gives equal weight to stores, irrespective of review count. Table 4 shows that for independent restaurants, the average store-level star rating (3.56) is lower than the average review-level star rating (3.8). Thus, "good" independent restaurants receive disproportionately many reviews relative to "bad" independent restaurants. In contrast, the average store-level star rating of national established chains (2.46) is higher than the average review-level star rating (2.34). Thus, "good" restaurants that belong to established chains receive disproportionately fewer reviews relative to "bad" restaurants of established chains. This difference suggests a differential propensity to review chains and independent restaurants, conditional on bad or good experiences.

Table 4 also shows that review-level average ratings for independent restaurants tend to be higher (3.8) compared to national established chains (2.3) or less established chains

(3.1). Figure 5 depicts the full distribution of ratings for restaurants with varying degrees of brand strength. National established chains which have the strongest brands receive a large number of 1-star reviews. Independent stores (for which consumers have no prior brand associations), mostly receive 4 and 5 star reviews. The distribution for less established chains is somewhere in between — they receive a mix of positive and negative WOM. This difference in star ratings is unlikely to be only attributed to quality differences. First, many chain restaurants repeatedly ranked higher on customer satisfaction according to the American Customer Satisfaction Index Survey (ACSI).[24] Second, many of these chains have continued to show revenue and profitability growth over the years according to the Quick Service Restaurants Reports 2009-2018. Finally, the number of years that the restaurants are active in the data (in Table 4) is comparable across segments, suggesting that exit of low quality independent restaurants cannot explain the high average reviews of independent restaurants.

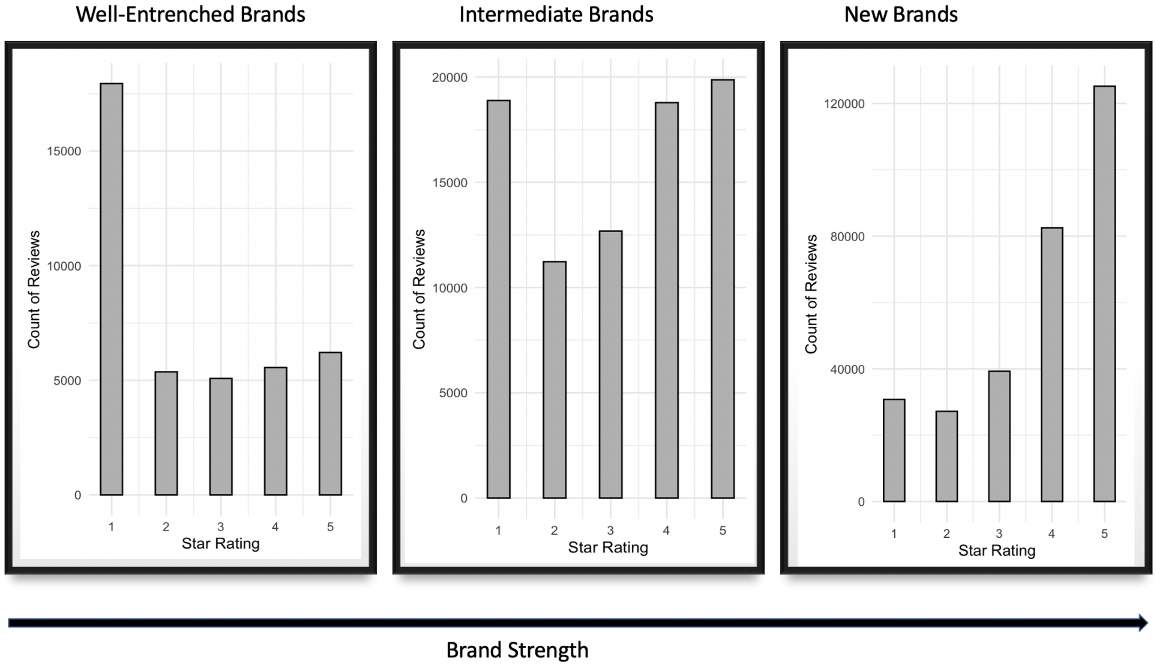
Table 4: Summary Statistics by Restaurant Brand Strength

	Brand Strength								
	Well-Entrenched (National Chains)			Less Known (Local/New Chains)			New Brand (Independent)		
	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Age of Brand (Yrs)	63	62	13	50	48	18		NA	
Stores in US ('000)	15.8	15	5.9	1.9	1	2.5		NA	
Age of Store (Yrs)	2.9	3	2.3	3.9	4	2.8	3.4	3	2.9
Store Rating	2.46	2.27	0.7	2.9	3.1	0.6	3.54	3.91	0.55
Review Rating	2.3	2	1.5	3.1	3	1.5	3.8	4	1.3
No of Stores in Data		2834			2913			6228	
No of Reviews in Data		30,419			86,359			307,622	

Differences in means are statistically significant ($p < 0.00001$).

[24]The American Customer Satisfaction Index (ACSI) measures the satisfaction of U.S. household consumers with the quality of products and services by surveying roughly 300,000 consumers—https://www.theacsi.org/about-acsi. ACSI index and revenue data for some of the largest chains are summarized in Table A.8 in the appendix.

Figure 5: Histograms of Star Ratings by Strength of Brand Image



Support for Instrumentality Motive: We now provide some further evidence for the instrumentality motive to write reviews. We perform a textual analysis of reviews to see what drives positive or negative WOM, and to uncover systematic differences (if any) in the motivations to review well-entrenched versus weak brands. Our hypothesis is that for well-entrenched brands, people have strong common expectations about product quality. Hence, there is greater propensity to report deviations from this *expectation* especially in negative reviews. However, for weak brands, we expect people to mainly review to report instances of *novelty* or *surprise* especially for the positive reviews.

We examine the textual content of a subset of randomly selected 750 reviews. In order to analyze how the review text differs for positive (4-5 stars), negative (1-2 stars) and neutral (3-star) reviews of national established chains (well-entrenched brands) and independent restaurants (weak brands), we create a custom dictionary of *expectation* words and use it to look for instances when reviewers mention prior beliefs and expectations in their reviews. Some examples of these words are “expect”, “past”, “improve”, “decline.” We also use the pre-built LIWC dictionary (Pennebaker, 1997) to identify mentions of discrepancies which capture deviation from expectations.[25] LIWC is a widely-used dictionary in psychology and marketing and examples of discrepancy words include "should", "could", "would have." Together, our custom dictionary of expectation and the LIWC discrepancy keyword list are used to identify mentions of common notions/ expectations and deviations from beliefs. We also construct a custom dictionary of *novelty* to identify mentions of "novel experiences" and being "surprised." Finally, we count the instances in which employees are mentioned.

Table 5 shows the proportion of reviews, by restaurant type and valence, that contain mentions of expectation, novelty and discrepancy. We can see that negative reviews of chains are most likely to have *expect* words (33% of all negative chain reviews). However, positive reviews of chains are also more likely to have *expect* words in comparison to independent restaurants (25% versus 16-18% in independent restaurants). This is consistent with our assumption of a strong brand image and well-entrenched beliefs for national established chains. Neutral reviews for independent restaurants contain more *expect* words—which is not surprising, as a 3-star most often means that the restaurant met expectations. Similarly, *discrepancy* words are more likely found in negative reviews and more common for independent restaurants. This suggests that reviewers are evaluating quality relative to their individual brand associations. *Novel* words are most often found in positive reviews of independent restaurants and negative reviews of chains, which means that people generally report positive surprises for independent restaurants (weak brands), but only negative surprises for national established chains (well-entrenched brands). Interestingly, employees are mentioned most in negative chain reviews. We interpret this as further suggestive evidence supporting our theory: Inspite of precise beliefs about aspects like food and ambiance for a well-entrenched brand, there is still residual uncertainty about service in a particular outlet. Overall, this linguistic analysis of review text suggests that instrumentality is indeed, an important motive to review on Yelp.com.

[25] See Appendix Table A.9 for our dictionaries of expectation, novelty and employee words.

Table 5: Presence of Expectation, Novelty and Discrepancy Words

	Chain (Proportion of Reviews)					Independent (Proportion of Reviews)				
	N	Expect	Novel	Employee	Discrep	N	Expect	Novel	Employee	Discrep
Negative	42	0.33	0.33	0.35	0.73	90	0.18	0.21	0.25	0.75
Positive	70	0.25	0.28	0.24	0.47	378	0.16	0.33	0.2	0.61
Neutral	46	0.3	0.41	0.3	0.65	125	0.27	0.3	0.19	0.78

Expect stands for presence of expect words. Likewise, Novel, Employee and Discrep stands for presence of novelty, staff-related and discrepancy words.

Impact of Strength of Brand Image on Restaurant Ratings: Finally, we perform a regression analysis to quantify the relationship between strength of brand image and review valence. We exclude chains that have an ACSI rating of less than 78 (the average rating for independent restaurants). This is to make both types of restaurants comparable in quality.

In the first specification we estimate the impact of being part of a chain on the overall rating of a restaurant, controlling for several restaurant and user characteristics — price range,[26] cuisine, city, reviewer's platform experience, Elite program membership and reviewer-specific rating leniency.

$$R_{ijt} = \beta_0 + \beta_1 \text{ Chain}_j + \beta_2 X_j + \beta_3 U_i + \epsilon_{ijt} \tag{5}$$

where R_{ijt} denotes the rating of restaurant j by reviewer i at time t, $Chain_j$ captures whether restaurant j is a chain or not, X_j includes restaurant characteristics, and U_i captures reviewer-specific variables.Table 6 (1) shows that being a chain restaurant results in getting about 1 star less than a comparable independent restaurant.[27]

In the second specification we focus on branded restaurants, and use the variables *brand age* and *number of stores* as proxies for the strength of the brand image:

$$R_{ijt} = \beta_0 + \beta_1 \text{ Brand age}_{jt} + \beta_3 \text{ No of stores}_{jt} + \beta_2 X_j + \beta_3 U_i + \epsilon_{ijt}, \tag{6}$$

where Brand age_{jt} measures the age of chain j at time t, and no of stores_{jt} is the number of

[26]Price is not the absolute price but rather a user's perception of restaurant's price range.
[27]We also ran the same regression (5) with only first-year reviews and the coefficients remain similar.

stores of the chain j in US at time t. Table 6 (2) shows that the propensity to write a negative review increases with age of brand and number of stores; a 50 year old brand with thousands of stores will receive 0.5 less stars as compared to a new chain with very few stores. The chain and brand age effects are quite resilient controlling for different user characteristics (4 and 5 in Table 6), though the magnitude of the chain effect is slightly reduced when we account for reviewer-specific leniency (average of reviewer's ratings on other restaurants).[28] In the appendix, section B.4.1 (robustness check), we repeat the same regression with only verified reviewers (i.e. Yelp elite). We again find that the effects are resilient which means that fake reviews alone cannot explain these differences.

Conclusion

We propose a model of strategic WOM that explains how positive and negative WOM arises in equilibrium. We show that the selection of positive versus negative WOM can depend on (i) the strength of the brand image as measured by the dispersion of brand associations, and (ii) the public information available about the new product. The brand image and advertising affects the propensity to review after a good/bad experience. A practical implication is that since the propensity to review varies after good or bad experiences, average reviews are not a reliable measure to compare quality across restaurants. On platforms like Yelp.com however, average ratings are highlighted (in addition to recent reviews).Our analysis implies that WOM should be interpreted differently depending on restaurant type, and quality comparisons solely based on ratings can be problematic. Solutions can be to incentivize all consumers to write reviews, or to present more sophisticated aggregated ratings that control for the systematic selection. Finally, our research sheds light on the link between "conversational motives" and outcomes like valence. We show how the review text can help identify the reviewer's motivation. We leave the optimal design of review aggregation mechanisms and a broader understanding of WOM motives for future research.

[28] There could be an impact of local competition. However, it is not straightforward to define the competition set for a restaurant. So instead, we control for location(city) that captures some of this effect.

Table 6: Impact of Chain Dummy and Brand on Star Ratings

	Dependent variable: Overall Rating				
	rev_stars				
	(1)	(2)	(3)	(4)	(5)
Chain Dummy	-1.061***	-1.062***	-1.021***		
	(0.0641)	(0.0633)	(0.00958)		
Brand Age(Yrs)				-0.0104***	-0.0106***
				(0.000166)	(0.000159)
No of Stores (US)				-0.0000435***	-0.0000380***
				(0.000000871)	(0.000000831)
Age of Store(Yrs)	-0.0103**	-0.00834*	-0.0113***	-0.0119***	-0.0142***
	(0.00329)	(0.00365)	(0.000813)	(0.000865)	(0.000807)
Price Range					
$$	-0.144***	-0.143***	-0.130***	-0.148***	-0.133***
	(0.0338)	(0.0334)	(0.00585)	(0.00563)	(0.00543)
$$$	0.0477	0.0434	0.0725***	0.170***	0.186***
	(0.0627)	(0.0628)	(0.0128)	(0.0131)	(0.0126)
$$$	0.161*	0.153*	0.196***	0.267***	0.293***
	(0.0755)	(0.0751)	(0.0149)	(0.0151)	(0.0146)
Price Range × Chain					
$$ × Chain	0.255***	0.258***	0.255***		
	(0.0595)	(0.0588)	(0.0122)		
$$$ × Chain	0.704**	0.700**	0.654***		
	(0.245)	(0.242)	(0.0556)		
$$$ × Chain	0.724***	0.732***	0.614***		
	(0.162)	(0.156)	(0.0679)		
Select Cuisines					
burger	-0.0531	-0.0548	-0.0694**	0.184***	0.136***
	(0.0905)	(0.0895)	(0.0242)	(0.0259)	(0.0247)
chicken	-0.0563	-0.0549	-0.0232	0.0328	0.0459
	(0.0837)	(0.0831)	(0.0282)	(0.0296)	(0.0283)
chinese	-0.165*	-0.162*	-0.167***	-0.470***	-0.482***
	(0.0741)	(0.0736)	(0.0354)	(0.0378)	(0.0357)
coffee	0.296***	0.293***	0.275***	0.534***	0.482***
	(0.0732)	(0.0724)	(0.0245)	(0.0261)	(0.0249)
dessert	0.407***	0.401***	0.376***	0.314***	0.291***
	(0.0659)	(0.0654)	(0.0277)	(0.0288)	(0.0275)
pizza	-0.116	-0.114	-0.0855***	0.157***	0.143***
	(0.0750)	(0.0744)	(0.0244)	(0.0261)	(0.0249)
sandwich	0.107	0.103	0.111***	0.243***	0.218***
	(0.0668)	(0.0655)	(0.0245)	(0.0262)	(0.0251)
Reviewer characteristics					
Yelp Experience		0.0000620		-0.000126	
		(0.000229)		(0.000110)	
Elite Years		0.0263***		0.0245***	
		(0.00194)		(0.00209)	
N	418653	415423	418653	415423	418653
adj. R-sq	0.096	0.097		0.106	
User Fixed Effect	N	N	Y	N	Y

Note: Standard errors clustered by business $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Note: Restaurant controls include restaurant price range, cuisine and city, where the price range is calculated from user perceptions of a restaurant's price range. User controls include user experience in years, an Elite dummy and reviewer average rating from other reviews. To account for competition we further control for the city location of the restaurant. Specification (1) measures the chain effect without reviewer controls, (2) with reviewer controls, (3) with reviewer fixed effect, (4) and (5) measures the differential impact of brand age and no of stores for a chain brand. (4) and (5) establish that the chain effect is mainly driven by brand strength.

References

Angelis, Matteo De, Andrea Bonezzi, Alessandro M Peluso, Derek D Rucker, and Michele Costabile, "On braggarts and gossips: A self-enhancement account of word-of-mouth generation and transmission," *Journal of Marketing Research*, 2012, *49* (4), 551–563.

Baker, Andrew M, Naveen Donthu, and V Kumar, "Investigating how word-of-mouth conversations about brands influence purchase and retransmission intentions," *Journal of Marketing Research*, 2016, *53* (2), 225–239.

Bass, Frank M, "A new product growth for model consumer durables," *Management science*, 1969, *15* (5), 215–227.

Berger, Jonah, "Word of mouth and interpersonal communication: A review and directions for future research," *Journal of Consumer Psychology*, 2014, *24* (4), 586–607.

Biel, Alexander L, "How brand image drives brand equity," *Journal of advertising research*, 1992, *32* (6), RC6–RC12.

Biyalogorsky, Eyal, Eitan Gerstner, and Barak Libai, "Customer referral management: Optimal reward programs," *Marketing Science*, 2001, *20* (1), 82–95.

Bruhn, Manfred, Verena Schoenmueller, and Daniela B Schäfer, "Are social media replacing traditional media in terms of brand equity creation?," *Management research review*, 2012.

Campbell, Arthur, "Word-of-mouth communication and percolation in social networks," *American Economic Review*, 2013, *103* (6), 2466–98.

_ , Dina Mayzlin, and Jiwoong Shin, "Managing buzz," *The RAND Journal of Economics*, 2017, *48* (1), 203–229.

Carman, Ashley, "Why do you leave restaurant reviews?," *The Verge*, 2018. accessed July 27, 2020.

Chakraborty, Ishita, Minkyung Kim, and K Sudhir, "Attribute Sentiment Scoring with Online Text Reviews: Accounting for Language Structure and Attribute Self-Selection," *Cowles Foundation Discussion Paper*, 2019.

Chen, Yubo and Jinhong Xie, "Online consumer review: Word-of-mouth as a new element of marketing communication mix," *Management Science*, 2008, *54* (3), 477–491.

Chevalier, Judith A and Dina Mayzlin, "The effect of word of mouth on sales: Online book reviews," *Journal of marketing research*, 2006, *43* (3), 345–354.

Chintagunta, Pradeep K, Shyam Gopinath, and Sriram Venkataraman, "The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets," *Marketing Science*, 2010, *29* (5), 944–957.

Dhar, Vasant and Elaine A Chang, "Does chatter matter? The impact of user-generated content on music sales," *Journal of Interactive Marketing*, 2009, *23* (4), 300–307.

Duan, Wenjing, Bin Gu, and Andrew B Whinston, "Do online reviews matter?—An empirical investigation of panel data," *Decision support systems*, 2008, *45* (4), 1007–1016.

Fradkin, Andrey, Elena Grewal, Dave Holtz, and Matthew Pearson, "Bias and reciprocity in online reviews: Evidence from field experiments on airbnb," in "Proceedings of the Sixteenth ACM Conference on Economics and Computation" ACM 2015, pp. 641–641.

Galeotti, Andrea, "Talking, searching, and pricing*," *International Economic Review*, 2010, *51* (4), 1159–1174.

_ and Sanjeev Goyal, "Influencing the influencers: a theory of strategic diffusion," *The RAND Journal of Economics*, 2009, *40* (3), 509–532.

Godes, David, "Product policy in markets with word-of-mouth communication," *Management Science*, 2016, *63* (1), 267–278.

_ , Dina Mayzlin, Yubo Chen, Sanjiv Das, Chrysanthos Dellarocas, Bruce Pfeiffer, Barak Libai, Subrata Sen, Mengze Shi, and Peeter Verlegh, "The firm's management of social interactions," *Marketing Letters*, 2005, *16* (3), 415–428.

Grice, H Paul, Peter Cole, Jerry Morgan et al., "Logic and conversation," *1975*, 1975, pp. 41–58.

He, Sherry, Brett Hollenbeck, and Davide Proserpio, "The market for fake reviews," *Available at SSRN*, 2020.

Hollenbeck, Brett, "Online reputation mechanisms and the decreasing value of chain affiliation," *Journal of Marketing Research*, 2018, *55* (5), 636–654.

Johnson, Justin P and David P Myatt, "On the simple economics of advertising, marketing, and product design," *American Economic Review*, 2006, *96* (3), 756–784.

Joshi, Yogesh V and Andres Musalem, "When consumers learn, money burns: Signaling quality via advertising with observational learning and word of mouth," *Marketing Science*, 2021, *40* (1), 168–188.

Kamada, Yuichiro and Aniko Öry, "Contracting with Word-of-Mouth Management," *Cowles Foundation Discussion Paper No. 2048*, 2017.

Ke, T Tony, Jiwoong Shin, and Jungju Yu, "A Theory of Brand Positioning: Product-Portfolio View," *Available at SSRN 3688098*, 2020.

Kornish, Laura J. and Qiuping Li, "Optimal Referral Bonuses with Asymmetric Information: Firm-Offered and Interpersonal Incentives," *Marketing Science*, 2010, *29* (1), 108–121.

Kotler, Philip, *Marketing Management: The Millennium Edition*, Prentice Hall, Inc., 2000.

Leduc, Matt V, Matthew O Jackson, and Ramesh Johari, "Pricing and referrals in diffusion on networks," *Games and Economic Behavior*, 2017, *104*, 568–594.

Lee, J Lucy, Jeffrey D James, and Yu Kyoum Kim, "A reconceptualization of brand image," *International Journal of Business Administration*, 2014, *5* (4), 1.

Liu, Xiao, Dokyun Lee, and Kannan Srinivasan, "Large-scale cross-category analysis of consumer review content on sales conversion leveraging deep learning," *Journal of Marketing Research*, 2019, *56* (6), 918–943.

Lovett, Mitchell J, Renana Peres, and Ron Shachar, "On brands and word of mouth," *Journal of marketing research*, 2013, *50* (4), 427–444.

Luca, Michael, "Reviews, reputation, and revenue: The case of Yelp.com," *Com (March 15, 2016). Harvard Business School NOM Unit Working Paper*, 2016, (12-016).

_ **and Georgios Zervas**, "Fake it till you make it: Reputation, competition, and Yelp review fraud," *Management Science*, 2016, *62* (12), 3412–3427.

Luo, Xueming, "Quantifying the long-term impact of negative word of mouth on cash flows and stock prices," *Marketing Science*, 2009, *28* (1), 148–165.

Mayzlin, Dina, "Promotional chat on the Internet," *Marketing science*, 2006, *25* (2), 155–163.

_ **and Jiwoong Shin**, "Uninformative advertising as an invitation to search," *Marketing Science*, 2011, *30* (4), 666–685.

_ **, Yaniv Dover, and Judith Chevalier**, "Promotional reviews: An empirical investigation of online review manipulation," *The American Economic Review*, 2014, *104* (8), 2421–2455.

Mittal, Vikas, William T Ross Jr, and Patrick M Baldasare, "The asymmetric impact of negative and positive attribute-level performance on overall satisfaction and repurchase intentions," *Journal of marketing*, 1998, *62* (1), 33–47.

Nelson, Phillip, "Information and Consumer Behavior," *Journal of Political Economy*, 1970, *78* (2), 311–329.

Newman, Joseph W, "New insight, new progress, for marketing," *Harvard business review*, 1957, *35* (6), 95–102.

Nosko, Chris and Steven Tadelis, "The limits of reputation in platform markets: An empirical analysis and field experiment," Technical Report, National Bureau of Economic Research 2015.

Onishi, Hiroshi and Puneet Manchanda, "Marketing activity, blogging and sales," *International Journal of Research in Marketing*, 2012, *29* (3), 221–234.

Pennebaker, James W, "Writing about emotional experiences as a therapeutic process," *Psychological science*, 1997, *8* (3), 162–166.

Reimers, Imke C and Joel Waldfogel, "Digitization and pre-purchase information: the causal and welfare impacts of reviews and crowd ratings," *American Economic Review (forthcoming)*, 2021.

Schoenmüller, Verena, Oded Netzer, and Florian Stahl, "The extreme distribution of online reviews: Prevalence, drivers and implications," *Journal of Marketing Research*, forthcoming.

Seiler, Stephan, Song Yao, and Wenbo Wang, "Does Online Word of Mouth Increase Demand?(And How?) Evidence from a Natural Experiment," *Marketing Science*, 2017.

Sun, Monic, "How does the variance of product ratings matter?," *Management Science*, 2012, *58* (4), 696–707.

Sundaram, DS and Cynthia Webster, "The role of brand familiarity on the impact of word-of-mouth communication on brand evaluations," *ACR North American Advances*, 1999.

Toubia, Olivier and Andrew T Stephen, "Intrinsic vs. image-related utility in social media: Why do people contribute content to twitter?," *Marketing Science*, 2013, *32* (3), 368–392.

A Appendix: Proofs

A.1 Properties of $\hat{\theta}(a, b_i, \emptyset)$

We use the following property of the first moment of truncated normal distributions: If $X \sim \mathcal{N}(\mu, \sigma)$, then $\mathbb{E}(X|\underline{s} < X < \bar{s}) = \mu - \sigma \frac{\phi\left(\frac{\bar{s}-\mu}{\sigma}\right) - \phi\left(\frac{\underline{s}-\mu}{\sigma}\right)}{\Phi\left(\frac{\bar{s}-\mu}{\sigma}\right) - \Phi\left(\frac{\underline{s}-\mu}{\sigma}\right)}$. This allows us simplify $\hat{\theta}(a, b_i, \emptyset)$ further for equilibria with $\mathcal{S}_{a,p} = (-\infty, \underline{s}]$, $\mathcal{S}_{a,p} = [\bar{s}, \infty)$, and $\mathcal{S}_{a,p} = (-\infty, \underline{s}] \cup [\bar{s}, \infty)$:

- If $\mathcal{S}_{a,p} = (-\infty, \underline{s}]$, then

$$\hat{\theta}(a, b_i, \emptyset) = \mu_\theta(a, b_i) + \beta \frac{\sigma_a^2 \sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2 + \sigma_\theta^2)}} \frac{\phi\left(\frac{\underline{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}}\right)}{1 - \Phi\left(\frac{\underline{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}}\right)} \tag{7}$$

- If $\mathcal{S}_{a,p} = [\bar{s}, \infty)$, then

$$\hat{\theta}(a, b_i, \emptyset) = \mu_\theta(a, b_i) - \beta \frac{\sigma_a^2 \sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2 + \sigma_\theta^2)}} \frac{\phi\left(\frac{\bar{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}}\right)}{\Phi\left(\frac{\bar{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2 2}}}\right)} \tag{8}$$

- If $\mathcal{S}_{a,p} = (-\infty, \underline{s}] \cup [\bar{s}, \infty)$, then

$$\hat{\theta}(a, b_i, \emptyset) = \mu_\theta(a, b_i) - \beta \frac{\sigma_a^2 \sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2 + \sigma_\theta^2)}} \frac{\phi\left(\frac{\bar{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}}\right) - \phi\left(\frac{\underline{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}}\right)}{\Phi\left(\frac{\bar{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}}\right) - \Phi\left(\frac{\underline{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}}\right)}. \tag{9}$$

A.2 Proof of Lemma 1

The fact that $\mathcal{S}_{a,p}$ is of the form $\emptyset$, $(-\infty, \underline{s}]$, $[\bar{s}, \infty)$, or $(-\infty, \underline{s}] \cup [\bar{s}, \infty)$ follows directly from the monotonicity of $t(p, s, a)$ in s and Equations (1) and (2). It remains to show that $\bar{s}$ and $\underline{s}$ satisfy Equations (3) and (4), respectively. To this end, we examine full, positive and negative WOM equilibria in turn:

Let us first consider $\mathcal{S}_{a,p} = (-\infty, \underline{s}] \cup [\bar{s}, \infty)$. First, we show that $\bar{s} > \underline{s}$, i.e., $\mathcal{S}_{a,p} \neq \mathbb{R}$, by contradiction. If $\mathcal{S}_{a,p} = \mathbb{R}$, then $\hat{\theta}(a, b_i, \emptyset) = \mu_\theta(a, b_i)$ and $\tilde{t}(p, \mathcal{S}_{a,p}, a, \beta) = p$. Further, for an early adopter with $s = \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a$, $t(p, s) = p + \frac{\sigma_a^2 \sigma_\theta^2}{\Sigma} \left(\frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a - s \right) = p$, so she does not have an incentive to engage in WOM. Hence, it must be that $\bar{s} > \underline{s}$.

An early adopter with a signal s with $t(p, s, a) \leq \tilde{t}(p, \mathcal{S}_{a,p}, a, \beta)$ weakly prefers to engage in WOM if and only if $\Phi_{\mu_\theta}(\tilde{t}(p, \mathcal{S}_{a,p}, a, \beta)) - \Phi_{\mu_\theta}(t(p, s, a)) \geq \xi$ by (1). If an early adopter with such a signal s (weakly) prefers to engage in WOM, then all early adopters with $s' > s$ strictly prefer to engage in WOM as well because $t(p, s', a) < t(p, s, a)$. So, if $t(p, \underline{s}, a) \leq \tilde{t}(p, \mathcal{S}_{a,p}, a, \beta)$, then $[\underline{s}, \infty) \subset \mathcal{S}_{a,p}$ which is a contradiction to $\mathcal{S}_{a,p} = (-\infty, \underline{s}] \cup [\bar{s}, \infty)$, $\underline{s} < \bar{s}$. Hence, it must be that $t(p, \underline{s}, a) > \tilde{t}(p, \mathcal{S}_{a,p}, a, \beta)$. Analogously, it follows that $t(p, \bar{s}, a) < \tilde{t}(p, \mathcal{S}_{a,p}, a, \beta)$.

Further, an early adopter with signal s such that $t(p, \underline{s}, a) > t(p, s, a) > \tilde{t}(p, \mathcal{S}_{a,p}, a, \beta)$ weakly prefers not to engage in WOM if and only if $\Phi_{\mu_\theta}(\tilde{t}(p, \mathcal{S}_{a,p}, a, \beta)) - \Phi_{\mu_\theta}(t(p, s, a)) \leq \xi$. By continutity of $t(p, s, a)$, it follows that (4) must hold to make sure that early adopters with signals $s \geq \underline{s}$ weakly prefer not to engage in WOM and early adopters with $s \leq \underline{s}$ weakly prefer to engage in WOM. Analogously, for $\bar{s}$ it follows that (3) must hold.

For $\mathcal{S}_{a,p} = (-\infty, \underline{s}]$, $t(p, \underline{s}, a) > \tilde{t}(p, \mathcal{S}_{a,p}, a, \beta)$ and (4) follow analogously. For $\mathcal{S}_{a,p} = [\bar{s}, \infty)$, $t(p, \bar{s}, a) < \tilde{t}(p, \mathcal{S}_{a,p}, a, \beta)$ and (3) follow analogously.

A.3 Proof of Proposition 1

We first discuss necessary conditions for a negative WOM equilibrium, a positive WOM equilibrium, a full WOM equilibrium, and a no WOM equilibrium to arise. Note that no other type of WOM equilibrium can be sustained by an analogous argument to Lemma 1. Then, we compare the induced profits in order to determine the profit-maximizing price.

The following are necessary conditions for different types of WOM equilibria to arise. Note that with a well-entrenched brand, the firm and early adopter observe $\hat{\theta}$ as there is no uncertainty about brand association.

1. Let us consider a candidate equilibrium in which the firm sets a price after which the target customer plays a negative WOM equilibrium, i.e., $\mathcal{S}_{a,p} = (-\infty, \underline{s}]$. After a message $m = \emptyset$, the target customer is willing to buys if and only if $\hat{\theta}(a, b_0, \emptyset) \geq p$. For a sufficiently small signal s,

$$\hat{\theta}(a, b_0, s) \;=\; \bar{\mu}_\theta(a, b_0) + \frac{\sigma_a^2 \sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2 + \sigma_\theta^2)}} \frac{s - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}} < \hat{\theta}(a, b_0, \emptyset),$$

so early adopter with such a signal s is instrumental (with an incentive to talk) only if $p \in (\hat{\theta}(a, b_0, s), \hat{\theta}(a, b_0, \emptyset))$. Thus, in a negative WOM equilirbium, all target consumers must not buy if they see a signal $s \leq \underline{s}$ and buy if they saw a signal $s > \underline{s}$. Thus,

$$p = \hat{\theta}(a, b_0, \underline{s}) \leq \hat{\theta}(a, b_0, \emptyset) = \mu_\theta(a, b_i) + \beta \frac{\sigma_a^2 \sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2 + \sigma_\theta^2)}} \frac{\phi\left(\frac{\underline{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}}\right)}{1 - \Phi\left(\frac{\underline{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}}\right)}.$$

Thus, it must be that $\frac{\underline{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}} \leq \beta \frac{\phi\left(\frac{\underline{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}}\right)}{1 - \Phi\left(\frac{\underline{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}}\right)}$. Note that $\frac{x(1-\Phi(x))}{\phi(x)}$ is increasing in x for $x \geq 0$ where $\frac{0(1-\Phi(0))}{\phi(0)} = 0$ and $\lim_{x \to \infty} \frac{x(1-\Phi(x))}{\phi(x)} = \infty$. Thus there is a unique $\underline{x}(\beta)$ so that $\frac{\underline{x}(\beta)(1-\Phi(\underline{x}(\beta)))}{\phi(\underline{x}(\beta))} = \beta$.[29] Thus, it must be that $\frac{\underline{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}} \leq \underline{x}(\beta)$. Profits are then

[29]Note that it also follows that $\underline{x}(\beta)$ is increasing in β which is used in the corollary.

given by

$$\Pi^*_{\text{neg}}(\beta) := \underbrace{\left((1-\beta)+\beta\left(1-\Phi\left(\frac{\underline{s}-\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2}a}{\sqrt{\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}}\right)\right)\right)}_{\text{prob. that the consumer sees } m=\emptyset}\left(\bar{\mu}(a,b_0)+\frac{\sigma_a^2\sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2+\sigma_\theta^2)}}\frac{\underline{s}-\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2}a}{\sqrt{\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}}\right).$$

2. Next consider a candidate equilibrium in which the target customer plays a positive WOM equilibrium, i.e., $\mathcal{S}_{a,p} = [\bar{s}, \infty)$, then early adopters have an instrumental incentive to talk if they see $s \geq \bar{s}$ but no incentive to talk when $s < \bar{s}$ only if

$$\hat{\theta}(a,b_0,\emptyset) \leq p = \hat{\theta}(a,b_0,\bar{s}) = \bar{\mu}_\theta(a,b_0)+\frac{\sigma_a^2\sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2+\sigma_\theta^2)}}\frac{\bar{s}-\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2}a}{\sqrt{\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}}.$$

Thus, it must be that $\frac{\bar{s}-\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2}a}{\sqrt{\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}} \geq -\beta\frac{\phi\left(\frac{\bar{s}-\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2}a}{\sqrt{\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}}\right)}{\Phi\left(\frac{\bar{s}-\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2}a}{\sqrt{\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}}\right)}$. Note that $-\frac{x\Phi(x)}{\phi(x)}$ is decreasing in x for $x \leq 0$ where $-\frac{0\Phi(0)}{\phi(0)} = 0$ and $\lim_{x\to-\infty} -\frac{x\Phi(x)}{\phi(x)} = \infty$. Thus there is a unique $\bar{x}(\beta) < 0$ so that $-\frac{\bar{x}(\beta)\Phi(\bar{x}(\beta))}{\phi(\bar{x}(\beta))} = \beta$. Then, in any positive WOM equilibrium it must be that $\frac{\bar{s}-\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2}a}{\sqrt{\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}} \geq \bar{x}(\beta)$. Profits are given by

$$\Pi^*_{\text{pos}}(\beta) := \beta\left(1-\Phi\left(\frac{\bar{s}-\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2}a}{\sqrt{\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}}\right)\right)\left(\bar{\mu}_\theta(a,b_0)+\frac{\sigma_a^2\sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2+\sigma_\theta^2)}}\frac{\bar{s}-\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2}a}{\sqrt{\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}}\right).$$

Note that $\overline{x}(\beta) = -\underline{x}(\beta)$ and $\bar{x}(\beta) < 0 < \underline{x}(\beta)$

3. If the firm sets a price after which the target customer plays a full WOM equilibrium, i.e., $\mathcal{S}_{a,p} = (-\infty, \underline{s}] \cup [\bar{s}, \infty)$, then by an analogous argument to above it must be that $\hat{\theta}(a,b_0,\emptyset) = \hat{\theta}(a,b_0,\bar{s}) = \hat{\theta}(a,b_0,\underline{s})$, which implies $\bar{s} = \underline{s}$, i.e., all early adopters engage in WOM. If all early adopters talk, then $\hat{\theta}(a,b_0,\emptyset) = \bar{\mu}(a,b_0) = b_0 + \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2}a$. Thus, if

consumers who see a message $m = \emptyset$ do not buy at a price p, then then consumers who see $s < \bar{\mu}(a, b_0)$ do not buy, so early adopters with $s < \bar{\mu}(a, b_0)$ do not have an incentive to talk. Similarly, if consumers buy after a message $m = \emptyset$, then early adopters with $s > \bar{\mu}(a, b_0)$ do not have an incentive to talk. Thus, a full WOM equilibrium can be sustained only if target consumers play a mixed strategy.

If a target customer who sees $m = \emptyset$ buys with probability $\alpha \in (0, 1)$ and does not buy with probability $1 - \alpha$, then the price must be $p = \bar{\mu}(a, b_0)$ and early adopters with $s < \bar{\mu}_\theta(a, b_0)$ have an incentive to talk if and only if $\alpha > \xi$ and early adopters with $s > \bar{\mu}_\theta(a, b_0)$ have an incentive to talk if and only if $(1 - \alpha) > \xi$, i.e., it must be that $\xi < \alpha < 1 - \xi$. Thus, a full WOM equilibrium exists if and only if the price is equal to $p^*_{\text{full}} = \bar{\mu}(a, b_0)$ and $\frac{1}{2} \geq \xi$. Maximal profits in that case are equal to

$$\Pi^*_{\text{full}}(\beta) := \left((1 - \beta)\big(1 - \xi\big) + \beta\frac{1}{2}\right) \bar{\mu}(a, b_0).$$

4. If a firm sets a price after which the target customer play a no WOM equilibrium, i.e., $\mathcal{S}_{a,p} = \emptyset$, then $\hat{\theta}(b_0, \emptyset) = \bar{\mu}_\theta(a, b_0)$. A no WOM equilibrium always exists if $\xi > 1$. Let us next assume $\xi \leq 1$. If $p < \bar{\mu}_\theta(a, b_0)$, target customers buy if $m = \emptyset$. Thus, all early adopters with sufficiently small s have an incentive to talk. If $p > \bar{\mu}_\theta(a, b_0)$, all early adopters with large s have an incentive to talk. Finally, consider $p = \bar{\mu}_\theta(a, b_0)$, and assume that target customers who see $m = \emptyset$ buy with probability α. Then, all early adopters do not talk if and only if and only if $\xi \geq \max\{\alpha, 1 - \alpha\}$. Thus, a no WOM equilibrium exists if and only if $\xi \geq \frac{1}{2}$ and $p = \bar{\mu}_\theta(a, b_0)$. Maximal profits in that case are $\Pi^*_{\text{no}} := \min\{1, \xi\}\bar{\mu}_\theta(a, b_0)$.

Finally, we can compare profits to determine the equilibrium outcome for small β:

1. If $\xi \leq \frac{1}{2}$, only positive, negative or full WOM equilibria can arise. For small β we have

$$\begin{aligned}\lim_{\beta\to 0}\Pi^*_{\text{neg}}(\beta) &= \lim_{\beta\to 0}\left((1-\beta)+\beta\left(1-\Phi\left(\underline{x}(\beta)\right)\right)\right)\left(\bar{\mu}(a,b_0)+\frac{\sigma_a^2\sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2+\sigma_\theta^2)}}\underline{x}(\beta)\right)\\ &> \lim_{\beta}\Pi^*_{\text{pos}}(\beta)=0.\end{aligned}$$

   Further,

$$\left((1-\beta)+\beta\left(1-\Phi\left(\underline{x}(\beta)\right)\right)\right)\left(\bar{\mu}(a,b_0)+\frac{\sigma_a^2\sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2+\sigma_\theta^2)}}\underline{x}(\beta)\right) > \left((1-\beta)(1-\xi)+\beta\frac{1}{2}\bar{\mu}(a,b_0)\right),$$

   so for small β, $\Pi^*_{\text{neg}}(\beta) > \Pi^*_{\text{full}}$. Thus, for small β there is a unique equilibrium, in which the firm sets a price of $p^*_{\text{neg}}(\beta) := \bar{\mu}_\theta(a,b_0)+\frac{\sigma_a^2\sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2+\sigma_\theta^2)}}\underline{x}(\beta)$, and early adopters talk if and only if $\underline{s} \leq \underline{s}^*(a,\beta) = \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2}a+\sqrt{\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}\underline{x}(\beta)$.

2. If $\frac{1}{2} < \xi \leq 1$, only negative, positive or no WOM equilibria can be sustained. Further, for the same reason to above, for small β, $\Pi^*_{\text{neg}}(\beta) > \Pi^*_{\text{pos}}(\beta)$. Finally, for sufficiently small β we have

$$\left((1-\beta)+\beta\left(1-\Phi\left(\underline{x}(\beta)\right)\right)\right)\left(\bar{\mu}(a,b_0)+\frac{\sigma_a^2\sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2+\sigma_\theta^2)}}\underline{x}(\beta)\right) > \xi\bar{\mu}_\theta(a,b_0),$$

   so the unique equilibrium is a negative WOM equilibrium with price

$$p^*_{\text{neg}}(\beta) := \bar{\mu}_\theta(a,b_0)+\frac{\sigma_a^2\sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2+\sigma_\theta^2)}}\underline{x}(\beta),$$

   and early adopters talk if and only if $s \geq \bar{s}^*(a,\beta) = \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2}a+\sqrt{\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}\bar{x}(\beta)$.

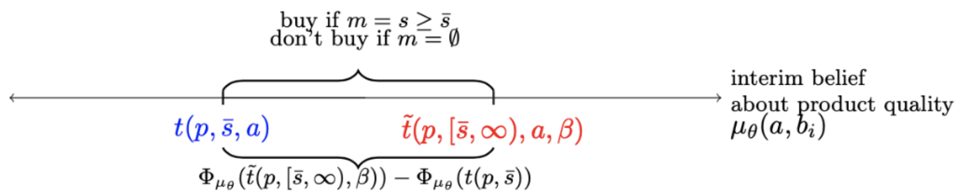
3. If $\xi > 1$, the firm sets the price equal to $\bar{\mu}_\theta(a,b_0)$ and induces a no-WOM equilibrium.

A.4 Proof of Lemma 2

We first show existence of equilibria and then uniqueness.

1. First, we show that with a pessimistic target customer, a positive WOM equilibrium exists. By Lemma 1, in a positive WOM equilibrium, we have that $\mathcal{S}_{a,p} = [\bar{s}, \infty)$, $b(p, \bar{s}) \leq \tilde{t}(p, \mathcal{S}_{a,p}, \beta)$ and (3) must hold. We first show that such a $\bar{s}$ exists. Then, we check that all early adopters with signals in $s \geq \bar{s}$ indeed want to talk, while early adopters with signals $s < \bar{s}$ do not want to talk. We know from (8) that

Figure A.1: Positive WOM equilibrium



$$\tilde{t}(p, [\bar{s}, \infty), a, \beta) = p + \beta \frac{\sigma_a^2 \sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2 + \sigma_\theta^2)}} \frac{\phi\left(\frac{\bar{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}}\right)}{\Phi\left(\frac{\bar{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2 2}}}\right)} \geq p - \frac{\sigma_a^2 \sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2 + \sigma_\theta^2)}} \frac{\bar{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}} = t(p, \bar{s}, a).$$

Analogously to the proof of Proposition 1 it follows that it must be that

$$\frac{\bar{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}} \geq \bar{x}(\beta) \quad \Leftrightarrow \quad s \geq \bar{s}^*(a, \beta).$$

Moreover, note that $x \mapsto \frac{\phi(x)}{\Phi(x)}$ is decreasing and $\lim_{x \to \infty} \frac{\phi(x)}{\Phi(x)} = 0$. Hence, with a pessimistic target customer we know that $\Phi_{\mu_\theta}(p) > \xi$, so

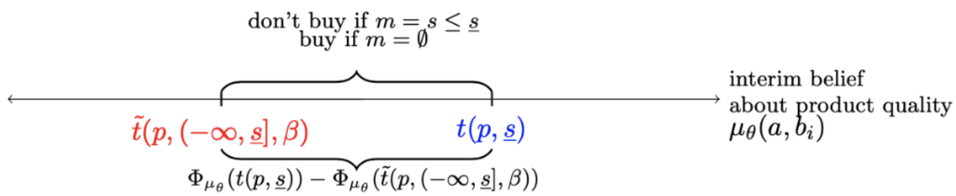
$$\begin{aligned} \lim_{\bar{s} \to \infty} \Phi_{\mu_\theta}(\tilde{t}(p, [\bar{s}, \infty), a, \beta)) - \Phi_{\mu_\theta}(t(p, \bar{s}, a)) = \Phi_{\mu_\theta}(p) \qquad & > \xi > \\ 0 = \lim_{\bar{s} \to \bar{s}^*(a,\beta)} \Phi_{\mu_\theta}(\tilde{t}(p, [\bar{s}, \infty), a, \beta)) - \Phi_{\mu_\theta}(t(p, \bar{s})) \qquad & . \end{aligned}$$

Thus, there exists a $\bar{s}$ with (3) by continuity of $\Phi_{\mu_\theta}(\tilde{t}(p, [\bar{s}, \infty), \beta)) - \Phi_{\mu_\theta}(t(p, \bar{s}))$ in $\bar{s}$ and the intermediate value theorem.

Finally, note that for each such $\bar{s}$, monotonicity of $t(p, s)$ implies that early adopters with $s < \bar{s}$ do not want to talk while early adopters with $s > \bar{s}$ want to talk.

2. Next we show that with optimistic target consumers, a negative WOM equilibrium exists. By Lemma 1, in a negative WOM equilibrium, we have that $\mathcal{S}_{a,p} = (-\infty, \underline{s}]$, $b(p, \underline{s}) \geq \tilde{t}(p, \mathcal{S}_{a,p}, \beta)$ and (4) must hold. We first show that such a $\underline{s}$ exists. Then, we check that all early adopters with signals in $s \leq \underline{s}$ indeed want to talk, while early adopters with signals $s > \underline{s}$ do not want to talk. We know from (7) that

Figure A.2: Negative WOM equilibrium



$$\tilde{t}(p, (-\infty, \underline{s}], a, \beta) = p - \beta \frac{\sigma_a^2 \sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2+\sigma_\theta^2)}} \frac{\phi\left(\frac{\underline{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}}\right)}{1 - \Phi\left(\frac{\underline{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}}\right)} \leq p - \frac{\sigma_a^2 \sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2+\sigma_\theta^2)}} \frac{\underline{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}} = t(p, \underline{s}, a).$$

Analogously to the proof of Proposition 1 it follows that it must be that

$$\frac{\underline{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2+\sigma_\theta^2}}} \leq \underline{x}(\beta) \quad \Leftrightarrow \quad s \leq \underline{s}^*(a, \beta).$$

Moreover, note that $x \mapsto \frac{\phi(x)}{1-\Phi(x)}$ is increasing and $\lim_{x \to -\infty} \frac{\phi(x)}{1-\Phi(x)} = 0$. Hence, with optimistic target consumers we know that $1 - \Phi_{\mu_\theta}(p) > \xi$, so

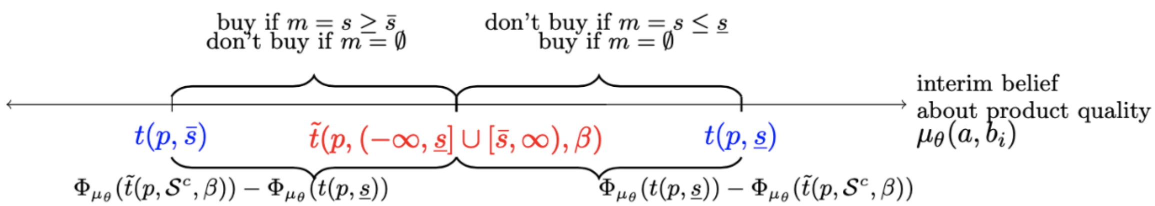
$$\begin{aligned} \lim_{\underline{s} \to -\infty} \Phi_{\mu_\theta}(t(p, \underline{s})) - \Phi_{\mu_\theta}(\tilde{t}(p, (-\infty, \underline{s}], \beta)) = 1 - \Phi_{\mu_\theta}(p) \qquad & > \xi > \\ 0 = \lim_{s \to \underline{s}^*(a,\beta)} \Phi_{\mu_\theta}(t(p, \underline{s})) - \Phi_{\mu_\theta}(\tilde{t}(p, (-\infty, \underline{s}], \beta)) \qquad & . \end{aligned}$$

Thus, there exists a $\underline{s}$ with (4) by continuity of $\Phi_{\mu_\theta}(t(p, \underline{s})) - \Phi_{\mu_\theta}(\tilde{t}(p, (-\infty, \underline{s}], \beta))$ in $\underline{s}$ and the intermediate value theorem.

Finally, note that for each such $\underline{s}$, monotonicity of $t(p, s)$ implies that early adopters with $s > \underline{s}$ do not want to talk while early adopters with $s < \underline{s}$ want to talk.

3. Next we show that with uninformed consumers, a full-WOM equilibrium exists. By Lemma 1, in a full-WOM equilibrium, we have that $\mathcal{S}_{a,p} = (-\infty, \underline{s}] \cup [\bar{s}, \infty)$, $b(p, \underline{s}) \geq \tilde{t}(p, \mathcal{S}_{a,p}, \beta) \geq b(p, \bar{s})$, and (3) and (4) must hold. We first show that such $\bar{s}$ and $\underline{s}$ exist. Then, we check that all early adopters with signals in $s \geq \bar{s}$ and $s \leq \underline{s}$ indeed want to talk, while those with signals $s \in (\underline{s}, \bar{s})$ do not want to talk. We know from (9) that

Figure A.3: Full-WOM equilibrium



$$\begin{aligned}
t(p, \bar{s}, a) &= p - \frac{\sigma_a^2 \sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2 + \sigma_\theta^2)}} \frac{\bar{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}} \leq \\
\tilde{t}(p, (-\infty, \underline{s}] \cup [\bar{s}, \infty) a, \beta) &= p + \beta \frac{\sigma_a^2 \sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2 + \sigma_\theta^2)}} \frac{\phi\left(\frac{\bar{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}}\right) - \phi\left(\frac{\underline{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}}\right)}{\Phi\left(\frac{\bar{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}}\right) - \Phi\left(\frac{\underline{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}}\right)} \\
&\leq p - \frac{\sigma_a^2 \sigma_\theta^2}{\sqrt{\Sigma(\sigma_a^2 + \sigma_\theta^2)}} \frac{\underline{s} - \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2} a}{\sqrt{\frac{\Sigma}{\sigma_a^2 + \sigma_\theta^2}}} = t(p, \underline{s}, a).
\end{aligned}$$

First, note that because $\Phi_{\mu_\theta}(p), 1 - \Phi_{\mu_\theta}(p) > \xi$, $\xi < \frac{1}{2}$. Hence, there exist constants y, z so that $1 - \Phi_{\mu_\theta}(y) = 2\xi$ and $\Phi_{\mu_\theta}(z) = 2\xi$. For any value of $\bar{s}$ with $t(p, \bar{s}, a) < y$, there exists a unique $\underline{S}(\bar{s})$ so that $\Phi_{\mu_\theta}(t(p, \underline{S}(\bar{s}), a)) - \Phi_{\mu_\theta}(t(p, \bar{s}, a)) = 2\xi$ because $t(p, s)$ is strictly decreasing in s. Further, $\underline{S}(\bar{s})$ is strictly increasing in $\bar{s}$.

As $b(\bar{s},a) \to y$, $\underline{S}(\bar{s}) \to -\infty$ and $\tilde{t}(p, (-\infty, \underline{S}(\bar{s})] \cup [\overline{s}, \infty)a, \beta) \to \tilde{t}(p, [\overline{s}, \infty), a, \beta) > p$, so $\lim_{b(\bar{s},a)\to y} \Phi_{\mu_\theta}(t(p, \underline{S}(\bar{s}))) - \Phi_{\mu_\theta}(\tilde{t}(p, (-\infty, \underline{S}(\bar{s})] \cup [\overline{s}, \infty)a, \beta)) > 1 - \Phi_{\mu_\theta}(p) > \xi$. As $b(\bar{s},a) \to -\infty$, $b(\underline{S}(\bar{s}), a) \to z$ and $\tilde{t}(p, (-\infty, \underline{s}] \cup [\overline{s}, \infty)a, \beta)$ converges to a number greater than p, so

$$\begin{aligned} & \lim_{b(\bar{s},a)\to -\infty} \Phi_{\mu_\theta}(\tilde{t}(p, (-\infty, \underline{S}(\bar{s})] \cup [\overline{s}, \infty)a, \beta)) - \Phi_{\mu_\theta}(t(p, \bar{s})) > \Phi_{\mu_\theta}(p) > \xi \\ \Leftrightarrow \quad & \lim_{b(\bar{s},a)\to y} \Phi_{\mu_\theta}(t(p, \underline{S}(\bar{s}))) - \Phi_{\mu_\theta}(\tilde{t}(p, (-\infty, \underline{S}(\bar{s})] \cup [\overline{s}, \infty)a, \beta)) < \xi. \end{aligned}$$

Hence, by the intermediate value theorem and continuity of $\Phi_{\mu_\theta}(\tilde{t}(p, (-\infty, \underline{S}(\bar{s})] \cup [\overline{s}, \infty)a, \beta)) - \Phi_{\mu_\theta}(t(p, \bar{s}))$ in $\bar{s}$ for $t(p, \bar{s}, a) < y$, there exists a $\bar{s}$ satisfying (3) and (4).

Finally, we show uniqueness of equilibria as follows: Note that for sufficiently small β, for any WOM equilibrium we have that $\tilde{t}(p, \mathcal{S}_{a,p}, \beta)$ is close to p. It follows immediately, that with pessimistic or well-informed target customers negative WOM cannot arise in equilibrium and with optimistic or well-informed target consumers positive WOM cannot arise in equilibrium. At the same time, with pessimistic/uninformed (optimistic/uninfi) target customers, an early adopter with a very good (bad) experience always has an incentive to engage in WOM. Hence, the WOM equilibrium is unique for sufficiently small β.

A.5 Proof of Proposition 2

First, note that for sufficiently small β, the profit-maximizing price p^* converges to $\arg\max(1-\Phi_{\mu_\theta}(p))p$ and induces

1. a full WOM equilibrium if $\Phi_{\mu_\theta}(p^*), 1 - \Phi_{\mu_\theta}(p^*) > \xi$;
2. a no WOM equilibrium if $\Phi_{\mu_\theta}(p^*), 1 - \Phi_{\mu_\theta}(p^*) < \xi$;
3. a negative WOM equilibrium if $\Phi_{\mu_\theta}(p^*) < \xi < 1 - \Phi_{\mu_\theta}(p^*)$;
4. a positive WOM equilibrium if $\Phi_{\mu_\theta}(p^*) > \xi > 1 - \Phi_{\mu_\theta}(p^*)$.

Thus, there exists a $\bar{\beta}$ so that consumers are uninformed, well-informed, optimistic, pessimistic if 1., 2., 3., 4. are satisfied, respectively and so that the WOM equilibrium is unique by Lemma 2. Next, consider limiting properties of p^* and $\Phi_{\mu_\theta}(p^*)$, recalling that

$$\mu_\theta(a, b_i) \sim \mathcal{N}\left(\bar{\mu}(a, b_0), \sigma_b^2\right), \text{ where } \bar{\mu}(a, b_0) := b_0 + \tfrac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_a^2}\ a$$

and that p^* is the unique solution of the FOC:

$$\underbrace{\frac{1 - \Phi_{\mu_\theta}(p^*)}{\phi_{\mu_\theta}(p^*)}}_{\substack{\text{decreasing in } p^* \\ \text{increasing in } \sigma_b^2 \\ \text{increasing in } a}} = p^*.$$

Hence, p^* is increasing in σ_b^2 and a, so $\Phi_{\mu_\theta}(p^*)$ is increasing in σ_b and a. Further $\Phi_{\mu_\theta}(p^*) > \frac{1}{2}$ for sufficiently large σ_b^2 and a. Hence, there exists a $\bar{\xi}$ such that if costs $\xi \geq \bar{\xi}$, then there is only a positive WOM in the unique equilibrium. If $\xi < \bar{\xi}$, then the unique equilibrium is a full WOM equilibrium. Further, for large a we expect a positive WOM equilibrium, for intermediate a a full or no WOM equilibrium and for low a a negative WOM equilibrium.

B WEB APPENDIX

B.1 Experimental Evidence from Yelp.com

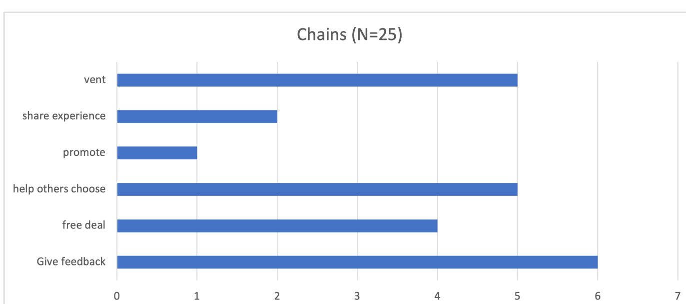
Here we summarize the two main findings from the survey for online restaurant portals. We find that the propensity to both read and write reviews is much higher for independent restaurants than chains (See A.7). Also we find that, among the people who decided to leave. a review, the top 2 motivations to write reviews are *giving feedback to the business* and *helping others make the right choice.* Among chain restaurants, venting is an important motive to review while in independent restaurants, reviewers care about promoting a new product(See Figures A.4 and A.5).

Table A.7: Likely to read/write reviews

	Read Reviews	Write Reviews
Chains	35%	19%
Independent	82%	46%

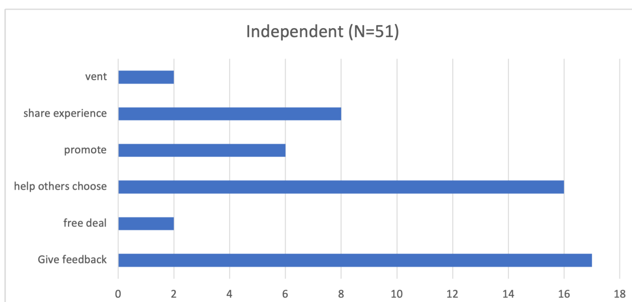
There are equal number of participants (approx 100) in both the chain as well as independent restaurant condition. The propensity to read as well as write reviews is higher for participants in the independent restaurant condition.

Figure A.4: Motivations to review Chains



B.2 Top National Chains (2017)

Figure A.5: Motivations to review Independent



We had also given an option of *other incentives* in the survey, but none of the participants used it.

Table A.8: Revenue, Satisfaction and Review Valence

Name	Stores (US)	Estd	Revenue (USD bn)	ACSI score	Brand (USD mn)	Star (Avg)	Negative WOM	Positive WOM	Neutral
Subway	25908	1965	11.3	80	18766	2.76	48%	38%	13%
McDonald's	14027	1955	37.64	69	126044	2.07	68%	20%	13%
Starbucks	13930	1971	17.65	78	44503	3.2	32%	50%	18%
Dunkin'	12538	1950	8.46	78	NA	2.6	54%	33%	13%
Pizza Hut	7522	1958	5.51	80	7372	2.19	66%	25%	9%
Burger King	7226	1953	9.65	76	6555	2.16	65%	20%	14%
Taco Bell	6446	1962	9.79	74	5213	2.64	53%	36%	12%
Wendy's	5769	1969	9.31	77	NA	2.29	62%	25%	13%
Domino's Pizza	5587	1960	5.93	79	7446	2.63	54%	37%	9%
KFC	4109	1952	NA	77	15131	1.78	77%	13%	9%
Arby's	3415	1964	3.63	79	NA	2.84	46%	40%	14%
Papa John's	3314	1984	1.78	80	NA	2.38	61%	30%	9%
Chipotle	2364	1993	4.48	79	4422	3.03	41%	46%	13%
Chick-fil-A	2100	1946	9	87	NA	3.74	23%	66%	11%

All data is for the year 2017. Negative WOM stands for share of negative reviews (1-2 star reviews), PWOM is the share of positive reviews (4-5 stars) and Neutral is share of 3-star reviews. The table shows that while there are some chains like McDonald's that have both lower ACSI scores as well as higher proportion of negative reviews, most other chains like Subway, Domino's Pizza, Papa John's and Pizza Hut have a very high proportion of negative reviews inspite of having good ACSI scores. The regression results in Table 6 remain similar if we exclude McDonald's. The chain dummy in that case is -0.91 and significant.

B.3 Dictionary for expectation and novelty words

Table A.9 is the dictionary we used to count occurrences of compare and novel words.

Table A.9: Custom Dictionaries

Expect	Novel	Employee
anticipate	curiosity	back office
belief	curious	bartender
brand	fresh	boy
change	innovative	desk
changed	learn	employee
consistent	new	front desk
contrary	novel	girl
declined	now	reception
deteriorate	offbeat	receptionist
exceed	recent	staff
expect	surprised	waiter
expectation	unique	waitress
expected	unusual	wait-staff
image	weird	
improve		
improved		
inconsistent		
met		
notion		
past		
prior		
recall		
remember		
reputation		
standard		
standards		
unexpected		
worsen		

B.3.1 Robustness: Verified Reviewers

It is widely known that fake reviews are common on many review platforms including Yelp.com (see e.g. Luca and Zervas (2016)). While the long-term negative impact of fake reviews seems to be limited (see He et al. (2020)), Mayzlin et al. (2014) document that, in the hotel industry, chains have a higher propensity of receiving fake negative reviews when the neighborhood includes more independent hotels.

Our dataset excludes all reviews that were identified by the Yelp filter to be fake, but the filter is likely not able to filter out all fake reviews. To show that the observed differences in valence cannot be completely attributed to chains receiving more fake negative reviews, we re-run our analysis with a subset of reviews written by Yelp-verified Elite Reviewers, which are guaranteed to be genuine. First note that Table A.10 shows that there are no major differences between Elites and Non-Elites in the type of restaurants they reviewed: Both groups review an almost equal proportion of chains and high-end restaurants (for the cuisines we are studying, i.e., “sandwich,” “pizza,” “burger,” “delis,” “coffee” etc. mentioned earlier). Elites, do write slightly more positive reviews (average Elite rating is 3.7 compared to 3.4 for Non Elites) and tend to review newer restaurants (average age of restaurant reviewed by Elites is 3.1 years whereas for Non Elites it is 3.6 years). Table A.11 summarizes the analogous results to Table 6 with only Elite reviews. In spite of the fact that Elites write more positive reviews, the reviews for chain restaurants are still 0.5 stars lower than comparable independent restaurants. Thus, our results are robust and remain qualitatively similar for a pool of verified reviewers.

Table A.10: Elite and Non Elite Reviewers

	Rating (Mean)	StoreAge (Mean)	Experience (Mean)	% Chains	% High-End
Elite	3.7	3.1	82	26.3%	10.7%
Non Elite	3.4	3.6	61	27.1%	10.4%

Table A.11: Impact of Chain Dummy and Brand on Star Ratings (For Elites)

	Dependent variable:			
	rev_stars			
	OLS			
	(1)	(2)	(3)	(4)
Chain Dummy	−0.539***	−0.538***		
	(0.012)	(0.012)		
Brand Age (Yrs)			−0.001**	−0.001**
			(0.0005)	(0.0005)
No of Stores			−0.00003***	−0.00003***
			(0.00000)	(0.00000)
Age of Store (Yrs)	0.011***	0.005***	0.008***	0.002
	(0.001)	(0.001)	(0.001)	(0.001)
Price Range				
$$	−0.107***	−0.108***	−0.106***	−0.107***
	(0.010)	(0.010)	(0.010)	(0.010)
$$$	0.147***	0.158***	0.185***	0.198***
	(0.021)	(0.021)	(0.021)	(0.021)
$$$$	0.334***	0.356***	0.373***	0.396***
	(0.024)	(0.024)	(0.024)	(0.024)
Observations	90,588	90,588	90,588	90,588
R^2	0.064	0.069	0.073	0.077
User Characteristics	*N*	*Y*	*N*	*Y*
Note:				*p<0.1; **p<0.05; ***p<0.01