The Recommendation Effect of Niche Products - How Consumer Learning Leads to Differentiation*

Maximilian Conze† Michael Kramm‡

October 27, 2016

Abstract

The recommendation effect introduces a new rationale for product differentiation other than the usual motivation to reduce price competition. We introduce consumer learning in a version of Hotelling's model (1929) of spatial competition with sequential consumer purchases and a second dimension of variation, quality, about which the consumers have differential information.

With consumer learning, firms are confronted with two offsetting effects: differentiation decreases the likelihood that a product is bought in earlier periods, but, by making inference more valuable, it also increases the likelihood that later consumers buy the differentiated good. We show that there exists a unique equilibrium in which the second effect dominates, so that the market incumbent locates in the center of the market, while the entrant differentiates by producing an ex-ante niche product.

Due to consumer learning, uninformed consumers are unambiguously better off in the equilibrium with differentiation than in the equilibrium of minimum differentiation which occurs without consumer learning. Informed consumers are better off in the latter equilibrium, so that the overall effect on consumer welfare depends on the parameters and we can show that in some cases transparency enhancing policies are welfare decreasing.

Keywords: Product Differentiation, Social Learning, Herding, Bayesian Observational Learning, Hotelling, Spatial Competition, Principle of Minimum Differentiation, Diversity

JEL codes: L13, L15, D83

---

*We would like to thank (in alphabetical order) Rodolphe Dos Santos Ferreira, Jörg Franke, Dennis Gärtner, Eugen Kovàč, Wolfgang Leininger, Lars Metzger and Ignacio Monzon for valuable comments and suggestions. This paper also benefited from discussions and remarks at the World Congress of the Game Theory Society GAMES 2016, the European Economic Association Annual Conference 2015, the Royal Economic Society Annual Conference 2014, the International Conference on Game Theory in Stony Brook 2014, the Spain-Italy-Netherlands Meeting on Game Theory 2014, and the Augustin-Cournot Doctoral Days 2015.

†University of St. Gallen, Email: maximilian.conze@unisg.ch

‡Corresponding Author, Ruhr Graduate School in Economics and Technical University of Dortmund, Email: michael.kramm@tu-dortmund.de

1 Introduction

> *10th rule of building a successful business: Swim upstream. Go the other way. Ignore the conventional wisdom. If everybody is doing it one way, there's a good chance you can find your niche by going exactly in the opposite direction. (Sam Walton, founder of Walmart)*[1]

Before a firm introduces a new product in a market it typically has to consider how it will design its product with respect to the products already offered by its competitors and with respect to consumer taste: shall it produce more of a mainstream product or shall it occupy a niche in the market, i.e. offer a product differentiated from its competitors and preferred ex-ante only by a minority of the consumers?

This question seems to gain even more importance as firms nowadays attribute more and more attention to the behavior of early adopters (agents consuming in earlier periods), as these find a growing number of opportunities to publicly announce their choice behavior using Internet platforms such as Yelp or the recommendation opportunities on the online market place Amazon for instance. Platforms like foursquare, Google and Facebook explicitly keep track of 'check-ins' in restaurants, bars and many other venues to provide this data online for undecided consumers. These sources of information obviously influence the choice behavior of laggards (agents consuming in later periods).

Such effects are not considered in earlier research dealing with the incentives to offer differentiated products. It has instead focused on the fact that price competition may yield incentives to differentiate one's product: by offering a differentiated product, firms are able to set a price above marginal costs and thus obtain a positive profit, although they serve a smaller market share. In those situations the competition effect, i.e. the ability to raise prices because of the 'local monopoly power' obtained by offering a differentiated product, dominates the market size effect, i.e. the possibility to serve a larger market share when offering a product similar to that of the competitor (see e.g. d'Aspremont et al., 1979). It is also well known that under different specifications of the model this dominance is reversed and firms offer non-differentiated products as the Principle of Minimum Differentiation suggests

[1] See `http://corporate.walmart.com/our-story/history/10-rules-for-building-a-business` (last accessed: December 4, 2015).

(see Hotelling, 1929).

We establish a fundamentally different rationale for offering differentiated products. The effect driving our result is one of informational nature and arises due to the possibility of consumer learning. In our model, laggards can observe the behavior of the early adopters. When adding the dimensions *time* and *quality*, about the latter of which the consumers have different knowledge, into a model of (spatial) product differentiation, the choice behavior of early adopters contains information influencing the choice behavior of laggards. A firm can influence and exploit consumer learning using its location choice (mainstream vs. niche). This may yield incentives to offer a differentiated (niche) product, as from a laggard's perspective, a purchases of a niche product by an early adopter is more likely based on its high quality than on a good match of consumer taste and product characteristic. We call this the 'recommendation effect'.

In Hotelling's original model, two firms compete on a linear and bounded market. Firms first choose their locations in this market and then set their prices. Firms' locations can be understood as representing their products' characteristics. Consumers' tastes (or types) are uniformly distributed over the market. Hotelling argued that in equilibrium both firms locate at the center and set the same price. It has later been shown by D'Aspremont et al. (1979) that the celebrated Principle of Minimum Differentiation is only valid when prices are exogenously fixed and equal; we will refer to this setup as Hotelling's 'pure spatial competition model'.[2]

We modify the pure spatial competition model in several aspects. Assuming sequential location choice of the firms constitutes the first departure. The second and more important change is the introduction of a second dimension of differentiation, which we interpret as quality although other aspects of a good fit this model aspect, too. In general this can be any characteristic of a good, where all consumers agree on the ranking of its different manifestations. Firms thus offer products that are vertically and (potentially) horizontally differentiated. Consumers are either completely informed or uninformed about the goods' qualities, which is the third modification. Without additional information, uninformed con-

[2] With fixed prices the firms' goals narrow down to serving the largest possible market share. To this end, given the other firm's position, a firm locates on the longer side of the market as close as possible to the other firm. Given that firms are located directly next to each other, the only situation without any incentive to relocate is the one where both firms are located at the center.

sumers perceive the quality of both goods to be the same ex ante. The fourth modification is that consumers purchase sequentially. Finally, consumers observe previous choice behavior. Our model thus introduces an aspect similar to the herding literature in that consumers might base their decisions on observable actions of others which may lead to 'wrong' decisions. Since we are interested in the effect of the possibility to learn from other consumers' purchases, we also present a benchmark model where consumer learning is not possible and we demonstrate that without consumer learning there would be no product differentiation in our model.

If laggards are uninformed, they use the observed choices of early adopters to infer information about the goods' qualities. This updating process crucially depends on the firms' locations. For an uninformed early adopter, both products have the same expected quality so she puts more weight on the match between her taste and the firms' locations than an informed consumer. To build intuition suppose that both firms are located at the center and consider the incentives to deviate to another location. A firm which moves away from the center decreases the probability of being chosen by uninformed early adopters, while it increases the probability to be chosen by uninformed laggards. This is because, whenever the deviating firm in this setup is chosen in the first period, it is more likely that this purchase was made by an informed consumer because of superior quality. Hence, uninformed laggards will now tend to follow the early adopters' decisions more often, this is the aforementioned 'recommendation effect'. If the additional expected demand from laggards outweighs the lost expected demand from early adopters, the total expected demand is increased by moving away from the opponent located at the center.

In the model at hand we assume that prices are exogenous.[3] Usually prices are considered flexible and are seen as an endogenously chosen component of the market competitors' strategies, but there are nonetheless examples where the assumption of fixed prices seems plausible. This can be the case either if prices are actually fixed or if price differences among products are perceived as too small to influence the consumption decision (see for example Courty, 2000) or in markets where prices are chosen in a long term perspective.[4] Generally, the exogeneity of prices can be seen as an abstraction of the real world in the sense that we

[3] In a companion paper (Conze and Kramm, 2016) we show that similar effects also arise in a setup of non-spatial differentiation with endogenous prices.

[4] We will discuss the fixed-price assumption in more detail further below.

consider markets which are characterized better by competition in market shares than by competition in prices.

Consider the movie industry where the entrance fees for blockbusters of the same length at cinemas are usually the same.[5] A recent event in this industry very well fits our model. The movie 'The Artist', which aired in cinemas in 2011, was a major success of that year and, in addition to receiving mainly positive critique, it won numerous prizes, including five Oscars.[6] It brought in almost $133.5M worldwide, while being produced with a $15M budget.[7] So on both counts - artistically and economically - it was a major success. What makes this movie especially interesting for our case, is that, compared to the advanced techniques commonly used in the movie industry nowadays with its 3D-effects and 'Dolby Surround', the means used for the shooting of 'The Artist' were rather unconventional: it was entirely shot in black-and-white and mainly abstracted from dialogues, almost making it a silent movie. Thus, one can say that, compared to the other blockbusters at that time, this movie was more of a 'niche product'. Yet, it may well be that the high popularity of this unconventional movie among the early adopters in the first weeks of broadcasting induced the laggards to attribute the reason for that choice behavior to the high cinematographic quality of 'The Artist'. Probably the producers anticipated just this reasoning and therefore decided to dive into this unorthodox project. Indeed, the director of 'The Artist', Michel Hazanavicius, said that

[5] See for instance Orbach and Einav (2007). Many people arguably decide on which movie to watch before seeing the prices. Furthermore, they most probably do not revise their decision when finding out that prices are slightly different than expected. De Vany (2006) discusses the three different pricing levels of the movie industry (producers, distributors, box offices) extensively and shows that empirically box office prices are fixed - which indeed is an economic puzzle. Additionally, it is shown that the producers obtain a contractually regulated share of the revenues generated by the box offices. This implies that the only way producers can influence their revenue is by generating a larger audience.

Another research field in which exogenous prices are frequently assumed is health economics, as medical treatment is reimbursed to consumers by their health insurance. Several research projects in this area use models of spatial competition with fixed prices and discuss differences in quality among hospitals, see e.g. Brekke et al. (2006), Brekke et al. (2011) and Gravelle and Sivey (2010).

[6] See `http://articles.economictimes.indiatimes.com/2012-02-27/news/31104573_1_oscars-foreign-language-category-actor-race` and `http://www.theguardian.com/film/2011/dec/08/artist-silent-film-michel-hazanavicius` (last accessed: December 4, 2015).

[7] See `http://www.boxofficemojo.com/movies/?id=artist.htm` and `http://www.imdb.com/title/tt1655442/business?ref_=ttrel_ql_4` (last accessed: December 4, 2015).

when he presented his idea "[he'd] only get an amused reaction - no one took this seriously".[8]

Note that we use the term 'niche product' in the sense that such a product is of relatively low appeal to uninformed consumers. As our model shows, and the example of the movie 'The Artist' illustrates, a niche product according to this definition can still generate a larger demand than a mainstream product ex-post. In our model this is due to the recommendation effect, which generates a higher demand by laggards for the 'niche product' than for the mass product.

Our contribution to the literature and the main goal of this paper is to show in a theoretical model how the firms' incentives to differentiate are affected by social learning among heterogeneous consumers.

The paper at hand is structured as follows. In Section 2 we review the related literature. Section 3 presents a short and simplified example, while the full model is introduced in Section 4. A benchmark model and the main model are solved in Sections 5 and 6 respectively, welfare comparisons are made in Section 6.5. Finally, Section 7 concludes. The proofs are relegated to the Appendix.

2 Literature

In his seminal paper on *spatial competition* and product differentiation, Hotelling (1929) proposed that, when choosing locations (which can be interpreted as representing the product's characteristics) on a linear bounded market - where consumers are uniformly distributed - before setting prices, firms choose the same location, namely the center, and set the same price in equilibrium. D'Aspremont et al. (1979, p. 1145) argued that this "so-called Principle of Minimum Differentiation [...] is invalid", by showing that whenever the distance between firms' locations is small, they have an incentive to slightly undercut the rival's price. As in any model of spatial competition with endogenous prices, there exist two offsetting effects in Hotelling's setup. On the one hand, firms have an incentive to increase the distance, thereby relaxing competition ('competition effect'). On the other hand, decreasing the distance allows

[8] See page 5 of the official press kit at: `http://www.festival-cannes.com/assets/Image/Direct/041859.PDF` (last accessed: 2015.12.04). Additionally, the success of the movie was called 'surprising' by the media, see e.g. `http://www.theguardian.com/film/2012/feb/04/hollywood-nostalgia-chaplin-valentino` (last accessed: 2015.12.04).

to serve a larger share of the market ('market size effect').[9]

In contrast to most models of spatial competition, we find asymmetric pure strategy equilibria, for both cases - where firms differentiate and where they do not. Tabuchi and Thisse (1995) assume a non-uniform distribution of consumers, sequential location choice and simultaneous price setting and also find asymmetric pure strategy equilibria. However, differing from their results, serving a smaller ex-ante market share - that is being a niche producer - need not be disadvantageous in our model, i.e. there is no second-mover disadvantage as it happens to be the case in Tabuchi and Thisse (1995).

Among others,[10] Economides (1989) combines *price setting, horizontal and vertical differentiation.* In both versions of his model - price competition followed by quality choice, or both choices of these strategic variables happening simultaneously (location choice occurs at the first stage in both versions) - maximum horizontal differentiation and minimal differentiation in quality and prices is obtained in equilibrium. Bester (1998) differs from Economides (1989) in assuming quadratic instead of linear transport costs (which strengthens incentives to differentiate) and, more importantly, in the consumer's imperfect knowledge about qualities. He shows that this imperfect knowledge mitigates product differentiation: as consumers associate low prices with a low quality, there is an endogenous lower bound to prices. Thus, price competition is already relaxed, making it less necessary to horizontally differentiate in order to decrease price competition.

In the literature on *social learning*, Bikhchandani et al. (1992) and Banerjee (1992) are the first to examine the phenomena of *information cascades* and *herding.* They show that with sequential consumer choice, rational Bayesian inference from the previous behavior of others may guide consumers to ignore their own (imperfect) private signal on the quality of a firm; a behavior which in the end may result in herding, driving all subsequent consumers to buy only from one firm. Smith and Sørensen (2000) deliver the most complete analysis of this setup of social learning.

Ridley (2008) combines the ideas of *Hotelling and herding.* Nevertheless, his research question is fundamentally different to ours: he models two firms with different information levels about market demand and - as they sequentially decide about entering the market -

[9] See for example Belleflamme and Peitz (2010, Chapter 5.2).

[10] See e.g. Gabszewicz and Thisse (1986), Dos Santos Ferreira and Thisse (1996) or Gabszewicz and Wauthy (2011).

the second mover can possibly deduce information from the other firm's decision.

Our model is also related to the recent literature on *Bayesian persuasion* as introduced by Kamenica and Gentzkow (2011): they analyze in which way the sender (in our case the firms) can influence the updating of the receiver (in our case the laggards) in their favor by choosing the 'sender-optimal' information structure of the signal. As in our model, the sender does not know the realization of the state (in our case the product's quality), when taking his decision. In the analysis at hand the early adopters are so to speak 'used' by the firm to signal their quality and differentiation in some cases is advantageous for the market entering firm, as via the Bayesian updating mechanism producing a niche product distorts the signal (i.e. the purchase decision of the early adopter) in its favor.

The strand of literature that is closest to our approach has taken a look at the *impact of consumers' social learning on competition among firms* producing horizontally and vertically differentiated products. In Caminal and Vives (1996) two firms compete for homogeneous consumers by setting prices. The authors formulate two models, and in one of these firms do not know the quality of their product, just as it is the case in our paper. Consumers have different information about the products' qualities and observe the history only partially. Given the incomplete observation of the history, consumers are led to believe that a good is of higher quality whenever its market share is high. The authors show that this leads to a strategic incentive for the firms to generate a higher demand in early periods by setting a low price. However, Caminal and Vives do not analyze incentives to differentiate.

Miklos-Thal and Zhang (2013) model a *monopolistic market with consumer learning* showing that "demarketing [i.e. visibly toning down the marketing efforts] lowers expected sales ex ante but improves the product quality image ex post, as consumers attribute good sales to superior quality" (Miklos-Thal and Zhang, 2013, p. 55). In the same vein, Parakhonyak and Vikander (2016) show that a monopolist may have an incentive to reduce its capacity in order to make laggards infer a high product quality from sold out capacities in earlier periods and thus induce herding in its favor.

Tucker and Zhang (2011)[11] show in an *empirical* paper that - in line with the intuition of our theoretical results - popularity information (indicated by the choice of previous con-

[11] In a working paper version, they also include a theoretical model in which location, however, is given exogenously.

sumers) is especially beneficial for niche products, because for the same popularity, niche products are more likely to be of superior quality than mainstream products.

Another empirical paper is even more suitable for our analysis - and especially our example on the movie industry with 'The Artist': Moretti (2011) is among the first researchers to empirically analyze real world data on social learning. He investigates in how far it influences movie sales. The results show that social learning indeed matters and that 'surprise' in the early demand increases later demand for a movie.[12] That is, if a movie was seen by surprisingly many consumers (compared to the prior) in the first weeks of airing, this will have the additional (indirect) effect of a social multiplier: while it immediately increases profits to the cinemas, it also generates a higher demand in the following periods. We can infer that this yields an incentive for movie producers to create 'surprising' movies in the sense that they are very successful in the first weeks compared to the expectations. This may just be the reason to produce a black and white silent movie nowadays.

3 Illustrative Example With Discrete Strategy Spaces

Before presenting the full model, we demonstrate the effects at force in a small illustrative example with discrete action spaces.

There are two firms producing an ex-ante homogeneous good: firm A produces a good with a deterministic value of $v = 20$€, and firm B produces a good of value $v_B = 30$€ with probability 0.5, and of value $v_B = 10$€ with probability 0.5. Neither firm knows the realized value of firm B's product. The price of all goods is 5€. Firms A and B sequentially choose their locations a and b along a road at one of three locations: kilometer 0, 0.5 or 1. Firm A is the first mover, and as $a = 0$ is equivalent $a = 1$, we restrict firm A to the right part of the interval. There are two consumers (each independently located at $0, 0.5$ or 1 with probability 1/3): an early adopter who with probability 0.5 is either completely informed or uninformed, and a laggard, who is completely uninformed about product B's value and the early adopter's location, but observes the choice behavior of the early adopter. A consumer has to pay 4.5€

[12] While Sorensen (2007) and Chen (2008) support the social learning argumentation in another setting, Gilchrist and Sands (2016) attribute the fact, that a positive 'shock' to early demand for cinema movies (e.g. by bad weather) increases demand in later periods, to network externalities, that is, 'people have something to talk about'.

for traveling the 0.5 km to the next location. Consumers maximize their expected utility and the firms compete over the two consumers by their location choice.

Let β denote the consumer's belief that firm B's product is superior. The expected difference in value between B's and A's product is then given by

$$E[v_B - v_A] = \beta \cdot 30€ + (1-\beta) \cdot 10€ - 20€ = (2\beta - 1) \cdot 10€. \tag{1}$$

Consumers compare this difference to the different transport costs between the firms. An uninformed early adopter perceives the goods' values to be the same ($\beta = 0.5$), and so chooses the closer firm. We assume that if firms locate at the same position, an uninformed early adopter at the same position chooses each firm with probability 0.5 and chooses B (A) if he is located left (right) of the firms. Obviously, there are several other tie breaking rules that are also compatible with rationality of consumers, but any other tie breaking rule would still lead to the result of one firm not positioning at the center of the market, which is a result driven solely by the recommendation effect described later. The tie breaking rule at hand is the discrete analogue of the one we employ in our model with a continuous action space. In this example the tie breaking rules out equilibria arising only due to the discrete action space.

For an informed consumer, i.e. $\beta \in \{0, 1\}$, the sure gain of buying the superior product (10€) is always higher than possible transport costs (9€ at the maximum), so she always buys at the better firm, no matter where she is located.

The belief of the laggard is of more interest, since she uses Bayes' rule to calculate the probability of each firm offering the superior product as follows. If firm B was chosen in the first period it is given by

$$\beta_B^u := Pr(v_B = 30 \mid C_1 = B) = \frac{Pr(C_1 = B \mid v_B = 30)}{Pr(C_1 = B)} \cdot Pr(v_B = 30).$$

The probability that B is superior, given A was bought in the first period β_A^u is derived analogously as

$$\beta_A^u := Pr(v_B = 30 \mid C_1 = A) = \frac{Pr(C_1 = A \mid v_B = 30)}{Pr(C_1 = A)} \cdot Pr(v_B = 30).$$

In general, these probabilities depend on the firms' locations, and because of the possibility that the first period consumer was informed, the probability that a product is bought is always higher if it is superior, so that updating is informative and $\beta_B^u > 0.5$ (and $\beta_A^u < 0.5$). When

firms are located at the same spot, every consumer has to incur the same transport costs for both firms, and so the laggard always follows the decision of the early adopter. If firms are not located at the same spot, consumers compare the expected additional value of the goods, as stated in equation (1) to the additional transport costs. For all symmetric positions of the two firms, the choice probabilities and thus the beliefs are the same and calculated as $\beta_B^u = 0.75 = 1 - \beta_A^u$.

If however $b = 0$ and $a = 0.5$, i.e. the firms locations are asymmetric, $\beta_B^u = \frac{0.5 \cdot 1 + 0.5 \cdot 1/3}{0.5 \cdot 0.5 + 0.5 \cdot 1/3} \cdot 0.5 = 4/5$ and $\beta_A^u = \frac{0.5 \cdot 0 + 0.5 \cdot 2/3}{0.5 \cdot 0.5 + 0.5 \cdot 2/3} \cdot 0.5 = 2/7$, so that in this case $\beta_B^u = 4/5 > 1 - \beta_A^u = 5/7$. We have seen that a choice of B in period 1 increases the laggards confidence in this product more than a choice of product A, in particular a laggard is willing to travel an additional distance of 0.5 to obtain product B instead of A if B was chosen in period 1, but she is unwilling to travel the same distance to buy A instead of B if A was chosen. This is easily seen by using (1) to compare the expected additional valuation to the additional transport costs:

$$(2\beta_B^u - 1) \cdot 10€ = 7€ > 4.5€ > (2 \cdot (1 - \beta_A^u) - 1) \cdot 10€ \approx 4.29€$$

All other beliefs can directly be obtained using these calculations because of the symmetry of the model. The induced asymmetry in the beliefs and the consequences for the behavior of the consumer are the driving effects for the results to follow.

The beliefs and the resulting behavior of the laggard partition the action space of the firms as visualized in Figure 1. In situations labeled $\mathfrak{D}^2$ firms locate at the maximal distance from each other, and it will never be the case that all types of laggards follow the behavior of the early adopter, as $(2\beta_B^u - 1) \cdot 10€ = 7€ < 9€$. In situations labeled $\mathfrak{D}^4$ firms locate at the same position and a laggard always follows the behavior of the early adopter.

In situation $\mathfrak{D}^{3B}$ ($\mathfrak{D}^{3A}$) the firms positions are different and asymmetric, furthermore B (A) is the niche firm here, thus the notation. As shown above, the laggards behavior now depends on the history: If, for instance, B is the niche firm, i.e. $(a, b) \in \mathfrak{D}_L^{3B}$, all consumers located at $x = 0.5$ will consume at firm B after observing it was chosen in the first period.[13] The same holds true for consumers located at $x = 1$, as they have the same (additional) travel costs as the consumers located at $x = 0.5$, once they traveled to location 0.5. Consumers

[13] Note, that we use subscripts L and R to indicate whether firm B is positioned left or right of firm A.

located at $x = 0$ will still consume at firm B, even if they observed $C_1 = A$. Consumer updating is beneficial for firm B, as it is the niche firm. The reverse is true in situations labeled with $\mathfrak{D}^{3A}$.

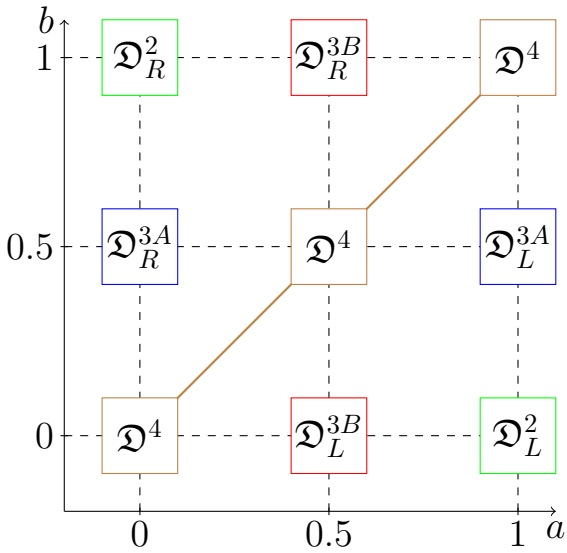


Figure 1: Action space for the discrete example.

Let $D(a, b)$ denote B's demand if locations are (a, b). Firm A serves the remainder of the market, so its demand is $\tilde{D}(a, b) = 2 - D(a, b)$. To obtain the equilibrium we need to calculate $D(0.5, 0) = \tilde{D}(1, 0.5)$, $D(1, 1) = D(0, 0)$ and $D(0.5, 0.5) = D(1, 0)$. It is easy to see that both firms split the market equally in the latter cases and the resulting demand equals $D(0.5, 0.5) = 1$. Additionally, we have

$$D(0.5, 0) = \underbrace{Pr(C_1 = B)}_{\text{1st period}} + \underbrace{P(C_1 = B) \cdot 1 + P(C_1 = A) \cdot 1/3}_{\text{2nd period}} = 5/12 + 5/12 + 7/36 = 37/36$$

$$D(1, 1) = 1/2 \cdot 1/2 + 1/2 \cdot 5/6 + P(C_1 = B) \cdot 1 + P(C_1 = A) \cdot 0 = 2/3 + 2/3 \cdot 1 = 1.33.$$

This implies that firm B's best response to a, $b^*(a)$, is given by $b^*(1) = 1$ and $b^*(0.5) = 0$. Firm A then chooses its best point of B's best response function, and will locate at $a^* = 0.5$, so that resulting equilibrium locations are $a^* = 0.5, b^* = 0$, i.e. an equilibrium with differentiation.[14] Without consumer learning, both firms would locate at the market center, i.e. in an equilibrium with symmetric minimum differentiation.

[14] Obviously, the same reasoning applies for situations indexed with R in the partition of the action space (i.e. $b \geq a$), and thus another equilibrium with differentiation is given by $(a^*, b^*) = (0.5, 1)$.

4 Model Setup

The following describes the model setup, which is an extension of the aforementioned pure spatial competition model and generalizes the results obtained in the example with a discrete action space.

Firms The model has two firms A and B (both: 'it') which produce (potentially) differentiated goods. Both firms produce at zero costs, and the retail price is regulated and set to $p > 0$. The firms' locations describing their products' characteristics are confined to the unit interval and are denoted by a and b for firm A and B respectively, so that $a \in \mathcal{A} := [0, 1]$, and $b \in \mathcal{B} := [0, 1]$. The location choice of the firms occurs sequentially, with firm A choosing its location first, and firm B following. The firm that is closer to any of boundaries of the interval $[0, 1]$ will be called the niche firm. Thus, B is the niche firm whenever $b < a$ and $b < 1 - a$, or $b > a$ and $b > 1 - a$

A firm's profit simply is the number of consumers served, multiplied with the regulated price p. We make the usual assumption of firms being risk neutral. It should be noted that discounting future profits would not alter the results qualitatively, and it is thus left out for simplicity.

Besides the horizontal differentiation as measured by the firms' locations, the goods are also of different 'value' to consumers. This value can be thought of as representing a good's quality. There is uncertainty about the quality differential between the firms' products, which is randomly determined after the firms have chosen their locations: the value of firm A's product is common knowledge and given by $v_A = v > 0$, while the second firm's quality v_B is either $v_B = v + \delta$ or $v_B = v - \delta$ with $\delta > 0$, both of which occur with probability 0.5, i.e. there are two different states of nature.[15] The realized value of v_B is unknown to both firms. Thus, producers possess the following information when choosing their location: firm A has no information, so its information set is given by $\mathcal{I}_A = \{\emptyset\}$, and firm B knows $\mathcal{I}_B = \{a\}$.

Firms play pure strategies. The strategy of firm A is the choice of its location a, while the strategy of the second mover B maps the location a of its competitor into its own location, i.e. $b(a) : \mathcal{A} \to \mathcal{B}$.

[15] This modeling of the supply side can be considered as the first mover A being the market incumbent with a known quality, while firm B with an unknown quality is a new entrant to the market.

While it is usually not without loss of generality to assume $b \leq a$, all cases with $b > a$ can be described with the formulas derived for $b \leq a$, so that in the following we focus on this case and specifically analyze situations with $b > a$ only when necessary.

Consumers On the other side of the market, there are two consumers (both: 'she') with heterogeneous preferences, who sequentially make their purchase decisions in periods $t = 1$ and $t = 2$. Consumers are exogenously sorted into being a laggard or an early adopter. Each consumer buys at most one good and will be referred to by the period she has the opportunity to make a purchase (an early adopter in period $t = 1$ and a laggard in period $t = 2$).

Heterogeneity is modeled by assuming that each consumer t is described by a location on the unit interval. In every period, the location of consumer t, x_t is independently drawn from a uniform distribution on $\mathcal{X} := [0, 1]$. It measures consumer t's preference towards a good of a firm $F \in \{A, B\}$ located at $f \in [0, 1]$. The closer the location of the consumer to the firm where she buys (holding everything else constant), the higher is the resulting utility. The ex-post utility of a consumer located at x when buying the product from firm F located at f is given by

$$u(x, F) = v_F - p - \tau|x - f|,$$

where the last term with $\tau > 0$ captures the transport costs. We normalize utility to zero for the case in which a consumer does not buy any of the two goods. Note that with v_B being stochastic this Bernoulli utility function implies risk-neutrality in money. As long as preferences are quasilinear and the ex-ante expected utility of both products (gross of transportation costs) is the same, the results would not change if consumers were risk-averse.

While it is generally possible that a consumer abstains from buying, we will make the following assumption for convenience:

Assumption 1. *Every consumer prefers to buy one good to not buying any good, i.e.* $v - \delta > p + \tau$.

In addition to the heterogeneous preferences, consumers differ in their expertise $\phi \in \{u, i\}$ about firm B's product. Informed consumers ($\phi = i$) observe the realization of v_B, whereas uninformed consumers ($\phi = u$) only have the prior information that $v_B = v + \delta$ or $v_B = v - \delta$, each with probability 0.5. In each period, the consumer is informed with probability $q \in (0, 1)$

and uninformed with probability $(1-q)$.[16] A consumer's expertise is independent of her location x and of the expertise of the other consumer. A consumer's type in period t is thus given by (x_t, ϕ_t).

In the second period, the laggard observes the action taken by the early adopter, but neither the early adopter's location nor whether she was informed. Formally, let $C_0 = \emptyset$ and let $C_1 \in \{A, B\}$ be the choice of an early adopter, then the information set of an uninformed consumer is given by $\mathcal{I}_t^u = \{a, b, x_t, v_A, C_{t-1}\}$ and that of an informed consumer by $\mathcal{I}_t^i = \mathcal{I}_t^u \cup \{v_B\}$, for $t = 1, 2$.

Consumers form beliefs β about the probability of firm B offering the product of higher quality by mapping the available information $\mathcal{I}$ into the probability space, $\beta := Pr(v_B > v|\mathcal{I}) \in [0, 1]$. The belief of an uninformed consumer is the function $\beta^u : \mathcal{A} \times \mathcal{B} \times \{\emptyset, A, B\} \to [0, 1]$, and that of an informed consumer is the function $\beta^i : \mathcal{A} \times \mathcal{B} \times \{\emptyset, A, B\} \times \{v-\delta, v+\delta\} \to [0, 1]$. Note that we assume that a consumer's location does not influence her belief.

The expected utility of a consumer with location type x and belief β is given by $u(x, \beta, a, b|A) = v - p - \tau|a - x|$, if she buys from firm A, and by $E[u(x, \beta, a, b|B)] = v + (2\beta - 1)\delta - p - \tau|x - b|$, if she buys from B. Clearly, the expected utilities depend on a consumer's belief β and location x. For any consumer type $x \in (b, a)$, the expected utility from B's (A's) product decreases (increases) in x. For all types x that are not located between the firms, changing x affects both expected utilities in exactly the same way, so that the difference of expected utilities is constant. The reason for this is that for all these consumers the difference in distances to the two firms is the same and so is the difference in transportation costs between the firms. This means that all consumers with the same belief β located left of b or right of a must have the same preferences, i.e. prefer the same firm or are indifferent.

These observations, imply that whenever there exists an unique indifferent consumer type, it must be located in the interval (b, a). A consumer located at x holding belief β is indifferent between the products of A and B if

$$E[u(x, \beta, a, b|B)] - u(x, \beta, a, b|A) = (2\beta - 1)\delta - \tau(|x - b| - |a - x|) = 0. \tag{2}$$

[16] In contrast to the example from above with a discrete action space, we allow for informed laggards in the second period in the model with continuous action spaces. This minor modeling difference in the example was merely introduced to simplify calculations.

Let us define

$$\bar{x}(\beta) := \frac{a+b}{2} + \frac{\delta}{\tau}\Big(\beta - \frac{1}{2}\Big), \tag{3}$$

which, for a given belief β, coincides with the consumer type x solving equation (2) whenever (2) has a solution $x \in (b, a)$. In this case there exists an unique indifferent consumer type $x \in (b, a)$ and it is given by $\bar{x}(\beta)$. Then, all consumers with the same belief β and a location left (right) of $\bar{x}(\beta)$ must prefer B (A).

If $\bar{x}(\beta) = b$ ($\bar{x}(\beta) = a$) the consumer with belief β and type $x = b$ ($x = a$) is indifferent between both products and so are all types with the same belief located left of b (right of a).

$\bar{x}(\beta) < b$ means that the consumer with belief β and located at $x = b$ prefers good A over B and thus this must be true for all consumer located left of $x = b$ and in fact for any consumer type with belief β. Similarly, if $\bar{x}(\beta) > a$, all consumer types x with belief β prefer good A.

We impose the following assumption on the behavior of indifferent consumers:

Assumption 2. *If $a \neq b$, then all indifferent consumer types buy at the firm that is located closer to them. If $a = b \geq 1/2$ ($a = b < 1/2$), then types $x \leq b$ ($x \geq b$) of indifferent consumers buy from firm B and the remaining indifferent consumers buy from firm A.*

With this assumption and with $b \leq a$ the above observations imply that there must be one highest consumer type for a given belief that purchases from firm B. We denote by $\tilde{x}(\beta)$ the threshold, such that all consumers with location $x \leq \tilde{x}(\beta)$ and belief β choose the product of firm B. Threshold $\tilde{x}(\beta)$ equals the indifferent consumer type in (b, a) with belief β whenever it exists. Otherwise, $\tilde{x}(\beta) = 0$, if all consumers with belief β prefer A, and in the analogous case, when all consumers with β prefer to buy from B, $\tilde{x}(\beta) = 1$. Thus, the threshold type can be calculated as

$$\tilde{x}(\beta) = \begin{cases} 0 & \text{if} \quad \beta < \frac{1}{2} - \frac{\tau}{\delta} \cdot \frac{a-b}{2} & \Leftrightarrow \bar{x}(\beta) < b, \\ 1 & \text{if} \quad \beta > \frac{1}{2} + \frac{\tau}{\delta} \cdot \frac{a-b}{2} & \Leftrightarrow \bar{x}(\beta) > a, \\ \bar{x}(\beta) & \text{if} \quad \beta \in \left[\frac{1}{2} - \frac{\tau}{\delta} \cdot \frac{a-b}{2}, \frac{1}{2} + \frac{\tau}{\delta} \cdot \frac{a-b}{2}\right] & \Leftrightarrow \bar{x}(\beta) \in [b, a]. \end{cases} \tag{4}$$

As consumers are uniformly distributed over $[0, 1]$, $\tilde{x}(\beta)$ is constructed such that it equals the probability of a consumer with belief β buying product B .

The strategy of a consumer is a mapping $C_t : \mathcal{A} \times \mathcal{B} \times \mathcal{X} \times [0,1] \to \{A, B\}$ from public and her private information into a purchase decision, where $C(a, b, x, \beta)$ is the choice of a consumer with location $x \in \mathcal{X}$ and belief $\beta \in [0,1]$ given $a \in \mathcal{A}$ and $b \in \mathcal{B}$. The consumer's optimal strategy with location x and belief β is always characterized by

$$C(x, \beta) = \begin{cases} B & \text{if} \quad x \leq \tilde{x}(\beta), \\ A & \text{if} \quad x > \tilde{x}(\beta). \end{cases}$$

Thus, to obtain the optimal consumer behavior we just need to find the relevant threshold types $\tilde{x}(\beta)$.

Solution Concept and Timing Because of the uniform distribution of consumers, the situation in which first-mover A chooses $a \geq 0.5$ is equivalent to a situation where A chooses $a' = 1 - a$ instead. Therefore, we focus on identifying equilibria with $a \geq 0.5$ in the following and keep in mind that for each of these equilibria an analogous equilibrium exists for the case that $a \leq 0.5$.[17]

We employ the concept of a Perfect Bayesian Nash Equilibrium in pure strategies to solve the game. We assume that there are only 'second order effects' of the firms' locations on the consumers' belief β via the interpretation of the early adopter's choice C_1. This assumption fixes off-equilibrium beliefs and is plausible, as firms have no information about the quality differential.

The timing of the game is depicted in the figure below:

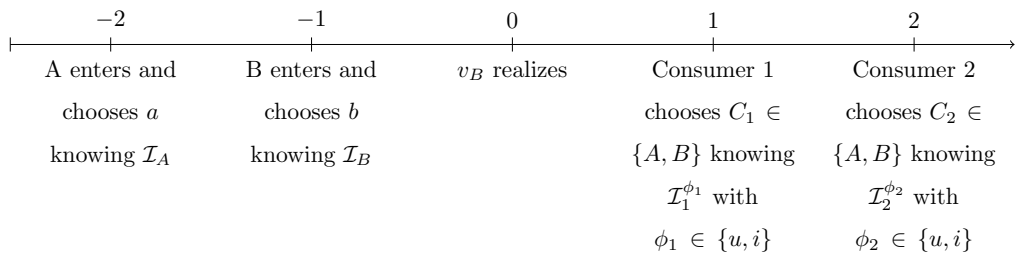


Figure 2: Timing of the Game

[17] We do not impose any further restrictions, such as the usual assumption $b \leq a$. The coordination issue of this assumption discussed Bester et al. (1996) does not arise in our setup due to the sequential location choice.

The game among the firms is an infinite two-player constant-sum game and firm A plays a minimax strategy.

Discussing the Assumptions

Before solving the model, some of the assumptions deserve additional thought. Similar effects as in this model are also be obtained in a model with endogenous prices, simultaneous choice of product differentiation, a continuously distributed quality differential and a continuous information structure (see Conze and Kramm, 2016).

Many other unit demand models, where the consumer's type is drawn from some distribution $G(\cdot)$, are equivalent to models with a continuum of consumers of mass one distributed according to $G(\cdot)$, because both formulations yield the same probabilistic demand and thus lead to the same behavior of the firms. However, since observing choices of the whole population is completely informative in our model, this equivalence does not hold. Nevertheless, introducing additional uncertainty to the model for instance by assuming that the distribution of the first period consumers and the value of q is uncertain reestablishes this connection even if the laggard observes choices of the whole population. Alternatively, we could reinterpret our modeling assumption as each laggard being drawn from a unit mass of consumers observing only one particular early adopter also drawn from a (different) unit mass of consumers.

In the standard model of spatial competition à la Hotelling and in our model, situations where sets of consumer types are indifferent may emerge. Assumption 2 deals with those cases. In the standard Hotelling model such a situation can only arise if firms locate at the same place. In contrast, in our model multiple consumer types may be indifferent if the distance between firms is sufficiently small. The behavior described in the assumption is obtained as the limiting case of situations in which the distance between the firms' locations is marginally greater. The second part of the assumption seems plausible in that light: as firm B is the second mover, it can always choose to locate infinitesimally close to firm A on each side, so that B intuitively chooses which side to position itself on, even if both firms are located at the same spot. If indifferent consumers behave otherwise than assumed, best responses may not be defined and a pure-strategy equilibrium may cease to exist, so that only this behavior is compatible with equilibrium. This argument is also put forward by Simon and Zame (1990) for general discontinuous games involving 'sharing rules' such as Assumption 2.

Thus, we only postulate it here to avoid complications due to off-equilibrium-path behavior.[18]

In order to simplify the updating of an uninformed laggard, we restrict ourselves to binary signals. One could well assume more than two possible levels of expertise so that consumers would not either be completely informed or completely uninformed. We leave out such specifications as this complicates the Bayesian updating and distracts from the main issue under consideration.

Since the updated probabilities are different for each history of the game, the (sets of) indifferent consumer types are also potentially different for each history, meaning that in each period $t > 1$, 2^{t-1} indifferent consumers have to be determined, quickly making the model intractable. The effects we wish to characterize are already apparent with one period of updating, i.e. with two consumers, which is why we concentrate on this case.

Different cost functions than the linear one applied here do not eliminate the underlying effects of our model, as long as costs are increasing in distance.[19] In the common Hotelling model, quadratic costs enhance the incentive to differentiate and would thus probably make the detection of the driving forces yielding the differentiation result in our model more complicated.

As mentioned above, the assumption of exogenous prices is applicable for markets in which the main endogenous determinant of the firms' profit is the market share. We show that, while without consumer learning the Principle of Minimum Differentiation prevails, consumer learning is sufficient for the existence of equilibria with differentiated products - even when prices are fixed. From a modeling perspective the assumption of exogenous prices has two advantages: on the one hand it assures the existence of a pure-strategy equilibrium - non-existence is a common problem in the original Hotelling model with linear transport costs - and on the other hand it eliminates the competition effect, which allows to more clearly identify the source of differentiation and is thus in favor of our focus on the effect of consumer learning.

The exogeneity of the quality differential seems plausible in many cases. For instance,

[18] In the light of our results, the recommendation effect, i.e. firm B's incentive to differentiate, is independent of Assumption 2, which merely determines the equilibrium behavior of firm A.

[19] We show later that Bayesian updating is unaffected by the cost function. Obviously, the demand functions - and thus in our model the best response functions of the firms - are not discontinuous for quadratic costs. Nevertheless, they entail regions, in which the recommendation effect makes it profitable to differentiate.

concerning the example of the movie 'The Artist' and the movie industry in general, it may well be that producers can not completely influence the (perceived) quality of a movie (see De Vany (2006) on this aspect of the movie industry). Also, note that movies - and many other goods - are experience goods (see e.g. De Vany, 2006), whose value is revealed to consumers only after their consumption.

Thus, in many cases even firms arguably do not know their product's relative quality (or the consumers' perceived quality) ex ante. This directly implies that the firms can not signal information about the realized quality to consumers. If their location choice would signal information to consumers, i.e. in a separating equilibrium, the information would already be revealed before the first consumer's choice and social learning would not occur. Since the effects of social learning by the consumers on the firms' location choices are exactly what we are interested in, situations with separating equilibria are not of our primary concern. We thus impose the assumption that firms are unaware of their quality in order to make sure that social learning is possible.

In the following section it will become obvious that the assumed setup is the most conservative one leading to product differentiation: abstracting from consumer learning leads to the usual result of (symmetric) minimal differentiation.

5 Equilibrium Analysis Without Consumer Learning (Benchmark)

In our benchmark model, consumers are unable to infer information from the other consumer's action. This is essentially the same as having two independent consumers purchasing in period one. In all other aspects, the model remains unchanged. Proposition 1 shows that this leads to the 'symmetric minimum differentiation' result also obtained by Hotelling. Our benchmark model differs from Hotelling's original model in that prices are fixed, firms move sequentially, products are of different quality and some consumers possess information about the quality differential. The result is easily obtained by the following steps.

5.1 Consumer Behavior

Let us start with assuming $b \leq a$. With the assumptions from above, an uninformed consumer's belief in the first period must equal the prior, and we denote it by $\beta^u_\emptyset := Pr(v_B = v+\delta) = 1/2$. As there is no updating involved in the benchmark model, according to equation (4) the uninformed indifferent consumer type

$$\tilde{x}(\beta^u_\emptyset) = \frac{a+b}{2},$$

i.e. the midpoint between the firms, characterizes the behavior of uninformed consumers. Note that, given $b \leq a$, whenever $\tilde{x}(\beta^u_\emptyset) > 0.5$ ($\tilde{x}(\beta^u_\emptyset) < 0.5$), firm B (A) is the niche firm, i.e. it serves the smaller ex-ante market share.

Informed consumers possess all relevant information and hence their beliefs are the same in all periods, depending only on which firm's product is of higher quality. Thus, from now on we write $\beta^i_A := Pr(v_B = v+\delta | v_B < v_A) = 0$ to denote the belief of an informed consumer when A is the better firm, and $\beta^i_B := Pr(v_B = v+\delta | v_B > v_A) = 1$ analogously for the case where B sells the superior product. Using equation (4) and defining

$$b_1(a) := a - \frac{\delta}{\tau},$$

which solves the equations

$$b \quad s.t. \quad \bar{x}(\beta^i_A) = b \qquad \text{and} \qquad b \quad s.t. \quad \bar{x}(\beta^i_B) = a, \tag{5}$$

threshold types for informed consumers can easily be calculated as

$$\tilde{x}(\beta^i_A) = \begin{cases} \bar{x}(\beta^i_A) := \frac{a+b}{2} - \frac{\delta}{2\tau} & \text{if} \quad b \leq b_1(a) \Leftrightarrow \bar{x}(\beta^i_A) \in [b,a], \\ 0 & \text{if} \quad b > b_1(a) \Leftrightarrow \bar{x}(\beta^i_A) < b, \end{cases} \tag{6}$$

if A is the superior product, and

$$\tilde{x}(\beta^i_B) = \begin{cases} \bar{x}(\beta^i_B) := \frac{a+b}{2} + \frac{\delta}{2\tau} & \text{if} \quad b \leq b_1(a) \Leftrightarrow \bar{x}(\beta^i_B) \in [b,a], \\ 1 & \text{if} \quad b > b_1(a) \Leftrightarrow \bar{x}(\beta^i_A) > a, \end{cases} \tag{7}$$

in the case that B's product is of higher quality. Intuitively, if $b > b_1(a)$, i.e. firm B locates relatively close to firm A, the additional transport costs when traveling to the better firm matter less than the additional value to informed consumers and thus all of them buy

according to their signal. It can directly be seen that $\bar{x}(\beta_A^i) \in [b, a] \Leftrightarrow \bar{x}(\beta_B^i) \in [b, a]$, and furthermore that $\bar{x}(\beta_A^i) < b \Leftrightarrow \bar{x}(\beta_B^i) > a$. These two cases distinguish whether firms are sufficiently close to each other or not, such that all informed consumers follow their signal.

5.2 Firm Behavior and Equilibrium

Combining the thresholds from above, firm B's demand, given $b \leq a$, calculates as

$$\begin{aligned} D_L(a, b) &= 2 \cdot \left[\frac{q}{2}(\tilde{x}(\beta_A^i) + \tilde{x}(\beta_B^i)) + (1-q)\tilde{x}(\beta_\emptyset^u)\right] \\ &= \begin{cases} a + b & \text{if} \quad b \leq b_1(a), \\ q + (1-q)(a+b) & \text{if} \quad b > b_1(a), \end{cases} \end{aligned}$$

where the subscript L denotes that firm B is located left of A. As stated in the model setup, it is without loss of generality to concentrate on situations where $a \geq 0.5$. Clearly, B's demand is increasing in both parts of the demand, so that with $a \geq 0.5$, B can never profit from choosing $b > a$, implying that B's best response $b^*(a)$ is given by one of the two maximal points of each segment, i.e. $b^*(a) \in \{b_1(a), a\}$, where $b_1(a)$ is a feasible choice whenever $b_1(a) \geq 0$.

Firm A's goal is to minimize B's demand by choosing its optimal point of B's best response function. In order for B to prefer $b = b_1(a)$ to $b = a$, a must be sufficiently high, in particular, as shown in Appendix A it must exceed 0.5. For any $a > 0.5$, B's demand is higher than 1. By choosing $a = 0.5$, A induces B to choose $b = a$, which leads to a demand of 1 for each firm, the highest demand A can generate in this model. The result in the benchmark model is thus as follows:

Proposition 1 (Symmetric Minimum Differentiation)**.** *In the unique equilibrium of the model without consumer learning, firms do not differentiate their products and equilibrium locations are $a = b = 0.5$.*

Proof. See Appendix A. □

6 Equilibrium Analysis With Consumer Learning

This section shows that the possibility to learn from previous consumers' actions can drastically change the outcome of the game compared to the benchmark model. With the con-

sumer's ability to observe her predecessor's action, new effects arise in the model and Proposition 1 will not generally hold. Instead, our main result, Proposition 2, shows that for particular values of the parameters at least one firm moves away from the center, and differentiation can arise in equilibrium. We postpone this result to the end of the section in order to now guide the reader through its construction. We again let the subscript L indicate that firm B positions left of firm A, i.e. $b \leq a$, whereas the subscript R represents the opposite case. We then make use of the fact that all situations $b > a$ can be described using the formulas obtained for $b \leq a$.

6.1 Informed Consumers and Uninformed Early Adopters

The decision of the consumer in the first period does not differ from the benchmark model, so that their behavior is fully characterized by $\tilde{x}(\beta^u_\emptyset)$ if uninformed and - depending on which firm is superior - by $\tilde{x}(\beta^i_A)$ or $\tilde{x}(\beta^i_B)$ if informed. As mentioned above, because informed consumers already have perfect information about both goods, an informed consumer in $t = 2$ behaves as one in period $t = 1$. In what follows, we discuss peculiarities only occurring in the model with consumer learning.

6.2 Uninformed Laggards: Updating and the Recommendation Effect

An uninformed laggard uses her information to update her belief $\beta^u(a, b, C_1) : \mathcal{A} \times \mathcal{B} \times \{A, B\} \to [0, 1]$. Let $\beta^u_{C_1} := \beta^u(a, b, C_1)$ denote the belief of an uninformed laggard given that $C_1 \in \{A, B\}$ was chosen in the first period. Although we assumed that the beliefs do not depend on the firms' locations directly, a and b have an indirect effect via the interpretation of the predecessor's action, C_1. Observing C_1 becomes useful for uninformed consumers, because of the possibility that the previous consumer was informed. Hence, history C_1 can now contain information that allows an uninformed laggard to update her estimate of which firm produces the good of higher value. Using Bayes' Rule she will calculate her belief $\beta^u_{C_1}$ of the probability that firm B is the higher quality firm as

$$\beta^u_{C_1} = Pr(v_B > v|C_1) = \frac{Pr(C_1|v_B > v) \cdot Pr(v_B > v)}{Pr(C_1)}.$$

In the first period, the products of both firms have the same expected utility (gross transportation costs) for uninformed consumers. This however is not the case in the second

period, as the updated probability $\beta^u_{C_1}$ must be used to calculate expected utilities when comparing the utility of buying good A to the expected utility from purchasing firm B's product. The updating introduces an asymmetry in expected valuations of the products, implying that in contrast to period $t = 1$, it is possible that no type of uninformed consumer is indifferent between the products.

It is thus necessary to distinguish three cases for any given belief β^u_{C1}. Either there is an unique indifferent consumer type, meaning it is in the interval $[b, a]$,[20] or the consumer located at a prefers B or the consumer type at b prefers A. In the latter two cases the same holds for all types right of a, respectively left of b; the intuition behind this was described in the model setup. Using equation (4) we have

$$\tilde{x}(\beta^u_{C_1}) = \begin{cases} \bar{x}(\beta^u_{C_1}) := \tilde{x}(\beta^u_{\emptyset}) + \frac{\delta}{\tau}\left(\beta^u_{C_1} - \beta^u_{\emptyset}\right) & \text{if} \quad \bar{x}(\beta^u_{C_1}) \in [a, b], \\ 0 & \text{if} \quad \bar{x}(\beta^u_{C_1}) < b, \\ 1 & \text{if} \quad \bar{x}(\beta^u_{C_1}) > a. \end{cases} \tag{8}$$

The uninformed indifferent consumer for the case that it is in $[b, a]$, i.e. $\bar{x}(\beta^u_{C_1}) \in [a, b]$, can nicely be interpreted, in that it is the first period's uninformed indifferent type, shifted to the left (right) by a term that weighs the product of the additional likeliness that B is the superior firm, if the choice in the first period was firm B (firm A), and the excess utility from choosing the better product against the additional transport costs.

In the literature on social learning (see e.g. Smith and Sørensen, 2000) 'herding' is defined as a behavior, where an agent's action is independent of her private signal: all information she uses comes from the (possibly updated) public belief derived from the behavior of others. The situation where an uninformed laggard always follows the early adopter can be viewed from a similar perspective: an agent chooses to buy from one firm using information which only comes from the observed behavior of other consumers. Thus, in our model herding does not mean that imitation dominates private information, but rather that imitation dominates own (ex-ante) tastes. We could extend our model to the case where signals are not completely informative and 'herding' consumers additionally ignore the information revealed by their own private signal.

The results that follow crucially depend on the behavior of an uninformed laggard, which

[20] For the case of sets of indifferent consumers their behavior is determined by Assumption 2.

in turn is dictated by her belief, and explicitly calculating this belief can greatly help with intuition for much of the firm behavior that follows. Bayes' rule is used to calculate the updated probability that B is the superior firm given $C_1 = B$ as

$$\begin{aligned}\beta_B^u = Pr(v_B > v|C_1 = B) &= \frac{Pr(C_1 = B|v_B > v) \cdot Pr(v_B > v)}{Pr(C_1 = B)} \\ &= \frac{q\tilde{x}(\beta_B^i) + (1-q)\tilde{x}(\beta_\emptyset^u)}{q(\tilde{x}(\beta_A^i) + \tilde{x}(\beta_B^i)) + (1-q)2\tilde{x}(\beta_\emptyset^u)}.\end{aligned} \tag{9}$$

The updated probability that $v_B > v_A$ after observing $C_1 = A$ is calculated similarly, and given by

$$\begin{aligned}\beta_A^u = Pr(v_B > v|C_1 = A) &= \frac{Pr(C_1 = A|v_B > v) \cdot Pr(v_B > v)}{Pr(C_1 = A)} \\ &= \frac{q[1-\tilde{x}(\beta_B^i)] + (1-q)[1-\tilde{x}(\beta_\emptyset^u)]}{q\{[1-\tilde{x}(\beta_A^i)] + [1-\tilde{x}(\beta_B^i)]\} + (1-q)\cdot 2 \cdot [1-\tilde{x}(\beta_\emptyset^u)]}.\end{aligned} \tag{10}$$

It was shown before, that $\tilde{x}(\beta_A^i) < \tilde{x}(\beta_B^i)$, so that we can see from equations (9) and (10) that $\beta_B^u > 0.5 > \beta_A^u$, meaning that observing $C_1 = B$ ($C_1 = A$) increases (decreases) the probability that B sells the good of higher value, just as one would expect.

Letting the fraction of informed consumers approach zero, that is $q \to 0$, the 'updated' probabilities approach the prior: $\beta_A^u, \beta_B^u \to \frac{1}{2}$. Overall, we can order all relevant beliefs according to $0 = \beta_A^i < \beta_A^u < \beta_\emptyset^u = 0.5 < \beta_B^u < \beta_B^i = 1$.

An interesting observation that can be made with regard to the updated probabilities, is that product differentiation has two effects for a firm. Suppose that both firms are symmetrically positioned around 0.5, i.e. $a + b = 1$. In this case a purchase of each good is equally informative to uninformed laggards as $\beta_B^u = 1 - \beta_A^u$. Now consider firm B's' incentives to increase the differentiation to A's product. If $b \leq a$, this means that B considers decreasing b. Increasing the product differentiation, which means that B is now producing a 'niche product', makes it less likely that product B is chosen by uninformed consumers in the first period, thus $Pr(C_1 = B)$ and $Pr(C_1 = B|v_B > v)$ and therefore the nominator and the denominator of equation (9) get smaller. Since b affects all threshold types $\tilde{x}(\cdot)$ in this equation in the same way, the effect on the denominator is twice as large as the one for the nominator and the updated probability that B's product is superior given it was chosen in the first period, β_B^u, increases, i.e. $\partial\beta_B^u/\partial b < 0$. This mechanism lays the foundation for the 'recommendation effect'. Intuitively, since a niche product is a good match to relatively few

consumer types (compared to a mainstream product), if it was chosen in $t = 1$, it is more likely that this was due to superior information about the quality than due to a better match of the product's characteristic and the consumer's taste.

An opposing effect is created regarding the updated probability, β_A^u. With A being the mainstream product (compared to B), an observed choice of it in the first period was more likely induced by a good match (a consumer located close to a) than by an informed consumer.

Put differently, the increased confidence that a product is of superior quality after having observed that it was bought in the first period, is higher for a niche product than for a mainstream product. Those are precisely the effects for which Tucker and Zhang (2011) find empirical evidence by examining the usefulness of popularity information for what they call products of 'narrow' and 'broad appeal'.

Although the 'recommendation effect' and the already mentioned 'competition effect', have similar implications on the firm's location, they are fundamentally different. With the 'competition effect', firms seek to lessen competition to increase their market power and markup. The 'recommendation effect' in contrast, exploits the way consumers conduct inference after observing earlier choices.

6.3 Firms' Expected Demand

Having calculated the threshold types of the consumers, B's demand for $b \leq a$ is given by

$$\begin{aligned} D_L(a,b) = q\Big[\tilde{x}(\beta_A^i) + \tilde{x}(\beta_B^i)\Big] \\ + (1-q)\Big[\tilde{x}(\beta_\emptyset^u) + Pr(C_1 = A)\tilde{x}(\beta_A^u) + Pr(C_1 = B)\tilde{x}(\beta_B^u)\Big]. \end{aligned} \tag{11}$$

It was shown in the previous sections that the threshold for a given belief β, $\tilde{x}(\beta)$ can either be at an interior value, meaning in the interval $[b, a]$ or it equals 0 or 1. As $\tilde{x}(\beta)$ is shifted away from $\tilde{x}(\beta_\emptyset^u) = (a+b)/2$, whether the threshold for some belief β is at an interior level or not depends on the distance between the two firms. Only for a sufficiently large distance, can $\tilde{x}(\beta)$ be at an interior level. Equations (4) and (8) show that the necessary distance between firms' locations implying $\tilde{x}(\beta)$ to be interior is increasing with the belief, and by equation (3) also $\bar{x}(\beta)' > 0$. Hence the threshold type for informed consumers is always shifted further away from $\tilde{x}(\beta_\emptyset^u)$ than the one of the uninformed laggards, meaning that $\bar{x}(\beta_A^u) \geq \bar{x}(\beta_A^i)$ and $\bar{x}(\beta_B^u) \leq \bar{x}(\beta_B^i)$. Thus, $\bar{x}(\beta_B^u) < a$ or $\bar{x}(\beta_A^u) < b$ directly imply $\bar{x}(\beta_A^i) < b, \bar{x}(\beta_B^i) > a$. It was

already argued that $\bar{x}(\beta_A^i) \notin [b,a] \Leftrightarrow \bar{x}(\beta_B^i) \notin [b,a]$ and because a first period purchase from one firm increases the belief that it offers the superior product, a situation with $\bar{x}(\beta_B^u) = b$ or $\bar{x}(\beta_A^u) = a$ can not occur.

This leaves us with five qualitatively different combinations of threshold types induced by different tuples (a,b). More precisely, in Proposition B.1 in Appendix B we show that for each configuration of the values of the parameters δ, τ and q there is a unique partition $\mathfrak{D}_L$ of $\{(a,b) \in \mathcal{A} \times \mathcal{B} | b \leq a\}$ given by

1) $\bar{x}(\beta_A^i), \bar{x}(\beta_B^i) \in [b,a], \quad \bar{x}(\beta_B^u),\ \bar{x}(\beta_A^u) \in [b,a],$

2) $\bar{x}(\beta_A^i), \bar{x}(\beta_B^i) \notin [b,a], \quad \bar{x}(\beta_B^u),\ \bar{x}(\beta_A^u) \in [b,a],$

3A) $\bar{x}(\beta_A^i), \bar{x}(\beta_B^i) \notin [b,a], \quad \bar{x}(\beta_B^u) \in [b,a],\ \bar{x}(\beta_A^u) \notin [b,a],$

3B) $\bar{x}(\beta_A^i), \bar{x}(\beta_B^i) \notin [b,a], \quad \bar{x}(\beta_B^u) \notin [b,a],\ \bar{x}(\beta_A^u) \in [b,a],$

4) $\bar{x}(\beta_A^i), \bar{x}(\beta_B^i) \notin [b,a], \quad \bar{x}(\beta_B^u),\ \bar{x}(\beta_A^u) \notin [b,a].$

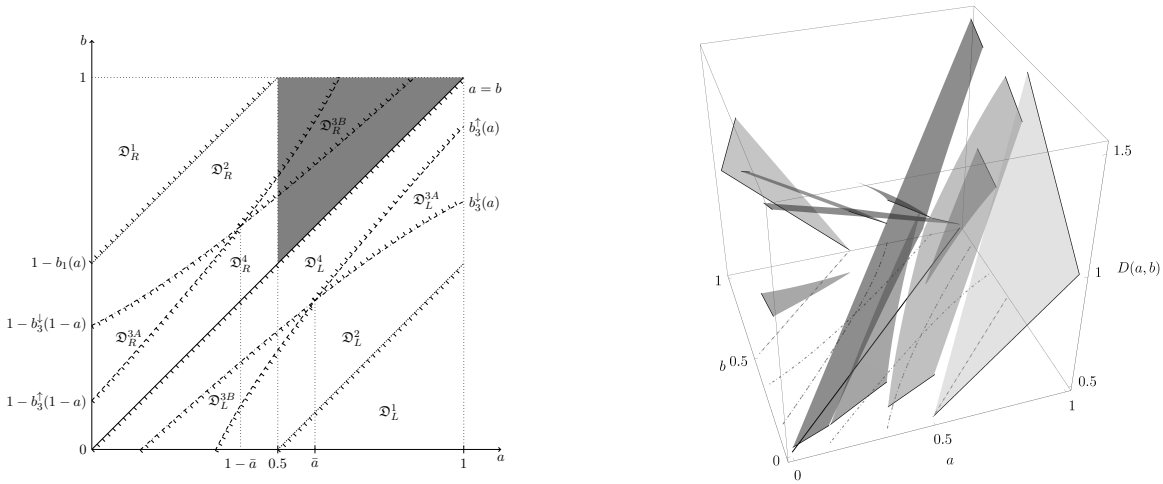
We use the notation $\mathfrak{D}_L$ to symbolize B being located left of A and $\mathfrak{D}^{3F}$ to imply that firm F is the niche firm in this region.

For $b > a$ there also exist five qualitatively different and mutually exclusive situations described by the threshold types $\bar{x}(\cdot)' := 1 - \bar{x}(\cdot)$, which induce a partition $\mathfrak{D}_R$ of $\{(a,b) \in \mathcal{A} \times \mathcal{B} | b > a\}$, where the R symbolizes B being located right of A. The partition is given by $\mathfrak{D}_R^j$ obtained from the definition of $\mathfrak{D}_L$, by replacing all a, b and $\bar{x}(\cdot)$ by $(a', b') := (1-a, 1-b)$ and $\bar{x}(\cdot)' = 1 - \bar{x}(\cdot)$. This is the case, as any situation with $b > a$ is equivalent to $a' = 1 - a$ and $b' = 1 - b$ in our model, because both situations generate the same demand. Thus, $\mathfrak{D} := \mathfrak{D}_L \cup \mathfrak{D}_R$ describes a partition of the whole action space $\mathcal{A} \times \mathcal{B}$. Further below we will describe the boundaries of the respective elements of the partition in more detail.

Each element of the partition leads to a specific form of B's demand, and we will denote B's demand in any given part by $D(a,b) = D_S^j(a,b)$ iff $(a,b) \in \mathfrak{D}_S^j \in \mathfrak{D}$ for $j \in \{1, 2, 3A, 3B, 4\}$ and $S \in \{L, R\}$.

If firm B chooses a position such that $b > a$ and additionally $a \geq 0.5$, we say that it positions on the short side of the market. Figure 3 depicts the partition of the action space and visualizes many results we obtain in the following.[21]

[21] The figure depicts the generic partition of the action space, i.e. for any particular configuration of the



(a) Generic partition of the action space. Note that the graphs for $a < 0.5$ are point reflection at $(0.5, 0.5)$ of the graphs for $a \geq 0.5$. The gray area depicts the short side. The little ticks indicate which region the boundaries belong to.

(b) Firm B's demand as an area over the action space.

Figure 3: Visualization of the partition of the action space and the implied form of firm B's demand. (Parameters: $q = 0.4, \tau = 2, \delta = 1$)

For a fixed a, we can calculate the boundaries between the different regions by choosing b such that the threshold consumer types of the respective parts of the partition are at the location of one of the firms. In addition to $b_1(a)$ as defined in the benchmark model, there are two further points of discontinuity of B's demand, namely the two points where, for each history C_1 and for a given a, the unique indifferent uninformed laggard ceases to exist. Whereas $\tilde{x}(\beta_A^i) = 0$ implies $\tilde{x}(\beta_B^i) = 1$ and vice versa, this generally is not the case for $\tilde{x}(\beta_A^u)$ and $\tilde{x}(\beta_B^u)$. Here, one threshold type may still be interior while the other already is at a corner value. This is the case, because whenever firms are not symmetrically positioned around 0.5, the updating is asymmetric, so that the thresholds' distances from $(a+b)/2$ are not the same. Those two discontinuity points are implicitly characterized by the following equations stating that the indifferent type in period 2 after A (B) was chosen in the first period is located at b

values of the parameters the partition qualitatively looks the same as the one depicted. The implication of the changes in parameters is described in Lemma 6 in Appendix B. Note that given Assumption 1, the figure does not qualitatively depend on v or δ.

(a):

$$b \quad \text{s.t.} \quad \bar{x}(\beta_A^u) = b \quad \Leftrightarrow \quad b = \frac{a+b}{2} + \frac{\delta}{2\tau}(2\beta_A^u - 1)$$
$$\Leftrightarrow \quad \frac{\delta q}{\tau} = (2-q)(a-b) - (1-q)\left(a^2 - b^2\right), \tag{12}$$

and

$$b \quad \text{s.t.} \quad \bar{x}(\beta_B^u) = a \quad \Leftrightarrow \quad a = \frac{a+b}{2} + \frac{\delta}{2\tau}(2\beta_B^u - 1)$$
$$\Leftrightarrow \quad \frac{\delta q}{\tau} = q(a-b) + (1-q)\left(a^2 - b^2\right). \tag{13}$$

Since both equations are quadratic in b, they both have two solutions. As shown in Lemma 2 in Appendix B, at most one of those solutions, for each equation, lies in the permissible range of $[0, 1]$. Those permissible solutions will in the following be referred to by $b_3^{\downarrow}(a)$ for equation (12) and $b_3^{\uparrow}(a)$ for equation (13). They are the discontinuity points in B's demand, so they mark the 'borders' between demand parts D_L^2, D_L^{3B} and D_L^4 or between D_L^2, D_L^{3A} and D_L^4.

To illustrate the underlying mechanism of the discontinuity in the demand of firm B, suppose that b is chosen such that $\bar{x}(\beta_B^u) = a$, meaning that firm B positions itself at the location $b_3^{\uparrow}(a)$, and all uninformed laggards to the right of a are indifferent between both firms. If B instead chose a location slightly smaller than $b_3^{\uparrow}(a)$, this would increase the transport costs of all consumers located at or to the right of a, meaning that those uninformed laggards would then prefer A's product. On the other hand, a location b slightly larger than $b_3^{\uparrow}(a)$ would induce all those uninformed laggards to buy the product from firm B. Taken together this implies that at $b_3^{\uparrow}(a)$, B's demand has an upward jump. Thus, we use the notation "$\uparrow$".

Similar reasoning leads to the observation that B's demand jumps downward at $b_3^{\downarrow}(a)$ as at this point the whole mass of consumers to the left of b switches from preferring B's good to preferring the one of firm A, if $C_1 = A$. Thus, overall we have

$$\tilde{x}(\beta_A^u) = \begin{cases} \bar{x}(\beta_A^u) := \frac{a+b}{2} + \frac{\delta}{2\tau}(2\beta_A^u - 1) & \text{if} \quad b \leq b_3^{\downarrow}(a) \Leftrightarrow \bar{x}(\beta_A^u) \in [a, b], \\ 0 & \text{if} \quad b > b_3^{\downarrow}(a) \Leftrightarrow \bar{x}(\beta_A^u) < b, \end{cases} \tag{14}$$

if A is the superior product, and

$$\tilde{x}(\beta_B^u) = \begin{cases} \bar{x}(\beta_B^u) := \frac{a+b}{2} + \frac{\delta}{2\tau}(2\beta_B^u - 1) & \text{if} \quad b \leq b_3^{\uparrow}(a) \Leftrightarrow \bar{x}(\beta_B^u) \in [a, b], \\ 1 & \text{if} \quad b > b_3^{\uparrow}(a) \Leftrightarrow \bar{x}(\beta_B^u) > a. \end{cases} \tag{15}$$

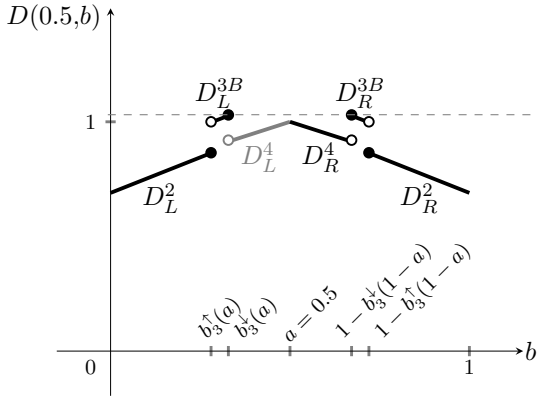
The derivatives of the discontinuity points can be calculated and it can be seen that $\partial b_3^{\uparrow}(a)/\partial a > 1 > \partial b_3^{\downarrow}(a)/\partial a > 0$ (see Lemma 3 in Appendix B). The calculations in Appendix B show that for small a we have $b_3^{\uparrow}(a) < b_3^{\downarrow}(a)$ and $b_3^{\uparrow}(a) > b_3^{\downarrow}(a)$ for larger a. By $a = \bar{a}$ we denote the location of firm A, so that $b_3^{\downarrow}(a) = b_3^{\uparrow}(a)$. In the case of $b_3^{\uparrow}(a) < b_3^{\downarrow}(a)$, B's 'middle' demand is characterized by D_L^{3B} and "starts" with an 'upward jump'. In the case of $b_3^{\uparrow}(a) > b_3^{\downarrow}(a)$, B's 'middle' demand is characterized by D_L^{3A} and "starts" with a 'downward jump'. To understand these points intuitively, recall, that in order for $\tilde{x}(\beta_B^u)$ to be at an interior value, the distance between a and b must be large enough and that updating is more favorable for a niche firm, meaning that $\tilde{x}(\beta_B^u)$ is shifted further from $(a+b)/2$ than $\tilde{x}(\beta_A^u)$ whenever firm B is the niche firm. Depending on a, it can be the case that the distance between a and b when $\tilde{x}(\beta_B^u)$ stops being at an interior level is obtained with B being the niche firm, i.e. $b_3^{\uparrow}(a) < 1 - a$, or not. Whenever $\tilde{x}(\beta_B^u) = 0$ and B is the niche firm, we are in region $\mathfrak{D}_L^{3B}$ with resulting demand D_L^{3B}, this will happen if a is sufficiently small. If a is large enough so that $\tilde{x}(\beta_B^u)$ stays at interior levels as long as A is the niche firm, $\tilde{x}(\beta_A^u)$ changes to its corner value before $\tilde{x}(\beta_B^u)$ does, and the applicable region is $\mathfrak{D}_L^{3A}$ with the corresponding demand. The transition between those two situations is obtained for $a = \bar{a}$ which plays a crucial role in the following equilibrium characterization. At this point, $b_3^{\downarrow}(a) = b_3^{\uparrow}(a)$, meaning that (only with this location a) $\tilde{x}(\beta_A^u) = 0$ implies $\tilde{x}(\beta_B^u) = 1$ and vice versa, hence, no firm is a niche firm when either $\tilde{x}(\beta_F^u)$ stops being at an interior level. If $a = \bar{a}$, then $(a, b) \notin \mathfrak{D}_L^{3B}$ and $(a, b) \notin \mathfrak{D}_L^{3A}$ for all $b \leq a$, and so B's demand does not contain demand part 3, for smaller (larger) a, $(a, b) \in \mathfrak{D}_L^{3B}$ ($(a, b) \in \mathfrak{D}_L^{3A}$) for any $b \leq a$.

The behavior of consumer types $\tilde{x}(\cdot)$ determines which part of the demand functions the discontinuity points belong to (see Lemma 6 in Appendix B for a complete description of the boundary points in the unique partition of the action space). By construction these types are indifferent and their behavior is chosen in Assumption 2 to guarantee the existence of equilibrium.

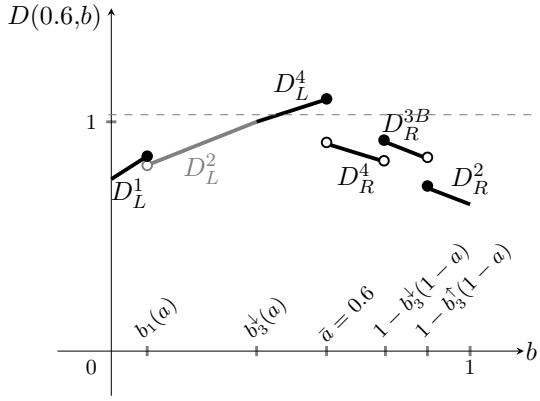
From equation (11) it is easy to see that the demand for $b \leq a$ is increasing in a and b as long as no threshold type changes from being interior to 0 or 1, which implies that the demand in each part $D_L^j(a, b)$ is increasing in a and b (see also Lemma 5 in Appendix B). By symmetry, demand in each part $D_R^j(a, b)$ is then decreasing in a and b.

With these observations and the above description of the demand parts, we can depict

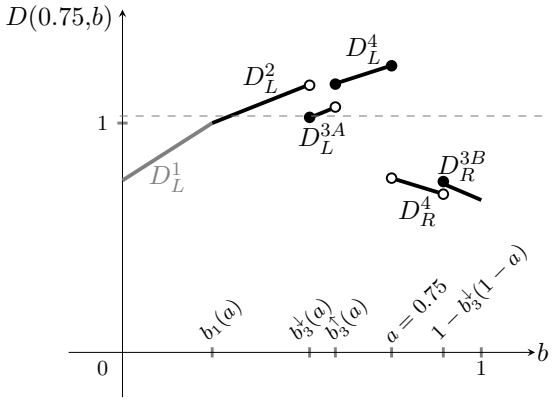
B's demand for a fixed a and varying choices of b, as is done in Figure 4, which shows the three generic situations that may occur in our model. Each of the different panels in this figure can be viewed as 'slice' cut out of the demand depicted in the right panel of Figure 3 for a fixed a.



(a) B's expected demand given $a = 0.5$ (exploiting the recommendation effect is advantageous for firm B). Note: $b_1(0.5) = 0$, i.e. D^1 is not available for B.



(b) B's expected demand given $a = \bar{a} = 0.6$.



(c) B's expected demand given $a = 0.75$ (exploiting the recommendation effect is not advantageous for firm B).

Figure 4: B's expected demand as a function of the chosen location b, for three different locations of A. In panel (a), $b_3^{\downarrow} > b_3^{\uparrow}$ so that D^3 of the demand is an upward step. Reversed situation in panel (c) and no jump in panel (b). The parameters yield an equilibrium with asymmetric central differentiation (see Proposition 2), which is visualized also by the dashed line. (Parameters: $q = 0.4, \tau = 2, \delta = 1$)

6.4 Firms' Best Responses and Equilibrium

In the description of the different parts of the demand it was argued that B's demand is increasing in b in each part. Thus, for a given $a \geq 0.5$, the demand of firm B is maximized by setting b equal to one of the points $b_1(a), b_3^{\downarrow}(a), b = a$ or by $b = 1 - b_3^{\downarrow}(1-a) > a$. Point $b = 1 - b_3^{\downarrow}(1-a) > a$ is the only point right of a that can ever be optimal, since for the highest points in the other demand parts with $b > a$, there is a corresponding point left of a yielding a higher demand given that $a \geq 0.5$. For ease of notation we define the value functions $V_S^j(a)$ with $j \in \{1, 2, 3, 4\}, S \in \{L, R\}$ to equal B's demand if it locates optimally in part j of its demand left ($S = L$) or right ($S = R$) of a. Let the function $V_S(a)$ denote the maximum of all value functions $V_S^j(a)$ and let $V(a)$ be the overall maximum, that is the maximum of $V_R^3(a)$ and $V_L(a)$.[22] B's best response to any a can then be written as

$$b^*(a) \in \underset{b \in \{b_1(a), b_3^{\downarrow}(a), a, 1-b_3^{\downarrow}(1-a)\}}{\arg\max} V(a)$$

We are interested in equilibria where the outcome is not symmetric minimal differentiation, in particular when firm <u>B</u> prefers to <u>d</u>ifferentiate from the <u>c</u>enter, which follows whenever

$$V_L^3(0.5) > V_L^4(0.5). \tag{BDC}$$

Given that demand is shared equally when both firms locate at the center (i.e. $V_L^4(0.5) = 1$), equation (BDC) implies that, whenever A locates at the center, B's demand is higher if $b = b_3^{\downarrow}(0.5)$ than if $b = a = 0.5$. In a simultaneous model both firm's reaction function would be given by the one of firm B in our model. For any division of market shares when both locate at the center, at least one firm has a total demand that is not greater than one, so that at least this firm has an incentive to deviate to $b_3^{\downarrow}(0.5)$. Thus, the result of 'symmetric minimum differentiation' would also not be obtained in a model where the two firms choose their locations simultaneously. In such a model no equilibrium in pure strategies exists, which is why we concentrate on the model with sequential location choice of the firms.

Firm B's demand for all $b \leq a$ consists of the same parts for any $a < \bar{a}$ or for any $a > \bar{a}$, and since every single $V_L^j(a)$ is increasing in a, the same must be true for $V_L(a)$ for any point but $\bar{a}$. Clearly, $V_R^3(a)$ is decreasing in a. Firm A's goal is given by $\min_a V(a)$, which is either

[22] As indicated above, we show in Lemma 7 in Appendix C, that B's demand on the short side obtains its maximal value via $V_R^3(a)$. The formal definition of $V(\cdot)$ can be found in Corollary 3 in Appendix C.

obtained by equalizing $V_L(a)$ and $V_R^3(a)$ or by setting $a = \bar{a}$ where $V_L(a)$ potentially has a downward jump. Setting $a = 0.5$ equalizes $V_L(a)$ and $V_R^3(a)$ but there might also be some $a > \bar{a}$ equalizing the two value functions, so that,

$$a^* \in \underset{a \in \{\bar{a}, a'\}}{\arg\min}\ V(a).$$

where $a' = \max_a \text{s.t. } V_L(a) = V_R^3(a)$.

Firm A thus decides between inducing two qualitatively different situations. By setting $a = a'$, it induces B to differentiate and exploit the recommendation effect, either left of a, then $a = a' = 0.5$, or right of a, then $a = a' > \bar{a}$. For some particular values of the parameters, A can keep B from differentiating by granting a relatively high (ex-ante) market share to firm B, which might be obtained by $a = \bar{a}$.

Which of the situations is preferred by firm A depends on whether

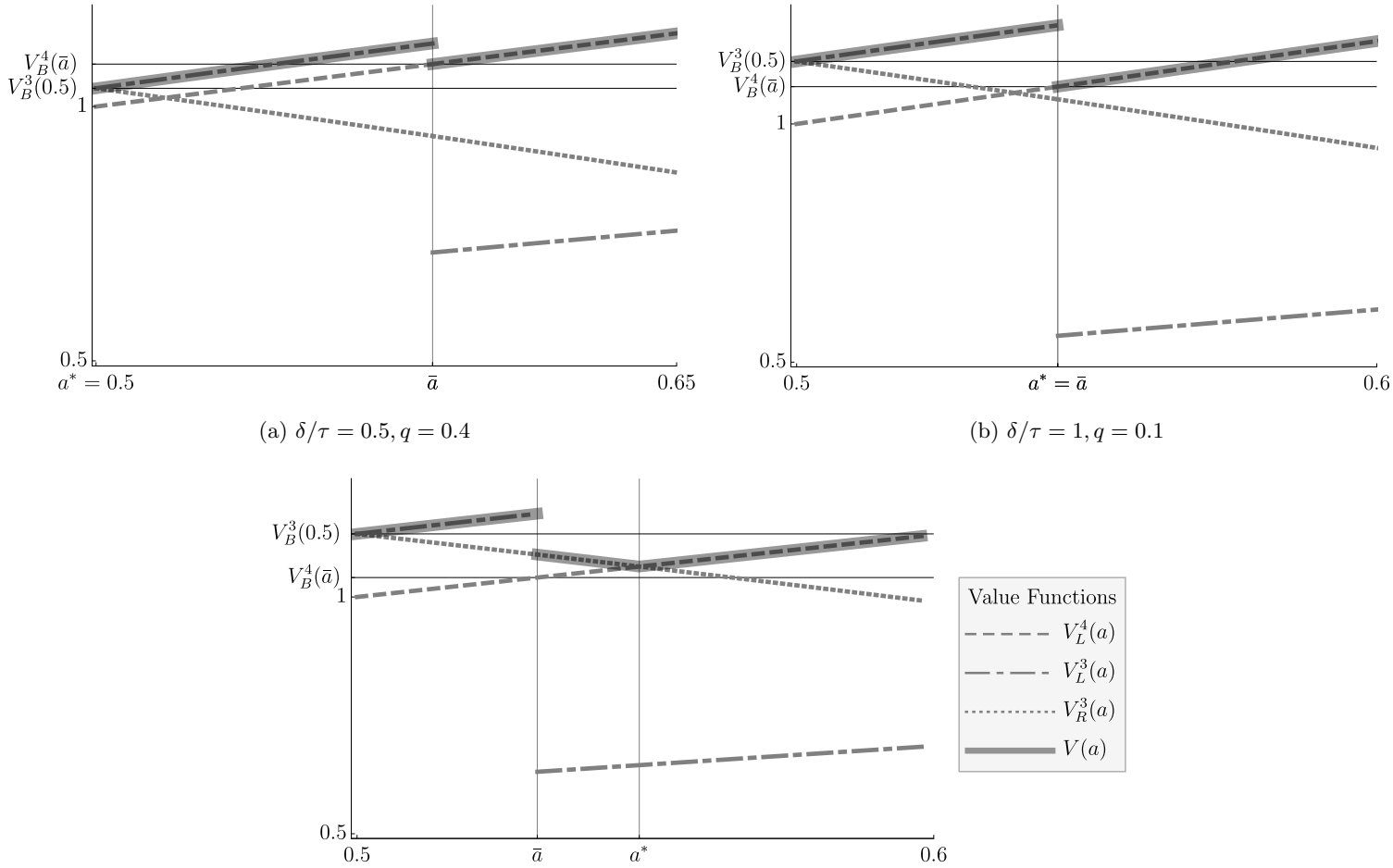
$$V_L^3(0.5) \leq V_L^4(\bar{a}), \tag{ADC}$$

which states that firm <u>A</u> prefers to induce <u>d</u>ifferentiation from the <u>c</u>enter. If firm A does not prefer to induce differentiation from the center, i.e. $a^* > 0.5$, firm <u>B</u> might still want to <u>d</u>ifferentiate on the <u>s</u>hort side:

$$V_L^4(\bar{a}) \leq V_R^3(\bar{a}), \tag{BDS}$$

states that firm B prefers to locate at $1 - b_3^{\downarrow}(\bar{a})$ instead of $b = a = \bar{a}$.

The three different situations are depicted in Figure 5. The complete parameter space is characterized in Figure 6 further below.



(a) $\delta/\tau = 0.5, q = 0.4$

(b) $\delta/\tau = 1, q = 0.1$

(c) $\delta/\tau = q = 0.25$

Figure 5: Combinations of the value functions induced by particular values of the parameters. In panels (a) and (b), B never profits from differentiating on the short side of the market. In (a), $V_L^4(\bar{a}) > V_L^3(0.5)$ so that $a^* = 0.5$. The reversed situation in (b), and $a^* = \bar{a}$. In (c), $V_R^3(\bar{a}) > V_L^4(\bar{a})$, and $a^* = \max a$ s.t $V_R^3(a) = V_L^4(a)$.

Equations (BDC), (ADC) and (BDS) distinguish the equilibria and are used to derive the respective parameter restrictions stated in Proposition 2. Note that (BDS) implies that (ADC) is violated, which in turn implies (BDC).

Proposition 2. *In the model with consumer learning we obtain the following results.*

1. *If equation* (BDC) *is violated, the strategies from the benchmark model constitute the unique equilibrium characterized by* $a^* = b^* = 0.5$ *(Symmetric Minimum Differentiation Equilibrium).*

2. *Equations* (BDC) *and* (ADC) *are necessary and sufficient conditions so that the locations are* $a^* = 0.5$ *and* $b^* < 0.5$ *in the unique equilibrium (Central Differentiation Equilibrium).*

3. *If equations* (ADC) *and* (BDS) *do not hold the locations are* $a^* = b^* > 0.5$ *in the unique equilibrium (Asymmetric Minimum Differentiation Equilibrium).*

4. *Equation* (BDS) *is a necessary and sufficient condition so that there exist an equilibrium (not necessarily unique) with locations* $b^* > a^* > \bar{a} > 0.5$ *(Short Side Differentiation Equilibrium).*

Uniqueness is up to symmetry, as to any equilibrium with (a^*, b^*) *there exists an analogous equilibrium with* $(1 - a^*, 1 - b^*)$.

Proof. See Appendix C. □

Sufficient conditions for the second and fourth equilibrium can be calculated as follows (see Corollary 4 of Appendix C for details):

$$\frac{\delta q}{\tau} < 0.192 \Rightarrow V_L^3(0.5) > V_L^4(0.5) \tag{BDC}$$

$$\frac{\delta q}{\tau} > 0.166, q < 0.4 \Rightarrow V_L^3(0.5) \leq V_L^4(\bar{a}) \tag{ADC}$$

If both those sufficient conditions are fulfilled, equilibrium locations are $b^* = b_3^{\downarrow}(0.5) < a^* = 0.5$. If

$$\frac{\delta q}{t} < 0.079 \Rightarrow V_R^3(\bar{a}) > V_L^4(\bar{a}) \tag{BDS}$$

holds, there is an equilibrium with $b^* = 1 - b_3^{\downarrow}(1 - a) > a > \bar{a}$. Note that we do not derive sufficient conditions for the equilibrium with asymmetric minimum differentiation, i.e. $b^* = a^* = \bar{a}$, as these would be "too small" in the parameter space. We only state the sufficient conditions here, as there exist no closed form solutions to the necessary and sufficient conditions, which are implied by the inequalities.

The following figure depicts the necessary and sufficient conditions on the values of the parameters inducing the equilibria, as well as the (weaker) sufficient conditions for equations BDC and (ADC), and additionally (BDS).

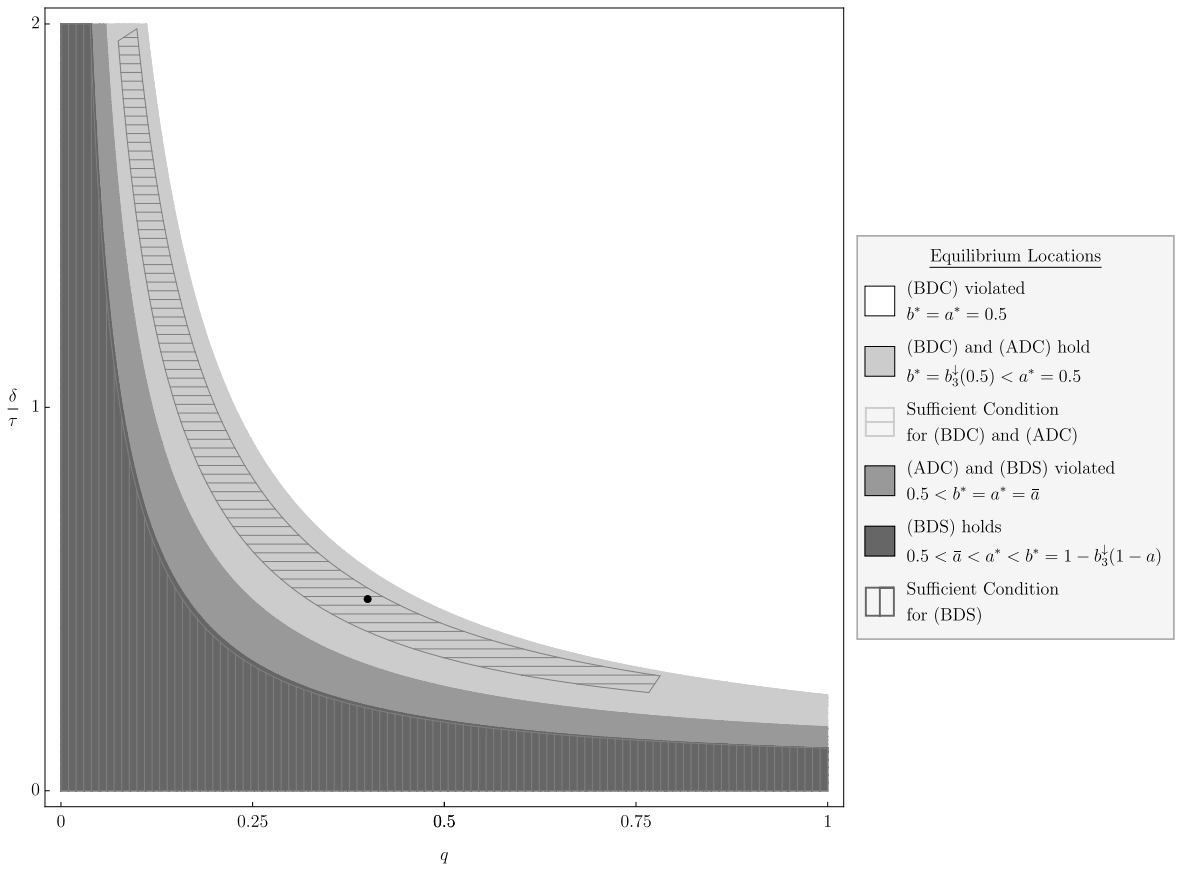


Figure 6: The figure depicts the necessary and sufficient parameter restrictions of the equilibria stated in Proposition 2. The black dot shows that the parameter combinations used in Figure 4 fulfills the conditions equations (BDC) and (ADC).

We now describe the intuition underlying the different equilibria. Figure 6 shows that $\frac{\delta q}{\tau}$ has to be sufficiently small for either of the 'new' equilibria to exist. The reason is that for all equilibria, firm B must want to deviate from a situation with symmetric minimum differentiation, which is formalized by condition (BDC). For differentiation to matter, $\frac{\delta q}{\tau}$ must not be too large. A large fraction implies that either the relative gain from choosing the higher quality product, $\frac{\delta}{\tau}$, or the likelihood that the first consumer is informed, q, is high. This makes it especially promising for uninformed consumers to follow the previous consumers' behavior even for large distances between the two firms' locations, in turn making differentiation unattractive for the firms.

As $\frac{\delta q}{\tau}$ decreases, $\bar{a}$ approaches 0.5, which has two effects. First it makes it less costly for firm A to locate at $\bar{a}$ thereby granting a relatively high ex-ante market share to firm B.

Secondly, as $\bar{a} \to 0.5$, the sides left and right of $\bar{a}$ are getting more and more alike, implying that it is more likely that B also wants to differentiate on the short side of the market, given that it prefers to do so if $a = 0.5$.

In contrast to Hotelling's result of 'symmetric minimum differentiation' where both firms choose to locate at the center, the firms' positions are not symmetric in any of the 'new' equilibria from above. The doubly sequential nature of the game clearly makes firm A worse off compared to the situation where consumers decided simultaneously (or were unable to observe the others' decisions). This also contrasts Tabuchi and Thisse (1995), who find asymmetric pure strategy equilibria with a first-mover advantage.[23]

6.5 Welfare

We use an utilitarian approach to compare the welfare induced by the three new equilibria (central differentiation $[b^* < a^* = 0.5]$, short side differentiation $[0.5 < a^* < b^*]$ and asymmetric minimum differentiation $[0.5 < a^* = b^*]$) in the model with consumer learning with that of the equilibrium with symmetric minimum differentiation $[a^* = b^* = 0.5]$ in the benchmark model.

We start by comparing the equilibrium with central differentiation under consumer learning with the symmetric minimum differentiation result in the benchmark model. First note that producer surplus is the same in both equilibria. When analyzing consumer surplus it is convenient to distinguish agents according to their different information. An informed consumer will have the same expected gain in both equilibria, which is given by $g^i = v + \delta$. However, in the differentiation equilibrium she has additional expected transport costs due to the fact that she might need to travel to niche firm B instead of the market center with probability 0.5.[24] These additional costs are given by $\Delta c^i = M/2$ with M as visualized in Figure 7 below.

[23] Bester et al. (1996) find asymmetric equilibria in mixed strategies.

[24] Note that when comparing expected transport, any location different from the center of the market, i.e. $f \neq 0.5$, implies additional expected transport costs compared to those of a firm located at the center, i.e. $f = 0.5$: the decrease in expected costs for the consumers located closer to the firm with $f \neq 0.5$ does not outweigh the increase in expected costs of the consumers located further away from that firm.

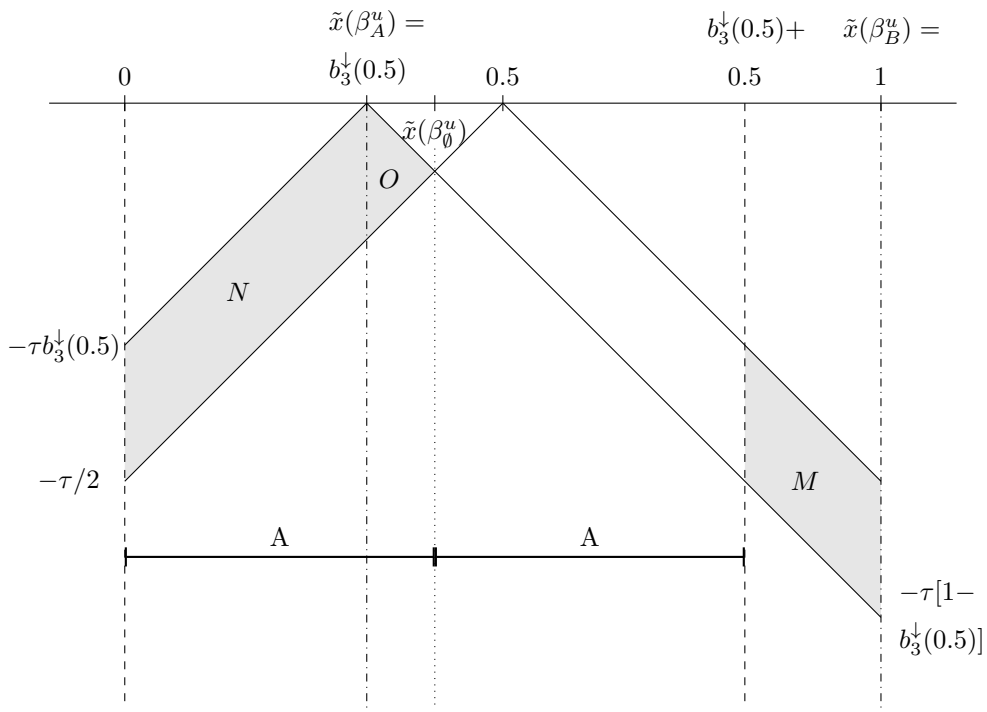


Figure 7: The figure depicts transport costs in the equilibrium of the benchmark model ($b^* = a^* = 0.5$) and the central differentiation result of the model with consumer learning ($b^* = b_3^{\downarrow}(0.5) < a^* = 0.5$). Indifferent types are depicted for the model with learning.

In the equilibrium with central differentiation, firm A is at the same location as in the benchmark model, namely at $a = 0.5$. For (completely) uninformed consumers, both firms have the same expected quality and firm A's quality is deterministic, and thus the same in both models, regardless of the available information. This implies that <u>uninformed consumers</u> can obtain the same expected utility in the model with learning as in the model without learning, whenever they choose to purchase good A in the model with learning. A revealed preference argument then implies that all uninformed consumers must be (weakly) better off in the model with learning.

Although the simple argument is enough to show that uninformed consumers (weakly) benefit from the outcome of the model with learning, we can go into further detail, by examining the uninformed consumers in the two periods and the different possible histories separately.

The uninformed early adopter ($t = 1$) unambiguously benefits in the central differentiation equilibrium: while her expected gain remains the same ($g_1^u = v$), the expected transport costs

decrease by $\Delta c_1^u = \tilde{x}(\beta_\emptyset^u) \cdot [N + O]$ with N and O as visualized in Figure 7 due to the fact that differentiation allows some consumer types to travel to the niche firm which is located closer to them and has the same expected quality as the main stream firm.

Additionally, for an uninformed laggard ($t = 2$), we can distinguish three cases.[25] Either she buys good A (then A must have been bought in period 1), or she buys B which can occur in both occasions, when A or B was bought in the first period.

Whenever the uninformed laggard buys A her utility is exactly equal to the one in the benchmark model. When she buys B after observing a purchase of A she also obtains exactly the same utility as in the benchmark model. This is because of the way the equilibrium is constructed: firm B chose its location precisely to make the uninformed laggard after history $C_1 = A$ indifferent. For the last possible situation, that is, a purchase of B in both periods, the revealed preference argument again implies that all consumers must be weakly better off when compared to the benchmark model. But now some consumer types $0 < x < \frac{a+b}{2}$ are strictly better off in terms of expected utility. Take the consumer located at the same spot as firm B for example. She has less distance to cover and since B was bought in the first period, she expects the product B to be of better quality. Thus she has less transport costs and a higher expected valuation when compared to her situation in the benchmark model.

Hence, uninformed consumers are unambiguously better off in the equilibrium with central differentiation and benefit from the fact that observing informed consumers provides additional information. Informed consumers, on the other hand, prefer the result of the benchmark model without consumer learning. Which of these opposing effects on the consumers dominates, depends on the share of informed consumers, q, and the excess utility when consuming the superior product, δ/τ. Since the explicit solution for the question under which parameter restrictions the welfare is enhanced is too complex, the numerical condition on the parameters is depicted in Figure 8 below.

[25] Overall, there are two effects on the expected transport costs of uninformed laggards, which stems from the fact that, given the history of the game, either the threshold type satisfies $\tilde{x}(\beta_B^u) = 1$ (just as the threshold type of informed consumers given $v_B > v$) or the threshold type $\tilde{x}(\beta_A^u)$ is in the interval (b, a) (just as the threshold type of uninformed early adopters). In the first case, expected costs increase, as she might have to travel further to the firm, which she perceives to be superior. In the second case, expected costs decrease. Overall, we have $\Delta c_2^u = P(C_1 = A) \cdot \tilde{x}(\beta_A^u) \cdot N + P(C_1 = B) \cdot M$.

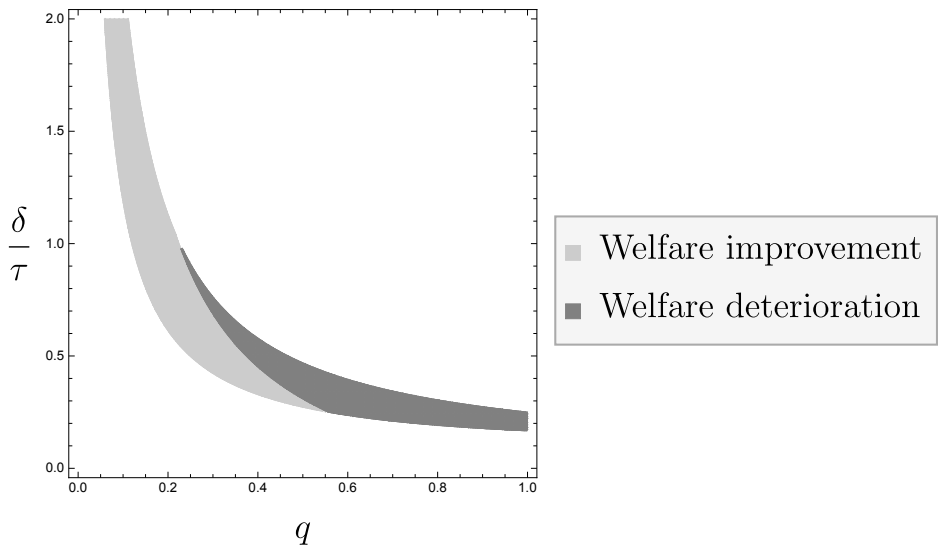


Figure 8: The shaded area depicts values of the parameters implying the existence of the central differentiation equilibrium ($b^* < a^* = 0.5$) in the model with consumer learning and under which this type of equilibrium is a welfare improvement / deterioration compared to the symmetric minimum differentiation result ($b^* = a^* = 0.5$) in the model without consumer learning.

A social planner might consider two different transparency enhancing policies: on the one hand, she could force the firms to provide more information on the product, so that q increases.[26] Figure 8 then shows that welfare is decreasing, simply because the share of informed consumers (which bear higher expected transportation costs in the model with learning than uninformed consumers) increases. On the other hand, the social planner could increase market transparency by making the firms provide information about previous sales, i.e. it could induce a switch from a world without consumer learning to a world with consumer learning. In this case, the result is ambiguous, i.e. welfare might increase or decrease depending on the particular values of the parameters $\frac{\delta}{\tau}$ and q, as Figure 8 shows.

The welfare analysis of the equilibrium with differentiation on the short side is similar to the equilibrium with central differentiation. However, expected transport costs (for informed and uninformed consumers) now increase even more due to the fact that $a \neq 0.5$.

In the equilibrium with asymmetric minimum differentiation of the model with consumer

[26] Even if firms do not know the quality (differential), the information provided by them might be helpful to evaluate the products' quality, as it might be the case for experience goods.

learning all consumers incur higher expected transportation costs and uninformed early adopters and informed consumers obtain the same expected gain as in the equilibrium of the benchmark model. Because of the updating the expected gain of uninformed laggards is higher. Overall, the comparison of consumer surplus depends on which effect dominates.

7 Conclusion

This paper has given an information-related explanation for why a firm may want to produce a product which appeals to relatively few consumers ex-ante.

In our variant of the classical model of spatial competition due to Hotelling (1929), the effect emerges because consumers, who are heterogeneous with respect to their preferred good and with respect to the level of information they possess, make their purchase decisions sequentially and are able to observe which good previous consumers bought. Uninformed consumers in later periods rationalize the choice of other consumers by considering that earlier consumers possibly made their decision because they were better informed about the quality of the different goods. An uninformed consumer thus updates her estimate about the difference in the good's quality after observing previous consumer choices using Bayes' Rule.

This updating is especially favorable for niche products. Niche products are not as appealing as mainstream products to a broad range of consumers. Therefore, later consumer's reasoning after having observed the purchase of a niche product puts more weight on the possibility that this purchase was due to the consumer being informed instead of being due to a good match of the earlier consumer's preference and the good's characteristic.

When deciding about the good's characteristics, i.e. how much to differentiate from the opponent's product, a firm has to take into account two offsetting effects. On the one hand producing a niche product decreases the product's overall appeal to consumers, hence the expected demand in early periods is decreased. But on the other hand, exactly because the overall appeal is decreased, an early purchase of the niche product leads to a higher boost of later uninformed consumers' confidence in the niche products superior quality. As this paper shows, the second effect can dominate, leading to an equilibrium with differentiated goods. This effect, the 'recommendation effect', is different from what is generally called the 'competition effect' which goes into a similar direction as it makes differentiation profitable for firms, but in the latter the driving force is that it relaxes price competition, thus increasing

possible markups.

On a broader view, our research is related to the literature which emphasizes the (evolutionary) value of diversity in the need for innovation (see e.g. Page, 2008 or Surowiecki, 2005): when 'solutions' (e.g. products) to 'problems' are more diverse, then the process of finding out which is the 'better' one is more effective.[27] Our model shows that this 'social' rationale for diversity can be individually rational for firms.

In biology there is an effect similar to the one described in this paper: the Handicap Principle (see e.g. Zahavi, 1975) explains why some animals have certain features which at first sight seem to be an evolutionary disadvantage. A popular example is the tail of the peacock. This tail is a huge obstacle when being hunted by predators. But if one such peacock survives and is chosen by a mate to pass on his genes, then the (probably even bigger) tail of the offspring can work as a strong signal for its high (evolutionary) quality.[28]

Many of the simplifying assumptions made in our model are not necessary and similar effects also emerge in more general settings (see 'Discussion of the Assumptions' in Section 4 and Conze and Kramm, 2016).

It would be worthwhile to relax the symmetry of the model. For example, a similar intuition as the one for why early purchases are especially valuable for niche products might suggest that an ex-ante inferior firm potentially benefits more from early adopters choosing its product. Other extensions could include an endogeneity of the timing of consumption or replacing observational learning by word-of-mouth learning.

Appendices

A Proof of Proposition 1 for the Benchmark Model

Proof. In the following we start with assuming $b \leq a$, but similar arguments apply to $a' := 1 - a$ and $b' := 1 - b$ with $\tilde{x}' := 1 - \tilde{x}$ if $b > a$.

In the benchmark model there are two qualitatively different positions firm B can choose for

[27] The related issue (and the importance) of the speed of social learning and its convergence 'to the truth' is discussed for instance in Gale (1996).

[28] In contrast to our model, this theory however attributes the effect to the underlying mechanism of costly signaling.

a given a: either it chooses a small b which implies a high degree of differentiation and also that both, $\bar{x}(\beta_A^i)$ and $\bar{x}(\beta_B^i)$ are in the interval $[b, a]$ and $\bar{x}(\beta_A^i) + \bar{x}(\beta_B^i) = a + b$, or it chooses a location relatively close to a so that $\bar{x}(\beta_A^i) > b$ and $\bar{x}(\beta_B^i) < a$. We define $b_1(a)$ to be the largest b where the first case occurs. It is calculated as $b_1(a) = a - \frac{\delta}{\tau}$ from b s.t. $\bar{x}(\beta_B^i) = a$. Note that if $b_1(a) < 0$, firm B cannot induce a case in which all informed consumers follow their signal and the result follows immediately.

Let $D_L(a, b)$ denote B's expected demand depending on firm A's location a and firm B's location b whenever $b \leq a$. In the benchmark model it is given by

$$\begin{aligned} D_L(a,b) &= 2 \cdot \left[\frac{q}{2}(\tilde{x}(\beta_A^i) + \tilde{x}(\beta_B^i)) + (1-q)\tilde{x}(\beta_\emptyset^u)\right] \\ &= \begin{cases} a + b & \text{if} \quad b \leq b_1(a), \\ q + (1-q)(a+b) & \text{if} \quad b > b_1(a). \end{cases} \end{aligned}$$

As B's demand is increasing in b (there are no 'information effects' in the model without consumer learning), a profit-maximizing firm B chooses between the locations $b = b_1(a)$ and $b = a$. B will prefer $b = b_1(a)$ to $b = a$ if

$$\begin{aligned} D_L(a, b_1(a)) &\geq D_L(a, a) \\ \Leftrightarrow 2 \cdot \frac{a + b_1(a)}{2} &= 2a - \frac{\delta}{\tau} \geq 2[(1-q)a + \frac{q}{2}] \\ \Leftrightarrow a &\geq \frac{1}{2} + \frac{\delta}{2\tau q} =: \check{a}. \end{aligned}$$

Thus, firm B's best response is given by

$$b^*(a) = \begin{cases} b_1(a) & \text{if} \quad a \geq \check{a}, \\ a & \text{if} \quad a < \check{a}. \end{cases}$$

Note that firm A could always guarantee a demand of $2 - D_L(0.5, 0.5) = 1$ to itself. We have established that A would have to choose a location further to the right ($a = \check{a}$ instead of $a = 0.5$), if it would want to induce $b = b_1(a)$. The smallest a that would induce $b_1(a)$ is given by $\check{a}$. Firm A's expected demand $\tilde{D}_L(a, b) := 2 - D_L(a, b)$ anticipating the behavior of firm B is

$$\tilde{D}_L(a, b^*(a)) = \begin{cases} 2(1-a) + \frac{\delta}{\tau} & \text{if} \quad a \geq \check{a}, \\ 2 - q - 2a(1-q) & \text{if} \quad a < \check{a}. \end{cases}$$

For firm A to prefer inducing $b = b_1(a)$, we would need:

$$\begin{aligned} \tilde{D}_L(\check{a}, b_1(\check{a})) &\geq \tilde{D}_L(0.5, 0.5) \\ \Leftrightarrow 2\left[1 - \frac{\check{a} + b_1(\check{a})}{2}\right] &\geq 1 \\ \Leftrightarrow 1 &\leq 1 + \frac{\delta}{\tau} - \frac{\delta}{\tau q} \end{aligned}$$

which never holds. Thus, it will always be the case that $\bar{x}(\beta_A^i) < b$ and $\bar{x}(\beta_B^i) > a$ in equilibrium. It also follows immediately, that in equilibrium we have $a = b = 0.5$. □

B Generic Properties of Firm B's Expected Demand

Proposition B.1. *For each particular configuration of the values of the parameters δ, τ and q, the demand of both firms is characterized by the constellation of the different threshold types $\bar{x}(\cdot)$. For $b \leq a$, five qualitatively different and mutually exclusive cases can occur:*

1) $\bar{x}(\beta_A^i), \bar{x}(\beta_B^i) \in [b, a], \quad \bar{x}(\beta_B^u),\ \bar{x}(\beta_A^u) \in [b, a],$

2) $\bar{x}(\beta_A^i), \bar{x}(\beta_B^i) \notin [b, a], \quad \bar{x}(\beta_B^u),\ \bar{x}(\beta_A^u) \in [b, a],$

3A) $\bar{x}(\beta_A^i), \bar{x}(\beta_B^i) \notin [b, a], \quad \bar{x}(\beta_B^u) \in [b, a],\ \bar{x}(\beta_A^u) \notin [b, a],$

3B) $\bar{x}(\beta_A^i), \bar{x}(\beta_B^i) \notin [b, a], \quad \bar{x}(\beta_B^u) \notin [b, a],\ \bar{x}(\beta_A^u) \in [b, a],$

4) $\bar{x}(\beta_A^i), \bar{x}(\beta_B^i) \notin [b, a], \quad \bar{x}(\beta_B^u),\ \bar{x}(\beta_A^u) \notin [b, a].$

For $b > a$, there also exist five qualitatively different and mutually exclusive cases described by the threshold types $\bar{x}(\cdot)' := 1 - \bar{x}(\cdot)$, and replacing all a, b and $\bar{x}(\cdot)$ by $(a', b') := (1-a, 1-b)$ and $\bar{x}(\cdot)'$ in the above cases.

These cases translate to the following unique partition $\mathfrak{D} := \mathfrak{D}_L \cup \mathfrak{D}_R$ of the action space

$\mathcal{A} \times \mathcal{B}$ with (possibly empty) elements $\mathfrak{D}_L^j$ and $\mathfrak{D}_R^j$:

$$
\begin{aligned}
\mathfrak{D}_L^1 &= \{(a,b) \in \mathcal{A} \times \mathcal{B} \mid b \leq b_1(a)\},\\
\mathfrak{D}_L^2 &= \{(a,b) \in \mathcal{A} \times \mathcal{B} \mid b_1(a) < b \leq b_3^{\uparrow}(a)\} \cap \{(a,b) \in \mathcal{A} \times \mathcal{B} \mid b_1(a) < b < b_3^{\downarrow}(a)\},\\
\mathfrak{D}_L^{3A} &= \{(a,b) \in \mathcal{A} \times \mathcal{B} \mid b_3^{\downarrow}(a) \leq b < b_3^{\uparrow}(a)\},^{29}\\
\mathfrak{D}_L^{3B} &= \{(a,b) \in \mathcal{A} \times \mathcal{B} \mid b_3^{\uparrow}(a) < b \leq b_3^{\downarrow}(a)\},\\
\mathfrak{D}_L^4 &= \{(a,b) \in \mathcal{A} \times \mathcal{B} \mid b_3^{\downarrow}(a) < b_3^{\uparrow}(a) \leq b \leq a\} \cup \{(a,b) \in \mathcal{A} \times \mathcal{B} \mid b_3^{\uparrow}(a) \leq b_3^{\downarrow}(a) < b \leq a\} \cup \{b_3^{\downarrow}(\bar{a}), \bar{a}\},\\
\mathfrak{D}_R^4 &= \{(a,b) \in \mathcal{A} \times \mathcal{B} \mid a < b \leq 1 - b_3^{\uparrow}(a) < 1 - b_3^{\downarrow}(a)\}\\
&\quad \cup \{(a,b) \in \mathcal{A} \times \mathcal{B} \mid a < b < 1 - b_3^{\downarrow}(a) \leq 1 - b_3^{\uparrow}(a)\},\\
\mathfrak{D}_R^{3B} &= \{(a,b) \in \mathcal{A} \times \mathcal{B} \mid 1 - b_3^{\downarrow}(a) \leq b < 1 - b_3^{\uparrow}(a)\},\\
\mathfrak{D}_R^{3A} &= \{(a,b) \in \mathcal{A} \times \mathcal{B} \mid 1 - b_3^{\uparrow}(a) < b \leq 1 - b_3^{\downarrow}(a)\},\\
\mathfrak{D}_R^2 &= \{(a,b) \in \mathcal{A} \times \mathcal{B} \mid 1 - b_3^{\uparrow}(a) \leq b < 1 - b_1(a)\} \cap \{(a,b) \in \mathcal{A} \times \mathcal{B} \mid 1 - b_3^{\downarrow}(a) < b < 1 - b_1(a)\},\\
\mathfrak{D}_R^1 &= \{(a,b) \in \mathcal{A} \times \mathcal{B} \mid 1 - b_1(a) \leq b\}.
\end{aligned}
$$

The generic partition is depicted in Figure 3. The functional form of demand is different in each element of the partition, but in each element, firm B's demand is increasing in a and b if $b \leq a$ and decreasing in a and b otherwise.[30]

Proof. The proof is constructed using a succession of lemmata.

Lemma 1. *For $b \leq a$, firm B's demand, and thus also firm A's demand, consists of five qualitatively different and mutually exclusive cases described by the threshold types $\bar{x}(\cdot)$.*

Proof. For $b \leq a$, firm B's demand is given by

$$
\begin{aligned}
D_L(a,b) = q\Big[\tilde{x}(\beta_A^i) + \tilde{x}(\beta_B^i)\Big]&\\
+ (1-q)\Big[\tilde{x}(\beta_\emptyset^u) + Pr(C_1 = A)\tilde{x}(\beta_A^u) + Pr(C_1 = B)\tilde{x}(\beta_B^u)\Big]&.
\end{aligned}
$$

Case distinctions depending on whether the thresholds $\bar{x}(\cdot)$, determining the value of $\tilde{x}(\cdot)$, are at interior or corner values, which in turn depends on the locations a and b, have to be

[29] As $b_3^{\uparrow}(a)$ might be complex valued for values of a where $b_3^{\downarrow}(a) > 0$, the precise formulation of this conditions reads $\mathfrak{D}_L^{3A} = \{(a,b) \in \mathcal{A} \times \mathcal{B} \mid 0 > b > b_3^{\downarrow}(a)\} \cap \{(a,b) \in \mathcal{A} \times \mathcal{B} \mid \neg b_3^{\downarrow}(a) < b_3^{\uparrow}(a)\}$

[30] We deal with the fact that $\mathfrak{D}^{3A}$ and $\mathfrak{D}^2$ have no maximizers by defining the value function in these parts in terms of limits.

made. As a first period purchase from one firm increases the belief that it offers the superior product, a case with $\bar{x}(\beta_B^u) < b$ or $\bar{x}(\beta_A^u) > a$ can not occur. It can directly be seen that $\bar{x}(\beta_A^i) \in [b,a] \Leftrightarrow \bar{x}(\beta_B^i) \in [b,a]$, and furthermore that $\bar{x}(\beta_A^i) \notin [b,a] \Leftrightarrow \bar{x}(\beta_B^i) \notin [b,a]$. Another crucial insight is that the threshold type for informed consumers is always shifted further apart from $\bar{x}(\beta_\emptyset^u)$ than the one of the uninformed laggards, meaning that $\bar{x}(\beta_A^u) \geq \bar{x}(\beta_A^i)$ and $\bar{x}(\beta_B^u) \leq \bar{x}(\beta_B^i)$. This can easily be seen in equation (8) with 0 or 1 put in place of the belief $\beta_{C_1}^u$. Thus, $\bar{x}(\beta_B^u) \notin [b,a]$ or $\bar{x}(\beta_A^u) \notin [b,a]$ directly imply $\bar{x}(\beta_A^i) \notin [b,a]$ and $\bar{x}(\beta_B^i) \notin [b,a]$. These five cases describe all cases. Note that $\bar{x}(\beta_\emptyset^u) = (a+b)/2$ in any case. □

Lemma 2. *Equations* (5)*, implicitly defining the boundaries between regions* $\mathfrak{D}_L^1$ *and* $\mathfrak{D}_L^2$ *have the solution* $b_1(a) = a - \frac{\delta}{\tau}$*. Equations* (12) *and* (13)*, implicitly defining the boundaries between regions* $\mathfrak{D}_L^2, \mathfrak{D}_L^3$ *and* $\mathfrak{D}_L^4$ *have at most one solution in* $[0,1]$*. Call these solutions* $b_3^{\downarrow}(a)$ *and* $b_3^{\uparrow}(a)$ *respectively. If* a *is s.t.* $b_3^{\uparrow}(a), b_3^{\downarrow}(a) \in [0,1]$*, then the respective function is continuous in* a*. Function* $b_1(a)$ *is continuous in* a*. For any given* a*, the function values* $b_3^{\uparrow}(a)$*,* $b_3^{\downarrow}(a)$ *and* $b_1(a)$ *correspond to the discontinuity point of the demand.*

Proof. The first statement can be proven by straightforward calculations. The possibly valid solutions to equations (12) and (13) are given by

$$b_3^{\downarrow}(a) = \frac{(2-q) - \sqrt{(2-q)^2 - 4(1-q)\left[(2-q)a - (1-q)a^2 - \frac{\delta q}{\tau}\right]}}{2(1-q)}$$

for (12), and

$$b_3^{\uparrow}(a) = \frac{q - \sqrt{q^2 + 4(1-q)\left[qa + (1-q)a^2 - \frac{\delta q}{\tau}\right]}}{-2(1-q)}$$

for (13). A bit of calculation shows that both discontinuity points exist for a such that

$$b_3^{\downarrow}(a) \in [0,1] \;\Leftrightarrow\; a \in \left[\frac{(2-q) - \sqrt{(2-q)^2 - 4(1-q)\frac{\delta q}{\tau}}}{2(1-q)}, \frac{(2-q) - \sqrt{(2-q)^2 - 4(1-q)\frac{\delta q + \tau}{\tau}}}{2(1-q)}\right]$$

and

$$b_3^{\uparrow}(a) \in [0,1] \;\Leftrightarrow\; a \in \left[\frac{-q + \sqrt{q^2 + 4(1-q)\frac{\delta q}{\tau}}}{2(1-q)}, \frac{-q + \sqrt{q^2 + 4(1-q)\frac{\delta q + \tau}{\tau}}}{2(1-q)}\right].$$

Note that these two restrictions on a imply that the radicands in the definition of $b_3^{\uparrow}(a)$ and $b_3^{\downarrow}(a)$ are in the interval $[q^2, (2-q)^2] \subset \mathbb{R}_{++}$, which in turn implies that $b_3^{\uparrow}(a)$ and $b_3^{\downarrow}(a)$ are real valued if they are in $[0,1]$.[31] Continuity can easily be seen and the last statement in the lemma follows by construction of $b_1(a)$, $b_3^{\downarrow}(a)$ and $b_3^{\uparrow}(a)$. □

Lemma 3. $\frac{\partial b_3^{\uparrow}(a)}{\partial a} > 1 > \frac{\partial b_3^{\downarrow}(a)}{\partial a} > 0$ *for* $b \leq a$ *and* $a, b \in [0,1]$. *Furthermore,* $b_3^{\downarrow}(a)$ *and* $b_3^{\uparrow}(a)$ *cross only once at* $\bar{a} = \frac{1}{2} + \frac{\delta q}{2\tau}$, *and* $b_3^{\downarrow}(a)$ *is concave. Hence* $a < \bar{a} \Leftrightarrow b_3^{\uparrow}(a) < b_3^{\downarrow}(a)$ *and* $a > \bar{a} \Leftrightarrow b_3^{\uparrow}(a) > b_3^{\downarrow}(a)$. *In addition* $b_3^{\downarrow}(\bar{a}) = b_3^{\uparrow}(\bar{a}) = 1 - \bar{a}$.

Proof. Derivatives can be directly calculated from equations (12) and (13) via implicit differentiation as:

$$\frac{\partial b_3^{\downarrow}(a)}{\partial a} = \frac{(2-q) - 2(1-q)a}{(2-q) - 2(1-q)b_3^{\downarrow}(a)}$$

and

$$\frac{\partial b_3^{\uparrow}(a)}{\partial a} = \frac{q + 2(1-q)a}{q + 2(1-q)b_3^{\uparrow}(a)}.$$

Simple calculations show that the stated inequalities hold, given that $b_3^{\downarrow}(a), b_3^{\uparrow}(a) < a$. Monotonicity and continuity of $b_3^{\uparrow}(a)$ and $b_3^{\downarrow}(a)$ allow the application of the intermediate value theorem, which implies the uniqueness of the intersection point $\bar{a}$. The second derivative of $b_3^{\downarrow}(a)$ is negative whenever

$$0 > -2(1-q)[2-q-2(1-q)b_3^{\downarrow}(a)] + 2(1-q)\frac{\partial b_3^{\downarrow}(a)}{\partial a}[2-q-2(1-q)a].$$

The right hand side of this expression is smaller than

$$-2(1-q)[2-q-2(1-q)b_3^{\downarrow}(a)] + 2(1-q)[2-q-2(1-q)a] = 2(1-q)2(1-q)(b_3^{\downarrow}(a) - a) < 0,$$

so that $b_3^{\downarrow}(a)$ is indeed concave.

The rest of the lemma is obtained by simply equalizing equations (12) and (13), which gives the condition $b_3^{\downarrow}(\bar{a}) = b_3^{\uparrow}(\bar{a}) = 1 - \bar{a}$. Plugging in one of the values of $b_3^{\downarrow}(a)$ or $b_3^{\uparrow}(a)$ for b allows to calculate $\bar{a}$ as stated in the lemma. □

[31] To handle situations in which they are smaller than zero and complex-valued, which is a relevant situation for the definition of the partition in the proposition, we could define $b_3^{\uparrow}(a)$ and $b_3^{\downarrow}(a)$ as functions with the same positive slope intersecting the horizontal axis at $b_3^{\uparrow}(a) = 0$, and $b_3^{\downarrow}(a) = 0$ (as defined above), respectively.

Lemma 4. *Let* $\underline{b_3^{\downarrow}}(a) := a - \frac{\delta q}{\tau}$ *and* $\overline{b_3^{\downarrow}} := b_3^{\downarrow}(\bar{a}) = \frac{1}{2} - \frac{\delta q}{2\tau}$. *For any* $a \leq \bar{a}$, $b_3^{\downarrow}(a)$ *lies in the interval* $[\underline{b_3^{\downarrow}}(a), \overline{b_3^{\downarrow}}]$. *If* $a > \bar{a}$, *then* $b_3^{\downarrow}(a) \in [\overline{b_3^{\downarrow}}, \underline{b_3^{\downarrow}}(a)]$. *Furthermore, it holds that* $b_3^{\uparrow}(a) > b_1(a)$ *and* $b_3^{\uparrow}(a) < a$, *if* $b \leq a$.

Proof. Lemma 3 implies that the distance $a - b_3^{\downarrow}(a)$ is increasing in a for all $a \leq \bar{a}$, hence

$$a - b_3^{\downarrow}(a) \leq \bar{a} - b_3^{\downarrow}(\bar{a}) = 2\bar{a} - 1$$

which, for any $a \leq \bar{a}$ gives a lower bound on $b_3^{\downarrow}(a)$ as $\underline{b_3^{\downarrow}}(a) := a - \frac{\delta q}{\tau}$. Since $b_3^{\downarrow}(a)$ is increasing in a, $\overline{b_3^{\downarrow}} := b_3^{\downarrow}(\bar{a})$ is an upper bound on $b_3^{\downarrow}(a)$ for all $a \leq \bar{a}$.
Define a'' as a s.t. $b_3^{\uparrow}(a) = 0$ and a''' as a s.t. $b_1(a) = 0$. As $\frac{\partial b_3^{\uparrow}(a)}{\partial a} > 1 = \frac{\partial b_1(a)}{\partial a}$, we need to show that $a'' < a'''$, in order to show $b_3^{\uparrow}(a) > b_1(a)$.

$$b_3^{\uparrow}(a) = 0 \Leftrightarrow a = \sqrt{\frac{\delta q}{\tau(1-q)} + \frac{q^2}{4(1-q)^2}} - \frac{q}{2(1-q)} =: a'',$$

and

$$a'' < a''' := \frac{\delta}{\tau} \Leftrightarrow 0 < \frac{\delta}{\tau},$$

which always holds under the assumptions.
As $\frac{\partial b_3^{\uparrow}(a)}{\partial a} > 1$, and as the slope of the diagonal through the action space is 1, we need to show that $b_3^{\uparrow}(1) < 1$ in order to show $b_3^{\uparrow}(a) < a$.

$$b_3^{\uparrow}(1) = \frac{q - \sqrt{q^2 + 4(1-q)(1 - \frac{\delta q}{\tau})}}{2q - 2} < 1 \Leftrightarrow \frac{4\delta q}{\tau}(q-1) < 0,$$

which always holds under the assumptions. □

Corollary 1. $b_1(a) < b_3^{\uparrow}(a) < a$ *and* $b_1(a) < b_3^{\downarrow}(a) < a$, $\forall (a,b) \in \mathcal{A} \times \mathcal{B}$ *with* $b \leq a$.

Proof. This follows from Lemma 3 and 4 . □

Corollary 2. *There is an equivalence between the five cases of Proposition B.1 and the partition described therein.*

Proof. The result follows immediately from equations (6), (7), (14) and (15) and the previous lemmata. □

Lemma 5. *If* $b \leq a$, *firm B's expected demand is strictly increasing in* a *and* b *in each element of partition* $\mathfrak{D}_L$. *The opposite is true for* $\mathfrak{D}_R$, *i.e. when* $b > a$.

Proof. The following reformulation of the second period parts of demand of an uninformed consumer will prove to be useful:

$$\begin{aligned} Pr(C_1)\tilde{x}(\beta^u_{C_1}) &= Pr(C_1)\left(\frac{a+b}{2} + \frac{\delta}{2\tau}\left(2\frac{Pr(C_1|v_L > v)\cdot Pr(v_L > v)}{Pr(C_1)} - 1\right)\right) \\ &= Pr(C_1)\tilde{x}(\beta^u_\emptyset) + \frac{\delta}{2\tau}\Big(Pr(C_1|v_L > v) - Pr(C_1)\Big). \end{aligned}$$

Note that a enters each of the demand parts in the same way as b (this will become even more obvious below) and thus whenever demand is increasing in b it also increases in a.

- Part 1: $\mathfrak{D}^1_L(a,b) := \{(a,b) \in \mathcal{A}\times\mathcal{B} \mid \tilde{x}(\beta^u_B) \in (b,a), \tilde{x}(\beta^u_A) \in (b,a), \tilde{x}(\beta^i_A)+\tilde{x}(\beta^i_B) = a+b\}$ With all indifferent types being at interior levels, they are symmetrically spread around $\tilde{x}(\beta^u_\emptyset)$ and $D^1_L(a,b)$ simplifies to:

  $$D^1_L(a,b) = q[2\tilde{x}(\beta^u_\emptyset)] + (1-q)[2\tilde{x}(\beta^u_\emptyset)] = 2\tilde{x}(\beta^u_\emptyset) = a+b.$$

  Obviously, $D^1_L(a,b)$ is increasing in b.

- Part 2: $\mathfrak{D}^2_L(a,b) := \{(a,b) \in \mathcal{A}\times\mathcal{B} \mid \tilde{x}(\beta^u_B) \in (b,a),\ \tilde{x}(\beta^u_A) \in (b,a),\ \tilde{x}(\beta^i_A)+\tilde{x}(\beta^i_B) = 1\}$ If the uninformed indifferent laggard and thus $\tilde{x}(\beta^u_{C1})$ is always in between the location of both firms, demand can be written as follows:

  $$\begin{aligned} D^2_L(a,b) = q + (1-q)\Bigg\{&\tilde{x}(\beta^u_\emptyset) + Pr(C_1 = B)\tilde{x}(\beta^u_\emptyset) + \frac{\delta}{2\tau}\Big(Pr(C_1 = B|v_L > v) - Pr(C_1 = B)\Big) \\ &+Pr(C_1 = A)\tilde{x}(\beta^u_\emptyset) + \frac{\delta}{2\tau}\Big(Pr(C_1 = A|v_L > v) - Pr(C_1 = A)\Big)\Bigg\} \end{aligned}$$

  With $Pr(C_1 = A) = 1-Pr(C_1 = B)$ and $Pr(C_1 = A|v_L > v) = 1-Pr(C_1 = B|v_L > v)$, this simplifies to

  $$D^2_L(a,b) = q + (1-q)(2\cdot\tilde{x}(\beta^u_\emptyset)) = q + (1-q)(a+b).$$

  Hence, for the case where $\tilde{x}(\beta^u_{C1})$ is between both firms' locations, B's demand increases linearly in b.

- Part 3A: $\mathfrak{D}^{3A}_L(a,b) := \{(a,b) \in \mathcal{A}\times\mathcal{B} \mid \tilde{x}(\beta^u_B) \in (b,a), \tilde{x}(\beta^u_A) = 0, \tilde{x}(\beta^i_A)+\tilde{x}(\beta^i_B) = 1\}$ In this part, uninformed laggards follow the choice if the early adopter chose A, if $C_1 = B$, the indifferent consumer in period 2 lies between the two firm's locations. Hence, B's

demand calculates as

$$\begin{aligned} D_L^{3A}(a,b) &= q + (1-q)\Big[\tilde{x}(\beta_\emptyset^u) + Pr(C_1 = B)\tilde{x}_2(B)\Big] \\ &= q + (1-q)\Big[\tilde{x}(\beta_\emptyset^u)\left(1 + \frac{q + (1-q)(a+b)}{2}\right) + \frac{q\delta}{4\tau}\Big] \end{aligned}$$

and the following derivative shows that $D_L^{3A}(a,b)$ is strictly increasing in b:

$$\frac{\partial D_L^{3A}}{\partial \tilde{x}(\beta_\emptyset^u)} = (1-q)\Big[\tfrac{1}{2}q + 1 + 2(1-q)\tfrac{a+b}{2}\Big].$$

- Part 3B: $\mathfrak{D}_L^{3B}(a,b) := \{(a,b) \in \mathcal{A} \times \mathcal{B} \mid \tilde{x}(\beta_B^u) = 1,\ \tilde{x}(\beta_A^u) \in (b,a),\ \tilde{x}(\beta_A^i) + \tilde{x}(\beta_B^i) = 1\}$

  If a purchase of B in the first period is always followed by an uninformed laggard, but not a purchase of A, demand of B is given by

$$\begin{aligned} D_L^{3B}(a,b) &= q + (1-q)\Big[\tilde{x}(\beta_\emptyset^u) + Pr(C_1 = B) + Pr(C_1 = A)\tilde{x}(\beta_A^u)\Big] \\ &= q + (1-q)\Big\{\tilde{x}(\beta_\emptyset^u) + \frac{q}{2} + (1-q)\tilde{x}(\beta_\emptyset^u) + Pr(C_1 = A)\tilde{x}(\beta_\emptyset^u) \\ &+ \frac{\delta}{2\tau}\Big(Pr(C_1 = A|v_L > v) - Pr(C_1 = A)\Big)\Big\} \\ &= q + (1-q)\left\{\tilde{x}(\beta_\emptyset^u)\Big[2 - \tfrac{q}{2} + (1-q)(1-\tilde{x}(\beta_\emptyset^u))\Big] - \frac{q\delta}{4\tau} + \frac{q}{2}\right\}, \end{aligned}$$

  which is quadratic in $\tilde{x}(\beta_\emptyset^u)$ and thus in b. Nevertheless, the derivative:

$$\frac{\partial D_L^{3B}}{\partial \tilde{x}(\beta_\emptyset^u)} = \Big[3 - \tfrac{3}{2}q - 2(1-q)\tfrac{a+b}{2}\Big](1-q)$$

  shows that it is strictly increasing in b for the relevant values of a and b.

- Part 4: $\mathfrak{D}_L^4(a,b) := \{(a,b) \in \mathcal{A} \times \mathcal{B} \mid \tilde{x}(\beta_B^u) = 1,\ \tilde{x}(\beta_A^u) = 0,\ \tilde{x}(\beta_A^i) + \tilde{x}(\beta_B^i) = 1\}$

  $\tilde{x}(\beta_B^u) = 1$ and $\tilde{x}(\beta_A^u) = 0$ means that an uninformed laggard always follows the lead of the early adopter. The demand in such a case is described by

$$\begin{aligned} D_L^4(a,b) &= q + (1-q)\Big[\tilde{x}(\beta_\emptyset^u) + Pr(C_1 = B)\Big] \\ &= q + (1-q)\Big[(2-q)\Big(\frac{a+b}{2}\Big) + \frac{q}{2}\Big] \end{aligned}$$

  Demand in this case is linear, and increasing in b.

  Inspection of the different demand parts shows that updating of the uninformed laggards and thus the shifting of the indifferent consumer types is symmetric in parts D^2 and D^4 and asymmetric in parts D^{3B} and D^{3A}. Only in the latter cases does the demand depend on the parameters δ and τ. Furthermore, $\partial D^j/\partial b \downarrow$ as $j \uparrow$ with $j \in \{1, 2, 3B, 3A, 4\}$.

□

Combining the lemmata yields the result.

□

Lemma 6. *$b_3^{\downarrow}(a)$ and $b_3^{\uparrow}(a)$ converge to $b_1(a)$, as $q \to 1$. Also, $b_3^{\downarrow}(a)$ and $b_3^{\uparrow}(a)$ converge to a, as $q \to 0$. Thus, increasing (decreasing) $\frac{\delta}{\tau}$ "stretches" ("compresses") the graphs of $b_3^{\uparrow}(a)$ and $b_3^{\downarrow}(a)$ (compare Figure 3).*

Proof. Straightforward calculations show that the derivative of the numerator of $b_3^{\downarrow}(a)$ w.r.t. q is given by

$$-1 - \frac{(2a-1)[2a(q-1)-q+2]\tau + \delta(2-4q)}{\tau\sqrt{\frac{\tau[q-2-2a(q-1)]^2 - 4\delta(q-1)q}{\tau}}}$$

Similarly, the derivative of the numerator of $b_3^{\uparrow}(a)$ w.r.t. q can be calculated as

$$1 - \frac{(2a-1)[2a(q-1)-q]\tau + \delta(4q-2)}{\tau\sqrt{\frac{4(q-1)\{a\tau[a(q-1)-q]+\delta q\}}{\tau} + q^2}}$$

Applying l'Hopital's rule then yields the results for $q \to 1$. The limits for $q \to 0$ can be obtained directly.

Using these results, and as $b_1(a) = a - \frac{\delta}{\tau}$, $b_1(a) < b_3^{\uparrow} < a$ and $b_1(a) < b_3^{\downarrow} < a$, one can easily observe that increasing (decreasing) $\frac{\delta}{\tau}$ "stretches" ("compresses") the graphs of $b_3^{\uparrow}(a)$ and $b_3^{\downarrow}(a)$. □

C Proof of Proposition 2 (Main Result)

Remember that for any equilibrium with $b^* < a^*$ there exists an analogous equilibrium with $1 - b^* > 1 - a^*$. Thus, when we assume $a \geq 0.5$ in the following, this is without loss of generality.

Lemma 7. *B's best response $b^*(a)$ to any $a \in [0.5, 1]$ is a subset of $\{b_1(a), b_3^{\downarrow}(a), a, 1 - b_3^{\downarrow}(1-a)\}$.*

Proof. By Lemma 5 all parts of B's demand are increasing in b. The highest demand in each part is thus obtained at the highest possible value, belonging to this part. Clearly it can never be optimal to choose $b = b_3^{\uparrow}(a)$, since B's demand has an upward jump at this point,

so any slightly larger b will increase the demand. For $0.5 \leq a \leq \bar{a}$, the demands $D_L(a,b)$ and $D_R(a,b)$ consist of the same parts, so the optimum of B's demand must be obtained for some $b \in \mathfrak{D}_L$, i.e. $b^*(a) \leq a \leq \bar{a}$, as B's demand increases in a. If, however, $0.5 \leq \bar{a} \leq a$, then $D_L(a,b)$ and $D_R(a,b)$ do not consist of the same parts anymore. On the left side, part 3 of B's demand is given by $D_L^{3A}(a,b)$, but on the right, the demand part 3 consists of $D_R^{3B}(a,b)$. Hence, the demand maximizing location for firm B might be at the optimum of $D_R^{3B}(a,b)$, which is calculated as $1 - b_3^{\downarrow}(1-a)$. □

Corollary 3. *The value function of firm B in the respective regions of the action space is given by*

$$V_L^1(a) = \begin{cases} 0 & \textit{if} \quad b_1(a) < 0, \\ D_L^1(a, b_1(a)) & \textit{else.} \end{cases}$$

$$V_L^2(a) = \begin{cases} D_L^2(a, b_3^{\uparrow}(a)) & \textit{if} \quad a \leq \bar{a}, \\ \lim\limits_{b \to b_3^{\downarrow}(a)} D_L^2(a,b) & \textit{if} \quad a > \bar{a}, \\ 0 & \textit{if} \quad \min\{b_3^{\uparrow}(a),\ b_3^{\downarrow}(a)\} < 0. \end{cases}$$

$$V_L^3(a) = \begin{cases} D_L^{3B}(a, b_3^{\downarrow}(a)) & \textit{if} \quad a \leq \bar{a}, \\ \lim\limits_{b \to b_3^{\uparrow}(a)} D_L^{3A}(a,b) & \textit{if} \quad a > \bar{a}, \\ 0 & \textit{if} \quad \max\{b_3^{\uparrow}(a),\ b_3^{\downarrow}(a)\} < 0. \end{cases}$$

$$V_L^4(a) = D_L^4(a,a)$$

$$V_R^3(a) = \begin{cases} 0 & \textit{if} \quad 1 - b_3^{\downarrow}(a) > 1, \\ D_R^{3B}(a, 1 - b_3^{\downarrow}(1-a)) & \textit{else.} \end{cases}$$

Lemma 8. *For $a \leq \bar{a}$, $V_L^4(a) > V_L^2(a) > V_L^1(a)$. This implies that for $a \leq \bar{a}$, firm B will choose a location such that $(a,b) \notin \mathfrak{D}_L^1$ and $(a,b) \notin \mathfrak{D}_L^2$.*

Proof. Remember that all demand parts are increasing in a and b. With $b_3^{\downarrow}(\bar{a}) = 1 - \bar{a}$, we can calculate $V_L^1(\bar{a}) = D_L^1(\bar{a}, b_1(\bar{a})) = 2\bar{a} - \delta/\tau = 1 + \delta q/\tau - \delta/\tau < 1$ and $V_L^2(\bar{a}) = D_L^2(\bar{a}, 1-\bar{a}) = q + (1-q)(\bar{a} + 1 - \bar{a}) = 1$. For V_L^4 we know $D_L^4(\bar{a}, 1-\bar{a}) = q + (1-q)(\bar{a} + 1 - \bar{a}) =$

$1 < D_L^4(\bar{a}, \bar{a}) = V_L^4(\bar{a})$. As $\partial D_4(a,b)/\partial a = (1-q)(2-q)/2 < \partial D_1(a,b)/\partial a = 1$, and $V_L^2(0.5) = D_L^2(0.5, b_3^{\uparrow}(0.5)) < D_L^2(0.5, 0.5) = D_L^4(0.5, 0.5) = V_L^4(0.5)$, the result follows. □

Lemma 9. *B's demand at the optimum of any part of D_L is increasing in a, i.e. $\partial V_L^j(a)/\partial a > 0$ $\forall j \in \{1,2,3,4\}$. Furthermore, $D_L^{3B}(a, b_3^{\downarrow}(a))$ is concave, while $V_L^j(a)$ is linear in a for $j \in \{1,4\}$.*

Proof. By Lemma 5 all parts of $D_L(a,b)$ are increasing in a and b. Since all potentially optimal locations of B, i.e. $b_1(a), b_3^{\downarrow}(a), b_3^{\uparrow}(a)$ and a, are increasing in a, the first part of the lemma follows. For B's demand $D_L^{3B}(a, b_3^{\downarrow}(a))$, we can calculate the derivative w.r.t. a as follows

$$\frac{\partial D_L^{3B}(a, b_3^{\downarrow}(a))}{\partial a} = (1-q)\left(1 + \frac{\partial b_3^{\downarrow}(a)}{\partial a}\right)\left[\frac{4-q}{4} + (1-q)\frac{1-a-b_3^{\downarrow}(a)}{2}\right].$$

This derivative is positive whenever

$$b_3^{\downarrow}(a) \leq \frac{6-3q}{2-2q} - a$$

which always holds for $a \in [0,1]$.

The second derivative of $D_L^{3B}(a, b_3^{\downarrow}(a))$ w.r.t a is negative if

$$\frac{\partial^2 b_3^{\downarrow}(a)}{\partial a^2}\left(\frac{4-q}{4} + (1-q)\frac{1-a-b_3^{\downarrow}(a)}{2}\right) + \left(1 + \frac{\partial b_3^{\downarrow}(a)}{\partial a}\right)\left(-\frac{1-q}{2}\left(1 + \frac{\partial b_3^{\downarrow}(a)}{\partial a}\right)\right) < 0$$

which can seen to be the case using the results of Lemma 3. Linearity can directly be seen from the linearity of $b_1(a), b = a$ and from the demand $D_L^j,\ j \in \{1,4\}$ calculated in Lemma 5. □

Lemma 10. *For $a \geq 0.5$, $V_L(a) := \max V_L^j(a), j \in \{1,2,3,4\}$ is increasing in a for all $a \in \mathcal{A} \setminus \bar{a}$ and the function $V_R^{3B}(a)$ is decreasing in a for all $a \geq 0.5$.*

Proof. By Lemma 9, given $b \leq a$, B's demand is increasing in a at the maximizing b in any single part. Below and above $a = \bar{a}$, the maximum taken over increasing functions must thus be increasing in a. Note that, whenever $V_L^4(\bar{a}) \geq V_L^{3B}(\bar{a})$, $V_L(a)$ is increasing in a for all $a \in \{\mathcal{A} \cup \bar{a} \mid a \geq b\}$. Since $\bar{a} > 0.5$ symmetry implies that $V_R^{3B}(a)$ is decreasing for all $b > a$. □

Lemma 11.

$$V_L^4(0.5) < V_L^{3B}(0.5), \tag{BDC}$$

implies that $\bar{a} < 1$ *and* $b_3^{\downarrow}(0.5) > 0$.

Proof. An upper bound of $V_L^{3B}(0.5) = D_L^{3B}(0.5, b_3^{\downarrow}(0.5))$ is given by

$$\begin{aligned} D_L^{3B}(0.5, 0.5(1-\delta q/\tau)) &= q + (1-q)\left\{\left(\frac{1}{2} - \frac{\delta q}{4\tau}\right)\left(2 - \frac{q}{2} + (1-q)\left(\frac{1}{2} + \frac{\delta q}{4\tau}\right)\right) - \frac{\delta q}{4\tau} + \frac{q}{2}\right\} \\ &= q + (1-q)\left\{\frac{5}{4} - \frac{\delta q}{4\tau}\left(3 - \frac{q}{2} + (1-q)\frac{\delta q}{4\tau}\right)\right\}. \end{aligned}$$

In order for this upper bound to be larger than 1 for some q, $\delta q/\tau$ must be smaller than $2/5$. By Lemma 4, $b_3^{\downarrow}(a) \in \left[a - \frac{\delta q}{\tau}, \frac{1}{2} - \frac{\delta q}{2\tau}\right]$, meaning that, given $\delta q/\tau < 2/5$, $b_3^{\downarrow}(a) > 0$ if $a > 2/5$ and. Additionally, $\bar{a} = 0.5 + \delta q/(2\tau) < 7/10$.

Lemma 12. *If* (ADC) *does not hold, that is,*

$$V_L^{3B}(0.5) > V_L^4(\bar{a}),$$

$\bar{a} < 2/3 < 1$ *and* $1 - b_3^{\downarrow}(1-\bar{a}) < 1$. *Furthermore,* $\exists a' < 2/3$ *s.t.* $V_L^4(a') = V_R^{3B}(a')$.

If (ADC) is violated, it must be that

$$\begin{aligned} V_L^4(\bar{a}) = D_L^4(\bar{a}, \bar{a}) = q + (1-q)\left(1 + (1-0.5q)\frac{\delta q}{\tau}\right) &< V_L^3(0.5) = D_L^{3B}(0.5, b_3^{\downarrow}(0.5)) \\ &< D_L^{3B}(0.5, 0.5) = q + (1-q)\left\{\frac{1}{2}\left(2 - \frac{q}{2} + (1-q)\frac{1}{2}\right) - \frac{\delta q}{4\tau} + \frac{q}{2}\right\} \\ &= q + (1-q)\left\{\frac{5}{4} - \frac{\delta q}{4\tau}\right\} \end{aligned}$$

which implies

$$1 + \frac{\delta q}{\tau}(1-0.5q) < \frac{5}{4} - \frac{\delta q}{4\tau} \Leftrightarrow \frac{\delta q}{\tau}\left(\frac{5}{4} - \frac{q}{2}\right) < \frac{1}{4}.$$

For this equation to be fulfilled for some q, it must be that $\frac{\delta q}{\tau} < \frac{1}{3}$. In this case, $\bar{a} = 0.5 + \delta q/2t < 2/3$ and $b_3^{\downarrow}(a) > 0$ if $a > 1/3$, so that $1 - b_3^{\downarrow}(1-\bar{a}) < 1$.

The demand $D_L^4(2/3, 2/3)$ calculates as

$$\begin{aligned} D_L^4(2/3, 2/3) = q + (1-q)\left((2-q)\frac{2}{3} + \frac{q}{2}\right) &= q + (1-q)\left(\frac{4}{3} - \frac{q}{6}\right) \\ > q + (1-q)\left(\frac{5}{4} - \frac{\delta q}{4\tau}\right) &= D_L^{3B}(0.5, 0.5) > D_L^{3B}(1 - 2/3, b_3^{\downarrow}(1 - 2/3)) \end{aligned}$$

As $V_L^4(a) = D_L^4(a, a)$ is increasing and $V_R^{3B}(a) = D_L^{3B}(1-a, b_3^{\downarrow}(1-a))$ is decreasing in a, the remainder of the lemma follows. □

Proposition 2. *In the model with consumer learning we obtain the following results.*

1. *If equation* (BDC) *is violated, the strategies from the benchmark model constitute the unique equilibrium characterized by* $a^* = b^* = 0.5$ *(Symmetric Minimum Differentiation Equilibrium).*

2. *Equations* (BDC) *and* (ADC) *are necessary and sufficient conditions so that the locations are* $a^* = 0.5$ *and* $b^* < 0.5$ *in the unique equilibrium (Central Differentiation Equilibrium).*

3. *If equations* (ADC) *and* (BDS) *do not hold the locations are* $a^* = b^* > 0.5$ *in the unique equilibrium (Asymmetric Minimum Differentiation Equilibrium).*

4. *Equation* (BDS) *is a necessary and sufficient condition so that there exist an equilibrium (not necessarily unique) with locations* $b^* > a^* > \bar{a} > 0.5$ *(Short Side Differentiation Equilibrium).*

Uniqueness is up to symmetry, as to any equilibrium with (a^*, b^*) *there exists an analogous equilibrium with* $(1 - a^*, 1 - b^*)$.

Proof. A's objective is given by $\min_a V(a) := \max\{V_L(a), V_R(a)\}$. It is useful to notice, that whenever $V_R^{3B}(\bar{a}) \leq V_L^4(\bar{a})$, as the former is decreasing in a and the latter is increasing in a, it must be that $V(a) = V_L(a)$.

- Central Differentiation Equilibrium: $b^* = b_3^{\downarrow}(0.5) < a^* = 0.5$

  $V_L^{3B}(0.5) \leq V_L^4(\bar{a})$ implies $D_R^{3B}(\bar{a}) < V_L^{3B}(0.5) \leq V_L^4(\bar{a})$ and thus $V(a) = V_L(a)$. By Lemma 10, $V_L(a)$ is increasing in a if $V_L^{3B}(0.5) \leq V_L^4(\bar{a})$, so A's optimal choice is the smallest possible a, given by $a^* = 0.5$. As $V_L^4(0.5) < V_L^{3B}(0.5)$, $b^*(0.5) = b_3^{\downarrow}(0.5)$.

- Asymmetric Minimum Differentiation Equilibrium: $b^* = a^* = \bar{a}$

  $V_L^{3B}(0.5) > V_L^4(\bar{a}) > D_R^{3B}(\bar{a})$ again implies $V(a) = V_L(a)$. In contrast to the previous case $V_L(a)$ has a downward jump at $a = \bar{a}$. By Lemma 10 it is increasing in a at all other points. As $V_L^{3B}(0.5) > V_L^4(\bar{a})$, $a = \bar{a}$ is the unique minimizer of $V_L(a)$ in this case. Since demand part $D^{3B}(a, b)$ exists only for $a < \bar{a}$, B's best response is given by $b^*(\bar{a}) = \bar{a}$.

- Short Side Differentiation Equilibrium: $b^* = 1 - b_3^{\downarrow}(1-a) > a^* > \bar{a}$

  If $V_R^{3B}(\bar{a}) > V_L^4(\bar{a})$, there is some $a' > \bar{a}$ is such that $V(a') = V_R^{3B}(a') = V_L^4(a')$, as V_R^{3B} decreases in a and $V_L^4(a)$ increases in a. Clearly, a' minimizes $V(a)$. B's best response to this a' is not unique, since, by construction, B is indifferent between choosing either the $b \leq a$ maximizing demand or the $b > a$ maximizing demand. Furthermore, the optimal $b \leq a$ can be any of the set $\{b_1(a'), b_3^{\downarrow}(a'), a'\}$. The optimal $b > a$, however, is given by $1 - b_3^{\downarrow}(1-a')$, which by Lemma 12 exists in the action space.

Note that what distinguishes the last two equilibria, is whether $a' \leq \bar{a}$ or $a' > \bar{a}$. □

Corollary 4. *Sufficient conditions for the conditions of Proposition 2 can be calculated as follows:*

$$\frac{\delta q}{\tau} < 0.192 \Rightarrow D_L^{3B}(0.5, b_3^{\downarrow}(0.5)) > D_L^4(0.5, 0.5) \tag{BDC}$$

$$\frac{\delta q}{\tau} > 0.166, q < 0.4 \Rightarrow D_L^{3B}(0.5, b_3^{\downarrow}(0.5)) \leq D_L^4(\bar{a}, \bar{a}) \tag{ADC}$$

$$\frac{\delta q}{t} < 0.079 \Rightarrow D_L^{3B}(1-\bar{a}, b_3^{\downarrow}(1-\bar{a})) > D_L^4(\bar{a}, \bar{a}) \tag{BDS}$$

Proof. Using the upper and lower bound of $b_3^{\downarrow}(a)$ as calculated in Lemma 4, sufficient conditions for the conditions of Proposition 2 can be calculated as follows:

$$\begin{aligned}\frac{\delta q}{\tau} < 0.192 \Rightarrow \frac{\delta q}{\tau}\left((1-q)\frac{\delta q}{\tau} + 5 - q\right) < 1 \Leftrightarrow D_L^{3B}(0.5, \underline{b_3^{\downarrow}}(a)) > 1 \\ \Rightarrow D_L^{3B}(0.5, b_3^{\downarrow}(0.5)) > D_L^4(0.5, 0.5)\end{aligned} \tag{BDC}$$

$$\begin{aligned}\frac{\delta q}{\tau} > 0.166, q < 0.4 \Rightarrow 1 \leq \frac{\delta q}{2\tau}\left[\frac{\delta q}{2\tau}(1-q) + 14 - 5q\right] \Leftrightarrow D_L^{3B}(0.5, \overline{b_3^{\downarrow}}(a)) \leq D_L^4(\bar{a}, \bar{a}) \\ \Rightarrow D_L^{3B}(0.5, b_3^{\downarrow}(0.5)) \leq D_L^4(\bar{a}, \bar{a})\end{aligned} \tag{ADC}$$

If both those sufficient conditions are fulfilled, equilibrium locations are $b = b_3^{\downarrow}(0.5) < a = 0.5$.

$$\begin{aligned}\frac{\delta q}{t} < 0.079 \Rightarrow 1 > \frac{\delta q}{\tau}\left(13 - 4q + 4(1-q)\frac{\delta q}{\tau}\right) \Leftrightarrow D_L^{3B}(1-\bar{a}, 1-\bar{a}) > D_L^4(\bar{a}, \bar{a}) \\ \Rightarrow D_L^{3B}(1-\bar{a}, b_3^{\downarrow}(1-\bar{a})) > D_L^4(\bar{a}, \bar{a})\end{aligned} \tag{BDS}$$

If this holds, there is an equilibrium with $b = 1 - b_3^{\downarrow}(1-a) > a > \bar{a}$. Note that we do not derive sufficient conditions for the equilibrium with asymmetric minimum differentiation, i.e. $b = a = \bar{a}$, as these would be "too small" in the parameter space. □

References

Banerjee, Abhijit V. 1992. "A Simple Model of Herd Behavior." *The Quarterly Journal of Economics*, *107* (3): 797–817.

Belleflamme, Paul and Martin Peitz. 2010. *Industrial Organization: Markets and Strategies*, Cambridge: Cambridge University Press.

Bester, Helmut. 1998. "Quality Uncertainty Mitigates Product Differentiation." *The RAND Journal of Economics*, *29*, 828–844.

Bester, Helmut, André De Palma, Wolfgang Leininger, Jonathan Thomas, and Ernst-Ludwig Von Thadden. 1996. "A Noncooperative Analysis of Hotelling's Location Game." *Games and Economic Behavior*, *12*, 165–186.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades." *Journal of Political Economy*, *100* (5): 992–1026.

Brekke, Kurt R., Luigi Siciliani, and Odd Rune Straume. 2011. "Hospital Competition and Quality with Regulated Prices*." *Scandinavian Journal of Economics*, *113* (2): 444–469.

Brekke, Kurt, Robert Nuscheler, and Odd Rune Straume. 2006. "Quality and Location Choices under Price Regulation." *Journal of Economics and Management Strategy*, *15*, 207–227.

Caminal, Ramon and Xavier Vives. 1996. "Why Market Shares Matter: An Information-Based Theory." *RAND Journal of Economics*, *27*, 221–239.

Chen, Yi-Fen. 2008. "Herd behavior in purchasing books online." *Computers in Human Behavior*, *24*, 1977–1992.

Conze, Maximilian and Michael Kramm. 2013. "The Recommendation Effect in the Hotelling Game - A New Result for an Old Model." *Bonn Econ Discussion Papers*, *14*, 1–36.

Conze, Maximilian and Michael Kramm. "Consumer Learning and Incentives to Differentiate in Cournot and Bertrand Competition." 2016.

Courty, Pascal. 2000. "An Economic Guide to Ticket Pricing in the Entertainment Industry." *Louvain Economic Review*, *66* (1): 167–192.

d'Aspremont, Claude, Jean J. Gabszewicz, and Jacques F. Thisse. 1979. "On Hotelling's "Stability in Competition"." *Econometrica*, *47* (5): 1145–1150.

De Vany, Arthur. 2006. "The Movies." in Victor A. Ginsburg and David Throsby, eds., Vol. 1 of *Handbook of the Economics of Art and Culture*, Elsevier, 2006. pp. 615 – 665.

Dos Santos Ferreira, Rodolphe and Jacques François Thisse. 1996. "Horizontal and vertical differentiation: the Launhardt model." *International Journal of Industrial Organization*, *14*, 485–506.

Economides, Nicholas. 1989. "Quality Variations and Maximal Variety Differentiation." *Regional Science and Urban Economics*, *19*, 21–29.

Gabszewicz, Jean J. and Jacques François Thisse. 1986. "On the nature of competition with differentiated products." *The Economic Journal*, *96*, 160–172.

Gabszewicz, Jean J. and Xavier Y. Wauthy. 2011. "Nesting horizontal and vertical differentiation." *Regional Science and Urban Economics*, *42*, 998–1002.

Gale, Douglas. 1996. "What Have We Learned From Social Learning?" *European Economic Review*, *40*, 617–628.

Gilchrist, Duncan Sheppard and Emily Glassberg Sands. 2016. "Something to Talk About: Social Spillovers in Movie Consumption." *Journal of Political Economy*, *124* (5): 1339–1382.

Gravelle, Hugh and Peter Sivey. 2010. "Imperfect information in a quality-competitive hospital market." *Journal of Health Economics*, *29*, 524–535.

Hotelling, Harold. 1929. "Stability in Competition." *The Economic Journal*, *39* (153): 41–57.

Kamenica, Emir and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review*, *101*, 2590–2615.

Miklos-Thal, Jeanine and Juanjuan Zhang. 2013. "(De)marketing to Manage Consumer Quality Inferences." *Journal of Marketing Research*, *50*, 55–69.

Moretti, Enrico. 2011. "Social Learning and Peer Effects in Consumption: Evidence from Movie Sales." *Review of Economic Studies*, *78*, 356–393.

Orbach, Barak Y. and Liran Einav. 2007. "Uniform Prices for Differentiated Goods: The Case of the Movie-Theater Industry." *International Review of Law and Economics*, *27*, 129–153.

Page, Scott E. 2008. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*, Princeton: Princeton University Press.

Parakhonyak, Alexei and Nick Vikander. "Inducing Herding with Capacity Constraints." Department of Economics Discussion Paper Series 808, University of Oxford 2016.

Ridley, David B. 2008. "Herding versus Hotelling: Market Entry with Costly Information." *Journal of Economics & Management Strategy*, *17* (3): 607–631.

Simon, Leo K. and William R. Zame. 1990. "Discontinuous Games and Endogenous Sharing Rules." *Econometrica*, *58*, 861–872.

Smith, Lones and Peter Sørensen. 2000. "Pathological Outcomes of Observational Learning." *Econometrica*, *68*, 371–398.

Sorensen, Alan T. 2007. "Bestseller Lists and Product Variety." *Journal of Industrial Economics*, *55*, 715–738.

Surowiecki, James. 2005. *The Wisdom of Crowds*, New York: Anchor.

Tabuchi, Takatoshi and Jacques-Franqois Thisse. 1995. "Asymmetric equilibria in spatial competition." *International Journal of Industrial Organization*, *13*, 213–227.

Tucker, Catherine and Juanjuan Zhang. 2011. "How Does Popularity Information Affect Choices? A Field Experiment." *Management Science*, *57* (5): 828–842.

Zahavi, Amotz. 1975. "Mate Selection—A Selection for a Handicap." *Journal of Theoretical Biology*, *53* (1): 205–214.