# Essays on Information Transmission and its Effects in Markets with Imperfect Competition

Inaugural-Dissertation

zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

MAXIMILIAN CONZE

aus Stuttgart
2017

| | |
|---|---|
| Dekan: | Prof. Dr. Daniel Zimmer, LL.M. |
| Erstreferent: | Prof. Dr. Dennis Gärtner |
| Zweitreferent: | Prof. Dr. Hendrik Hakenes |
| Tag der mündlichen Prüfung: | 11.05.2017 |

# Contents

# List of Figures

# Acknowledgments

First and foremost, I would like to thank my supervisor Dennis Gärtner for his exceptional supervision and his steady support. I have greatly benefited from his advice and his valuable comments, particularly during the times when I started working on each of the different chapters of this thesis. Discussions with my second supervisor Hendrik Hakenes have been very insightful and I am grateful for his remarks.

I want to especially thank my co-author and friend Michael Kramm for his effort and for a great amount of time spent with discussions on and off the subjects of this thesis. I am also thankful to my fellow doctoral students, in particular to Anne-Katrin Roesler and Thomas Nebeling, without whom the last years would have been far less enjoyable.

Finally, I am deeply grateful to my family and friends for their never ending support and encouragement.

# Introduction

This thesis consists of four essays in the field of industrial organization. The essays explore questions of information transmission and its influence in markets characterized by imperfect competition.

The first part of this thesis, chapters one and two, examines the incentives of a manufacturer to communicate and reveal private information about his product to consumers in a setting where consumers are not served by the manufacturer directly but instead purchase the manufacturer's good from a retailer who resells the good bought from the producer. The second part, chapters three and four, which are joint work with Michael Kramm, analyze the effects of consumer learning on how firms design their products with respect to the products of the competitors.

In the first chapter, a situation where consumers consider purchasing a good which is of random and unknown quality is modeled. All consumers prefer goods of higher quality, but differ in how much so. As consumers are not aware of the quality of the good until after they have consumed it, they do not exactly know their valuation for the good when making their purchases. The good is produced by a monopolistic manufacturer who then sells it to a monopolistic retailer and the retailer finally offers the good to consumers. While it is long known, that the retail price might reveal some information about the good to consumers the novelty of this chapter is the introduction of an additional, direct way of communication between the manufacturer and consumers. When making his sales to the retailer, the manufacturer can send a costless and unverifiable cheap-talk signal about the quality to consumers. The costless and unverifiable nature of the communication might for example correspond to claims about a product a firm makes in an advertising campaign or it can be seen to represent retail price recommendations, which are prevalent in reality but whose informativeness seems questionable.

While it seems natural to assume that the manufacturer in this case would always exaggerate the true quality in order to maximize consumers' willingnesses to pay, the results show that the manufacturer will truthfully reveal information about the quality, that is if the quality of the good is low the manufacturer will also inform consumers about this fact. By doing so, the manufacturer manages consumers expectations and if consumers' expect the good to be of low quality, they don't leave much room for the retailer to charge high prices which is shown to be profitable for the manufacturer.

The second chapter applies a similar model of a vertical supply chain, but instead of using cheap-talk to communicate with consumers, the manufacturer here only has the costly possibility to truthfully reveal, or disclose, the quality of his product (or remain silent). This is meant to model situations where producers let third-parties test and certify the quality of their products.

In the previous literature on information disclosure, it has been shown to be the case that firms prefer to disclose the quality of their product whenever it is highest, as consumers would otherwise expect the quality to be worse. This motive is still in place in this model, but an additional motive is generated by the vertical chain setup of the model. By disclosing low quality levels, the manufacturer limits the retailers freedom in setting the retail price in order to mitigate the well known double-marginalization problem and to increase the manufacturers profits.

The third chapter introduces consumer learning in a version of Hotelling's model (1929) of spatial competition. Two firms produce goods that are vertically and horizontally differentiated, that is the goods differ in their characteristics and in their quality. The firms choose their product location, or their product's characteristic, on the unit interval in order to compete for consumers whose preferences are uniformly distributed on said interval and the quality difference of the two products is randomly determined. The two consumers have differential information about the goods' qualities and sequentially purchase one of the two goods.

With consumer learning, firms are confronted with two offsetting effects: producing a niche-product 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. It is shown 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.

The fourth chapter applies this reasoning to the differentiated duopoly model of Dixit (1979). The model of this chapter combines two extensions of the mentioned model to analyze the effect of consumer learning on firms' incentives to differentiate their products in models of Cournot and Bertrand competition.

Products are of different quality, consumers buy sequentially and are imperfectly informed about the quality of the goods. Before simultaneously competing in quantities, firms simultaneously choose their investment into differentiation. Investments have to be higher the higher the firms want to differentiate their products from each other or equivalently, the less substitutable they want their products to be. Late consumers can observe earlier consumers' decisions and extract information about the quality of the goods. This possibility influences the firms' incentives to differentiate. If firms compete in quantities, they are more likely to invest in differentiation with consumer learning than without. This is in line with implications of the recommendation effect introduced in the third chapter. The opposite is true if firms compete in prices. Here, the effect of consumer learning is reversed, so that differentiation is less likely with consumer learning.

# Chapter 1

# Informative Advertising in Vertical Relationships - Using Cheap-Talk to Reveal Unfavorable Information

Many advertisements, for example on TV, contain descriptions about quality aspects of a good. This is especially true for advertisements of manufacturing firms which (also) indirectly sell through retailers. A priori, it is not clear how valuable this information is for consumers. In contrast to much of the existing literature, interpreting the target of informative advertising to be horizontal, that is attracting consumers from competing firms, a different explanation is presented here.

A simple model of vertical relations, where consumers differ in their valuation of quality, is presented. It is shown that a monopolistic manufacturer can use advertising to transmit private information about the quality of his good to consumers, thereby influencing the incentives of a (monopolistic) retailer and reducing the severity of the well-known double marginalization problem. The cases of experience goods and search goods are distinguished and it is shown that information transmission is especially effective if consumers find it difficult to determine the quality of the good in question.

## 1.1 Introduction

In many situations, consumers are uncertain about aspects of a good they plan to purchase. This uncertainty about the characteristics of the good often means that consumers are not able to ascertain their valuations of the good in advance.

The connection between uncertainty about aspects of the good and consumers' valuation is especially obvious if the the quality of the good is the aspect the consumer is in doubt of.

What may make the assessment of a good's quality even harder for consumers is a situation in which the good is not sold by the manufacturer directly but instead through some kind of vertical chain. In such a setting every agent has its own incentives which the consumer has to take into consideration when trying to assess information about the quality of the good in question.

At the same time it seems prevalent that manufacturers send large amounts of information to consumers, for example through TV advertising where characteristics are explicitly stated or through using reviews carried out by third parties to convince consumers of the superiority of their good.

The problem in those situations, as mentioned above, is that the manufacturer clearly follows his own goals when sending out such information and there is good reason for the consumer to question the credibility of the information provided.

Building on this motivation, this chapter will model a situation where a monopolistic manufacturer produces a good with a quality that is unknown to consumers and sold by a monopolistic retailer. All consumers prefer goods of higher quality, but they differ in how much they prefer high quality goods over low quality ones. Since the quality is unknown to consumers, it is impossible for them to determine their exact valuation for the good they consider purchasing. The question then will be if the manufacturer can credibly reveal part of his private information to consumers via cheap-talk and why he would choose to do so.

It is easy to see that this can not happen in a model where the manufacturer directly sells his goods to consumers. In such a model the manufacturer gets the full retail price and since consumers' utility is decreasing in the charged price, when they buy the good, the interests of consumers' and the manufacturer are completely opposed. All consumers' utilities are increasing in the quality and they all possess the same information, so that if any signal the manufacturer directs to consumers would influence their willingness to pay (WTP), the manufacturer would always send the signal which maximizes consumers' WTP. In other words, because of the opposing interests and the costless nature of cheap-talk signals, there is no way to align the incentives of consumers and the manufacturer in a way that opens the possibility of informative signaling as in Crawford and Sobel (1982).

The situation is quite different if we are dealing with a vertical chain setting where the manufacturer sells his good through a retailer. For such a setup it will be shown that the manufacturer can indeed reveal part of his private information and he will choose to do so when he is given the opportunity. The interesting question here is why he would reveal that his good is of relatively low quality to consumers, even though their valuations are increasing in quality, especially since he has no incentive to do so in the direct selling setup. It will be shown that by generating two opposing effects, the vertical chain setup is crucial in enabling the manufacturer to credibly reveal parts of his private information.

While the effect in place in the direct selling setup, namely that signaling higher qualities increases consumers' WTP is still present in a vertical chain model, a second effect emerges. Since the manufacturer does not directly sell his good to consumers but does so only through a retailer, he has to make sure that the retailer has no incentive to deviate from the proposed strategy. Intuitively, both the retailer and the manufacturer have to decide between setting high prices and selling to few consumers or setting low prices and selling to many consumers. In the present model, the gap between consumers' valuations increases for higher qualities, which means that it will get harder for the manufacturer to make the retailer stick to selling to many consumers at a relatively low price the higher consumers expect the quality to be. Put differently, by signaling a low quality good to consumers, the manufacturer decreases the incentives of the retailer to deviate from a strategy that prescribes serving many customers at a low price, in fact, the manufacturer does not only decrease the incentives but also limits the deviation possibilities of the retailer.

In the terminology of Nelson (1970), the cases of experience goods and search goods will be distinguished. In the first case, consumers can only observe the good's quality after consuming it, whereas in the second case they can learn the quality at some costs, which are often interpreted as costs for visiting the point of sale.

For both kinds of goods the introduced possibility of transmitting information to consumers will be valuable for the manufacturer. Yet, the extent to which this is true, differs. The effect of the transmitted information is greater the harder it is for consumers to find out about the quality from other channels of information. In the case of experience goods, information transmission will in general be feasible and useful to consumers, whereas the informativeness of the sent information with search goods depends on how costly it is for consumers to learn about the quality on their own. Intuitively, information provided by the manufacturer will be less useful when a visit to the retailer where the quality of the good is observed gets cheaper.

The chapter proceeds as follows, the next section gives an overview of the related literature, Section 1.3 introduces the model which is solved for different information settings in Section 1.4. Section 1.5 extends the model to search goods and Section 1.6 concludes. The proofs are relegated to the appendix.

## 1.2 Related Literature

The model in this essay combines two popular topics in the industrial organization literature.

On the one hand, the manufacturer in this chapter transmits information revealing part of his private information about the quality of the good to the consumers. Information transmission in the industrial organization literature is often seen to take the form of advertising. The topic of informative advertising is quite common in the industrial organization literature (for a survey see Bagwell, 2007). In contrast to most of the papers discussing issues of advertising where it is costly, here the manufacturer uses cheap-talk like in Crawford and Sobel (1982) as the advertising device.

On the other hand, this chapter deals with the issue of double-marginalization first noticed by Cournot (1863) (see Spengler, 1950, for an early formal discussion of the issue). This issue arises whenever a monopolistic retailer sells a good to a retailer who then sells it to the consumer because both firms charge a markup on the prices they set. Here the manufacturer will use his private information to influence the markup charged by the retailer, influencing the double marginalization effect.

The remainder of this section first presents related work dealing with the questions on how information is transmitted in different market settings and then describes the existing work that deals with such questions in a manufacturer-retailer setup.

### 1.2.1 Information Transmission

In general, there is a distinction between the topics of 'Quality Disclosure' and 'Quality Signaling', the distinction being the fact that in the case of disclosure, information can not be misrepresented (but withheld) and in signaling situations, misrepresenting is possible.

For both cases there is a great amount of work in the economic literature (for a model unifying both strands see Daughety and Reinganum, 2007).

The disclosure literature generally argues that firms prefer to disclose their quality since if there were a pooling of firms with different qualities, the highest quality firms benefit from disclosing information, thereby increasing their profits (an early paper in this context is Milgrom, 1981). Given that the highest type would disclose its qualities, consumers reason that the quality must be lower if they do not observe that the quality is disclosed. This in turn leads the second highest quality firm to also disclose and so on. This result is commonly referred to as 'unraveling' and if disclosure is costless, firms of all quality types will disclose in equilibrium.

The literature about signaling usually concentrates on the price and/or on (costly) advertising as the instruments used to inform consumers about the quality. The idea that (seemingly uninformative) advertising may improve consumers' information about a good's quality was first put forward by Nelson (1970) and later formalized by Milgrom and Roberts (1986) who also added the possibility that the prices inform consumers. They find that in a model of repeated purchases of an experience good, a pricing scheme together with the optimal amount of advertising can signal quality to consumers. For more recent work see for example Guo and Zhao (2009) and Levin et al. (2009).

Apart from these articles there is some work on the effect of different forms or intensities of competition on the disclosure decision (Board, 2009 and Guo, 2009) or the ability to use the price as a signal of quality (Janssen and Roy, 2010, Adriani and Deidda, 2011). Although the studies mentioned come to somewhat different results, the general message is that the ability to inform consumers about a product's quality is higher, the higher the market power of the informing agent.

Besides using prices and advertising to signal quality other instruments found in the literature are specialization (Kalra and Li, 2008) and most prominently certification of another agent (for a survey see Dranove and Jin, 2010).

### 1.2.2 Vertical Chains

Surprisingly there is rather little literature on the issue of quality disclosure and signaling in a vertical trade setting, the exceptions being Chu and Chu (1994) and Guo (2009). Chu and Chu model a situation where a manufacturer of a high-quality product 'rents' the reputation of a retailer and thereby convinces the consumers that his product is of high quality. Guo models how a manufacturer would choose to disclose quality, either disclosing directly or leaving the decision to the retailer.

A relatively new area of work deals with the reasoning for so-called retail price recommendations (RPR). Work in this area emerged only recently and can be divided into three branches, all dealing with models of vertical trade and information transmission in such models.

Buehler and Gärtner (2013) model a situation where one retailer repeatedly sells the good of one manufacturer to consumers and the manufacturer has private information about his costs of production. Since the information about costs is necessary to maximize the joint profits of the manufacturer and the retailer, they show that this information can be communicated from the manufacturer to the retailer using RPRs.

The following articles consider the target of the information sent by the manufacturer to be the consumers, just as in the essay at hand. They can be divided by the degree of rationality imposed on consumers. Puppe and Rosenkranz (2011) and Fabrizi et al. (2010) employ models of behavioral economics and also consider a vertical trade setting, but in their model consumers have reference dependent preferences. They then assume that RPRs, sent from a manufacturer, are used by the consumers as their reference point and they show that this can increase the profits of the manufacturer.

Lubensky (2011) builds a model where a manufacturer possesses better information about the state of demand than rational searching consumers, who do not observe the price prior to visiting a retailer. He then shows that the manufacturer can send signals about the state of demand to consumers, thereby influencing their search decision and increasing his profits.

While the result in Lubensky's article, that a manufacturer can communicate private information to consumers, is similar to the results that will be presented later, there are some fundamental differences. First of all, the models differ since he assumes one manufacturer and a continuum of retailers whereas a vertical chain of two monopolists is assumed in this paper. However, the crucial difference lies in the kind of information that the manufacturer is better informed of. In Lubensky's model, the manufacturer possesses better information about demand than consumers do. This information is not directly relevant to consumers, but it gets valuable only through its influence on the retail price. As a consequence of this, his model does not allow to evaluate the different implications that experience and search goods have. In the present paper, consumers are uninformed about the quality of the good so that this distinction can easily be made.

## 1.3 Model Setup

The following describes the model setup, starting with the two firms who are producing and selling the product and then describing the preferences of the consumers.

The model consists of a manufacturer (M) who produces a good of random quality $q$ with the quality being distributed according to a distribution function $F(q)$ with full support on $Q = [0, \overline{q}]$. M then sells the good to a retailer (R) (who observes the realized quality before) at the wholesale price $w$ and the retailer in turn sells the good to consumers at the retail price $p$. The retailer doesn't face any costs for selling the good.

The manufacturer's unit costs of production depend on the quality and are given by $c(q)$ with $c(0) = 0 < k-(m^2-1)E(q) < c(\overline{q}) < k$, where $m$ denotes the total number of consumers and $k$ gives the consumers' valuation for a good with quality $q = 0$ (see below). Higher quality products are assumed to be more costly to produce, so that $c'(q) > 0$. Besides normalizing the costs of the lowest quality to zero, the assumption guarantees that separation is feasible ($k - (m^2 - 1)E(q) < c(\overline{q})$) as will be shown later and that the good can profitably be sold irrespectively of the realized quality ($c(\overline{q}) < k$).

The manufacturer can directly communicate with consumers, as he is able to send some costless (cheap-talk) signal $\tilde{q}$ to consumers. While this signal need not be restricted and potentially might be anything, it is useful to interpret the signal $\tilde{q}$ as a level of quality.

Assuming random quality is motivated by the idea that even big companies are not able to fully control the produced quality, especially so if they depend on parts they have to buy from other companies.

A prevalent example can be seen in the so-called "Antennagate" of Apple's iPhone 4. There the design of the Antenna led to problems with certain holding positions. Since Apple gave out free cases to complaining consumers, later reached a settlement that alternatively offered consumers 15$[1] and finally changed the antenna design for the model's successor, it seems reasonable to assume that the problems were unforeseen and we can talk about random quality here. In the manner of this example we could also interpret the random quality assumption as describing the uncertainty that remains with the manufacturer when he starts to produce the good. Technically the assumption is needed to make information transmission profitable for the manufacturer. If he were to choose the quality which is produced (and without intro-

[1] see the official site of the settlement at: `https://www.iphone4settlement.com/` (last accessed November 25, 2016)

ducing some other random element), there would generally be one optimal quality that M will prefer, thus taking away the need to signal information. Phrased differently, for information transmission to be valuable, there must be something that M can possibly inform consumers about. Without any randomness, the equilibrium would generally depend on parameters only so that consumers would be able to infer all equilibrium values and information transmission would be unnecessary and of no use.

The good in consideration is what is commonly called an 'Experience Good' (Nelson, 1970), which in this case means that consumers are not aware of the good's quality until after they consumed it.

There are two groups of consumers, the 'low types' ($i = \ell$) and the 'high types' ($i = h$). The groups are of size $n_\ell$ and $n_h$ respectively, all consumers have unit-demand. The two groups differ in how they value quality improvements of the good. Without loss of generality, $n_\ell$ will be set to 1, so that the total number of consumers equals $m = n_h + 1$.

The (ex-post) monetary valuations of the consumers for a good of quality $q$ take the form

$$v_i(q) = k + \theta_i q \qquad i \in \{\ell, h\}$$

where $k$, the "base-line" utility is a positive constant, and the type $\theta_i$ gives consumer $i$'s sensitivity to quality improvements, with $1 \geq \theta_h > \theta_\ell \geq 0$.

Thus, when consumer $i$ buys the good, his (ex-post) utility is given by

$$U_i(q) = k + \theta_i q - p \qquad i \in \{\ell, h\}$$

If any consumer decides not to buy the good, his utility is assumed to be zero. Since consumers do not observe the quality prior to consuming the good they have to form beliefs about it using the information they observe, namely retail prices and quality signals. A belief $\beta$ in the usual sense is a probability distribution over types of the manufacturer (quality realizations), that is, $\beta : (p, \tilde{q}) \to \Delta Q$. In the model at hand, consumers optimally buy the good whenever their expected valuation for it exceeds the retail price, i.e. only the expected quality for a given belief $\mu(p, \tilde{q}) := E_{\beta(p,\tilde{q})}(q)$ is relevant for their decision. This expectation will be referred to when making statements about consumers' beliefs. With the assumed valuations and because the quality level is at least equal to zero, it is optimal for any consumer to buy the good whenever $p < k$, independent of the expected quality. Therefore beliefs for prices lower than $k$ can be chosen arbitrarily.

In contrast to the quality, consumers are able to observe the price before making the purchase decision. Assuming that prices are directly observable simplifies the model and can be seen as a shortcut for a model in which consumers can learn the price only at some positive cost but where the retailer is able to commit to charging some retail price, for example by advertising it.

Because the retailer cannot distinguish consumers, he can potentially sell the good to all consumers at a lower price or to the high-type consumers only at a higher price. His profits, which will be denoted by $\pi_R$, then are $1+n_h$ and $n_h$ times the retail minus the wholesale price, respectively. Similarly the manufacturer's profits (denoted by $\pi_M$) are $1+n_h$ ($n_h$) times the wholesale price minus the unit cost $c(q)$, when selling to all (the high-type consumers only).

In this model a strategy of the manufacturer gives a pair of wholesale price and quality signal depending on the realized quality. The retailer's strategy is the retail price potentially depending on the realized quality, quality signal and the wholesale price. A consumer's strategy simply is the decision to buy or not to buy the good, given the price and quality signal he observed.

Figure 1.1 summarizes the timing and actions of all agents.

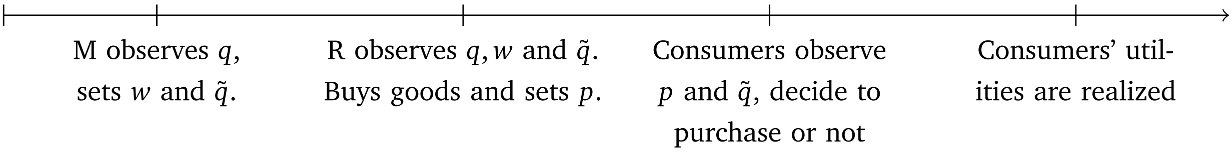



Figure 1.1: Model Timing

In the following, we will concentrate on (weak) perfect Bayesian equilibria (PBE) in pure strategies, which in this setting requires that

**Requirement 1.1.** *Firms maximize their profit given the other firm's and consumers' strategies. Consumers act optimally given the strategies of the firms and their beliefs* $\mu$.

**Requirement 1.2.** *Consumers' beliefs are derived from the firms' strategies and Bayes' rule whenever possible.*

Since this solution concept allows for off-equilibrium path beliefs that are not 'credible' (Sadanand and Sadanand, 1995), an equilibrium refinement based on perfect sequentiality (Grossman and Perry, 1986) as formulated in Gertner et al. (1988) will be imposed. In particular, this refinement rules out equilibria where prices are not monotone in the quality. Roughly, the refinement requires that there is no deviation which is profitable for some set of types (of manufacturers or retailers), given consumers attribute this deviation to exactly those types.

Formally, an interpretation $I(p,\tilde{q})$ of a price and quality signal pair not on the equilibrium path, is a set of qualities to which the consumer attributes this deviation, i.e. $I : (p,\tilde{q}) \rightarrow Q$. Using this interpretation, consumers form their belief given the interpretation, $\mu(p,\tilde{q}) = E[q|q \in I(p,\tilde{q})]$ and act optimally given this belief.

An interpretation is consistent if the types to which consumers attribute the deviation indeed prefer to deviate. In the vertical chain we have to distinguish two possible deviations that can be observed by consumers and thus need to be interpreted, the retailer can deviate by setting a retail price not on the equilibrium path or the manufacturer can send an off-equilibrium path quality signal.

Given an equilibrium with prices $w^*$ and $p^*$,signal $\tilde{q}^*$ and beliefs $\mu(p^*, \tilde{q}^*)$ an interpretation of the retailers' deviation to $p'$, $I(p', \tilde{q}^*) = T \subseteq Q$ is consistent if

$$\pi_R(p^*, w^*, \mu(p^*)) \leq \pi_R(p', w^*, I(p', \tilde{q}^*)) \quad \forall t \in T$$
$$\pi_R(p^*, w^*, \mu(p^*)) > \pi_R(p', w^*, I(p', \tilde{q}^*)) \quad \forall t \notin T$$

Since the manufacturers profits depend directly only on the wholesale price he earns, a deviation in the quality signal that is observed by consumers must be accompanied by a deviation in the wholesale pricing scheme if it is to be profitable for M. This different wholesale price as well as the consumers' interpretation of a different quality signal also make it necessary for the retailer to adjust his strategy. Thus, given a deviation of M from the equilibrium signal $\tilde{q}^*$ to $\tilde{q}'$ together with changing the wholesale price from $w^*$ to $w'$, the interpretation $I(p^{BR}, \tilde{q}') = T \subseteq Q$ where $p^{BR}$ is the retailer's best response to M's changed wholesale price and given consumers' interpretation, is consistent if:

$$\pi_M(w^*, \tilde{q}^*, p^*, \mu(p^*, \tilde{q}^*)) \leq \pi_M(w', \tilde{q}', p^{BR}, I(p^{BR}, \tilde{q}')) \quad \forall t \in T$$
$$\pi_M(w^*, \tilde{q}^*, p^*, \mu(p^*, \tilde{q}^*)) > \pi_M(w', \tilde{q}', p^{BR}, I(p^{BR}, \tilde{q}')) \quad \forall t \notin T$$

Because deviations with a consistent interpretation can be seen as a self-fulfilling prophecy, the following additional requirement is imposed on equilibrium strategies:

**Requirement 1.3.** *There is no deviation with consistent interpretation in any equilibrium.*

Equilibria that fulfill this requirement will be referred to as 'Perfect Sequential Equilibria'.

Without the cheap-talk signal $\tilde{q}$, the model is a standard vertical relations model. It will be shown that while some information about the quality can be transmitted to consumers via the retail price, the manufacturer benefits from the introduction of an additional possibility of direct communication to the consumers. We will later see that the diverging interests of the vertical chain, that were already mentioned in the introduction, will allow M to credibly signal some of his private information to consumers.

In particular the manufacturer uses his private information strategically in a way that decreases the severity of the double marginalization problem occurring in a vertical chain where both firms possess market power.

## 1.4 Results

Instead of directly working with the full model, we will first look at a simplified version of it, namely one with one consumer of each type (i.e. $n_\ell = n_h = 1$), where one consumer's valuation of the good is independent of the quality, that is $\theta_\ell = 0$ ('the quality ignoring consumer') and the second consumer's quality sensitivity is set to one, i.e. $\theta_h = 1$ ('the quality aware consumer').

Consumers' valuations therefore are given as:

$$v_\ell = k \qquad \text{and} \qquad v_h(q) = k + q$$

While the assumed difference in valuations may seem rather extreme (and as will be shown in the next section, is not necessary for the result) it is easy to imagine such situations. For example, think about the good being a mobile phone and the let $q$ measure the quality of the built-in camera. There certainly exist people who never use their mobile's camera and so they most probably don't care about its quality. At the same time many people are using the built-in camera and so are very well interested in the camera's specifications.

### 1.4.1 No Quality Signaling

Before dealing with the question if M can credibly transmit information about the quality of the good he produced to consumers, it is worthwhile to look at a situation where M does not possess this possibility.

This is crucial if we are interested to find out about M's incentives to signal information to consumers or not, and many of the arguments here can be used in the model with quality signaling with only minor changes

It is easy to see that in the given model, the retailer can not, on his own, inform consumers about the quality through the price. If, for a given wholesale price, the retailer would be charging different retail prices depending on the realization of the quality, the retailer would not act optimally. In this situation R would always be better off changing his strategy to one where he would always charge the higher retail price. Hence for a given wholesale price the retailer can only charge one retail price. However this does not rule out all possibilities that the retail price convey information about the quality. When setting his wholesale price, the manufacturer has to decide between setting a low wholesale price, trying to induce the retailer to serve all consumers, or setting a high wholesale price which in turn will induce the retailer to charge a high retail price only serving the quality aware consumer. As the following proposition shows separation is possible through the right choice of the wholesale pricing scheme. The retailer then passes on the information inherent in the wholesale price through his retail price.

To shorten notation, in everything that follows, the retailers' strategy will be represented by prices only, incorporating that he buys two goods from the manufacturer if his strategy is to set a retail price of $p_h$, he buys one good if $p = p_\ell$ and he buys no good if his strategy does not prescribe any retail price.

**Proposition 1.1.** *In the unique perfect sequential equilibrium in the simplified game without quality signaling, the type space Q is partitioned into* $Q_\ell = [0, \hat{q})$ *and* $Q_h = [\hat{q}, \overline{q}]$.[2]
*The partitioning must be such that:*

$$c(\hat{q}) = k - 3E(Q_h) \tag{1.1}$$

*where* $E(Q_i) := E(q|q \in Q_i)$*. Pricing schemes are given by:*

$$w = \begin{cases} w_\ell = k - E(Q_h) & \text{if} \quad q \in Q_\ell \\ w_h = k + E(Q_h) & \text{if} \quad q \in Q_h \end{cases}$$

$$p = \begin{cases} p_\ell = k & \text{if} \quad w \leq w_\ell \\ p_h = k + E(Q_h) & \text{if} \quad w \in (w_\ell, w_h] \end{cases}$$

*The equilibrium is supported by beliefs of the form* $\mu(p) = E(Q_h)\, \forall p > k$.

In equilibrium, optimal behavior implies that both consumers buy if the price is $p_\ell$ and only the high type does so when $p = p_h$.

It is easily verified that this indeed constitutes an equilibrium. Consumers act optimally since they either know that their valuation (weakly) exceeds the price (if $p = p_\ell$), or the price exactly equals the high type's expected valuation ($p = p_h$). Consumers' beliefs make it unprofitable for the retailer to set a price higher than $p_h$ and for prices between $p_\ell$ and $p_h$ the demand is constant so that such a price is always dominated by setting a price of $p_h$. Similarly all prices below $p_\ell$ are dominated by $p_\ell$ so that, given consumers' beliefs, the retailer will set a price of either $p_\ell$ or $p_h$. The specified off-equilibrium path beliefs yield a demand of zero for any price that is higher than $p_h$, which in turn implies that in this case the manufacturer can extract the full surplus by setting his wholesale price equal to the retail price $w = w_h = p_h$ making it impossible for the retailer to deviate. If the manufacturer wants to induce a retail price for which all consumers buy, the retailer has to be at least indifferent between setting this price $p_\ell$ and deviating to the high price $p_h$, which is exactly how $w_\ell$ is constructed. Finally, the manufacturer can now choose between setting a price of $w_h$ and selling two units or selling

[2]Note that this equilibrium requires $k > 3E(Q)$. Whenever this is not the case, the equation defining $\hat{q}$ has no solution and the resulting equilibrium is a pooling equilibrium which can be obtained from the proposition by setting $Q_\ell = \emptyset$ and $Q_h = Q$. Uniqueness of the equilibrium is up to the behavior of the two firms in the case that they are indifferent.

one unit at a price of $w_\ell$. As costs are increasing in the quality, the price-cost margin decreases for higher levels of $q$ and at a quality $\hat{q}$ the manufacturer is indifferent between selling two products with a small margin or selling one product with a relatively higher margin. For qualities below $\hat{q}$, the small margin when selling two products is more than compensated by the number of goods sold, and for levels above $\hat{q}$, only one good is sold.

### 1.4.2 Quality Signaling

Before going through the model explicitly, it is worthwhile to think about the incentives a manufacturer might have to actually reveal some information about the quality to consumers.

By the structure of the demand, we can split up M's problem of maximizing his profit into two parts. M first calculates the maximal possible wholesale prices when selling to both consumers or to the quality caring consumer only, and in a second step, he then chooses his targeted consumers.

To build up intuition first think of an imaginary situation where M sends a quality signal $\tilde{q} \in [0, \overline{q}]$ and consumers literally believe this signal, that is $\mu(p, \tilde{q}) = \tilde{q}$.

Suppose first that M's goal is to sell to both consumers. He then knows that this can only be done at a retail price of at most $k$, which at the same time of course is the upper bound for the wholesale price in this situation. In the hypothetical situation where consumers take M's signal literally, M's best choice would be to send a signal of $\tilde{q} = 0$.

This would take away all possibilities from R to set a retail price higher than $p = k$, the price that leaves both consumers with an expected utility (given the signal) of zero. For a higher retail price no consumer would visit R.

The situation is quite different if M wants to sell to the quality caring consumer only. For any price that is higher than $p = k$ only one consumer demands the good and he only does so if the price is at most $k + \tilde{q}$.

As before M can extract the full surplus by setting his wholesale price equal to the retail price if only the high type consumer is to buy the good. Because the maximal retail price in this hypothetical situation is obtained for the highest possible signal, M would always signal $\tilde{q} = \overline{q}$ if he intends to sell to the quality caring consumer only.

Put differently, if M targets both consumers, his incentives for signaling information about the quality (for a given wholesale price), coincide with the consumers' incentives, namely keeping the retail price down. We can also interpret the signaling in this situation as means of M to decrease the severity of the well-known double marginalization problem. The lower the quality signal, the smaller the part of the profit that has to be granted to the retailer in order to keep him from deviating from the proposed strategy.

If, in contrast, the manufacturer targets only the quality caring consumer, the structure of demand implies that M's incentives are perfectly aligned with those of R, namely setting the retail price as high as possible. In this case, M can set his wholesale price equal to the retail price that will emerge. The retailer then has no choice but to follow the strategy of M, otherwise he won't make any sales (for a higher retail price) or make negative profits (for a lower retail price).

Those observations and the intuition that because the gap in valuations between the two consumers increases with the produced quality so that selling to only the quality caring consumer is especially profitable for high realizations of the quality, give rise to the conjecture that the best the manufacturer could do is to find a cut-off quality level $\hat{q} \in [0, \overline{q}]$, so that M signals a low quality for qualities below this cut-off and a high quality for realizations of q above this level.

The following proposition shows that such a signaling scheme can indeed be part of an equilibrium of the simplified model, and in fact is used by the manufacturer in any perfect sequential equilibrium in this game.

**Proposition 1.2.** *In any perfect sequential equilibrium of the game with quality signaling, the type space Q is partitioned into $Q_\ell = [0, \hat{q})$ and $Q_h = [\hat{q}, \overline{q}]$.*
*The partitioning must be such that:*

$$c(\hat{q}) = k - 2E(Q_\ell) - E(Q_h) \tag{1.2}$$

*with $E(Q_i) := E(q|q \in Q_i)$. Pricing and signaling schemes are given by:*

$$(\tilde{q}, w) = \begin{cases} (\tilde{q}_\ell, w_\ell = k - E(Q_\ell)) & \text{if} \quad q \in Q_\ell \\ (\tilde{q}_h, w_h = k + E(Q_h)) & \text{if} \quad q \in Q_h \end{cases}$$

$$p = \begin{cases} p_\ell := k & \text{if} \quad w \le w_\ell \\ p_h := k + E(Q_h) & \text{if} \quad w \in (w_\ell, w_h] \end{cases}$$

*With beliefs of the following form:*

$$E(Q_h) = \mu(p, \tilde{q}_h) \ge \mu(p, \tilde{q}) \ge \mu(p, \tilde{q}_\ell) = E(Q_\ell) \quad \forall p, \tilde{q} \notin \{\tilde{q}_\ell, \tilde{q}_h\}$$

In equilibrium, both consumers buy if the price is $p_\ell$ and only the high type does so when $p = p_h$.

The equilibrium is not unique, as the behavior of indifferent firms is arbitrary. Additionally, as stated in the proposition, the perfect sequential equilibrium allows for many off-equilibrium beliefs of the consumers.

Comparing the equilibria from Propositions 1 and 2, the wholesale pricing scheme directly shows that M's equilibrium profit is increased by sending signals to the consumers. Although the prices look similar, the difference is in the expected quality levels. Where there are the same (conditional) expectations for both wholesale prices in the game without direct information transmission, the formulas for $w_\ell$ and $w_h$ contain different conditional expectations in an equilibrium with quality signaling and because it must be the case that $E(Q_\ell) \leq E(Q_h)$, the wholesale prices in informative equilibria are higher than those in non-informative ones.

In the separating equilibria without quality signaling, R could in principle always set a high price and make consumers believe that the good is of high quality. This made it necessary for the manufacturer to grant a relatively high share of the profits to the retailer in order to make him stick to setting a price that induces sales to all consumers. With the additional possibility of directly informing consumers, this is no longer true. By revealing that his good is of rather low quality, the deviation price that the retailer can set is decreased compared to a situation where M does not reveal such information. This decreased gain from a possible deviation makes it less costly for M to induce the retailer to set a price at which all consumers will buy, thereby increasing M's profits.

Theoretically, the messages $\tilde{q}_\ell$ and $\tilde{q}_h$ don't have to be specified, as long as they are understood. Nevertheless it might be worth to think about what they could refer to empirically.

Sending a high-quality signal could for example be interpreted as something along the lines of 'moon pricing' where in a model of reference dependent preferences, a manufacturer spoils consumers reference points by giving a high retail price recommendation (cf. Puppe and Rosenkranz, 2011 and Fabrizi et al., 2010). The situation here is similar if the message for a quality above the threshold is $\tilde{q}_h = \overline{q}$. Then the signaled quality is always (weakly) higher than the truly realized one, something which might also be observed in advertisings where quality (or other specifications) is often exaggerated in some way. The previous analysis nevertheless shows that such exaggerated advertisements can still contain valid information of interest to consumers.

### 1.4.3 General Model

The driving force behind the results of the last section is that the wholesale price the manufacturer can charge depends positively on the expected quality if the targeted consumer is the high type only and the situation is reversed if both types of consumers are targeted. While the first effect, that M can achieve a higher price the higher consumers expect their valuation to be, is intuitive, the effect that the wholesale price depends negatively on the expected quality if both consumers are targeted deserves some closer inspection.

In the simplified model, a higher expected quality translates into a higher willingness to pay of the high type. This in turn has two effects. On the one hand it increases the potential wholesale price for sales only to the high type. But on the other hand, it also increases the deviation incentives for the retailer whenever M intends to sell to both consumers. Since R's only meaningful deviation in such a situation is to sell to the high type only, this gets more profitable the higher this type's willingness to pay gets.

The goal of this section is to show how and under what restrictions the results of the simplified model, i.e. the possibility of credible quality signaling, can be transfered to the general model.

To briefly recap the general model: there are $n_\ell = 1$ low type and $n_h$ high type consumers, whose sensitivity is given by $\theta_\ell$ and $\theta_h$ respectively, with $\theta_\ell < \theta_h$. Consumers valuations are given as $v_i(q) = k + \theta_i q$.

For the proposition in this section we will also make the assumption that consumers trust the manufacturer's information, i.e. for any combinations $p, \tilde{q}$ and $p', \tilde{q}$ that are not on the equilibrium path, $\mu(p, \tilde{q}) = \mu(p', \tilde{q})$.

**Proposition 1.3.** *In any perfect sequential equilibrium of the game with quality signaling, where consumers differ sufficiently to fulfill $n_h\theta_h > m\theta_\ell$, the type space Q is partitioned into $Q_\ell = [0, \hat{q})$ and $Q_h = [\hat{q}, \overline{q}]$ with*

$$c(\hat{q}) = k - m(n_h\theta_h - m\theta_\ell)E(Q_\ell) - n_h\theta_h E(Q_h)$$

*and price, signaling schemes and supporting beliefs are given by:*

$$(\tilde{q}, w) = \begin{cases} (\tilde{q}_\ell, w_\ell(\tilde{q}_\ell) = (1+n_h)p_\ell(\tilde{q}_\ell) - n_h p_h(\tilde{q}_\ell) & \text{if} \quad q \in Q_\ell \\ (\tilde{q}_h, w_h(\tilde{q}_h) = p_h(\tilde{q}_h)) & \text{if} \quad q \in Q_h \end{cases}$$

$$p = \begin{cases} p_\ell(\tilde{q}_\ell) = k + \theta_\ell E(Q_\ell) & \text{if} \quad (\tilde{q}, w) = (\tilde{q}_\ell, w_\ell(\tilde{q}_\ell)) \\ p_h(\tilde{q}_h) = k + \theta_h E(Q_h) & \text{if} \quad (\tilde{q}, w) = (\tilde{q}_h, w_h(\tilde{q}_h)) \end{cases}$$

$$\mu(p, \tilde{q}_h) = E(Q_h) \geq \mu(p, q) \geq E(Q_\ell) = \mu(p, \tilde{q}_h) \quad \forall p, \tilde{q} \notin \{\tilde{q}_\ell, \tilde{q}_h\}$$

In equilibrium, both consumers buy if the price is $p_\ell$ and only the high type does so when $p = p_h$.

The effects that make credible signaling possible are the same as in the simplified model. M's setting of the wholesale price can be seen as the decision between setting a low wholesale price that induces (a retail price inducing) sales to all consumers and setting a high wholesale price that will only induce sales to the high type consumers. When M intends to induce sales to all consumers with a relatively low wholesale price, he has to make sure that the retailer has no incentive to deviate to another price at which only high type consumers purchase the good. In such a situation, a change in the expected quality of consumers which the manufacturer tries to influence with his signal, has two effects.

The first effect, which was already seen in the simplified model, is that an increased expected quality makes it more tempting for the retailer to only sell to high type consumers instead of to all consumers, which suggests trying to influence consumers' in a way that they decrease their expectation about the quality of the good.

An opposing effect, which before was ruled out by assuming that low types valuations are independent of the quality, arises since now all consumers' valuations are increasing for higher quality products. Hence a higher expected quality also raises the low types WTP, and thus the possible surplus that M can extract when low type consumers buy the good.

Whenever the first effect is stronger than the second, signaling a relatively low quality when inducing sales to all consumers is profitable, and the situation is similar to the simplified model. For this to be the case, consumers have to be sufficiently different as stated in the proposition.

## 1.5 Search Goods

The previous section dealt with the case of so-called experience goods. In many cases the assumption of such goods, that is the idea that consumers observe a good's quality only after actually consuming it may be too strict and modeling the good as a search good is more applicable.

A good is called a search good if some characteristics (here: quality) are unknown to consumers until they visit the point of sale. There they are able to observe the previously unknown characteristics of the good.

For this model to be different from one of perfect information, we have to introduce positive costs of visiting the retailer, and those shopping costs will be denoted by $s$. To rule out prohibitively large shopping costs it is assumed that they are smaller than the lowest possible valuation, that is $s < k$.

The change from experience to search goods greatly increases the necessary notation, so that this section will only use the simplified model from the previous section up to subsection 1.4.2. We thus again restrict the model to a situation where there are two consumers, one high-type and one low-type and the low-type's valuation for the good is independent of the realized quality, that is $m = 2, n_h = 1$ and $\theta_\ell = 0$ and $\theta_h = 1$.

The change of the nature of the good and the introduction of shopping costs are the only changes from the simplified model.

The following figure shows the adjusted timing for the model with a search good.

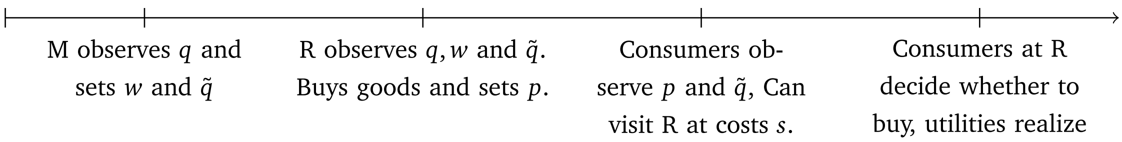



Figure 1.2: Model Timing for a Search Good

We will first look at a model without quality signaling of M, and then see if and how this can profitably be introduced.

The introduction of search costs makes it necessary for consumers to make two choices, first they have to decide whether to visit the retailer and if they did, they have to decide whether to buy the good or not. Clearly, consumers will decide to visit the retailer if they expect to gain a (weakly) positive utility from purchasing the good, i.e. consumer $i$ will visit R whenever $k+\theta_i\mu(p) \geq p+s$. If a consumer decided to visit the retailer, his search costs are sunk and he will buy the good whenever his valuation is higher than the price charged, that is $v_i(q) \geq p$.

Those two conditions lead to the following equilibrium in the search goods game without quality signaling:

**Proposition 1.4.** *In the search goods model, there exists a partly separating equilibrium where the type space Q is partitioned into $\{Q_N,...,Q_1\}$ as follows.*

*The partitions $Q_i$ are constructed recursively, starting with the highest one: $Q_1 = [q_1, \overline{q}]$ such that $q_1 = E(Q_1) - s$ and $Q_i = [q_i, q_{i-1}), i > 1$ such that $q_i = E(Q_i) - s$. The lowest one is given as $Q_N = [0, q_{N-1})$ such that $q_{N-1} > 0$ and $q_N = E(Q_N) - s \leq 0$.*

$$w(q) = \begin{cases} w_\ell(Q_i) = k - E(Q_i) - s & \text{if} \quad q \in Q_i, c(q) \leq k - 3E(Q_i) - s \\ w_h(Q_i) = k + E(Q_i) - s & \text{if} \quad q \in Q_i, c(q) > k - 3E(Q_i) - s \end{cases}$$

$$p = \begin{cases} k - s & \text{if} \quad w = w_\ell(Q_i) \\ k + E(Q_i) - s & \text{if} \quad w = w_h(Q_i) \end{cases}$$

$$\mu(p) = E(Q_i) \qquad \text{if} \quad p - k + s \in Q_i$$

*Sketch of Proof.* The strategies closely resemble the strategies from the experience good model and the crucial observation here is that in each interval $Q_i$ the situation essentially is the same as in the model of experience goods in that there are two possible retail and wholesale prices.

It is then easily checked that the proposed strategies indeed constitute and equilibrium.

For any proposed equilibrium price, the consumers are indifferent between visiting they retailer and not, and whenever they observe the quality at the retailer, they are always (weakly) better off by purchasing the good.

Given the beliefs, the wholesale pricing scheme is constructed in a way to keep the retailer from deviating, and given the retailer's and consumers' strategies, the manufacturer can not improve upon the proposed wholesale pricing scheme.

The reasoning behind the constructed intervals comes from observing that it can not happen in equilibrium that no consumer buys the good, since sales (and positive profits) are always achievable at a price of $p = k$. If it can not be that no consumer purchases the good, this means that at least one consumer must visit the retailer for any price on the equilibrium path, hence for all prices on the equilibrium path $E(v_h(q)|p) \geq p + c$ because whenever the low-type consumer buys the good, so does the high-type. This reasoning also implies that $v_h(q) > p$ for all prices on the equilibrium path. Putting those observations together with the fact that expectations have to be correct in equilibrium, the construction of the intervals assures that the retailer has no incentive to change the price, once consumers' search decision is sunk.

□

With the observation that, by partitioning the space of possible qualities, the retail price now already delivers more information about the quality than was possible in the experience good model, the question is if M still is able to profitably send information to consumers.

With the observation that in each interval $Q_i$, the situation is almost identical to the experience goods setup, there should also be a possibility that the manufacturer informs the consumers.

More precisely, as the following proposition shows, M can indeed inform consumers about quality by dividing some interval $Q_j$ into $\{Q_j^\ell, Q_j^h\}$, where a low wholesale price that induces sales to all consumers is set in $Q_j^\ell$ and a high wholesale price which equals the retail price is set in $Q_j^h$.

**Proposition 1.5.** *If $k - s \in (2q_j + E(Q_j) + c(\overline{q}_i), 2E(Q_j) + q_{j-1} + c(\overline{q}_{i+1}))$ for some interval $Q_j$, with intervals constructed according to the previous proposition, there is an equilibrium where the manufacturer reveals information to consumers via cheap-talk in this interval. Pricing schemes for all intervals different from $Q_j$ remain as in the previous proposition. In $Q_j$:*

$$\tilde{q} = \begin{cases} \tilde{q}_\ell & \text{if} \quad q \leq q' \\ \tilde{q}_h & \text{else} \end{cases}$$

$$w(q) = \begin{cases} w_\ell(Q_j^\ell) := k - E(Q_j^\ell) - s & \text{if} \quad q \in Q_j, q \leq q' \\ w_h(Q_j^h) := k + E(Q_j^h) - s & \text{if} \quad q \in Q_j, q > q' \end{cases}$$

$$p = \begin{cases} k - c & \text{if} \quad w = w_\ell(Q_j^\ell) \\ w & \text{if} \quad w = w_h(Q_j^h) \end{cases}$$

$$\mu(p) = E(Q_j^f) \qquad \text{if} \quad p - k + s \in Q_j^f, f \in \{\ell, h\}$$

*The threshold $q'$ is defined by:*

$$c(q') = k - s - 2E(Q_j^\ell) - E(Q_j^h) \tag{1.3}$$

*Sketch of Proof.* Since in the interval in question, the game is as in the previous section, the cheap-talk nature of the signal makes it necessary that M's profits are the same for for the quality realization $q'$ for both signals sent and corresponding wholesale prices. As the costs are increasing, whenever a threshold exists, the interval $Q_j$ is divided into $Q_j^{\ell} = \{q \in Q_j, q \leq q'\}$ and $Q_j^h = \{q \in Q_j, q > q'\}$. The threshold is similar to the one of Proposition 1.2, with additionally the shopping costs being taken into account. The condition from the beginning of the proposition ensures that the equation 1.3 defining the threshold has a solution.

If there is one interval, constructed according to the proposition such that this equation has a solution, M can credibly and profitably signal information to consumer even in the model of search goods.

□

While this proposition shows that signaling can also occur in a model of search goods it is clear that it's informativeness depends on the shopping costs $s$. The manufacturer can only reveal information that further divide one of the already existing intervals. And because smaller shopping costs induce a finer partition with smaller intervals, smaller shopping costs make the signal less informative. This can also be easily seen by considering two extreme cases. If the shopping costs go to zero, consumers can almost costlessly learn the quality of the good so that we approach a game of perfect information in which the price will either be $p = k + q - s$ or $p = k - s$. In this case information revelation of M is of no value. In the other extreme of very large shopping costs, the retail price can not partition the quality space, so that we are back in the situation of experience goods, where M's signaling is informative for a wide range of parameters.

This result also fits with the intuition, that the harder it is for consumers to get information (high $s$ or experience goods) the more valuable is the information that the manufacturer can transmit. It can also be interpreted as confirming the claim of Nelson (1970) that informative advertising is more prevalent for experience than for search goods.

## 1.6 Conclusion

Conventional wisdom might suggest that manufacturers should always exaggerate the quality of their products, questioning the usefulness of such statements for consumers. Nevertheless announcements which can be interpreted as informations about quality are abundantly observed in reality.

This paper showed that incentives to do the exact opposite might arise as well in a situation where a manufacturer sells his goods indirectly through a retailer to heterogeneous consumers who are uninformed about the good's quality.

The goal when exaggerating a product's quality is clear, namely increasing consumer's willingness to pay, but the reasoning for downplaying the quality is less apparent. As this paper demonstrated, a vertical chain structure might lead to exactly those incentives.

If the manufacturer in such a situation has to decide what information about the quality of his good he wants to transmit to consumers, he has to take into account two potentially opposing effects.

On the one hand, and in line with conventional wisdom, leading consumers to believe that a good is of higher quality increases their willingness to pay, thereby increasing the possible profits the manufacturer can extract.

On the other hand, if consumers get increasingly heterogeneous for higher quality realizations, inducing consumers to believe the produced good is of high quality increases the deviation incentives for the retailer for prices where the demand is elastic, so that it can become optimal to induce low quality expectations with consumers.

Phrased differently, the manufacturer can use strategic information transmission to consumers as means to control the retailers freedom in price setting, which directly influences his deviation incentives. If the retailer has less incentives to deviate, the manufacturer can charge higher wholesale prices for the same retail price thereby increasing his profits.

While most empirical examples that come to mind when thinking about information transmission from manufacturers to consumers seem to follow the theme of exaggeration there are some examples where informations sent from manufacturers could lead consumers to expect a good of a lower quality.

One of those examples are retail price recommendations. A (relatively) low retail price recommendation probably leads consumers to expect a good to be of rather low quality. Although the face value of such RPR might not provide much information (just as the signals in the model with cheap talk are not pinned down), the comparison of different RPRs gives consumers an orientation about what quality to expect. Besides RPRs, many manufacturers sell their goods under the name of different brands and those brands often carry specific projections about quality. The Taiwanese producer of computer parts 'Asustek' for example, founded a sub-company concentrating on the low-price segment called 'ASRock' and many car manufacturers (like VW) produce cars of a variety of brands.

## Appendix 1.A Proofs

This section provides the proofs of the theorems in the main text, which are restated here for convenience.

**Proposition 1.1.** *In the unique perfect sequential equilibrium in the simplified game without quality signaling, the type space Q is partitioned into* $Q_\ell = [0, \hat{q})$ *and* $Q_h = [\hat{q}, \overline{q}]$.[3]
*The partitioning must be such that:*

$$c(\hat{q}) = k - 3E(Q_h) \tag{1.1}$$

*where* $E(Q_i) := E(q|q \in Q_i)$*. Pricing schemes are given by:*

$$w = \begin{cases} w_\ell = k - E(Q_h) & \text{if} \quad q \in Q_\ell \\ w_h = k + E(Q_h) & \text{if} \quad q \in Q_h \end{cases}$$

$$p = \begin{cases} p_\ell = k & \text{if} \quad w \leq w_\ell \\ p_h = k + E(Q_h) & \text{if} \quad w \in (w_\ell, w_h] \end{cases}$$

*The equilibrium is supported by beliefs of the form* $\mu(p) = E(Q_h)\ \forall p > k$.

In order to prove this proposition, we will first establish some properties that any equilibrium in this game must have, then show what equilibria exist and in the last step argue that other equilibria are not sequentially perfect.

*Proof.*
**Step 1:** R's optimal strategy can at most be binary.
Fix a set of beliefs $\mu$ and let the sets $P_i, i \in \{\ell, h\}$ denote all prices that induce sales to a consumer of type $i$, that is: $P_i = \{p : p \leq k + \theta_i \mu(p)\}$. The sets $P_i$ are non-empty since $p = k$ must be part of both. Let $p_i$ be the highest price that induces sales to consumers of type $i$, $p_i = \max P_i$. In the simplified model, $\theta_\ell = 0$ so that $p_\ell = k$. Because any price that induces sales to the low-types will also induce sales to the high type, $P_\ell \subseteq P_h$ and $p_h \geq p_\ell$. The retailer's optimal strategy for any $w \leq p_h$ is

$$p^*(w) = \arg \max_{p \in \{p_\ell, p_h\}} \pi_R(p, w).$$

[3] Note that this equilibrium requires $k > 3E(Q)$. Whenever this is not the case, the equation defining $\hat{q}$ has no solution and the resulting equilibrium is a pooling equilibrium which can be obtained from the proposition by setting $Q_\ell = \emptyset$ and $Q_h = Q$. Uniqueness of the equilibrium is up to the behavior of the two firms in the case that they are indifferent.

**Step 2:** If $p^*(w) = p_h$ for some $w$, by setting $w = w_h := p_h$, M can extract the whole surplus when sales are made to the high type consumer only. With $w = p_h$, the retailer is indifferent between buying the good from M and setting $p = p_h$ or not buying the good, and so by assumption he buys the good and sets $p = p_h$.

**Step 3:** M can induce a retail price of $p = k$.

R will set a price of $p = p_\ell = k$ if this yields weakly higher profits than the best deviation. As Step 1 shows that the only deviation is to set a price of $p_h$ this means:

$$2(p_\ell - w) \geq p_h - w$$
$$w \leq w_\ell := 2k - p_h$$

**Step 4:** M's optimal strategy partitions $Q$ into $Q_\ell$ and $Q_h$.

By setting $w = w_h$, M can always make positive profits, implying it can't be the case that no sales are made for some realization of $q$. Step 1 showed that R's strategy is at most binary and Steps 2 and 3 showed that retail prices $p_\ell$ and $p_h$ can be induced by wholesales prices $w_\ell$ and $w_h$ respectively, since $w_i$ is the maximal wholesale price that will induce $p_i, i \in \{\ell, h\}$, M's optimal strategy $w^*(q)$ will only contain those two wholesale prices. Together this implies that M partitions $Q$ into $Q_\ell$ and $Q_h$ for which wholesale prices are $w_\ell$ and $w_h$ respectively. Thus $Q_\ell \cup Q_h = Q = [0, \overline{q}]$ and because we concentrate on pure strategies $Q_\ell \cap Q_h = \emptyset$, hence:

$$w^*(q) = \arg \max_{w \in \{w_\ell, w_h\}} \pi_M(w)$$

**Step 5:** Beliefs have to be correct in equilibrium, i.e. $\mu(p_\ell) = E(Q_\ell) := E(q|q \in Q_\ell)$ and $\mu(p_h) = E(Q_h) := E(q|q \in Q_h)$ and thus $w_h = p_h = k + E(Q_h)$ and $w_\ell = k - E(Q_h)$.

In an equilibrium, R will set prices $p_\ell = k$ and $p_h = k + \mu(p_h) = k + E(Q_h)$, and by Step 2, $w_h = p_h$ and $w_\ell = 2k - p_h = k - E(Q_h)$.

**Step 6:** M prefers to induce a price of $p_\ell$ by charging a wholesale price $w_\ell$ if $\pi_M(w_\ell) > \pi_M(w_h)$, that is whenever

$$2(k - E(Q_h) - c(q)) > k + E(Q_h) - c(q)$$
$$c(q) < k - 3E(Q_h).$$

If the converse is true, M is better off by setting $w_h$. For any fixed partition $Q_\ell, Q_h$, if the condition holds for some $q$ it must also hold for any smaller $q$, since the left-hand side is increasing in $q$. With these observations, there is a unique quality level $\hat{q}$ for which M is indifferent and as stated in the proposition, $\hat{q}$ is given by $c(\hat{q}) = k - 3E(Q_h)$. The partition must thus be of the form $\{Q_\ell = [0, \hat{q}), Q_h = [\hat{q}, \overline{q}]\}$.

**Step 7:** The previous steps established the existence of a separating equilibrium and showed that no other separating equilibrium could possibly exist. It remains to show that pooling equilibria are either non-existent or ruled out by the refinement.

Two pooling equilibria are generally possible, depending on which consumers are served.

Start with a candidate for a pooling equilibrium where only the high-type consumer is served, thus $p = w = k + \mu(p) = k + E(Q)$. Lower prices can not be sustained since the retailer would always have an incentive to deviate and the beliefs are on the equilibrium path and thus have to be correct. In this situation, M could induce a price of $p' = k$ where both consumers are served by choosing $w' = k - E(Q)$ which he prefers to do whenever:

$$\begin{aligned}\pi_M(w') = 2(w' - c(q)) &> w - c(q) = \pi_M(w)\\ 2(k - E(q) - c(q)) &> k + E(q) - c(q)\\ c(q) &< k - 3E(q)\end{aligned}$$

which, by assumption is true for small realizations of $q$. There can thus be no pooling equilibrium where only the high-type consumer is served.

It remains to show that a candidate pooling equilibrium where both consumers are served is not sequentially perfect. In such an equilibrium the previous steps imply that the wholesale price would have to equal $w = k - \mu(p_h)$ where again $p_h = \max p$ s.t. $p \leq k + \mu(p)$. Instead of pooling on the wholesale price $w$, M could always induce a price of $p_h$ by setting $w = w_h = p_h$, so for the desired pooling equilibrium it must be the case that

$$2(w - c(q)) = 2(k - \mu(p_h) - c(q)) > k + \mu(p_h) - c(q) = w_h - c(q)$$

for all realizations of $q$, in particular also for $\overline{q}$:

$$c(\overline{q}) < k - 3\mu(p_h)$$

From the assumption on the cost function, we know that $c(\overline{q}) > k - 3E(q)$ so that the above condition is violated whenever $\mu(p_h) > E(q)$.

Therefore suppose $\mu(p_h) < E(q)$ for the remainder. By construction $w$ makes the retailer indifferent between setting a price of $p = k$ and $p_h = k + \mu(p_h)$. If $\mu(p_h) < E(q)$ the retailer would be better off by setting a price of $p' = k + E(q)$ irrespectively of the realized quality. If consumers were to observe a deviation price of $p'$, they would have to reason that this deviation price is profitable for the retailer for all realized qualities so that they would attribute this deviation price to the whole type space. With the interpretation $I(p') = Q$ which leads to beliefs of $\mu(p') = E(q)$, the retailer would indeed deviate to this price for all $q$, so that $p'$ and $I(p') = Q$ constitute a deviation with a consistent interpretation.

This shows that there is no sequentially perfect pooling equilibrium in this game.

□

**Proposition 1.2.** *In any perfect sequential equilibrium of the game with quality signaling, the type space Q is partitioned into* $Q_\ell = [0, \hat{q})$ *and* $Q_h = [\hat{q}, \overline{q}]$.
*The partitioning must be such that:*

$$c(\hat{q}) = k - 2E(Q_\ell) - E(Q_h) \tag{1.2}$$

*with* $E(Q_i) := E(q|q \in Q_i)$. *Pricing and signaling schemes are given by:*

$$(\tilde{q}, w) = \begin{cases} (\tilde{q}_\ell, w_\ell = k - E(Q_\ell)) & \text{if} \quad q \in Q_\ell \\ (\tilde{q}_h, w_h = k + E(Q_h)) & \text{if} \quad q \in Q_h \end{cases}$$

$$p = \begin{cases} p_\ell := k & \text{if} \quad w \leq w_\ell \\ p_h := k + E(Q_h) & \text{if} \quad w \in (w_\ell, w_h] \end{cases}$$

*With beliefs of the following form:*

$$E(Q_h) = \mu(p, \tilde{q}_h) \geq \mu(p, \tilde{q}) \geq \mu(p, \tilde{q}_\ell) = E(Q_\ell) \quad \forall p, \tilde{q} \notin \{\tilde{q}_\ell, \tilde{q}_h\}$$

*Proof.*
**Step 1:** For a given quality signal $\tilde{q}$, R's optimal strategy can at most be binary.
Fix a set of beliefs $\mu$ and a signal $\tilde{q}$ and let the sets $P_i(\tilde{q}), i \in \{\ell, h\}$ denote all prices that induce sales to a consumer of type $i$, given signal $\tilde{q}$, that is: $P_i(\tilde{q}) = \{p : p \leq k + \theta_i \mu(p, \tilde{q})\}$. The sets $P_i$ are non-empty since $p = k$ must be part of both. Let $p_i(\tilde{q})$ be the highest price that induces sales to consumers of type $i$, $p_i(\tilde{q}) = \max P_i(\tilde{q})$. In the simplified model, $\theta_\ell = 0$ so that $p_\ell = k$, independent of $\tilde{q}$. Because any price that induces sales to the low-types will also induce sales to the high type, $P_\ell \subseteq P_h(\tilde{q})$ and $p_h(\tilde{q}) \geq p_\ell$ for any signal $\tilde{q}$. The retailer's optimal strategy for any $w \leq p_h(\tilde{q})$ is

$$p^*(w, \tilde{q}) = \arg \max_{p \in \{p_\ell, p_h(\tilde{q})\}} \pi_R(p, w)$$

**Step 2:** Given R's optimal strategy and for signal $\tilde{q}$, by setting $w = w_h(\tilde{q}) := p_h(\tilde{q})$, M can extract the whole surplus when sales are made to the high type consumer only.

**Step 3:** M can induce a retail price of $p = k$.
R will set a price of $p = p_\ell = k$ if this yields weakly higher profits than the best deviation. As Step 1 shows that the only deviation is to set a price of $p_h(\tilde{q})$, this means:

$$2(p_\ell - w) \geq p_h(\tilde{q}) - w$$
$$w \leq w_\ell(\tilde{q}) := 2k - p_h(\tilde{q})$$

**Step 4:** M's optimal strategy partitions $Q$ into $Q_\ell$ and $Q_h$.
By setting $w = w_h(\tilde{q})$ for some signal $\tilde{q}$, M can always make positive profits, implying it can't be the case that no sales are made for some realization of $q$.

Steps 2 and 3 showed that retail prices $p_\ell$ and $p_h(\tilde{q})$ can be induced by wholesales prices $w_\ell(\tilde{q})$ and $w_h(\tilde{q})$. Hence, if M is going to send a signal $\tilde{q}$ to consumers, it can only be the case that one signal is sent per corresponding wholesale price, i.e.:

$$w^*(q) = \arg \max_{w \in \{w_\ell(\tilde{q}_\ell), w_h(\tilde{q}_h)\}} \pi_M(w)$$

with

$$\tilde{q}_i = \arg\max_{\tilde{q}} w_i(\tilde{q}) \quad i \in \{\ell, h\}$$

**Step 5:** Beliefs on the equilibrium path have to be correct, i.e. $\mu(p_\ell, \tilde{q}_\ell) = E(Q_\ell) := E(q|q \in Q_\ell)$ and $\mu(p_h(\tilde{q}_h), \tilde{q}_h) = E(Q_h) := E(q|q \in Q_h)$ and thus $w_h = p_h(\tilde{q}_h) = k + E(Q_h)$.
**Step 6:** By the same arguments as in the previous proposition there is a cut-off $\hat{q}$ that equalizes M's profits for both combinations of signals and wholesale prices, hence $\hat{q}$ is defined through:

$$2(w_\ell(\tilde{q}_\ell) - c(\hat{q})) = w_h(\tilde{q}_h) - c(\hat{q})$$
$$2(k - \mu(p_h(\tilde{q}_\ell), \tilde{q}_\ell) - c(\hat{q})) = k + \mu(p_h(\tilde{q}_h), \tilde{q}_h) - c(\hat{q}) = k + E(Q_h) - c(\hat{q})$$
$$c(\hat{q}) = k - 2\mu(p_h(\tilde{q}_\ell), \tilde{q}_\ell) - E(Q_h)$$

**Step 7:** The combination of $(\tilde{q}_\ell, p_h(\tilde{q}_\ell))$ is off the equilibrium path and thus $\mu(p_h(\tilde{q}_\ell), \tilde{q}_\ell)$ is not pinned down by the requirements of a weak perfect Bayesian equilibrium. But, as this step will show, perfect sequentiality of the equilibrium requires $\mu(p_h(\tilde{q}_\ell), \tilde{q}_\ell) = \mu(p_\ell(\tilde{q}_\ell), \tilde{q}_\ell) = E(Q_\ell)$. Let $\mu_Q(\tilde{q}', p')$ be the expected quality after an observed deviation to $(\tilde{q}', p')$ formed according to the interpretation $I(\tilde{q}', p') = Q$.

Take a candidate equilibrium with price, wholesale and signaling scheme constructed according to the previous steps and the partitioning of $Q_\ell$ and $Q_h$ according to equation (1.2).

Remember that by construction of $\tilde{q}_\ell$ and $p_h$ it must be the case that $\mu(p_h(\tilde{q}_\ell), \tilde{q}_\ell) \geq E(Q_\ell)$ and assume $\mu(p_h(\tilde{q}_\ell), \tilde{q}_\ell) > E(Q_\ell)$ in the candidate equilibrium.

Given that the manufacturer can send a different signal to consumers the task is to find a deviation that involves a different wholesale price and a different signal with a consistent interpretation. Consider a deviation to $w', \tilde{q}'$ with $w_h > w' > w_\ell$ so that all consumers would purchase the good. With consumers estimated quality being $\mu_Q(\tilde{q}', p)$ after observing the deviation signal, M can induce a price of $p = k$ by setting $w' = k - \mu_Q(\tilde{q}', p)$. Assuming that $w' > w_\ell$ M would prefer this deviation for all $q \in Q_\ell$. For $q \in Q_h$ M would prefer $w'$ over $w_h$ if

$$\begin{aligned} 2(w' - c(q)) &> k + E(Q_h) - c(q) \\ c(q) &> k - 2\mu_Q(\tilde{q}', p) - E(Q_h) \end{aligned}$$

the interpretation $\mu_{Q'}(\tilde{q}', p) = E(Q')$ where $Q' = [0, q']$ is consistent if

$$\begin{aligned} c(q) &> k - 2E(Q') - E(Q_h) \quad \forall q \in Q' \\ c(q) &< k - 2E(Q') - E(Q_h) \quad \forall q \notin Q' \end{aligned}$$

so that a 'deviation threshold' $q'$ can be constructed as $c(q') = k - 2E(Q') - E(Q_h)$. It remains to check that indeed $w' > w_\ell$, i.e. $E(Q') < \mu(p_h(\tilde{q}_\ell), \tilde{q}_\ell)$. From combining the condition defining the 'original' threshold $\hat{q}$, $c(\hat{q}) = k - 2\mu(p_h(\tilde{q}_\ell), \tilde{q}_\ell) - E(Q_h)$ with the equation defining the 'deviation threshold' we get:

$$c(q') + 2E(Q') = c(\hat{q}) + 2\mu(p_h(\tilde{q}_\ell), \tilde{q}_\ell) > c(\hat{q}) + 2E(Q_\ell)$$

which implies $q' > \hat{q}$ and thus

$$c(q') + 2E(Q') = c(\hat{q}) + 2\mu(p_h(\tilde{q}_\ell), \tilde{q}_\ell)$$

means

$$E(Q') < \mu(p_h(\tilde{q}_\ell), \tilde{q}_\ell)$$

The above showed that if $\mu(p_h(\tilde{q}_\ell), \tilde{q}_\ell) > E(Q_\ell)$ in a weak perfect Bayesian equilibrium, this equilibrium is ruled out by perfect sequentiality. Intuitively, in an equilibrium that is not sequentially perfect, consumers are giving the retailer too much 'power' in forming their beliefs although a deviation of only the retailer can not contain any additional information about the quality, so that beliefs should not change in response to only price changes. This is exactly what the refinement of perfect sequentiality formalizes.

**Step 8:** The same reasoning of the last step in the previous proposition can be used to show that no pooling equilibrium can be sequentially perfect.

□

**Proposition 1.3.** *In any perfect sequential equilibrium of the game with quality signaling, where consumers differ sufficiently to fulfill $n_h\theta_h > m\theta_\ell$, the type space Q is partitioned into $Q_\ell = [0, \hat{q})$ and $Q_h = [\hat{q}, \overline{q}]$ with*

$$c(\hat{q}) = k - m(n_h\theta_h - m\theta_\ell)E(Q_\ell) - n_h\theta_h E(Q_h)$$

*and price, signaling schemes and supporting beliefs are given by:*

$$(\tilde{q}, w) = \begin{cases} (\tilde{q}_\ell, w_\ell(\tilde{q}_\ell) = (1+n_h)p_\ell(\tilde{q}_\ell) - n_h p_h(\tilde{q}_\ell) & \textit{if} \quad q \in Q_\ell \\ (\tilde{q}_h, w_h(\tilde{q}_h) = p_h(\tilde{q}_h)) & \textit{if} \quad q \in Q_h \end{cases}$$

$$p = \begin{cases} p_\ell(\tilde{q}_\ell) = k + \theta_\ell E(Q_\ell) & \textit{if} \quad (\tilde{q}, w) = (\tilde{q}_\ell, w_\ell(\tilde{q}_\ell)) \\ p_h(\tilde{q}_h) = k + \theta_h E(Q_h) & \textit{if} \quad (\tilde{q}, w) = (\tilde{q}_h, w_h(\tilde{q}_h)) \end{cases}$$

$$\mu(p, \tilde{q}_h) = E(Q_h) \geq \mu(p, q) \geq E(Q_\ell) = \mu(p, \tilde{q}_h) \quad \forall p, \tilde{q} \notin \{\tilde{q}_\ell, \tilde{q}_h\}$$

The proof is almost the same as in the previous proposition, with the biggest difference being that the retail price that induces sales to all consumers now also depends on consumers' beliefs.

*Proof.*
**Step 1:** For a given quality signal $\tilde{q}$, R's optimal strategy can at most be binary.
Fix a set of beliefs $\mu$ and a signal $\tilde{q}$ and let the sets $P_i(\tilde{q}), i \in \{\ell, h\}$ denote all prices that induce sales to a consumer of type $i$, given signal $\tilde{q}$, that is: $P_i(\tilde{q}) = \{p : p \leq k + \theta_i\mu(p, \tilde{q})\}$. The sets $P_i$ are non-empty since $p = k$ must be part of both. And $P_\ell \subseteq P_h(\tilde{q})$ and $p_h(\tilde{q}) \geq p_\ell(\tilde{q})$ for any signal $\tilde{q}$.

The retailer's optimal strategy for any $w \leq p_h(\tilde{q})$ is

$$p^*(w, \tilde{q}) = \arg \max_{p \in \{p_\ell(\tilde{q}), p_h(\tilde{q})\}} \pi_R(p, w)$$

**Step 2:** If for a fixed $\tilde{q}$, $p = p_h(\tilde{q})$, M can set $w = w_h(\tilde{q}) = p_h(\tilde{q})$.
**Step 3:** R will set $p = p_\ell(\tilde{q})$ if:

$$\begin{aligned} (1+n_h)(p_\ell(\tilde{q}) - w) &\geq n_h(p_h(\tilde{q}) - w) \\ w &\leq w_\ell(\tilde{q}) := p_\ell(\tilde{q}) - n_h(p_h(\tilde{q}) - p_\ell(\tilde{q})) \end{aligned}$$

As before prices will be such that consumers' expected utility equals zero, i.e. $p_i(\tilde{q}) = k + \mu(\tilde{q}, p_i(\tilde{q}))$.

Hence,

$$\begin{aligned} w_\ell(\tilde{q}) &= k + m\theta_\ell \mu(\tilde{q}, p_\ell(\tilde{q})) - n_h \theta_h \mu(\tilde{q}, p_h(\tilde{q})) \\ w_h(\tilde{q}) &= k + \theta_h \mu(\tilde{q}, p_h(\tilde{q})) \end{aligned}$$

$w_h(\tilde{q})$ is increasing in consumers expectation, thus, whenever $w_\ell(\tilde{q})$ is decreasing in $\mu$, M will optimally choose different signals to accompany each wholesale price. We will for now just assume that this is the case and come back to the conditions on the parameters that guarantee this later on.

M's optimal strategy then is.

$$w^*(q) = \arg \max_{w \in \{w_\ell(\tilde{q}_\ell), w_h(\tilde{q}_h)\}} \pi_M(w)$$

with

$$\tilde{q}_i = \arg\max_{\tilde{q}} w_i(\tilde{q}) \quad i \in \{\ell, h\}$$

**Step 4:** M's optimal strategy partitions $Q$ into $Q_\ell$ and $Q_h$.
By setting $w = w_h(\tilde{q})$ for some signal $\tilde{q}$, M can always make positive profits, implying it can't be the case that no sales are made for some realization of $q$.
**Step 5:** As in the two previous propositions, a unique threshold $\hat{q}$ that equalizes the profits from both combinations of signals and wholesale prices, hence $\hat{q}$ is defined through:

$$\begin{aligned} \pi_M(\tilde{q}_\ell, w_\ell(\tilde{q}_\ell)) &= \pi_M(\tilde{q}_h, w_h(\tilde{q}_h)) \\ m w_\ell(\tilde{q}_\ell) &= n_h w_h(\tilde{q}_h) \\ c(\hat{q}) &= k + m^2 \theta_\ell \mu(\tilde{q}_\ell, p_\ell(\tilde{q}_\ell)) - n_h \theta_h [m\mu(\tilde{q}_\ell, p_h(\tilde{q}_\ell)) + \mu(\tilde{q}_h, p_h(\tilde{q}_h))] \end{aligned}$$

**Step 6:** Beliefs on the equilibrium path have to be correct, i.e. $\mu(p_\ell, \tilde{q}_\ell) = E(Q_\ell)$ and $\mu(p_h(\tilde{q}_h), \tilde{q}_h) = E(Q_h)$, so that

$$\begin{aligned} w_\ell(\tilde{q}_\ell) &= k + m\theta_\ell E(Q_\ell) - n_h \theta_h \mu(\tilde{q}, p_h(\tilde{q})) \\ w_h(\tilde{q}_h) &= k + \theta_h E(Q_h) \end{aligned}$$

**Step 7:** As in Proposition 2, the on-equilibrium wholesale price $w_\ell(\tilde{q}_\ell)$ depends on an off-equilibrium belief, namely on $\mu(\tilde{q}, p_h(\tilde{q}))$. The same Arguments from Step 7 of Proposition 2 imply that in any perfectly sequential equilibrium $\mu(\tilde{q}, p_h(\tilde{q})) = E(Q_\ell)$ so that again a deviation by only the retailer can not contain information. Thus $w_\ell(\tilde{q}_\ell) = k + m\theta_\ell E(Q_\ell) - n_h\theta_h\mu(\tilde{q}, p_h(\tilde{q})) = k + E(Q_\ell)(m\theta_\ell - n_h\theta_h)$ is decreasing in the expected quality, as wanted, whenever $n_h\theta_h > m\theta_\ell$. The cut-off can then be rewritten as

$$c(\hat{q}) = k - (n_h\theta_h - m\theta_\ell)mE(Q_\ell) - n_h\theta_h E(Q_h)$$

□

# Chapter 2

# Quality Disclosure in a Supply Chain - Unraveling from the Top and the Bottom

For many products, consumers can not assess the quality in advance and a firm may rely on certification or disclosure to communicate the quality of its product to consumers. The standard unraveling result in such a setup states that a firm will disclose its quality for all but the lowest realizations. The reasoning is, that consumers 'pool' all realizations for which a firm does not disclose the quality, yielding incentives to defect from this pooling whenever the realized quality is high.

The objective of this chapter is to show that this unraveling result is fundamentally changed if instead of looking at a firm that directly sells its product to consumers, we model a manufacturer who indirectly sells through a retailer. If the manufacturer is given the (costly) opportunity to disclose the quality of his product, he will do so for all but some intermediate quality levels.

## 2.1 Introduction

Purchase decisions of consumers are often characterized by uncertainty about many aspects of the good they are considering and much of this uncertainty vanishes only when consumers have consumed the product in question. This is especially the case for the quality of a product. At the same time, it may well be that the producer of the good has superior knowledge about his product's quality. Such products are typically called experience goods (Nelson, 1970) and they have produced a large amount of literature. In particular the literature explores how information about the quality can credibly be transmitted from the producer to consumers in order to facilitate and influence their purchasing decision.

The literature is roughly separated into two distinct areas, one of information disclosure, and one of information signaling. Whereas the disclosure literature assumes that firms can choose between truthfully revealing the quality of their product or remain silent, the signaling literature models scenarios where consumers can rationally infer a good's quality from observed behavior of the firm. Disclosure can be understood as having a third party firm audition and certify the quality of an experience good. Signaling is mostly modeled as happening through the price, but may as well happen through costly advertising (a seminal paper in this regard is Milgrom and Roberts, 1986).

The seminal papers on quality disclosure (Milgrom, 1981; Grossman, 1981) assume that disclosure is costless and that production costs of the experience good are independent of the quality level. In such a setting the producer of the good will choose to disclose the quality of his product for any quality level. This result can intuitively be understood by considering the manufacturer's incentives to disclose for different levels of product quality. Without any disclosure, consumers must expect the quality to be average, so that the manufacturer will certainly want to disclose the quality of his product whenever it has the highest possible value. Consumers anticipate these incentives of the manufacturer and when they do not observe the manufacturer disclosing the quality, they reason that the good can not be of the highest quality so that the expected quality if no disclosure is observed is now lower than before. This gives the manufacturer incentives to disclose the quality for lower quality levels which again leads consumers to expect an even lower quality if no disclosure is observed. With costless disclosure, this continues for any quality of the product and the sketched process is often called 'unraveling'. If disclosure is not costless the same logic applies, only that now the firm will not disclose the quality for the lowest levels, as the disclosure costs at some point outweigh the gain from disclosing (see for example Jovanovic, 1982; Viscusi, 1978).

Daughety and Reinganum (2008) were the first to combine the two strands of literature. In their model, firms can disclose their products quality at positive costs, but unlike in the classical literature on disclosure, the alternative to disclosing is not pooling of the qualities for which the firm does not disclose. Instead the firm chooses whether to communicate the quality of its product to consumers via disclosure or via signaling through the retail price, so that consumers will in equilibrium know the quality of the good irrespectively of whether disclosure has taken place or not. Daughety and Reinganum show that there is a range of disclosure costs such that the firm chooses to reveal its quality via price-signaling for low quality levels and it chooses to disclose for relatively high quality levels. They emphasize that disclosure is less frequent than with models were the firm has no opportunity of revealing the quality by signaling.

In all the aforementioned papers, the firm that produces the good is also the one which is selling it to consumers. This chapter builds on the idea of Daughety and Reinganum (2008), namely that a firm may choose between using disclosure and signaling via the retail price to communicate its quality, and applies it to a manufacturer-retailer model, where the firm producing the good is not the one selling it to consumers. Both firms in the setup are monopolists and the occurrence of a double marginalization effect or 'problem' is well known in such a context and was already anticipated by Cournot (1863) and formally derived in Spengler (1950). The double-marginalization effect states that in such a setup the retail price is even higher than the one where a monopolist directly sells the good to consumers. This is the case as the monopolistic retailer now also exhibits market power and thus will charge an additional markup thereby inflating the retail price.

It will be shown that in this setup, and contrary to the existing literature, the manufacturer will decide to disclose the quality of his product for all but some intermediary levels. This will be seen to be profitable for him as he can use his disclosure decision to weaken the problem of double marginalization.

One could ask whether a firm in reality would disclose or certify low quality of its product. Interestingly this is precisely what a German manufacturer of mattresses did recently. The German testing institute ‘Stiftung Warentest’ (which translates as ‘foundation for product test’ and is similar to consumer reports in the United States) examined mattresses of five start-up firms.[1] The five mattresses received grades from 2.3 to 4.7 (where 1 is the best and equals the American ‘A’ and the worst grade, 5 equals ‘F’) and surprisingly one of the firms, called ‘muun’, with one of the worst test grades, namely 4.1, advertised this grade in a marketing campaign, calling their mattress the ‘most sufficient’ (sufficient describes the grade 4 in German schools).[2] The only other firm which advertised the result of the test, was the winner with the highest grade.[3] While the goal of muun seemingly was to put the test results into perspective, it is nevertheless interesting that only one of the worst and the test winner used the test results in their advertising. It might also be well the case that potential consumers only got to the know the test results because of the advertising campaign, and that they wouldn’t have found out about it otherwise.

Clearly, only the test winner and one of the firms with the worst test results advertising or disclosing the test results, is not compatible with any of the disclosure patterns of the previous literature, but as we will see it is more in line with the results of this chapter.

The chapter is structured as follows: The next section presents the model. Optimal price setting for a setting with full information and for one with asymmetric information are derived in section 2.3. In section 2.4, these results are compared order to assess disclosure incentives in the model. The closing section concludes.

[1] See for example (in German) `http://www.welt.de/icon/moebel/article158957032/Wie-ein-Matratzen-Start-Up-Stiftung-Warentest-vorfuehrt.html` (last accessed November 28, 2016)

[2] See (in German) `https://muun.co/stiftungwarentest` (last accessed November 28, 2016)

[3] See (in German) `https://www.home24.de/smood-shop/` (last accessed November 28, 2016)

## 2.2 Model Setup

The model resembles the one from the previous chapter with the main difference being that the focus is on disclosure instead of on cheap-talk.

The model features two firms (both "he"), a manufacturer (M) and a retailer (R). M produces an experience good of random quality $q$ with $q$ being distributed according to a distribution function $F(q)$ with full support on the interval $Q = [0, \overline{q}]$. The manufacturer has to pay $c(q)$ to produce one unit of the good, higher quality goods are more costly to produce, so we impose $c'(q) > 0$. Additionally we normalize $c(0) = 0$. After the production, M sells the good(s) to the retailer, who observes $q$, at the wholesale price $w$. The retailer then serves the customers at the retail price $p$. Besides paying the wholesale price $w$, R does not incur any costs when making sales.

At the same time M sets his wholesale price, he can decide whether to disclose the quality level or not. Disclosing, which comes at costs $d > 0$ for the manufacturer, leads to the realization of $q$ being public knowledge. $D \in \{0, 1\}$ denotes M's disclosure decision, with $D = 1$ if M disclosed the quality, and $D = 0$ otherwise. Disclosure and the costs it bears can be seen as a costly form of certification by a third party (not modeled here) like in the example given above, so that the manufacturer has no possibility to lie, he can only choose to reveal the quality of his product truthfully or to remain silent.

On the consumer side there are two types of consumers with unit demand, considering to purchase the good. Consumers (referred to as "she"), are either of a low-type ($i = \ell$) or of a high-type ($i = h$). There are $m$ consumers in total, $n_h$ of which are high-types and $n_\ell$ are low-types. For convenience, we normalize $n_\ell = 1$. The high-types utility from consumption of one unit of the good is always weakly higher than the utility of the low-type when consuming a unit of the same quality. More precisely, consuming one unit of the good leads to the following ex-post utility:

$$v_i(q) = k + \theta_i q$$

where $\theta_i \in \{\theta_\ell, \theta_h\}$ is the consumers type with $0 < m\theta_\ell < n_h\theta_h$.

Consumers may not know the quality of the good at the time they consider purchasing it, so they have to form beliefs, mapping the available information into an assessment of the good's quality. By construction of the demand, any consumer only cares about the expected value of the quality so that we can reduce the belief. Formally, let $\beta : (p, D) \to \Delta Q$ be a consumer's belief and let $\mu(p, D) = E_{\beta(p,D)}(q)$ be the expected quality derived from this belief. A consumer purchases the good whenever her expected utility is at least as large as the charged retail price, thus a consumer of type $\theta_i$ then purchases the good whenever

$$E(u_i(q)) = k + \theta_i \mu(p, D) \geq p. \tag{2.1}$$

As the manufacturer has no possibility to lie, all consumers know the realized quality if M decided to disclose so that $\mu(p, D = 1) = q$ for any price $p$.

Perfect Bayesian Equilibrium will be the employed solution concept, which here means:

**Requirement 2.1.** *Firms maximize their profit given the other firm's and consumers' strategies. Consumers act optimally given the strategies of the firms and their beliefs $\mu$.*

**Requirement 2.2.** *Consumers' beliefs are derived from the firms' strategies and Bayes' rule whenever possible.*

Regarding the relation of the different parameters, the condition $c(\overline{q}) \geq k - \overline{q}(n_h\theta_h - m\theta_\ell)$ will be imposed. It ensures that if consumers know that the quality is given by $\overline{q}$ the joint profits of the two firms are maximized with sales to the high-type consumers only and not by sales to all consumers. Additionally, we will assume that $k \geq n_h\theta_h E(q)$, which guarantees that the firms joint profits at a quality level of $q = 0$ if $q$ were known, are higher than selling to all consumers with the prior expectation of the quality. The above condition ensures that there is some value $q' \in Q$, such that joint firm profits are maximized by selling to all consumers for lower values and by selling to high-type consumers only, for higher values of $q$. We will see in the next sections that this condition will allow for separation to take place.

## 2.3 Optimal Price Setting

We are interested in the manufacturer's incentives to disclose his quality to consumers, to this end we need to analyze the subgames starting with the manufacturer's decision to disclose or not to disclose. In the next step we can then compare the results in order to obtain M's optimal disclosure decision. In each case we will employ backwards induction logic, starting with characterizing the behavior of the consumers, then looking at R's optimal decision, given the consumer behavior and in the last step we analyze M's optimal strategy, given the behavior of all other agents.

We will begin with a subgame that started with the manufacturer disclosing $q$. Given that consumers (and by assumption, the retailer) know $q$ after $M$ decided to disclose, the following subgame is one of perfect information.

### 2.3.1 Full Information

With the consumers being perfectly informed about the quality $q$, a consumer of type $\theta_i$ will purchase the good whenever $v_i(q) = k + \theta_i q \geq p$. The retailer anticipates this consumer behavior and optimally chooses his retail price $p$. Clearly, for a fixed quality $q$, the demand is constant and equal to $m$ for all prices $p \leq p_\ell^D(q) := k + \theta_\ell q$, it equals $n_h$ for prices $p$ such that $p_\ell^D(q) < p \leq p_h^D(q) := k + \theta_h q$ and is zero for higher prices.

R will set the price that maximizes his profit $\pi_R^D(p)$, and if $w$ is not prohibitively high, this is achieved by setting one of the prices $p_\ell^D(q)$ or $p_h^D(q)$. R will charge $p_\ell^D(q)$ if

$$\begin{aligned}
&\pi_R^D(p_\ell^D(q)) = m(k + \theta_\ell q - w) \geq n_h(k + \theta_h q - w) = \pi_R^D(p_h^D(q)) \\
\Leftrightarrow \quad & w \leq w_\ell^D(q) := k + (m\theta_\ell - n_h\theta_h)q
\end{aligned}$$

As the demand at R is a step function, the same is true for the quantities R demands from M, where R demands $m$ units if $w \leq w_\ell^D(q)$, $n_h$ units if $w_\ell^D(q) < w \leq p_h^D(q)$ and zero for higher wholesale prices. This implies that M can leave R with a profit of zero if the wholesale price is set to $w_h^D(q); = p_h^D(q)$. Anticipating R's behavior, M optimally chooses between setting either of the wholesale prices $w_\ell^D(q)$ and $w_h^D(q)$. M is indifferent between $w_\ell^D(q)$ and $w_h^D(q)$ if

$$\begin{aligned}
&\pi_M^D(w_\ell^D(q)) = m(k + (m\theta_\ell - n_h\theta_h) - c(q)) = n_h(k + \theta_h q - c(q)) = \pi_M^D(w_h^D(q)) \\
\Leftrightarrow \quad & c(q) = k + (m^2\theta_\ell - n_h(1+m)\theta_h)q
\end{aligned}$$

By assumption, the two consumer groups differ enough so that $n_h\theta_h > m\theta_\ell$ and thus, this equation has a unique solution, the solution will be denoted by $\hat{q}^D$. For values of $q < \hat{q}^D (q > \hat{q}^D)$, M will set a price of $w_\ell^D$ $(w_h^D)$.

The following summarizes the equilibrium in the game with full information:

**Lemma 2.1.** *In the game with full information, the equilibrium wholesale and retail pricing schemes are given as follows*

$$w = \begin{cases} w_\ell^D(q) = k + (m\theta_\ell - n_h\theta_h)q & \text{if} \quad q \leq \hat{q}^D \\ w_h^D(q) = p_h^D(q) = k + \theta_h q & \text{else} \end{cases}$$

$$p = \begin{cases} p_\ell^D(q) = k + \theta_\ell q & \text{if} \quad w \leq w_\ell^D(q) \\ p_h^D(q) = k + \theta_h q & \text{else} \end{cases}$$

*The threshold $\hat{q}^D$ is defined by the following equation*

$$c(\hat{q}^D) = k + (m^2\theta_\ell - n_h(1+m)\theta_h)\hat{q}^D$$

*If $p = p_\ell^D(q)$ all consumers buy, at the higher price only the high types purchase the good.*

### 2.3.2 Asymmetric Information

Let us next look at a situation where disclosure costs are prohibitively high, so that M will never find it profitable to disclose his products' quality. Much of the optimal behavior of firms and consumers is similar to the previous characterization, with the main difference being that instead of using the realized quality $q$ as before, we now have to take into account the beliefs of consumers.

If Consumers do not know the quality of M's product, they purchase whenever their expected utility is positive. Fix a set of beliefs $\mu$ and, for $i \in \{\ell, h\}$, define a set of prices $P_i$, so that consumers of type $i$ purchase the good for all prices $p \in P_i$. From equation (2.1) these sets are given by $P_i := \{p : p \leq k + \theta_i \mu(p, D = 0)\}$. Let $p_i(\mu(p_i, D))$ denote the maximal element of the set $P_i$, $p_i(\mu(p_i, D)) := \max P_i$. For ease of notation, let $\mu_i := \mu(p_i, D = 0)$, be the belief induced by the maximal prices consumers of type $i \in \{\ell, h\}$, are willing to pay. $p_i(\mu_i)$ thus gives the highest price for which consumers of type $i$ are willing to purchase the product. As high-type consumers purchase the good whenever low-types do, it must be the case that $p_\ell(\mu_\ell) \leq p_h(\mu_h)$

Clearly, the demand the retailer faces, again is a step function and equal to $m$ for all prices $p \leq p_\ell(\mu_\ell)$, it equals $n_h$ for prices $p_\ell(\mu_\ell) < p \leq p_h(\mu_h)$ and is zero for higher prices.

The goal of the retailer is to maximize his profit $\pi_R(p)$, which, if $w$, the wholesale price charged by the manufacturer, is not prohibitively high, is achieved by a retail price of either $p_\ell(\mu_\ell)$ or $p_h(\mu_h)$. R will set a price of $p_\ell(\mu_\ell)$ if

$$\begin{aligned} & \pi_R(p_\ell) = m(k + \theta_\ell \mu_\ell - w) \geq n_h(k + \theta_h \mu_h - w) = \pi_R(p_h) \\ \Leftrightarrow \quad & w \leq w_\ell(\mu_\ell)^4 := k + m\theta_\ell \mu_\ell - n_h \theta_h \mu_h \end{aligned}$$

As was the case with full information, because the demand at R is a step-function, the same is true for the quantities R demands from M and R demands $m$ units for $w \leq w_\ell(\mu_\ell)$, $n_h$ units if $w_\ell(\mu_\ell) < w \leq p_h(\mu_h)$ and zero units for higher values of the wholesale price. Given the structure of the demand M faces, he can leave R with zero profit by setting $w = w_h(\mu_h) := p_h(\mu_h)$ and M's optimal wholesale price must again be either $w_\ell(\mu_\ell)$ or $w_h(\mu_h)$. M is indifferent between $w_\ell(\mu_\ell)$ and $w_h(\mu_h)$ if

$$\begin{aligned} & \pi_M(w_\ell) = m(k + m\theta_\ell \mu_\ell - n_h \theta_h \mu_h - c(q)) = n_h(k + \theta_h \mu_h - c(q)) = \pi_M(w_h) \\ \Leftrightarrow \quad & c(q) = k + m^2 \theta_\ell \mu_\ell - n_h(1 + m)\theta_h \mu_h \end{aligned} \tag{2.2}$$

[4]Note that $w_\ell$ clearly depends on both beliefs $\mu_\ell$ and $\mu_h$, nevertheless we will in the following only use $\mu_\ell$ as an argument, as the region where a wholesale price of $w_\ell$ is set defines the belief $\mu_\ell$.

As it can not be the case that no sales are made in equilibrium, at least one of the beliefs $\mu_\ell$ and $\mu_h$ is on the equilibrium path. Depending on those beliefs $\mu_\ell$ and $\mu_h$, this equation may or may not have a solution. If the equation has a solution, it will be referred to as $\hat{q}$, and the properties of the cost function $c(q)$ imply that M must then prefer to set a wholesale price $w_\ell(\mu_\ell)$ for $q < \hat{q}$ and he will find it optimal to set $w_h(\mu_h)$ for $q > \hat{q}$.

Combining these observations, it is clear that, depending on the beliefs $\mu_\ell$ and $\mu_h$, three qualitatively different equilibria might emerge. Either we observe a pooling equilibrium where sales are only made to the low type consumers or a pooling equilibrium where only high type consumers buy, and the last possibility is given by a separating equilibrium constructed according to equation (2.1).

The three possibilities are summarized in the following lemma. For brevity, only the behavior on the equilibrium path is stated, off equilibrium behavior in each of the three equilibria is characterized by the above descriptions.

**Lemma 2.2.** *If disclosure costs are prohibitively high, depending on the beliefs of consumers $\mu_\ell$ and $\mu_h$, one of the following situations with stated wholesale and retail prices will emerge in equilibrium.*

1. *High type pooling equilibrium:*

$$w = w_h(E(q)) = p_h(E(q)) = k + \theta_h E(q)$$
$$p = p_h(E(q)) = k + \theta_h E(q)$$

2. *Low type pooling equilibrium:*

$$w = w_\ell(E(q)) = k + m\theta_\ell E(q) - n_h\theta_h\mu_h$$
$$p = p_\ell(E(q)) = k + \theta_\ell E(q)$$

3. *(Partly) separating equilibrium:*

$$w = \begin{cases} w_\ell\,(E(Q_\ell)) = k + m\theta_\ell E(Q_\ell) - n_h\theta_h E(Q_h) & \text{if} \quad q \leq \hat{q} \\ w_h\,(E(Q_h)) = p_h = k + \theta_h E(Q_h) & \text{else} \end{cases}$$

$$p = \begin{cases} p_\ell = k + \theta_\ell E(Q_\ell) & \text{if} \quad w = w_\ell \\ p_h = k + \theta_h E(Q_h) & \text{if} \quad w = w_h \end{cases}$$

*where $\hat{q}$ is defined by equation (2.2) and, $\mu_i = E(Q_i) := E(q|q \in Q_i)$ for $i \in \{\ell, h\}$ with $Q_\ell = [0, \hat{q}]$ and $Q_h = (\hat{q}, \overline{q}]$.*

*In any equilibrium, all consumers buy if $p = p_\ell$, at the higher price of $p = p_h$ only the high types purchase the good.*

## 2.4 Incentives to Disclose

If we want to investigate the manufacturer's incentives to disclose his quality, we need to compare his profit in a subgame after disclosure with his profit without disclosure. With sufficiently small disclosure costs, it is clear that in equilibrium, there must be realizations of $q$ for which M will prefer to disclose the quality and induce sales to all customers, as well as situations where he discloses and intends to have only high type consumers served. For example at $q = 0$ M will find disclosure and serving all consumers profitably if disclosure costs are low enough, similarly if $q = \overline{q}$ it is optimal for M to disclose and serve high type consumers only. We can thus use the separating equilibrium from the above lemma as the starting point.

Let us for the moment assume, that M already decided on how to divide the interval $Q$ with regards to what consumers he wants to serve, that is, let us assume that M will set a wholesale price of $w_\ell$ or $w_\ell^D(q)$ if $q \in Q_\ell = [0, \hat{q}_\ell^h]$ and the wholesale price will be $w_h$ or $w_h^D(q)$ if $q \in Q_h = (\hat{q}_\ell^h, \overline{q}]$ for some $\hat{q}_\ell^h \in (0, \overline{q})$. By $Q_i^D$ we will denote the region of the type space $Q$ where M chooses to disclose and sets the corresponding wholesale price $w_i^D(q)$ and similarly, $Q_i^{ND}$ denotes the a region without disclosure and corresponding price $w_i\left(E(Q_i^{ND})\right)$. As before, we will write $E(Q_i^j) = E(q|q \in Q_i^j)$ with $i \in \{\ell, h\}, j \in \{\mathrm{D}, \mathrm{ND}\}$.

Comparing the wholesale pricing schemes in the interval $Q_h$ with and without disclosure, it is clear that for sufficiently small disclosure costs, M has an incentive to disclose his quality for very high quality levels. This is easily seen by comparing the wholesale prices $w_h^D$ and $w_h$:

$$w_h^D(q) - w_h\left(E(Q_h)\right) = \theta_h(q - E(Q_h))$$

This difference clearly is negative for $q < E(Q_h)$ and (weakly) positive otherwise. The difference is maximized at $q = \overline{q}$, meaning that M has the greatest incentives to disclose his quality when his quality is the highest. Given that consumers know, M would disclose his quality whenever it is the highest, they reason that the quality must be lower if it is not revealed leading to M preferring to disclose his quality for marginally smaller values of q. This unraveling continues until the increase in revenue equals the costs for disclosure, that is, until

$$w_h^D(\hat{q}_h) - w_h\left(E\left(Q_h^{ND}\right)\right) = \theta_h[\hat{q}_h - E(Q_h^{ND})] = d \tag{2.3}$$

This process of unraveling is precisely the same identified in most of the previous literature on disclosure, and it is clear that with positive disclosure costs, M will never disclose for all quality realizations if he intends to serve all consumers. Clearly, for positive disclosure costs $d$, $\hat{q}_h > \hat{q}_\ell^h$ .

The more interesting, and different, situation is the one where M wants to induce R to serve all consumers. We have already seen that with disclosed quality levels, M's wholesale price is decreasing in the quality $q$, which as we will see, changes the unraveling behavior fundamentally. To this end, we again compare wholesale prices with and without disclosure:

$$\begin{aligned} w_\ell^D(q) - w_\ell\left(E(Q_\ell)\right) &= m\theta_\ell\left(q - E(Q_\ell)\right) + n_h\theta_h\left(E(Q_h) - q\right) \\ &= -q(n_h\theta_h - m\theta_\ell) - m\theta_\ell E(Q_\ell) + n_h\theta_h E(Q_h) \end{aligned}$$

For fixed values $E(Q_\ell)$ and $E(Q_h)$, the difference is decreasing in $q$, and since it is positive for $q = 0$, it is positive for any $q \in Q_\ell$, meaning that for very small disclosure costs $d$, M will choose to disclose the quality for all $q \in Q_\ell$. This also implies that it can not be the case that M chooses to disclose his quality for some $q' \in Q_\ell$ but he chooses not to disclose for some $q < q'$. Which means that it will be never the case that the manufacturer chooses to disclose his quality for all but the smallest quality levels as is commonly the case in the disclosure literature. Instead, M will either disclose the quality for all $q \in Q_\ell$ or for all values in $Q_\ell^D = [0, \hat{q}_\ell]$ where $\hat{q}_\ell$ is defined through

$$w_\ell^D(\hat{q}_\ell) - w_\ell\left(E\left(Q_\ell^{ND}\right)\right) = m\theta_\ell\left(\hat{q}_\ell - E\left(Q_\ell^{ND}\right)\right) + n_h\theta_h\left(E\left(Q_h^{ND}\right) - \hat{q}_\ell\right) = d. \tag{2.4}$$

Depending on the parameters, this equation may have a solution $\hat{q}_\ell \in Q_\ell$. In this case $Q_\ell$ consists of $Q_\ell^D$ and $Q_\ell^{ND}$, otherwise $Q_\ell = Q_\ell^D$. The above discussion implies that $Q_\ell$ will include $Q_\ell^{ND}$ whenever disclosure costs are at an intermediate level and for smaller disclosure costs, $Q_\ell = Q_\ell^D$.

Combining the disclosure incentives when selling to high type consumers only with the just derived incentives when M intends to sell to all consumers, we observe that if disclosure costs are sufficiently small, M will always disclose for the lowest and the highest quality realizations. What we have not yet calculated are the thresholds quality levels for which M switches from disclosure to non-disclosure and from selling to all consumers to serving only high type consumers. We have seen that $Q_h$ will always consist of $Q_h^D$ and $Q_h^{ND}$, and with the two possibilities for the interval $Q_\ell$ two situations might emerge in equilibrium. $\hat{q}_h$ is always defined by equation (2.3). For the rest of the interval $Q$, we have to distinguish the two different scenarios for the interval $Q_\ell$.

First, M may not disclose $q$ in regions where he intends to sell to all consumers, as well as in regions where sales are only made to high-type consumers, which as we know happens for intermediate disclosure costs, and in this case equation (2.4) has a solution. In this case, the type space will be partitioned as shown in the following picture:

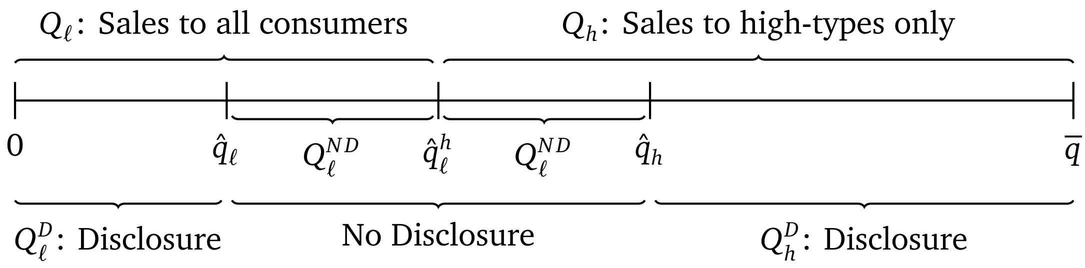



Figure 2.1: Partitioning of the Type Space for Intermediate Disclosure Costs

As $\hat{q}_\ell$ and $\hat{q}_h$ are defined by equations (2.3) and (2.4), the remaining threshold to be calculated is $\hat{q}_\ell^h$. At the threshold $\hat{q}_\ell^h$, M is in different between sales to all consumers and sales to the high-types only, both without disclosing his price. Let $\pi_m(w)$ denote M's profit, ignoring disclosure costs, where for ease of notation $w = w_i^D$ includes disclosure of M, and $w = w_i$ implies no-disclosure, with $i \in \{\ell, h\}$. $\hat{q}_\ell^h$ is then defined by:

$$\pi_M\left[w_h\left(E\left(Q_h^{ND}\right)\right)\right] = \pi_M\left[w_\ell\left(E\left(Q_\ell^{ND}\right)\right)\right] \tag{2.5}$$

Where the left hand site of this equation is increasing in $\hat{q}_\ell^h$ and the right hand side is decreasing in the value $\hat{q}_\ell^h$.

Alternatively the parameters might make M disclose for any quality when sales are made to all consumers, as outlined in the next figure

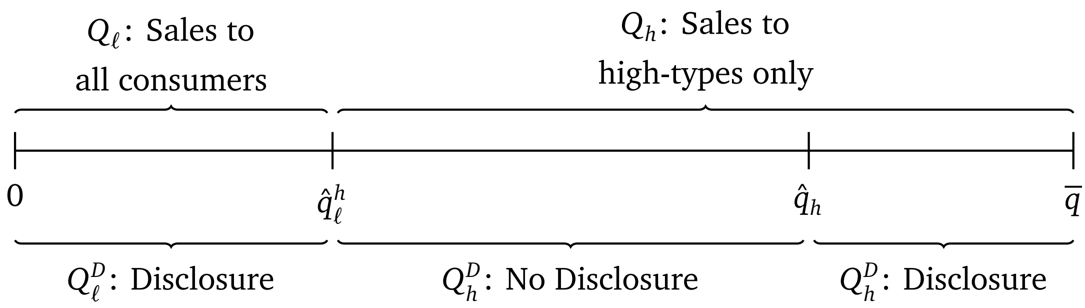



Figure 2.2: Partitioning of the Type Space for Small Disclosure Costs

In this case, M switches directly from selling to all consumers and disclosing, to serving only the high type consumers, which as we have seen before, must for some interval happen without disclosure given that $d > 0$. In such a situation, we only need to calculate one remaining threshold. The threshold $\hat{q}_h$ , characterizing the switch from serving the high type consumers only without and with disclosure, is again defined by (2.3). The remaining threshold, for a given $\hat{q}_h$ defining the value $\hat{q}_\ell^h$ in this setting, characterizes a situation where the manufacturer is indifferent between serving all consumers while disclosing the quality and serving the high type consumers only, then without disclosure. $\hat{q}_\ell^h$ is implicitly defined by the following equations.

$$\pi_M\left[w_\ell^D(\hat{q}_\ell^h)\right] - \pi_M\left[w_h\left(E\left(Q_h^{ND}\right)\right)\right] = d \tag{2.6}$$

As $\pi_M\left[w_\ell^D(\hat{q}_\ell^h)\right]$ is decreasing in $\hat{q}$ and $\pi_M\left[w_h\left(E\left(Q_h^{ND}\right)\right)\right]$ is increasing in $\hat{q}_\ell^h$, the assumption on the parameters imply that this equation has a solution for sufficiently small disclosure costs $d$.

The following proposition summarizes the findings of this section:

**Proposition 2.1.** *For sufficiently small disclosure costs, M discloses his products' quality for a all levels below* $\min\{\hat{q}_\ell^h, \hat{q}_\ell\}$ *and above* $\hat{q}_h$ *and he doesn't disclose q for intermediate levels.*
*If equation* (2.4) *has a solution, the type space Q is partitioned into* $Q_\ell^D = [0, \hat{q}_\ell]$, $Q_\ell^{ND} = (\hat{q}_\ell, \hat{q}_\ell^h]$, $Q_h^{ND} = (\hat{q}_\ell^h, \hat{q}_h]$, $Q_h^D = [\hat{q}, \overline{q}]$ *where the thresholds are defined trough equations* (2.3) *to* (2.5).
*Else the type space is partitioned into* $Q_\ell^D = [0, \hat{q}_\ell^h], Q_\ell^{ND} = \emptyset, Q_h^{ND} = (\hat{q}_\ell^h, \hat{q}_h], Q_h^D = [\hat{q}, \overline{q}]$ *with thresholds defined by equations* (2.3) *and* (2.6).
*Wholesale and retail pricing schemes are as follows:*

$$w = \begin{cases} w_\ell^D(q) & \text{if} \quad q \in Q_\ell^D \\ w_\ell\left(E\left(Q_\ell^{ND}\right)\right) & \text{if} \quad q \in Q_\ell^{ND} \\ w_h\left(E\left(Q_h^{ND}\right)\right) & \text{if} \quad q \in Q_h^D \\ w_h^D(q) & \text{if} \quad q \in Q_h^{ND} \end{cases} \quad \text{and} \quad p = \begin{cases} p_\ell^D(q) & \text{if} \quad w = w_\ell^D(q) \\ p_\ell\left(E\left(Q_\ell^{ND}\right)\right) & \text{if} \quad w = w_\ell\left(E\left(Q_\ell^{ND}\right)\right) \\ p_h\left(E\left(Q_h^{ND}\right)\right) & \text{if} \quad w = w_h\left(E\left(Q_h^{ND}\right)\right) \\ p_h^D(q) & \text{if} \quad w = w_h^D(q) \end{cases}$$

## 2.5 Conclusion

The chapter has shown that if a Manufacturer of an experience good in a vertical supply chain setup is given the opportunity to disclose his quality at positive costs, he will choose to do so for any possible realization of the quality except for some intermediate values.

This is in stark contrast to the usual unraveling result, where, in a direct-selling setup, the firm will disclose the quality for all but a range of the smallest possible values. The reasoning is that the firm can extract higher rents from consumers if they perceive the quality to be higher. Given that the firm has the largest incentives to disclose the quality of its product if the quality is at the highest level, consumers expect the quality to be lower than the highest level if the firm does not disclose. This generates disclosure incentives for slightly smaller realizations of the quality, which again reduces consumers expectations if no disclosure is observed. This unraveling process continues for all but the smallest quality realizations where the additional revenue after disclosure is outweighed by the disclosure costs.

In the presented model of a vertical supply chain, this effect and the unraveling from the top are also present. But additionally, the profit of the manufacturing firm is decreasing in the consumers' perceived quality whenever the manufacturer targets as many consumers as possible. This implies that the manufacturer now also has an incentive to disclose the product quality at the lowest level, which makes consumers reason that if many consumers are served and no disclosure is observed the quality must be higher than the lowest level. Keeping the served consumers constant, this reduces the manufacturers profit leading to disclosure incentives for quality realizations slightly above the lowest one, which then leads to consumers increasing their expectations if they do not observe disclosure, and as the unraveling from the top, this process continues until the additional revenue after disclosure is outweighed by the disclosure costs.

# Chapter 3

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

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 with sequential consumer purchases and a second dimension of variation, quality, about which the consumers have differential information. 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. In some equilibria firms differentiate and transparency enhancing policies may be welfare decreasing.

## 3.1 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?

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

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 (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 consumers 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' strate-

[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.

[3]In the next chapter we show that similar effects also arise in a setup of non-spatial differentiation with endogenous prices.

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

---

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

[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).

[8] See page 5 of the official press kit at: `http://www.festival-cannes.com/assets/Image/Direct/041859.PDF` (last accessed: December 4, 2015). Additionally, the success of the movie

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 chapter is to show in a theoretical model how the firms' incentives to differentiate are affected by social learning among heterogeneous consumers.

The remainder of this chapter is structured as follows. In Section 3.2 we review the related literature. Section 3.3 presents a short and simplified example, while the full model is introduced in Section 3.4. A benchmark model and the main model are solved in Sections 3.5 and 3.6 respectively, welfare comparisons are made in Section 3.6.5. Finally, Section 3.7 concludes. The proofs are relegated to the Appendix.

## 3.2 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 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).

was called 'surprising' by the media, see e.g. http://www.theguardian.com/film/2012/feb/04/hollywood-nostalgia-chaplin-valentino (last accessed: December 4, 2015).

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

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 - 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 model. Consumers have

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

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 consumers) 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.

## 3.3 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.

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

[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'.

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€ 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 $\beta$ 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

$$E[v_B - v_A] = \beta \cdot 30€ + (1-\beta) \cdot 10€ - 20€ = (2\beta - 1) \cdot 10€. \tag{3.1}$$

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 $\beta_A^u$ 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 (3.1) 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 (3.1) 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 3.1. 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

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

as the consumers located at $x = 0.5$, once they traveled to location 0.5. Consumers 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 $\mathfrak{D}^{3A}$.

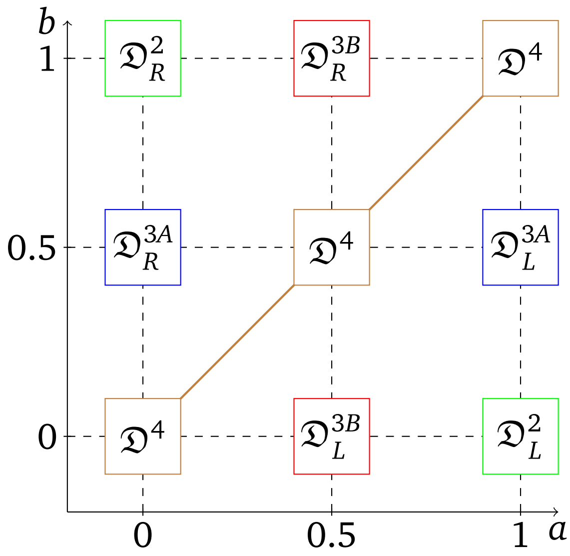



Figure 3.1: Action space for the discrete example.

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 $(a^*, b^*) = (0.5, 1)$.

## 3.4 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 $a \in \mathscr{A} := [0, 1]$, and $b \in \mathscr{B} := [0, 1]$. 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 $\mathscr{I}_A = \{\emptyset\}$, and firm B knows $\mathscr{I}_B = \{a\}$.

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. $b(a) : \mathscr{A} \to \mathscr{B}$.

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 $t = 1$

[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.

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 $(x_t, \phi_t)$.

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 $\beta$ about the probability of firm $B$ offering the product of higher quality by mapping the available information $\mathcal{I}$ into the probability space, $\beta := Pr(v_B >$

[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.

$v|\mathscr{I}) \in [0,1]$. The belief of an uninformed consumer is the function $\beta^u : \mathscr{A} \times \mathscr{B} \times \{\emptyset, A, B\} \to [0,1]$, and that of an informed consumer is the function $\beta^i : \mathscr{A} \times \mathscr{B} \times \{\emptyset, A, B\} \times \{v-\delta, v+\delta\} \to [0,1]$. 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 $\beta$ 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 $\beta$ 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 $\beta$ 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 $\beta$ is indifferent between the products of $A$ and $B$ if

$$E[u(x,\beta,a,b|B)] - u(x,\beta,a,b|A) = (2\beta - 1)\delta - \tau(|x-b| - |a-x|) = 0. \tag{3.2}$$

Let us define

$$\bar{x}(\beta) := \frac{a+b}{2} + \frac{\delta}{\tau}\Big(\beta - \frac{1}{2}\Big), \tag{3.3}$$

which, for a given belief $\beta$, coincides with the consumer type $x$ solving equation (3.2) whenever (3.2) 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 $\beta$ 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 $\beta$ 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 $\beta$ 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 $\beta$. Similarly, if $\bar{x}(\beta) > a$, all consumer types $x$ with belief $\beta$ 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 $\beta$ choose the product of firm B. Threshold $\tilde{x}(\beta)$ equals the indifferent consumer type in $(b, a)$ with belief $\beta$ whenever it exists. Otherwise, $\tilde{x}(\beta) = 0$, if all consumers with belief $\beta$ prefer $A$, and in the analogous case, when all consumers with $\beta$ prefer to buy from $B$, $\tilde{x}(\beta) = 1$. Thus, the threshold type can be calculated as

$$\tilde{x}(\beta) = \begin{cases} 0 & \text{if} \quad \beta < \frac{1}{2} - \frac{\tau}{\delta} \cdot \frac{a-b}{2} & \Leftrightarrow \bar{x}(\beta) < b, \\ 1 & \text{if} \quad \beta > \frac{1}{2} + \frac{\tau}{\delta} \cdot \frac{a-b}{2} & \Leftrightarrow \bar{x}(\beta) > a, \\ \bar{x}(\beta) & \text{if} \quad \beta \in \left[\frac{1}{2} - \frac{\tau}{\delta} \cdot \frac{a-b}{2}, \frac{1}{2} + \frac{\tau}{\delta} \cdot \frac{a-b}{2}\right] & \Leftrightarrow \bar{x}(\beta) \in [b, a]. \end{cases} \tag{3.4}$$

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

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

[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.

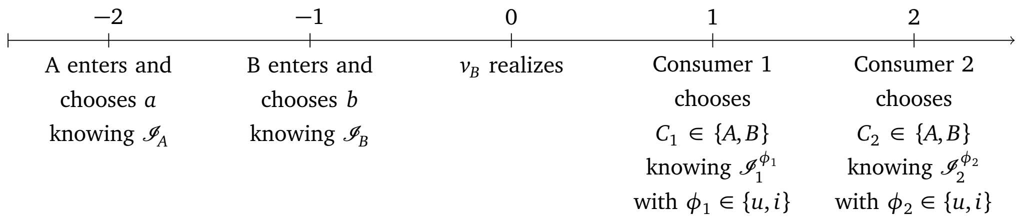



Figure 3.2: Timing of the Game

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 the next chapter).

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, 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.

[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.

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.

## 3.5 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 3.1 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.

### 3.5.1 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 (3.4) 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 $\beta_A^i := Pr(v_B = v + \delta | v_B < v_A) = 0$ to denote the belief of an informed consumer when $A$ is the better firm, and $\beta_B^i := Pr(v_B = v + \delta | v_B > v_A) = 1$ analogously for the case where $B$ sells the superior product. Using equation (3.4) and defining

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

which solves the equations

$$b \quad s.t. \quad \bar{x}(\beta_A^i) = b \qquad \text{and} \qquad b \quad s.t. \quad \bar{x}(\beta_B^i) = a, \tag{3.5}$$

threshold types for informed consumers can easily be calculated as

$$\tilde{x}(\beta_A^i) = \begin{cases} \bar{x}(\beta_A^i) := \frac{a+b}{2} - \frac{\delta}{2\tau} & \text{if} \quad b \leq b_1(a) \Leftrightarrow \bar{x}(\beta_A^i) \in [b, a], \\ 0 & \text{if} \quad b > b_1(a) \Leftrightarrow \bar{x}(\beta_A^i) < b, \end{cases} \tag{3.6}$$

if A is the superior product, and

$$\tilde{x}(\beta_B^i) = \begin{cases} \bar{x}(\beta_B^i) := \frac{a+b}{2} + \frac{\delta}{2\tau} & \text{if} \quad b \leq b_1(a) \Leftrightarrow \bar{x}(\beta_B^i) \in [b, a], \\ 1 & \text{if} \quad b > b_1(a) \Leftrightarrow \bar{x}(\beta_A^i) > a, \end{cases} \tag{3.7}$$

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.

### 3.5.2 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 3.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 3.1** (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 3.A. □

## 3.6 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 consumer's ability to observe her predecessor's action, new effects arise in the model and Proposition 3.1 will not generally hold. Instead, our main result, Proposition 3.2, 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 $b \leq a$.

### 3.6.1 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.

### 3.6.2 Uninformed Laggards: Updating and the Recommendation Effect

An uninformed laggard uses her information to update her belief $\beta^u(a, b, C_1) : \mathscr{A} \times \mathscr{B} \times \{A, B\} \to [0, 1]$. 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 $\beta^u_{C1}$. 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 (3.4) we have

$$\tilde{x}(\beta^u_{C_1}) = \begin{cases} \bar{x}(\beta^u_{C_1}) := \tilde{x}(\beta^u_{\emptyset}) + \frac{\delta}{\tau}\left(\beta^u_{C_1} - \beta^u_{\emptyset}\right) & \text{if} \quad \bar{x}(\beta^u_{C_1}) \in [a, b], \\ 0 & \text{if} \quad \bar{x}(\beta^u_{C_1}) < b, \\ 1 & \text{if} \quad \bar{x}(\beta^u_{C_1}) > a. \end{cases} \tag{3.8}$$

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

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

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 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

$$\begin{aligned}\beta_B^u = Pr(v_B > v | C_1 = B) &= \frac{Pr(C_1 = B | v_B > v) \cdot Pr(v_B > v)}{Pr(C_1 = B)} \\ &= \frac{q\tilde{x}(\beta_B^i) + (1-q)\tilde{x}(\beta_\emptyset^u)}{q(\tilde{x}(\beta_A^i) + \tilde{x}(\beta_B^i)) + (1-q)2\tilde{x}(\beta_\emptyset^u)}.\end{aligned} \tag{3.9}$$

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

$$\begin{aligned}\beta_A^u = Pr(v_B > v | C_1 = A) &= \frac{Pr(C_1 = A | v_B > v) \cdot Pr(v_B > v)}{Pr(C_1 = A)} \\ &= \frac{q[1 - \tilde{x}(\beta_B^i)] + (1-q)[1 - \tilde{x}(\beta_\emptyset^u)]}{q\{[1 - \tilde{x}(\beta_A^i)] + [1 - \tilde{x}(\beta_B^i)]\} + (1-q) \cdot 2 \cdot [1 - \tilde{x}(\beta_\emptyset^u)]}.\end{aligned} \tag{3.10}$$

It was shown before, that $\tilde{x}(\beta_A^i) < \tilde{x}(\beta_B^i)$, so that we can see from equations (3.9) and (3.10) 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 (3.9) 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, $\beta_B^u$, 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, $\beta_A^u$. 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.

### 3.6.3 Firms' Expected Demand

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

$$\begin{aligned} D_L(a,b) = & q\Big[\tilde{x}(\beta_A^i) + \tilde{x}(\beta_B^i)\Big] \\ & + (1-q)\Big[\tilde{x}(\beta_\emptyset^u) + Pr(C_1 = A)\tilde{x}(\beta_A^u) + Pr(C_1 = B)\tilde{x}(\beta_B^u)\Big]. \end{aligned} \tag{3.11}$$

It was shown in the previous sections that the threshold for a given belief $\beta$, $\tilde{x}(\beta)$ 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 $\beta$ 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 (3.4) and (3.8) show that the necessary distance between firms' locations implying $\tilde{x}(\beta)$ to be interior is increasing with the belief, and by equation (3.3) 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 3.3 in Appendix 3.B we show that for each configuration of the values of the parameters $\delta, \tau$ and $q$ there is a unique partition $\mathfrak{D}_L$ of $\{(a,b) \in \mathscr{A} \times \mathscr{B} | b \leq a\}$ 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)** $\bar{x}(\beta_A^i), \bar{x}(\beta_B^i) \notin [b,a], \quad \bar{x}(\beta_B^u),\, \bar{x}(\beta_A^u) \notin [b,a].$

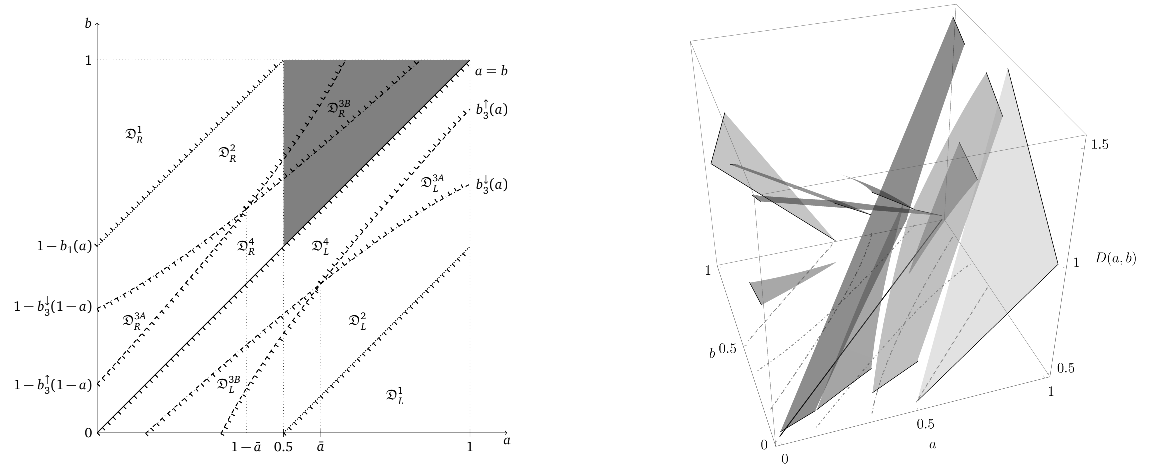
We use the notation $\mathfrak{D}_L$ to symbolize B being located left of A and $\mathfrak{D}^{3F}$ to imply that firm $F$ is the niche firm in this region.

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 $\{(a,b) \in \mathscr{A} \times \mathscr{B} | b > a\}$, 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 $\mathscr{A} \times \mathscr{B}$. 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 3.3 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 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 3.6 in Appendix 3.B. Note that given Assumption 1, the figure does not qualitatively depend on $\nu$ or $\delta$.




(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 3.3: 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 $b$ ($a$):

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

$$\Leftrightarrow \quad \frac{\delta q}{\tau} = (2-q)(a-b) - (1-q)\left(a^2 - b^2\right), \qquad (3.12)$$

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

$$\Leftrightarrow \quad \frac{\delta q}{\tau} = q(a-b) + (1-q)\left(a^2 - b^2\right). \qquad (3.13)$$

Since both equations are quadratic in $b$, they both have two solutions. As shown in Lemma 3.2 in Appendix 3.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 (3.12) and $b_3^{\uparrow}(a)$ for equation (3.13). 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 "↑".

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

$$\tilde{x}(\beta_A^u) = \begin{cases} \bar{x}(\beta_A^u) := \frac{a+b}{2} + \frac{\delta}{2\tau}(2\beta_A^u - 1) & \text{if} \quad b \leq b_3^{\downarrow}(a) \Leftrightarrow \bar{x}(\beta_A^u) \in [a,b], \\ 0 & \text{if} \quad b > b_3^{\downarrow}(a) \Leftrightarrow \bar{x}(\beta_A^u) < b, \end{cases} \tag{3.14}$$

if A is the superior product, and

$$\tilde{x}(\beta_B^u) = \begin{cases} \bar{x}(\beta_B^u) := \frac{a+b}{2} + \frac{\delta}{2\tau}(2\beta_B^u - 1) & \text{if} \quad b \leq b_3^{\uparrow}(a) \Leftrightarrow \bar{x}(\beta_B^u) \in [a,b], \\ 1 & \text{if} \quad b > b_3^{\uparrow}(a) \Leftrightarrow \bar{x}(\beta_B^u) > a. \end{cases} \tag{3.15}$$

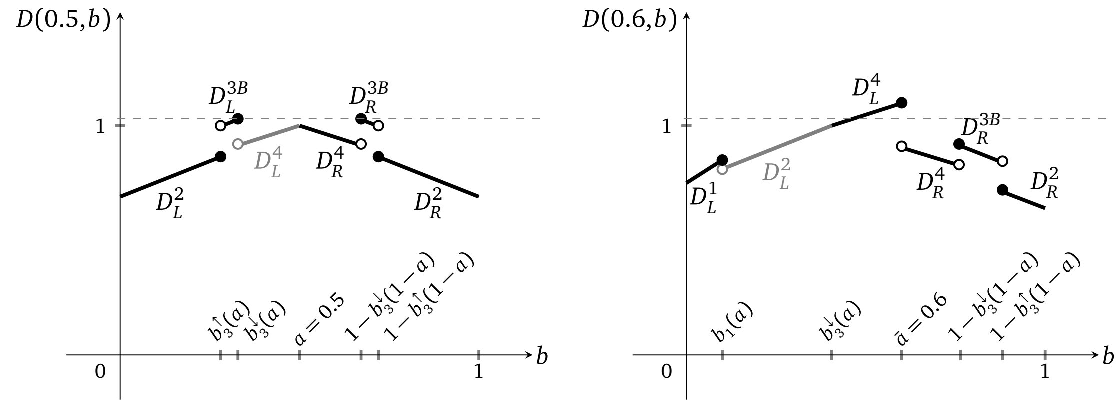
The derivatives of the discontinuity points can be calculated and it can be seen that $\partial b_3^{\uparrow}(a)/\partial a > 1 > \partial b_3^{\downarrow}(a)/\partial a > 0$ (see Lemma 3.3 in Appendix 3.B). The calculations in Appendix 3.B show that for small $a$ we have $b_3^{\uparrow}(a) < b_3^{\downarrow}(a)$ 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 3.6 in Appendix 3.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.

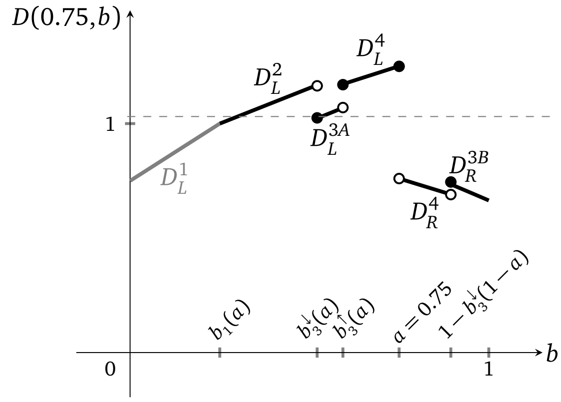
From equation (3.11) it is easy to see that the demand for $b \leq a$ is increasing in $a$ and $b$ as long as no threshold type changes from being interior to 0 or 1, which implies that the demand in each part $D_L^j(a, b)$ is increasing in $a$ and $b$ (see also Lemma 3.5 in Appendix 3.B). By symmetry, 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 3.4, 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 3.3 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 3.4: 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 3.2), which is visualized also by the dashed line. (Parameters: $q = 0.4, \tau = 2, \delta = 1$)

### 3.6.4 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 B prefers to differentiate from the center, which follows whenever

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

[22] As indicated above, we show in Lemma 3.7 in Appendix 3.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 3.3 in Appendix 3.C.

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 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

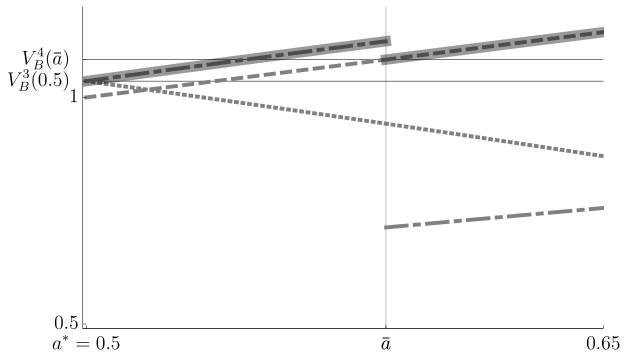
$$V_L^3(0.5) \leq V_L^4(\bar{a}), \tag{ADC}$$

which states that firm A prefers to induce differentiation from the center. If firm A does not prefer to induce differentiation from the center, i.e. $a^* > 0.5$, firm B might still want to differentiate on the short side:

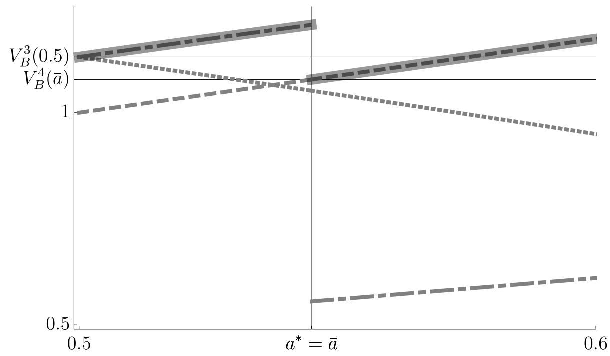
$$V_L^4(\bar{a}) \leq V_R^3(\bar{a}), \tag{BDS}$$

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

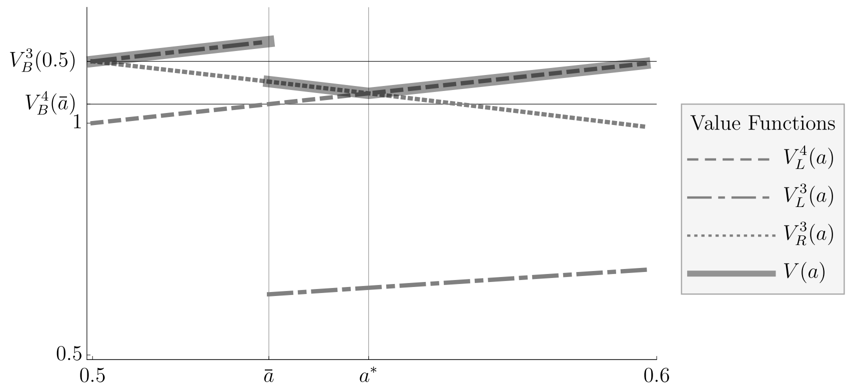
The three different situations are depicted in Figure 3.4. The complete parameter space is characterized in Figure 3.5 further below.




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




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




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

Figure 3.4: 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 3.2. Note that (BDS) implies that (ADC) is violated, which in turn implies (BDC).

**Proposition 3.2.** *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 3.C. □

Sufficient conditions for the second and fourth equilibrium can be calculated as follows (see Corollary 3.4 of Appendix 3.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.

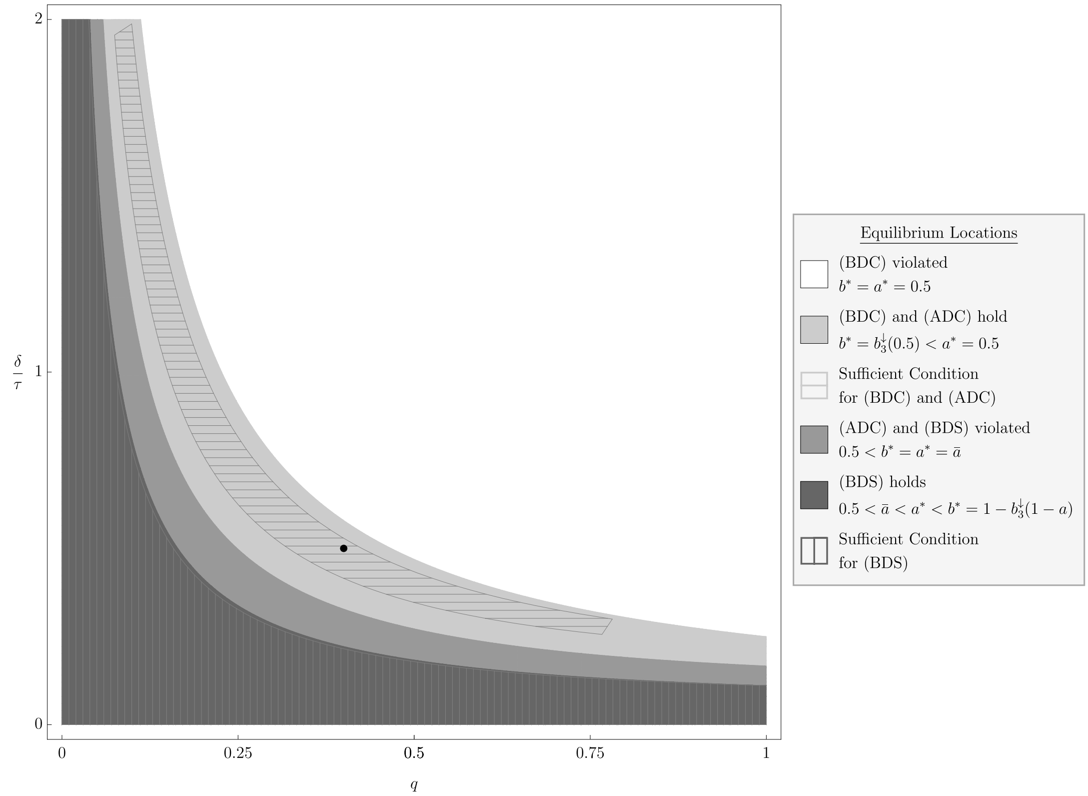



Figure 3.5: The figure depicts the necessary and sufficient parameter restrictions of the equilibria stated in Proposition 3.2. The black dot shows that the parameter combinations used in Figure 3.4 fulfills the conditions equations (BDC) and (ADC).

The above 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).

We now describe the intuition underlying the different equilibria. Figure 3.5 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]

### 3.6.5 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 3.6 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.

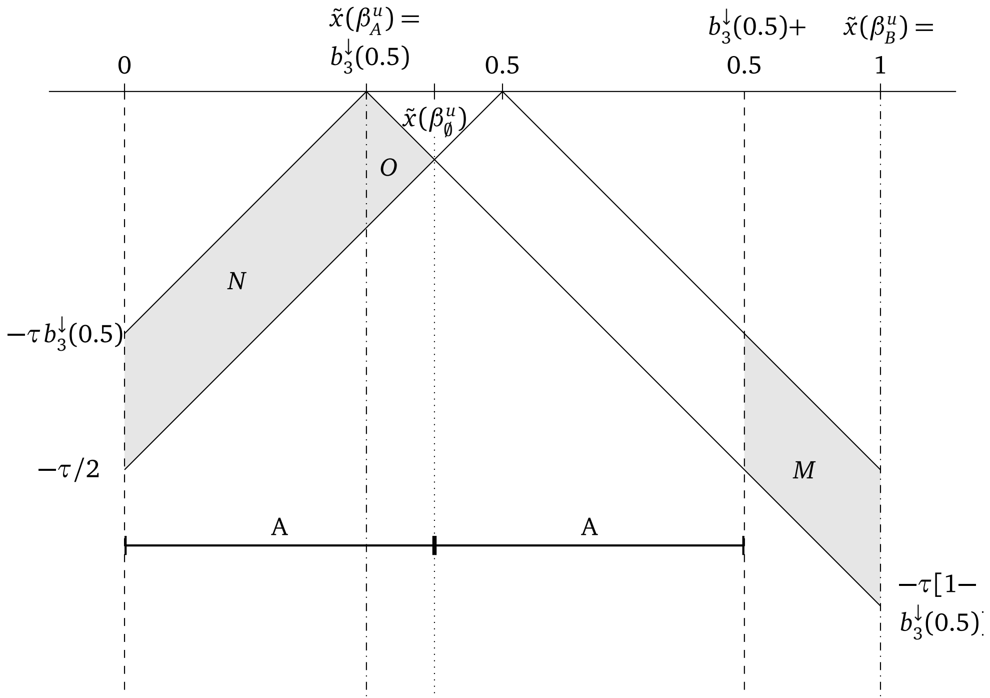



Figure 3.6: The figure depicts transport costs in the equilibrium of the benchmark model ($b^* = a^* = 0.5$) and the central differentiation result of the model with consumer learning ($b^* = b_3^{\downarrow}(0.5) < a^* = 0.5$). Indifferent types are depicted for the model with learning.

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 uninformed consumers 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 3.6 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, $\delta/\tau$. 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 3.7 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 $\Delta c_2^u = P(C_1 = A) \cdot \tilde{x}(\beta_A^u) \cdot N + P(C_1 = B) \cdot M$.

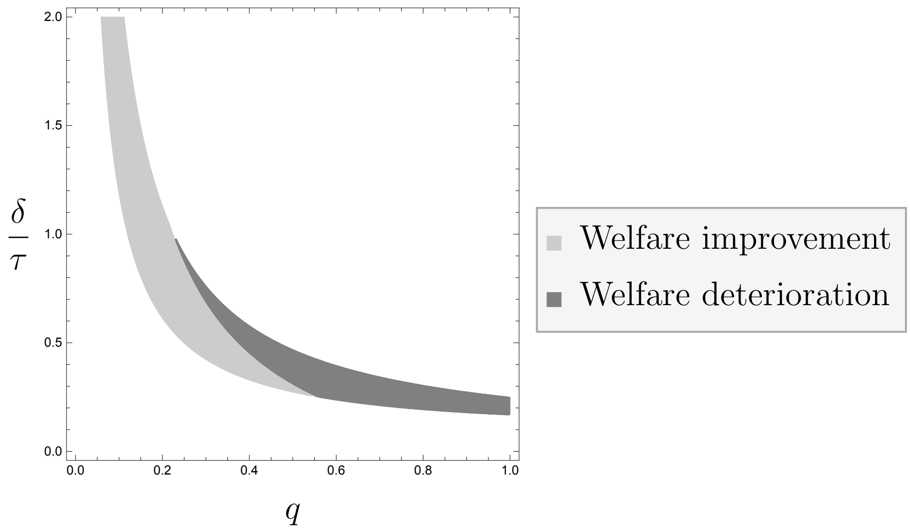



Figure 3.7: The shaded area depicts values of the parameters implying the existence of the central differentiation equilibrium ($b^* < a^* = 0.5$) in the model with consumer learning and under which this type of equilibrium is a welfare improvement / deterioration compared to the symmetric minimum differentiation 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 3.7 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 3.7 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 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.

[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.

## 3.7 Conclusion

This chapter 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 chapter 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 chapter: 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

[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).

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 3.4 and the next chapter).

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.

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

## Appendix 3.A Proof of Proposition 3.1 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 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

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

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

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

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 $a = b = 0.5$. □

## Appendix 3.B Generic Properties of Firm $B$'s Expected Demand

**Proposition 3.3.** *For each particular configuration of the values of the parameters $\delta, \tau$ 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 $\mathscr{A} \times \mathscr{B}$ with (possibly empty) elements $\mathfrak{D}_L^j$ and $\mathfrak{D}_R^j$:*

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

*The generic partition is depicted in Figure 3.3. 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 3.1.** *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 made. As a first period purchase from one firm increases the belief that it offers the superior product,

[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 $\mathfrak{D}_L^{3A} = \{(a,b) \in \mathscr{A} \times \mathscr{B} \mid 0 > b > b_3^{\downarrow}(a)\} \cap \{(a,b) \in \mathscr{A} \times \mathscr{B} \mid \neg b_3^{\downarrow}(a) < b_3^{\uparrow}(a)\}$

[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.

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 (3.8) 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 3.2.** *Equations* (3.5)*, 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* (3.12) *and* (3.13)*, 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 (3.12) and (3.13) 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 (3.12), 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 (3.13). A bit of calculation shows that both discontinuity points exist for $a$ such that

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

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 3.3.** $\frac{\partial b_3^{\uparrow}(a)}{\partial a} > 1 > \frac{\partial b_3^{\downarrow}(a)}{\partial a} > 0$ *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 (3.12) and (3.13) 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

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

so that $b_3^{\downarrow}(a)$ is indeed concave.
The rest of the lemma is obtained by simply equalizing equations (3.12) and (3.13), 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 3.4.** *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 3.3 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 3.1.** $b_1(a) < b_3^{\uparrow}(a) < a$ *and* $b_1(a) < b_3^{\downarrow}(a) < a$, $\forall (a, b) \in \mathscr{A} \times \mathscr{B}$ *with* $b \leq a$.

*Proof.* This follows from Lemma 3.3 and 3.4 . □

**Corollary 3.2.** *There is an equivalence between the five cases of Proposition 3.3 and the partition described therein.*

*Proof.* The result follows immediately from equations (3.6), (3.7), (3.14) and (3.15) and the previous lemmata. □

**Lemma 3.5.** *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: $\mathfrak{D}^1_L(a,b) := \{(a,b) \in \mathscr{A} \times \mathscr{B} \mid \tilde{x}(\beta^u_B) \in (b,a),\ \tilde{x}(\beta^u_A) \in (b,a),\ \tilde{x}(\beta^i_A) + \tilde{x}(\beta^i_B) = a + b\}$ 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: $\mathfrak{D}^2_L(a,b) := \{(a,b) \in \mathscr{A} \times \mathscr{B} \mid \tilde{x}(\beta^u_B) \in (b,a),\ \tilde{x}(\beta^u_A) \in (b,a),\ \tilde{x}(\beta^i_A) + \tilde{x}(\beta^i_B) = 1\}$ 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:

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

  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: $\mathfrak{D}^{3A}_L(a,b) := \{(a,b) \in \mathscr{A} \times \mathscr{B} \mid \tilde{x}(\beta^u_B) \in (b,a),\ \tilde{x}(\beta^u_A) = 0,\ \tilde{x}(\beta^i_A) + \tilde{x}(\beta^i_B) = 1\}$ 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

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

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

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

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: $\mathfrak{D}_L^4(a,b) := \{(a,b) \in \mathscr{A} \times \mathscr{B} \mid \tilde{x}(\beta_B^u) = 1,\ \tilde{x}(\beta_A^u) = 0,\ \tilde{x}(\beta_A^i) + \tilde{x}(\beta_B^i) = 1\}$ $\tilde{x}(\beta_B^u) = 1$ 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 $\delta$ and $\tau$. 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 3.6.** *$b_3^{\downarrow}(a)$ 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 3.3).*

*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 $b_3^{\downarrow}(a)$. □

## Appendix 3.C Proof of Proposition 3.2 (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 3.7.** *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 3.5 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 3.3.** *The value function of firm B in the respective regions of the action space is given by*

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

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

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

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

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

**Lemma 3.8.** *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 3.9.** *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 3.5 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 3.3. 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 3.5. □

**Lemma 3.10.** *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 $a \in \mathscr{A} \setminus \bar{a}$ and the function $V_R^{3B}(a)$ is decreasing in $a$ for all $a \geq 0.5$.*

*Proof.* By Lemma 3.9, 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 $a \in \{\mathscr{A} \cup \bar{a} \mid a \geq b\}$. Since $\bar{a} > 0.5$ symmetry implies that $V_R^{3B}(a)$ is decreasing for all $b > a$. □

**Lemma 3.11.**

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

*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 3.4, $b_3^{\downarrow}(a) \in \left[a - \frac{\delta q}{\tau}, \frac{1}{2} - \frac{\delta q}{2\tau}\right]$, 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 3.12.** *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 3.2.** *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 $V_R^{3B}(\bar{a}) \le V_L^4(\bar{a})$, 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: $b^* = b_3^{\downarrow}(0.5) < a^* = 0.5$

  $V_L^{3B}(0.5) \le V_L^4(\bar{a})$ implies $D_R^{3B}(\bar{a}) < V_L^{3B}(0.5) \le V_L^4(\bar{a})$ and thus $V(a) = V_L(a)$. By Lemma 3.10, $V_L(a)$ is increasing in $a$ if $V_L^{3B}(0.5) \le V_L^4(\bar{a})$, 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 3.10 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 $b \le a$ maximizing demand or the $b > a$ maximizing demand. Furthermore, the optimal $b \le a$ 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 3.12 exists in the action space.

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

**Corollary 3.4.** *Sufficient conditions for the conditions of Proposition 3.2 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 3.4, sufficient conditions for the conditions of Proposition 3.2 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. □

# Chapter 4

# Consumer Learning and Incentives to Differentiate in Cournot and Bertrand Competition

We combine two extensions of the differentiated duopoly model of Dixit (1979), namely Caminal and Vives (1996) and Brander and Spencer (2015a,b) to analyze the effect of consumer learning on firms' incentives to differentiate their products in models of Cournot and Bertrand competition.

Products are of different quality, consumers buy sequentially and are imperfectly informed about the quality of the goods. Before simultaneously competing in quantities, firms simultaneously choose their investment into differentiation. Late consumers can observe earlier consumers' decisions and extract information about the quality of the goods. This possibility influences the firms' incentives to differentiate. If firms compete in quantities, they are more likely to invest in differentiation with consumer learning than without. This is in line with implications of the recommendation effect introduced in the previous chapter in a model of spatial differentiation. We also examine the case in which firms compete in prices. Here, the effect of consumer learning is reversed, so that differentiation is less likely with consumer learning. Consumer learning thus increases the competition in the Bertrand setting and weakens it in the Cournot model.

## 4.1 Introduction

Most of the literature dealing with firms' incentives to differentiate characterizes two different and opposing effects. The competition effect induces firms to differentiate their products from each other, since they then obtain local monopoly power and are able to charge higher prices. On the other hand, differentiating decreases the market share and the amount of goods that the firm in question is able to sell. This is the so-called market-size effect (see e.g. Belleflamme and Peitz, 2010, Chapter 5.2).

In the related research project of the previous chapter we use a spatial model of product differentiation à la Hotelling (1929) to establish a new effect that may incentivize firms to differentiate. The effect arises because of the possibility of consumer learning, and is called the recommendation effect. Its intuition is as follows. In the model, two firms *A* and *B* compete by choosing their locations on the unit interval, representing the choice of the goods' characteristics. Two consumers sequentially choose between the two goods that are of different quality. Consumers are heterogeneous with respect to their preference towards the goods and to their information about the quality differential. The late consumer (laggard) observes the purchase decision of her predecessor (early adopter), which may contain valuable information. Neither the information, nor the preference of the early adopter is observed by the laggard. The laggard then uses Bayes' Rule to update her belief on the good's qualities. In this setup, it is the case that a purchase of a niche (i.e. a differentiated) product in the first period is more likely based on its high quality than on a good match of consumer taste and product characteristic. A firm can influence and exploit consumer learning using its location choice (mainstream vs. niche) which yields incentives to offer a differentiated (niche) product.

In order to make the above mentioned model tractable, it abstracts from endogenous prices by assuming that they are regulated to the same value for both firms. Additionally, there are only two consumers and two states of the world, either good *A*'s quality exceeds the one of good *B* by a fixed amount, or vice versa. Although there are situations plausibly described with these assumptions, the goal of the research at hand is to show that similar effects as described above also arise in a model where these assumptions are relaxed.[1]

Underlying the model in the paper at hand is the standard model of differentiated duopoly introduced by Dixit (1979). Dixit's model was extended in various ways. In particular, Caminal and Vives (1996) formulate a model of Bertrand competition with two firms who compete for a continuum of consumers in each of two periods by setting prices. The goods of the two firms are horizontally and vertically differentiated, but firms can not control either dimension. Firms also do not know the quality of their good when setting their prices. Consumers receive signals about the goods' qualities, and the consumers in the second period observe past market shares but not past prices. The authors show that such a situation leads to lower prices in the first period - compared to a situation without consumer learning - as each firm has an incentive to decrease its price in order to obtain a higher market share. This is because a high market shares serves as a signal of high quality to consumers in the second period. The authors state that "[...] an increase in the degree of product substitutability [...] increases the effectiveness of the manipulation by firms" (Caminal and Vives, 1996, p. 228). As the model does not allow

[1] More explicitly, the model at hand differs from the model of the previous chapter in that it entails endogenous prices, a simultaneous choice of product differentiation, a continuously distributed quality differential, a continuous information structure and the fact that firms may reset one of their choice variables (quantities in Cournot, prices in Bertrand competition), which would be equivalent to allowing firms to relocate in the model of the previous chapter.

for endogenous levels of differentiation or substitutability, Caminal and Vives do not elaborate on this insight any further.

This is where the model of Brander and Spencer (2015a,b), also based on Dixit's model, comes in. In their papers, the authors analyze the competition between two firms in the common differentiated duopoly setup without vertical differentiation. The goal of their article is to compare the firms' incentives to differentiate in Cournot and Bertrand competition setups. To endogenize the levels of differentiation, the authors assume that firms can make costly investments in order to increase the differentiation between the products. It is shown that differentiation is more likely to occur in Bertrand than in Cournot setups.

We combine the approaches of Caminal and Vives (1996) and Brander and Spencer (2015a) to explore how the possibility of consumer learnings influences firms incentives to differentiate their products. Our model thus makes the following changes to the differentiated duopoly setup of Dixit (1979): First, we allow the firms' products to be of different, random and a priori unknown quality, which introduces information asymmetries. The model consists of three stages. Before firms compete in quantities, they decide on their investment into differentiation. This stage is followed by two stages in which firms set quantities and consumers buy the goods, based on imperfect signals about the goods qualities. In the last stage, consumers additionally observe past prices (but not market shares).

Additionally, we analyze an analogous Bertrand model with consumer learning and compare the implications of Bayesian updating among consumers between the two models of price and quantity competition. We start with analyzing Cournot competition as in this model the endogenous variable of interest, namely the product differentiation, appears more intuitively in the utility function of the consumers.

## 4.2 Model Setup: Quantity Competition

In two periods, $t = 1, 2$, two firms, $j = A, B$, compete by producing quantities $x_t^j$. In each period there is a continuum of consumers with mass one uniformly distributed on $[0, 1]$ and indexed by $i$. Consumers have access to distinct information. The utility of the consumers when consuming any real-valued amount $x = (x^A, x^B, x^0) \in \mathbb{R}^3$ is given by

$$U(x^A, x^B, x^0) = (\alpha + q)x^A + (\alpha - q)x^B - 0.5\left[(x^A)^2 + 2\gamma x^A x^B + (x^B)^2\right] + x^0$$

with $\alpha > 0$. The idea is that all consumers agree on the utility of the goods and thus have the same utility function, which can be derived from quasilinear preferences.[2] Quantity $x^0$ captures the consumption of a composite good containing all other goods different from $x^A$ and $x^B$, and its price is normalized to $p^0 = 1$. The feasible range of the measure of product differentiation or substitutability is $\gamma \in (0, 1]$. The higher $\gamma$, the higher is the substitutability between the products. Goods are perfect substitutes if $\gamma = 1$ and they would be independent if $\gamma = 0$.

The (relative) quality of the goods is measured by the random variable $q$, which is normally distributed with mean zero and variance $1/\tau_q$.[3] Variable $q$ is unknown to firms and consumers, but each consumer receives a signal $s_t^i = q + \epsilon_t^i$ about it, where $\epsilon$ again is an independent and normally distributed random variable with zero mean and variance $1/\tau_\epsilon$. Both variances, $1/\tau_q$ and $1/\tau_\epsilon$, are known to the players in the model. We assume $\int_0^1 s_t^i di \xrightarrow{a.s.} q,\ t \in \{1, 2\}$, analogously to a version of the Strong Law of Large Numbers.[4] The consumers in period two also observe past prices, but not sold quantities or previous signals. This is the case for instance at online platforms on which consumers can observe the history of past prices and current quantities, but not quantities sold in previous periods. Let $I_t^i$ denote all available information to consumer $i$ in period $t$.

---

[2] In addition to heterogeneous information levels, we could introduce heterogeneity in the utility by choosing an individual parameter $\alpha_i$ with an appropriate distribution such that the results continue to hold. One could interpret each representative agent with a certain information level as representing a group of consumers with that same information level.

[3] As will become clear later on, when dealing with the rational updating of the consumers, it is easier to work with precisions $\tau$ than with variances.

[4] This assumption needs to be made due to a related issue pointed out by Judd (1985).

Rational consumers maximize their expected utility subject to the budget constraint $m = x^0 + p^A \cdot x^A + p^B \cdot x^B$, which yields the individual demand[5]

$$x_t^{A,i} = \frac{\alpha}{1+\gamma} + \frac{s_t^i}{1-\gamma} - \frac{p^A}{1-\gamma^2} + \frac{\gamma p^B}{1-\gamma^2},$$
$$x_t^{B,i} = \frac{\alpha}{1+\gamma} - \frac{s_t^i}{1-\gamma} - \frac{p^B}{1-\gamma^2} + \frac{\gamma p^A}{1-\gamma^2},$$
$$x_t^{0,i} = m - p_A \cdot x^{A,i} - p_A \cdot x^{B,i}.$$

Aggregating the rational consumers' demands via $x_t^j = \int_0^1 x_t^{j,i} di$ and inverting yields the total inverse demand for the product of firm $j \in \{A, B\}$ generated by rational consumers. In addition to the rational consumers, there are also consumers who ignore prices and whose utility for both products is the same irrespectively of the realized quality, in each period. Those consumers purchase from both firms randomly (see also Caminal and Vives, 1996). Their impact on the inverse demand at firm $j$ in period $t$ is the given by the random variable $u_t^j$. These random variables are i.i.d. draws from $\mathcal{N}(0, 1/\tau_u)$.[6]

Let $\eta_t := \int_0^1 E(q|I_t^i) di$ denote the aggregate belief on quality of all rational consumers in period $t$. Combining the demand of rational consumers and random shoppers leads to the following aggregated inverse demand functions

$$\begin{aligned} p_t^A &= \alpha + \eta_t - x_t^A - \gamma x_t^B + u_t^A, \\ p_t^B &= \alpha - \eta_t - x_t^B - \gamma x_t^A + u_t^B. \end{aligned} \tag{4.1}$$

The product substitutability $\gamma$ is endogenous and chosen by the firms via their investment decision. An example is the investment of Coca Cola and Pepsi into advertisement, in order to emphasize the differences between the two products, although their taste is indistinguishable

[5] As in many economic models, the application of normally distributed random variables has economically implausible consequences: depending on the received signal, individual demand might become negative or tend to infinity for a fixed set of prices. Analogous consequences can be found for individual inverse demand for a fixed set of quantities in Bertrand competition. To rule out such cases one could alternatively assume that $q$ is distributed according to a truncated normal distribution on an interval $[-Y, Y]$ with $0 < Y < \alpha/(1-\gamma^2)$. One would then only need to incorporate the changed variance of $q$, the rest of the updating remains unchanged. Note, that a truncated prior implies that the posterior is truncated at the same values independent of the distribution of the signal. Another possibility to decrease the probability of such 'unwanted' events would be to increase the precision of the random variables. Our results hold when we approach the model without asymmetric information and consumer learning, i.e. when $\tau_\epsilon \to \infty$.

If $\gamma = 1$, the stated individual demands become infinitely large or small. Furthermore, in the analogous model of Bertrand competition, the consumers' first order conditions are not invertible, such that there are problems with the microfoundation of the aggregate model described in Section 4.4. However, when starting directly with the standard aggregated form of the oligopoly models as represented in equations (4.1) and (4.7), such issues are avoided.

[6] The fact that irrational consumers may have a negative impact on (inverse) demand might be interpreted as some part of the rational consumers refraining from a purchase of the good, although rational utility optimization implied the opposite.

for consumers (see Brander and Spencer, 2015a,b). At the beginning of the game, in period $t = 0$, firms can make monetary investments $k^A$ and $k^B$ in order to increase the differentiation between their products, according to the following functional form

$$\gamma = e^{-\lambda(k^A + k^B)}.$$

If both firms make zero investments, then $\gamma = 1$ and the goods become perfect substitutes. Larger investments of any firm decrease $\gamma$, which approaches zero if the investments approach infinity. Of course there are cases in which firms invest into making their products complementary to each other, which however is not modeled in our setup. As $\gamma \in (0, 1]$, the goods are between the extremes of independence or perfect substitutability.[7] Cases where the goods are or can become complements are left out here, as our focus is on situations like the Coke-Vs-Pepsi story mentioned above. The parameter $\lambda \in \mathbb{R}_+$ measures the technology or effectiveness of how the firms' investments translate into increased differentiation. The higher $\lambda$, the smaller are the necessary investments to increase the differentiation by a certain amount.

Firms maximize their expected total profits and second period profits are discounted with the factor $\delta \in (0, 1)$. Note that firms do not know the quality (differential) when making their differentiation investments and choosing their quantities, which is plausible for instance in the case of experience goods. Thus, their information set in the first two periods is given by $I_0 = I_1 = \{\emptyset\}$, and by $I_2 = \{x_1, p_1\}$ in the last period.

As it is common for Cournot models, the process of price formation can be modeled via an auctioneer. She knows quality $q$, which allows her to calculate the consumers' aggregate belief on quality $\eta_t$ (see below), and she also knows the realization of $u_t^A$ and $u_t^B$ for $t = 1, 2$. In each period, the firms inform the auctioneer about the quantity produced and the auctioneer calculates $p_t^A$ and $p_t^B$. These prices are announced to consumers and each of them purchases his optimal quantity of each of the two goods. As announced prices contain information on the quality, we assume that consumers' beliefs are not affected by announced prices. This assumption can be justified by the fact that consumers do not understand the (informationally complex) process of price formation implemented by the auctioneer.

The alternative justification of Cournot models pioneered by Kreps and Scheinkman (1983), a model with a stage where both firms choose capacities before competing in prices, and which under certain assumptions leads to 'standard' Cournot outcomes, can not be applied here. This is the case as introducing demand uncertainty in a capacity-then-price-competition model leads to non-existence of a pure strategy equilibrium, in particular also to the absence of the equilibrium with Cournot quantities (see for example Hviid, 1991 and Behrens and Lijesen, 2012). As the case of homogeneous goods is nested in our model, the same applies here.

[7] Our model is a special case of a specification using the utility function $U(x^A, x^B, x^0) = (\alpha + q)x^A + (\alpha - q)x^B - 0.5\left[\beta(x^A)^2 + 2\gamma x^A x^B + \beta(x^B)^2\right] + x^0$ with $\beta = 1$. In the more general model goods may also be perfect complements, i.e. $\gamma = -1$.

For convenience we let firm specific variables without superscript denote the vector of the two variables of both firms and by $\Delta y$ we describe the difference of variable $y^A$ and $y^B$, so for example $x_1 = (x_1^A, x_1^B)$ and $\Delta u_1 = u_1^A - u_1^B$. The following graphic depicts the timing of the game:

$t = 0$
- Quality $q$ realizes.
- Firms make their differentiation investments $k^A$ and $k^B$ simultaneously knowing $I_0 = \emptyset$.

$t = 1$
- Firms set quantities $x_1^A, x_1^B$ knowing $I_1 = \emptyset$ .
- Consumer $i$ receives signal $s_1^i$

(• Auctioneer announces prices $p_1^A$ and $p_1^B$ knowing $\eta_1$ and $u_1$)
- Consumer $i$ purchases optimal quantity of goods knowing $I_1^i = \{s_1^i, x_1\}$.

$t = 2$
- Firms set quantities $x_2^A, x_2^B$ knowing $I_2 = \{p_1, x_1\}$ .
- Consumer $i$ receives signal $s_2^i$

(• Auctioneer announces prices $p_2^A$ and $p_2^B$ knowing $\eta_2$ and $u_2$)
- Consumer $i$ purchases optimal quantity of goods knowing $I_2^i = \{s_2^i, \Delta p_1, x_2\}$.

Figure 4.1: Timing of the Game

We employ the solution concept of Perfect Bayesian Equilibrium. To avoid complications off the equilibrium path, it is assumed that consumers' beliefs are constant with respect to observed current-period quantities, i.e. $\partial \eta_t^i / \partial x_t = 0$.

## 4.3 Solving the Model with Quantity Competition

We first replicate and adapt the results for Bertrand competition of Caminal and Vives (1996) for our Cournot framework. Then we analyze the effect of consumer learning on product differentiation in this framework.

### 4.3.1 Consumers

In order to characterize the optimal behavior of consumers, we only need to calculate how they use the available information to update their beliefs about the goods' qualities and can then use the aggregated inverse demand calculated in equation (4.1). We exploit the properties of the normal distribution, in particular the fact that the updating rules for both, mean and variance, are linear (see e.g. Section 2.2.2 of Chamley, 2004). Details on the calculations can be found in Appendix 4.A.

**First Period** Early adopter $i$'s belief on the quality is given by $\eta_1^i := E[q|I_1^i] = m_1 s_1^i$ where $m_1 := \frac{\tau_\epsilon}{\tau_\epsilon+\tau_q}$ weighs up the precisions of the distributions of the signal the consumers receive against the precision of the quality. The aggregate expectation in the first period can be calculated as

$$\eta_1 = \int_0^1 \eta_1^i di = m_1 q.$$

As $\mathrm{Var}(\eta_1) < \mathrm{Var}(q)$, some uncertainty is resolved in the aggregate and in aggregate the consumers' belief is closer to the true value of $q$ than the unconditional expectation $E[q]$. In equilibrium beliefs are correct, such that the consumers' first period equilibrium belief is $\eta_1^* = m_1 q$.

The utility maximization of the rational consumers and the behavior of the random shoppers results in the following inverse demand

$$p_1^A = \alpha + m_1 q - x_1^A - \gamma x_1^B + u_1^A,$$
$$p_1^B = \alpha - m_1 q - x_1^B - \gamma x_1^A + u_1^B.$$

**Second Period** In addition to the signal and the quantities $x_2$, the laggard $i$'s information set now also contains the observed price difference in period 1, i.e. her information set is $I_2^i = \{s_2^i, \Delta p_1, x_2\}$. The price differential $\Delta p_1$ contains information about $q$:

$$\Delta p_1 := p_1^A - p_1^B = 2m_1 q - (1-\gamma)\Delta x_1 + \Delta u$$
$$\Leftrightarrow q = [\Delta p_1 + (1-\gamma)\Delta x_1 - \Delta u_1]/2m_1.$$

As $\Delta x_1$ is not observed by consumers in the second period and $E[\Delta u_1] = 0$, a laggard's best estimate of the quality is $q^e = [\Delta p_1 + (1-\gamma)\Delta x_1^e]/2m_1$, where $\Delta x_1^e$ is the conjectured difference in quantities. Because actual first period quantities are not observed by consumers in the second period, they have to make conjectures about them, that is, they need to interpret past prices as signals of the chosen quantities, which is formalized by $\Delta x_1^e$. Thus, $q^e$ is obtained by solving the observed price difference $\Delta p_1$ for $q$ and replacing the unknown variables from the perspective of the consumer by their expected value ($\Delta u_1$) and the conjecture about the played strategy ($\Delta x_1$). Inserting the realized price difference $\Delta p_1$, it equals

$$q^e = q + [\Delta u_1 - (1-\gamma)(\Delta x_1 - \Delta x_1^e)]/2m_1. \tag{4.2}$$

This expression contains the two random variables $q$ and $\Delta u_1$, and the second summand captures the error the consumers make in conjecturing past market shares.

The laggard now combines his observation extracted from the price difference with his signal using Bayesian updating, so that her belief is given by $\eta_2^i := E[q|I_2^i] = m_2 s_2^i + n_2 q^e$, with $m_2 = \tau_\epsilon/\tau_2^i$, $n_2 = 2\tau_u m_1^2/\tau_2^i$ and $\tau_2^i = \tau_\epsilon + \tau_q + 2\tau_u m_1^2$. The aggregate belief then is given by

$$\eta_2 = \int_0^1 \eta_2^i di = m_2 q + n_2 q^e.$$

As in equilibrium beliefs are correct we obtain the equilibrium belief $\eta_2^* = m_2 q + n_2 \tilde{q}$ with $\tilde{q} = q + \Delta u_1/2m_1$. Note that $\tilde{q}$ equals $q^e$ with correctly conjectured first period quantities, i.e. with $\Delta x_1^e = \Delta x_1$.

Clearly,

$$\frac{\partial \eta_2}{\partial x_1^A} = -\frac{\partial \eta_2}{\partial x_1^B} = n_2 \cdot \frac{\partial q^e}{\partial x_1^A} = -\frac{(1-\gamma)n_2}{2m_1}. \tag{4.3}$$

The derivative shows that the effect of a change in the first period quantity on the consumer belief in the second period is higher, the smaller the substitutability between products, $\gamma$. Phrased differently, the more differentiated the goods are, the higher is the impact of a firm's change of its choice variable in the first period on the laggards' belief. The decreased price $p_1^A$ induced by a higher quantity $x_1^A$ decreases the belief that product A is of superior quality because first period quantities are not observed by laggards, so that consumers can not be certain whether the price decrease was due to a low quality product, or due to a high volume of sales. This reasoning is analogous to the recommendation effect as introduced in the previous chapter.

Similarly as in period one, the utility maximization of the rational consumers and the behavior of the random shoppers results in the following inverse demand in period two:

$$p_2^A = \alpha + m_2 q + n_2 q^e - x_2^A - \gamma x_2^B + u_2^A,$$
$$p_2^B = \alpha - m_2 q + n_2 q^e - x_2^B - \gamma x_2^A + u_2^B.$$

### 4.3.2 Firms

Firm behavior is analyzed via backward induction, but we start with analyzing the firms' information processing. Details on the calculations can be found in Appendix 4.B.

### Bayesian Updating

In order to optimally set quantities, firms need to forecast the consumers' beliefs on quality, so they form a belief about the consumers' (aggregate) belief on quality denoted by $\theta_t := E[\eta_t|I_t]$. Both firms have identical information and so cannot manipulate each other.

In period 1, the firms do not have any information about the consumers' belief and thus $\theta_1 = E[m_1 q|I_1] = 0$.

In period 2, in contrast to the consumers, the firms can extract $\tilde{q}$ (the consumers' second period estimate of the quality extracted from the price difference in the previous period with correctly conjectured quantities) from past prices and quantities, so that $\theta_2 = E[m_2 q + n_2 q^e|I_2] = m_2 E[q|\tilde{q}] + n_2 E[q^e|\tilde{q}]$. Equilibrium beliefs are defined to be $\theta_t^* = E[\eta_t^*|I_t]$.

### Optimal Quantities

Given the beliefs about consumers' beliefs, firms choose optimal quantities. In the second period, firms take the differentiation parameter and first period quantities and prices as given so that their optimization problem boils down to maximizing the profit $\pi_2^j = x_2^j \cdot p_2^j(x_2)$ by the choice of $x_2^j$. Best responses are given by $x_2^A(x_2^B) = \frac{\alpha+\theta_2-\gamma x_2^B}{2}$ and $x_2^B(x_2^A) = \frac{\alpha-\theta_2-\gamma x_2^A}{2}$. Equilibrium quantities are

$$x_2^{A*} = \frac{\alpha}{2+\gamma} + \frac{\theta_2^*}{2-\gamma},$$
$$x_2^{B*} = \frac{\alpha}{2+\gamma} - \frac{\theta_2^*}{2-\gamma}.$$

In the first period firms take into account the indirect effect their quantity choice has on the profit in period 2 via Bayesian updating among the consumers. Thus, the objective function of firm A is given by

$$\pi_1^A(x_1) = x_1^A(\alpha + \theta_1 - x_1^A - \gamma x_1^B) + \delta E[\pi_2^A|I_1], \tag{4.4}$$

where $\pi_2^A$ is firm A's second period profit and $\pi_1^A$ is the total revenue of firm A, that is the profit from periods one and two, ignoring potential investments in differentiation. Remember that $\theta_1 = 0$. Furthermore, note that $E[p_2^A|I_1] = \frac{\alpha}{2+\gamma} + \frac{\theta_2}{2-\gamma} = x_2^{A*}$, which implies that $E[\pi_2^A|I_1] = E[(x_2^{A*})^2|I_1]$. Firm A's best response is then $x_1^A(x_1^B) = \alpha/2 - x_1^B\gamma/2 + \frac{\delta x_2^{A*}}{2-\gamma} \cdot \frac{\partial\theta_2}{\partial x_1^A}$. Firm B's best response can be calculated analogously and, using equation (4.3), equilibrium quantities are

$$x_1^{A*} = x_1^{B*} = \frac{\alpha}{2+\gamma} \cdot \left(1 + \frac{2\delta}{4-\gamma^2}\frac{\partial\theta_2}{\partial x_1^A}\right).$$

Overall, we obtain the following result, which is analogous to the proposition for Bertrand competition in Caminal and Vives (1996).

**Lemma 4.1.** *In the equilibrium of our model, optimal quantities in period 2 are given by*

$$x_2^{A*} = \frac{\alpha}{2+\gamma} + \frac{\theta_2^*}{2-\gamma} \quad \textit{and} \quad x_2^{B*} = \frac{\alpha}{2+\gamma} - \frac{\theta_2^*}{2-\gamma}. \tag{4.5}$$

*Optimal quantities in period 1 are given by*

$$x_1^{A*} = x_1^{B*} = \frac{\alpha}{2+\gamma} \cdot \left(1 - \frac{\delta(1-\gamma)}{4-\gamma^2} \cdot \frac{n_2}{m_1}\right). \tag{4.6}$$

The optimal second-period quantity of firm $j$ is higher (lower) than in a standard differentiated Cournot model ($\alpha/(2+\gamma)$), if the expectation of the consumer belief is (not) in favor of firm $j$.[8] That is, the firm which is expected to be preferred by consumers sells a higher quantity.

As $\gamma \leq 1$, first period quantities are (weakly) lower than those without consumer learning, meaning that first period prices exceed those of standard differentiated Cournot model. This is due to consumers in period 2 only observing past prices but not quantities. A higher price thus leads them to expect the good to be of higher quality.

### Optimal Differentiation Investments

Forecasting the resulting optimal quantities, firms choose the investment into differentiation in period zero. There exist no closed-form solutions to derive the optimal investment in differentiation, $k^{j*}$, and furthermore, conventional comparative static tools such as the implicit function theorem or approaches via lattice theory involve calculations, which are too computationally complex. Thus, to compare the differentiation incentives without relying on the full solution, we use the technique of Brander and Spencer (2015a), who compare the minimal effectiveness of investments in differentiation needed to induce firms to invest, that is, we derive and compare the thresholds $\lambda$ so that firms investments become positive.

[8] Equilibrium quantities are positive whenever $n_2/m_1 < 4$, which is always fulfilled.

**Without Consumer Learning (Benchmark)** If second period consumers were to have a belief of $\eta_2 = 0$, the model in the second period is the same as the standard model of Dixit (1979), and the resulting optimal quantities would be $x_{NL}^{A*} = x_{NL}^{B*} = \alpha/(2+\gamma)$. The profit of a benchmark model with two periods without consumer learning is thus given by

$$\pi_{NL}^j = (1+\delta)\cdot E[\pi_2^j|I_1] - k^j = (1+\delta)\cdot(x_{NL}^{j*})^2 - k^j.$$

The derivative of the objective function is then

$$\begin{aligned}\partial\pi_{NL}^j/\partial k^j &= (1+\delta)\cdot\frac{d(x_{NL}^{j*})^2}{d\gamma}\cdot\frac{\partial\gamma}{\partial k^j} - 1 = (1+\delta)\cdot\frac{-2\alpha^2}{(2+\gamma)^3}\cdot(-\lambda\gamma) - 1\\ &= (1+\delta)\cdot\frac{2\lambda\gamma\alpha^2}{(2+\gamma)^3} - 1.\end{aligned}$$

Firm j will invest in differentiation in equilibrium if

$$\partial\pi_{NL}^j/\partial k^j\big|_{\gamma=1} > 0 \Leftrightarrow \lambda > \frac{27}{2(1+\delta)\alpha^2} := \bar{\lambda}_{NL}^C.$$

The threshold without learning can also be obtained as a corollary of Proposition 4 from Brander and Spencer (2015a) by extending their model to two periods. The threshold decreases in $\alpha$, as the positive effect of increased differentiation on profit is higher the higher $\alpha$, so that the necessary technology ($\bar{\lambda}$) is decreasing in $\alpha$. Additionally, differentiation incentives are stronger, as $\delta$ increases. This is because the gain from differentiation is higher than the costs compared to a situation with a lower $\delta$.

**With Consumer Learning** With $\pi_L^j(\cdot) := \pi_1^j(\cdot) - k_j$, $j \in \{A, B\}$, and using equation (4.4) and the results mentioned thereafter, the derivative of the objective function is given by

$$\partial\pi_L^j(x^*)/\partial k^j = \frac{\partial\left[x_1^{j*}\{\alpha - (1+\gamma)x_1^{j*}\}\right]}{\partial k^j} + \delta\cdot\frac{d(x_2^{j*})^2}{d\gamma}\cdot\frac{\partial\gamma}{\partial k^j} - 1.$$

Firm j will invest in differentiation in equilibrium if

$$\begin{aligned}\partial\pi_L^j/\partial k^j\big|_{\gamma=1} &= \frac{\alpha^2(n_2\delta + 2m_1(1+\delta))\lambda}{27m_1} - 1 > 0\\ \Leftrightarrow \qquad \lambda &> \frac{27m_1}{\alpha^2(2m_1 + 2\delta m_1 + \delta n_2)} := \bar{\lambda}_L^C.\end{aligned}$$

We can easily see that $\bar{\lambda}_L^C < \bar{\lambda}_{NL}^C$, as

$$\bar{\lambda}_L^C = \bar{\lambda}_{NL}^C \frac{2\alpha^2(1+\delta)}{2\alpha^2(1+\delta)+\alpha^2\delta n_2/m_1} < \bar{\lambda}_{NL}^C,$$

which leads to our first main result:

**Proposition 4.1.** *In the Cournot model, in equilibrium firms offer perfect substitutes for a smaller range of parameters $\lambda$ with consumer learning than without. That is, the threshold $\lambda$ above which firms invest in differentiation is lower with consumer learning than without, $\bar{\lambda}_L^C < \bar{\lambda}_{NL}^C$.*

The comparative statics of threshold $\bar{\lambda}_L$ with respect to $\alpha$ and $\delta$ are the same as those of threshold $\bar{\lambda}_{NL}^C$, but the extent of the changes on the threshold induced by changes in $\alpha$ and $\delta$ now depends on the parameters introduced by consumer learning, $m_1$ and $n_2$. Furthermore, as $\frac{\partial n_2}{\partial \tau_u} > 0$ and $\frac{\partial \bar{\lambda}_L}{\partial n_2} < 0$, the critical value $\bar{\lambda}_L$ decreases in $\tau_u$, $\frac{\partial \bar{\lambda}_L}{\partial \tau_u} < 0$. Intuitively, the more noise caused by the random shoppers is contained in the observed statistic about the quality, the smaller are the incentives to differentiate.

## 4.4 Solving the Model with Price Competition

As shown by Singh and Vives (1984), with a linear quadratic utility function, there is a close relationship between Cournot and Bertrand competition. In their words:

> Cournot (Bertrand) competition with substitutes is the dual of Bertrand (Cournot) competition with complements. This means that they share similar strategic properties. For example, with linear demand, reaction functions slope downwards (upwards) in both cases. It is a matter of interchanging prices and quantities. (Singh and Vives, 1984, p. 547)

Indeed, using the following utility function from Caminal and Vives (1996) which slightly differs from the one of the previous sections,

$$U^B(x^A, x^B, x^0) = (\alpha+(1-\gamma)q)x^A + (\alpha-(1-\gamma)q)x^B - 0.5\left[(x^A)^2 + 2\gamma x^A x^B + (x^B)^2\right] + x^0,$$

and including random shoppers, similar to the model before, we obtain the following set of demand functions

$$\begin{aligned} x_t^A &= a + \eta_t - b p_t^A + c p_t^B + u_t^A, \\ x_t^B &= a - \eta_t - b p_t^B + c p_t^A + u_t^B, \end{aligned} \tag{4.7}$$

where $a = \alpha/(1+\gamma)$, $b = 1/(1-\gamma^2)$ and $c = \gamma/(1-\gamma^2)$. Variable $\eta_t$ again is the aggregate belief about the quality of consumers in period $t$. Besides the slightly different utility function and induced demands, all variables remain as in the previous section.

Comparing the above system of direct demands in (4.7) to the inverse demand system from equation (4.1), we can obtain one from the other by simply exchanging quantities and prices and replacing $a$ by $\alpha$, $b$ by $\beta = 1$ and $c$ by $-\gamma$.

Using equation (4.3), this implies that the recommendation effect in the Bertrand model is formalized by

$$\frac{\partial \eta_2}{\partial p_1^A} = -\frac{\partial \eta_2}{\partial p_1^B} = n_2 \cdot \frac{\partial q^e}{\partial p_1^A} = -\frac{n_2}{2(1-\gamma)m_1}. \tag{4.8}$$

This shows, that the parameter of substitution ($\gamma$), has the inverse impact on the magnitude of the recommendation effect in Bertrand competition than in Cournot competition.

Additionally, using the above shortcut, we know from Lemma 4.1 that the equilibrium prices in this setting are given by:

**Lemma 4.2** (Caminal and Vives, 1996). *In the equilibrium of the model with price setting, optimal prices in period* 2 *are given by*

$$p_2^{A*} = \frac{a}{2b-c} + \frac{\theta_2^*}{2b+c} \quad \text{and} \quad p_2^{B*} = \frac{a}{2b-c} - \frac{\theta_2^*}{2b+c}. \tag{4.9}$$

*Optimal prices in period 1 are given by*

$$p_1^{A*} = p_1^{B*} = \frac{a}{2b-c} \cdot \left(1 - \frac{(b+c)b\delta}{4b^2-c^2} \cdot \frac{n_2}{m_1}\right). \tag{4.10}$$

Without learning, optimal prices of both firms are calculated as $p_{NL}^{A*} = p_{NL}^{B*} = \frac{a}{2b-c}$. We see that the firm with the higher perceived quality charges a higher price and in the first period both firms charge a lower price than in a model without learning.

As in the previous section, we can use the equilibrium prices to calculate equilibrium profits for a fixed $\gamma$ and solve the derivative of the profit with respect to the investment $k_j$ evaluated at $\gamma = 1$ for the $\lambda$ above which firms make their investments in differentiation. Details on the calculations can be found in Appendix 4.C. We obtain the following result:

**Proposition 4.2.** *In the Bertrand model, in equilibrium firms offer perfect substitutes for a smaller range of parameters $\lambda$ without consumer learning than with consumer learning. That is, the threshold $\lambda$ above which firms invest in differentiation with consumer learning ($\bar{\lambda}_L^B$) is higher than the threshold without learning ($\bar{\lambda}_{NL}^B$):*

$$\bar{\lambda}_L^B = \frac{2}{\alpha^2\big[(1+\delta) - 2\delta n_2/(3m_1)\big]} > \frac{2}{\alpha^2(1+\delta)} = \bar{\lambda}_{NL}^B.$$

## 4.5 Comparison of the Effect of Consumer Learning: Bertrand vs. Cournot

While we should keep in mind, that the parameter of substitution ($\gamma$) is incorporated in a different manner in the microfoundation of the Bertrand and the Cournot model,[9] it is nevertheless worthwhile to compare the influence of consumer learning on the incentives to differentiate of the two models. Comparing our findings in the different models yields our final main result:

**Proposition 4.3.** *The effect of consumer learning on the firms' incentives to differentiate their products is different in the Cournot model and in the Bertrand model. In contrast to quantity competition, consumer learning in a model with price setting decreases the firms' incentives to differentiate:*

$$\bar{\lambda}_L^B - \bar{\lambda}_{NL}^B > 0 > \bar{\lambda}_L^C - \bar{\lambda}_{NL}^C.$$

*Consumer learning thus tends to increase the competition in the Bertrand setting and it weakens it in the Cournot model.*

In order to understand this result in more detail, it is useful to compare the equilibrium choices from the models with learning to those without. From the perspective of period zero, where firms choose their differentiation investments, and given the equilibrium strategies for periods one and two, the expected optimal quantities in period 2 are the same in the models with and without learning as $E(\theta_2) = 0$. Thus, the second period affects the differentiation incentives only through its influence on the optimal first period choices of the firms. In the Cournot game, equilibrium quantities in the first period are given by

$$x_1^{j*} = x_{NL}^{j*} \cdot \left(1 - \frac{\delta(1-\gamma)}{4-\gamma^2} \cdot \frac{n_2}{m_1}\right) = \frac{\alpha}{2+\gamma} \cdot \left(1 - \frac{\delta(1-\gamma)}{4-\gamma^2} \cdot \frac{n_2}{m_1}\right)$$

and equilibrium prices in the Bertrand model are

$$\begin{aligned} p_1^{j*} &= p_{NL}^{j*} \cdot \left(1 - \frac{(b+c)b\delta}{4b^2-c^2} \cdot \frac{n_2}{m_1}\right) = \frac{a}{2b-c} \cdot \left(1 - \frac{(b+c)b\delta}{4b^2-c^2} \cdot \frac{n_2}{m_1}\right) \\ &= \frac{\alpha(1-\gamma)}{2-\gamma} \cdot \left(1 - \frac{(1+\gamma)\delta}{4-\gamma^2} \cdot \frac{n_2}{m_1}\right) \end{aligned}$$

for j $\in \{A, B\}$. In both models, the first factor, gives the optimal choices in a model without consumer learning. The second factor in both cases is (weakly) smaller than one, so that quantities in a Cournot style competition and prices in the Bertrand variant of our model are

[9]The different utility functions in the two models are employed, as they allow to compare the impact of $\gamma$ on the aggregated (inverse) demand systems in equations (4.1) and (4.7) more easily.

(weakly) decreased by the introduction of consumer learning. Only if $\gamma = 1$, the optimal choices in the models with learning and those without coincide.

Starting with the Cournot model and comparing the marginal profit of investing in differentiation (increasing $k^A$ or $k^B$) at a situation where $\gamma = 1$ ($k^A = k^B = 0$), the previous calculations showed that the marginal profit is higher with learning than without, leading to the lower threshold value in the model with learning compared to the model without.

The mechanics behind this difference are the following: as in the Cournot model without learning, and in any similar model, it is the case that the two firms could increase their first period profit by reducing their quantities. In the model without learning, decreasing one's quantity below $x_{NL}^{A*} = x_{NL}^{B*}$ is not individually rational. If $\gamma < 1$, consumer learning however introduces an incentive to decrease first period quantities below the level of a model without learning due to the recommendation effect, meaning that at $\gamma = 1$, consumer learning generates an additional incentive to invest in differentiation, as this enhances the impact of the recommendation effect.

The situation in the Bertrand setup is different in that prices are already too low in the model without learning if the goal is to maximize the joint first period profit of the firms. Firms could therefore increase their profits if they were to jointly raise their prices. With $\gamma = 1$ prices in the model with and without learning coincide and equal zero. Decreasing $\gamma$, that is increasing the differentiation, increases the optimal first period price, but the increase is smaller with consumer learning. The marginal profit of increasing $k^A = k^B = 0$, is thus higher in the model without learning than it is in the model with learning, explaining the ordering of the thresholds in this setup.

Finally, we can elaborate on the result of Brander and Spencer (2015a,b) who showed that firms are more likely to invest in differentiation in Bertrand than in Cournot competition. As our benchmark models without learning are two-period extensions of their models, we obtain the same result if we compare the models without learning, that is $\bar{\lambda}_{NL}^C > \bar{\lambda}_{NL}^B$. Consumer learning has been shown to decrease the threshold in the Cournot model and increase it in the Bertrand setting, but even then, the ranking of the two models is maintained, i.e. we also have $\bar{\lambda}_L^C > \bar{\lambda}_L^B$.

## 4.6 Conclusion

Differentiating one's product from those of a competitor results in a weaker competition and thus allows for higher prices and profits. By introducing consumer learning in a duopoly model with vertically differentiated goods and by endogenizing the horizontal differentiation between the products, we have shown that the incentives to differentiate are changed when consumer learning about the strength of the vertical differentiation is introduced. Furthermore, the effects created by consumer learning differ vastly between a model of quantity and a model of price competition.

In each of the two models, consumers learn from observed previous purchasing decisions. As their observations are not fully revealing all information, firms can manipulate the inference of late consumers by influencing the purchase decisions of early consumers. When setting their prices or quantities in early periods, firms take this effect of their choices on the inference of later consumers into account. Only when the two products are perfect substitutes, the presence of consumer learning does not change the firms optimal behavior compared to a model without consumer learning.

For quantity competition with differentiated products, firms optimally choose lower quantities in a model with consumer learning than in a model without. Low quantities lead to higher prices which tend to signal higher quality to later consumers. If firms compete in prices, optimal first period prices are below those of a model without consumer learning, as here higher sold quantities signal high quality to later consumers.

These 'distortions' of the optimal choices in early periods lead to different effects on the differentiation incentives of the firms induced by consumer learning. Because profits in a Cournot model can typically be increased by reducing the produced quantities, which is precisely the effect consumer learning has in our model of quantity competition, consumer learning increases the incentives to differentiate above the incentives generated by the desire to relax competition. The reverse is true in a model of Bertrand style price competition: here increasing prices would increase the profit of the firms, but consumer learning reduces the prices even further than the already strong competition in a Bertrand setup. The introduction of consumer learning thus decreases the incentives to invest in differentiation if firms compete in prices.

The presented results seem to support the notion that price competition, i.e. a game with strategic complements, leads to a stronger competition than quantity competition, that is competition with strategic substitutes, as products are more likely to be substitutes in the former oligopoly model.

## Appendix 4.A Bayesian Updating Among Consumers

Bayes' rule in our context can be formulated as

$$f(q|o) = \frac{\phi(o|q) \cdot f(q)}{\int \phi(o|q) \cdot f(q) dq},$$

where $f(\cdot)$ is the density of $q$ and $\phi(\cdot)$ is the density of some observation $o$ containing information on quality $q$, i.e. in our case signal $s_i^t$ or the estimate of $q$ extracted from the price difference in the first period, $q^e$. Gaussian models as the one at hand, i.e. Bayesian updating over normally distributed random variables and observations, are particularly tractable, as the posterior distribution is also normal and the updating rules for mean and variance are linear: the posterior mean is the weighted average of the prior mean and that of the observation weighted with the respective precisions, while the posterior variance is that of the prior increased by that of the observation.

In our model, consumers want to best estimate $q$ from their observations. Consumers have the prior knowledge that $q \sim N\left(\mu_q, \frac{1}{\tau_q}\right)$ and they make one or two additional observations $o_r$, $r \in \{1, 2\}$, with information about $q$. All consumers receive a signal about $q$ and consumers in period two observe past prices. Both, the signal and the information extracted from past prices can be reformulated to observation $o_{r,t}^i$ of consumer $i$ in period $t$ in the following form:

$$o_{r,t}^i = q + v_{r,t}^i \quad \text{where} \quad v_{r,t}^i \sim N\left(0, \frac{1}{\tau_{v_{r,t}^i}}\right)$$

Using Bayesian updating as described above, this leads to the following distribution of $q$ conditional on the available observations, for $t \in \{1, 2\}$

$$q|I_t^i \sim \left(\frac{\tau_q \mu_q + \sum_{r=1}^t \tau_{v_{r,t}^i} o_{r,t}^i}{\tau_q + \sum_{r=1}^t \tau_{v_{r,t}^i}}, \frac{1}{\tau_q + \sum_{r=1}^t \tau_{v_{r,t}^i}}\right)$$

Let $\eta_1^i := E[q|I_1^i]$ be the updated belief of consumer $i$ about $q$ in period 1 after receiving signal $s_1^i$, then

$$\eta_1^i \sim N\left((1 - m_1) \cdot 0 + m_1 \cdot s_1^i, \frac{1}{\tau_q + \tau_\varepsilon}\right).$$

with $m_1 := \frac{\tau_\varepsilon}{\tau_q + \tau_\varepsilon}$. Using the assumption on the average signal, the aggregate belief is given by

$$\eta_1 := \int_0^1 \eta_1^i di = \int_0^1 m_1 s_1^i di = m_1 \int_0^1 s_1^i di \to m_1 q.$$

The information a consumer $i$ in period 2 can extract about $q$ from the observed price difference only is given by

$$\begin{aligned} q^e &= [\Delta p_1 + (1-\gamma)\Delta x_1^e]/2m_1 \\ &= q + [\Delta u_1 - (1-\gamma)(\Delta x_1 - \Delta x_1^e)]/2m_1. \end{aligned} \tag{4.11}$$

This expression contains the two random variables $q \sim N(0, \frac{1}{\tau_q})$ and $\frac{\Delta u_1}{2m_1} \sim N(0, \frac{2}{4m_1^2\tau_u})$.

When combining this with the signal, $\eta_2^i := E[q|I_2^i]$, the updated belief of consumer $i$ about $q$ in period 2, is normally distributed with

$$\eta_2^i \sim N\left((1-m_2-n_2)\cdot 0 + m_2 s_2^i + n_2 q^e, \frac{1}{\tau_2^i}\right),$$

with $\tau_2^i = \tau_\epsilon + \tau_q + 2\tau_u m_1^2$, $m_2 = \tau_\epsilon/\tau_2^i$ and $n_2 = 2\tau_u m_1^2/\tau_2^i$ and thus the aggregate belief is given by

$$\eta_2 := \int_0^1 \eta_2^i di = \int_0^1 (m_2 s_2^i + n_2 q^e) di = m_2 \int_0^1 s_2^i di + n_2 q^e \to m_2 q + n_2 q^e,$$

again making use of the assumption on the average signal.

## Appendix 4.B Firm Behavior in the Cournot Model

Firm behavior is analyzed via backward induction.

### Quantity Setting in Stage $t = 2$

Firm $A$'s profit in stage $t = 2$ is given by

$$\pi_2^A = x_2^A \cdot p_2^A = x_2^A \cdot (\alpha + \theta_2 - x_2^A - \gamma x_2^B).$$

Best responses are obtained by the FOCs $\partial \pi_2^j / \partial x_2^j = 0$ with $j \in \{A, B\}$, which gives

$$x_2^A(x_2^B) = \frac{\alpha + \theta_2 - \gamma x_2^B}{2}, \qquad \text{and similarly} \qquad x_2^B(x_2^A) = \frac{\alpha - \theta_2 - \gamma x_2^B}{2}.$$

In equilibrium best responses intersect, so we obtain the equilibrium quantities

$$x_2^{A*} = \frac{\alpha}{2+\gamma} + \frac{\theta_2}{2-\gamma} \qquad \text{and} \qquad x_2^{B*} = \frac{\alpha}{2+\gamma} - \frac{\theta_2}{2-\gamma}.$$

### Quantity Setting in Stage $t = 1$

Firm $A$'s expected profit considered in stage $t = 1$ is given by

$$\pi_1^A = x_1^A p_1^A + \delta E[\pi_2^A | I_1] = x_1^A \cdot (\alpha + \theta_1 - x_1^A - \gamma x_1^B) + \delta E[p_2^A x_2^A | I_1].$$

In period 1, firms anticipate the equilibrium quantities from period 2, so that $E[p_2^A(x_2^*)|I_1] = \frac{\alpha}{2+\gamma} + \frac{\theta_2}{2-\gamma} = x_2^{A*}$, and thus $E[\pi_2^A|I_1] = (x_2^{A*})^2$. We can additionally use the observations that $\theta_1 = E[\theta_1^*|I_1] = 0$ and

$$\partial\theta_2/\partial x_1^A = -\partial\theta_2/\partial x_1^B = \partial\eta_2/\partial x_1^A = (\gamma-1)\frac{n_2}{2m_1}.$$

Using $\partial E[\pi_2^A|I_1]/\partial x_1^A = 2x_2^A \cdot \frac{1}{2-\gamma} \cdot \frac{\partial\theta_2}{\partial x_1^A}$, we obtain the FOC of firm A, given by

$$\frac{\partial \pi_1^A}{\partial x_1^A} = \alpha - 2x_1^A - \gamma x_1^B + \delta\left(\frac{2x_2^A}{2-\gamma} \cdot \frac{\partial\theta_2}{\partial x_1^A}\right) = 0.$$

This yields the best responses

$$x_1^A(x_1^B) = \frac{\alpha}{2} - \frac{\gamma x_1^B}{2} + \delta\left(\frac{2x_2^A}{2-\gamma} \cdot \frac{\partial\theta_2}{\partial x_1^A}\right),$$

and similarly

$$x_1^B(x_1^A) = \frac{\alpha}{2} - \frac{\gamma x_1^A}{2} + \delta\left(\frac{2x_2^B}{2-\gamma} \cdot \frac{\partial\theta_2}{\partial x_1^A}\right).$$

In equilibrium best responses intersect, and, using $E[\theta_2^*|I_1] = m_1 E[q] + n_2 E[\tilde{q}] = 0$, the equilibrium quantities are then given by

$$x_1^{A*} = x_1^{B*} = \frac{\alpha}{2+\gamma}\left(1 + \frac{\delta(\gamma-1)}{4-\gamma^2} \cdot \frac{n_2}{m_1}\right).$$

### Differentiation Investment in Stage $t = 0$

It holds that for $j \in \{A, B\}$

$$\frac{d(x_{NL}^{j*})^2}{d\gamma} \cdot \frac{\partial\gamma}{\partial k^j} = \frac{d(x_2^{j*})^2}{d\gamma} \cdot \frac{\partial\gamma}{\partial k^j} = \frac{2\delta\lambda\alpha^2}{27}.$$

Further helpful results for the calculation of the model with consumer learning are

$$\frac{dx_1^{j*}}{d\gamma}\cdot\frac{\partial\gamma}{\partial k^j}$$
$$=\left[\frac{-\alpha}{(2+\gamma)^2}+\frac{\alpha\delta n_2}{m_1}\left\{\frac{(2+\gamma)^2(2-\gamma)-(\gamma-1)[2(2+\gamma)(2-\gamma)-(2+\gamma)^2]}{(2+\gamma)^4(2-\gamma)^2}\right\}\right]\cdot(-\lambda\gamma)$$

and

$$\frac{\partial[\alpha-(1+\gamma)x_1^{j*}]}{\partial k^j}=\lambda\gamma x_1^{j*}-(1+\gamma)\left[\frac{dx_1^{j*}}{d\gamma}\cdot\frac{\partial\gamma}{\partial k^j}\right].$$

Evaluating at $\gamma=1$, using the sum rule in differentiation and the above results, we obtain

$$\left.\frac{\partial\{x_1^{j*}\cdot[\alpha-(1+\gamma)x_1^{j*}]\}}{\partial k^j}\right|_{\gamma=1}=\frac{\lambda\alpha^2}{27}\left(1-\frac{\delta n_2}{m_1}\right)+\frac{\lambda\alpha^2}{9}\cdot\left(1-\frac{2}{3}\left\{1-\frac{\delta n_2}{m_1}\right\}\right)$$
$$=\frac{\lambda\alpha^2}{9}\cdot\left(\frac{2}{3}+\frac{\delta n_2}{3m_1}\right).$$

Overall, the derivative of profit w.r.t. investment in differentiation, evaluated at $\gamma=1$, is

$$\partial\pi_L^j/\partial k^j\big|_{\gamma=1}=\left.\left(\frac{\partial\left[x_1^{j*}\{\alpha-(1+\gamma)x_1^{j*}\}\right]}{\partial k^j}+\delta\cdot\frac{d(x_2^{j*})^2}{d\gamma}\cdot\frac{\partial\gamma}{\partial k^j}-1\right)\right|_{\gamma=1}$$
$$=\frac{\lambda\alpha^2}{9}\cdot\left(\frac{2}{3}+\frac{\delta n_2}{3m_1}+\frac{2\delta}{3}\right)-1.$$

## Appendix 4.C Firm Behavior in the Bertrand Model

The calculations on the price setting in stages one and two of the Bertrand model can be done analogously to the quantity setting in the Cournot model (see Appendix 4.B), and thus we will only calculate the optimizing behavior for differentiation investment in stage $t=0$. In the following, the profit functions $\Pi_k^j$ represent the same profits as in the Cournot model, only adapted to the Bertrand setting.

### Differentiation Investment Without Consumer Learning (Benchmark)

No consumer learning implies $\theta_2 = 0$, such that the resulting optimal quantities and prices for firm A are

$$E[x_2^A(p_2^{A*})|I_1] = \frac{\alpha}{(2-\gamma)(1+\gamma)}$$
$$E[p_2^{A*}|I_1] = \frac{\alpha(1-\gamma)}{(2-\gamma)}.$$

The profit of firm A in a benchmark model with two periods without consumer learning is thus given by

$$\Pi_{NL}^j = (1+\delta)\cdot E[\Pi_2^A|I_1] - k^A = (1+\delta)E[x_2^A(p_2^{A*})|I_1]\cdot E[p_2^{A*}|I_1] - k^A.$$

The derivative of the objective function is then

$$\begin{aligned}\partial\Pi_{NL}^j/\partial k^j &= (1+\delta)\cdot\frac{d(E[x_2^A(p_2^{A*})|I_1]\cdot E[p_2^{A*}|I_1])}{d\gamma}\cdot\frac{\partial\gamma}{\partial k^j} - 1\\ &= (1+\delta)\cdot\frac{-\alpha^2(\gamma^3-3\gamma^2+4)+(3\gamma^2-6\gamma)(\alpha^2(1-\gamma))}{(\gamma^3-3\gamma^2+4)^2}\cdot(-\lambda\gamma)-1\end{aligned}$$

Firm A will invest in differentiation in equilibrium if

$$\partial\Pi_{NL}^j/\partial k^j\Big|_{\gamma=1} > 0 \Leftrightarrow \lambda > \frac{2}{\alpha^2(1+\delta)} = \bar{\lambda}_{NL}^B.$$

### Differentiation Investment With Consumer Learning

We can write

$$\begin{aligned}p_1^{A*} &= \frac{a}{2b-c}\cdot\left(1-\frac{(b+c)b\delta}{4b^2-c^2}\cdot\frac{n_2}{m_1}\right)\\ &= \frac{\alpha(1-\gamma)}{2-\gamma}\cdot\left(1-\frac{(1+\gamma)\delta}{4-\gamma^2}\cdot\frac{n_2}{m_1}\right)\\ &= \frac{\alpha(1-\gamma)}{2-\gamma}-\frac{(1-\gamma^2)\delta\alpha}{(4-\gamma^2)(2-\gamma)}\cdot\frac{n_2}{m_1}.\end{aligned}$$

The profit of firm A in the model with consumer learning is given by

$$\begin{aligned}\Pi_L^j &= p_1^{A*}\cdot\left[a+(c-b)p_1^{A*}\right]+\delta\cdot E[x_2^A(p_2^{A*})|I_1]\cdot E[p_2^{A*}|I_1]-k^A\\ &= p_1^{A*}\cdot\left[\frac{\alpha}{1+\gamma}+\frac{\gamma-1}{1-\gamma^2}\cdot p_1^{A*}\right]+\delta\cdot E[x_2^A(p_2^{A*})|I_1]\cdot E[p_2^{A*}|I_1]-k^A.\end{aligned}$$

Further helpful results are

$$\frac{\partial p_1^{A*}}{\partial \gamma} = \frac{-\alpha(2-\gamma)+\alpha(1-\gamma)}{(2-\gamma)^2} - \frac{(-2\gamma\delta\alpha)(2-\gamma)(4-\gamma^2)-(3\gamma^2-4\gamma-4)\alpha\delta(1-\gamma^2)}{(2-\gamma)^2(4-\gamma^2)^2}$$

$$\begin{aligned}
\frac{\partial\left[\frac{\alpha}{1+\gamma}+\frac{\gamma-1}{1-\gamma^2}\cdot p_1^{A*}\right]}{\partial k^A} &= -\alpha + \frac{(1-\gamma^2)+2\gamma(\gamma-1)}{(1-\gamma^2)^2}\cdot p_1^{A*} + \frac{\gamma-1}{1-\gamma^2}\cdot\frac{\partial p_1^{A*}}{\partial \gamma} \\
\left[\frac{\alpha}{1+\gamma}+\frac{\gamma-1}{1-\gamma^2}\cdot p_1^{A*}\right]\Bigg|_{\gamma=1} &= \alpha/2 \\
p_1^{A*}\Big|_{\gamma=1} &= 0 \\
\frac{\partial p_1^{A*}}{\partial \gamma}\Bigg|_{\gamma=1} &= -\alpha + \frac{2\alpha\delta n_2}{3m_1}
\end{aligned}$$

Overall, the derivative of profit w.r.t. investment in differentiation, evaluated at $\gamma = 1$, is

$$\begin{aligned}
\partial\Pi_L^j/\partial k^j\big|_{\gamma=1} &= \alpha\lambda\left(1-\frac{2\delta n_2}{3m_1}\right)\cdot\frac{\alpha}{2} + 0\cdot\frac{\partial\left[\frac{\alpha}{1+\gamma}+\frac{\gamma-1}{1-\gamma^2}\cdot p_1^{A*}\right]}{\partial k^A}(-\lambda) + \delta\lambda\frac{\alpha^2}{2} - 1 \\
&= \left[\frac{\alpha^2}{2}(1+\delta) - \frac{\alpha^2\delta n_2}{3m_1}\right]\lambda - 1.
\end{aligned}$$

Firm j will invest in differentiation in equilibrium if

$$\begin{aligned}
& \partial\Pi_L^j/\partial k^j\big|_{\gamma=1} > 0 \\
\Leftrightarrow \quad & \lambda > \frac{2}{\alpha^2\left[(1+\delta)-2\delta n_2/(3m_1)\right]} := \bar{\lambda}_L^B.
\end{aligned}$$

# Bibliography

**Adriani, Fabrizio and Luca G. Deidda.** 2011. "Competition and the signaling role of prices." *International Journal of Industrial Organization*, *29* (4): 412–425.

**Bagwell, Kyle.** 2007. "The Economic Analysis of Advertising." in Mark Armstrong and Robert Porter, eds., Vol. 3 of *Handbook of Industrial Organization*, Elsevier, 2007. pp. 1701–1844.

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

**Behrens, Christiaan and Mark Lijesen.** "Capacity Choice under Uncertainty with Product Differentiation." Tinbergen Institute Discussion Papers 12-113/VIII, Tinbergen Institute 2012.

**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.

**Board, Oliver.** 2009. "Competition and Disclosure." *The Journal of Industrial Economics*, *57* (1): 197–213.

**Brander, James A. and Barbara J. Spencer.** “Endogenous Horizontal Product Differentiation under Bertrand and Cournot Competition: Revisiting the Bertrand Paradox.” Working Paper 20966, National Bureau of Economic Research 2015.

**Brander, James A. and Barbara J. Spencer.** 2015. “Intra-industry trade with Bertrand and Cournot oligopoly: The role of endogenous horizontal product differentiation.” *Research in Economics*, *69*, 157–165.

**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.

**Buehler, Stefan and Dennis L. Gärtner.** 2013. “Making Sense of Nonbinding Retail-Price Recommendations.” *American Economic Review*, *103* (1): 335–59.

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

**Chamley, Christophe P.** 2004. *Rational Herds: Models of Social Learning*, Cambridge: Cambridge University Press.

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

**Chu, Wujin and Woosik Chu.** 1994. “Signaling quality by selling through a reputable retailer: An example of renting the reputation of another agent.” *Marketing Science*, *13* (2): 177–189.

**Cournot, Antoine-Augustin . A.** 1863. *Principes de la théorie des richesses*, Hachette.

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

**Crawford, Vincent P. and Joel Sobel.** 1982. “Strategic Information Transmission.” *Econometrica*, *50* (6): 1431–1451.

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

**Daughety, Andrew F. and Jennifer F. Reinganum.** 2007. “Competition and confidentiality: Signaling quality in a duopoly when there is universal private information.” *Games And Economic Behavior*, *58* (1): 94–120.

**Daughety, Andrew F. and Jennifer F. Reinganum.** 2008. "Communicating Quality: A Unified Model of Disclosure and Signaling." *Rand Journal of Economics*, *39* (4): 973–989.

**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.

**Dixit, Avinash.** 1979. "A Model of Duopoly Suggesting a Theory of Entry Barriers." *The Bell Journal of Economics*, *10*, 20–32.

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

**Dranove, David and Ginger Zhe Jin.** 2010. "Quality Disclosure and Certification: Theory and Practice." *Journal Of Economic Literature*, *48* (4): 935–963.

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

**Fabrizi, Simona, Steffen Lippert, Clemens Puppe, and Stephanie Rosenkranz.** "Suggested Retail Prices with Downstream Competition." Working Papers 1203, University of Otago, Department of Economics 2010.

**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.

**Gertner, R., R. Gibbons, and D. Scharfstein.** 1988. "Simultaneous signalling to the capital and product markets." *The RAND Journal of Economics*, pp. 173–190.

**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.

**Grossman, Sanford J.** 1981. "The Informational Role of Warranties and Private Disclosure about Product Quality." *The Journal of Law & Economics*, *24* (3): 461–483.

**Grossman, Sanford J. and Motty Perry.** 1986. "Perfect Sequential Equilibrium." *Journal of Economic Theory, 39* (1): 97–119.

**Guo, Liang.** 2009. "Quality Disclosure Formats in a Distribution Channel." *Management Science*, *55* (9): 1513–1526.

**Guo, Liang and Ying Zhao.** 2009. "Voluntary Quality Disclosure and Market Interaction." *Marketing Science*, *28* (3): 488–501.

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

**Hviid, Morten.** 1991. "Capacity constrained duopolies, uncertain demand and non-existence of pure strategy equilibria." *European Journal of Political Economy*, *7* (2): 183–190.

**Janssen, Maarten C. W. and Santanu Roy.** 2010. "Signaling quality through prices in an oligopoly." *Games And Economic Behavior*, *68* (1): 192–207.

**Jovanovic, Boyan.** 1982. "Truthful Disclosure of Information." *The Bell Journal of Economics*, *13* (1): 36–44.

**Judd, Kenneth L.** 1985. "The Law of Large Numbers with a Continuum of IID Random Variables." *Journal of Economic Theory*, *35*, 19–25.

**Kalra, Ajay and Shibo Li.** 2008. "Signaling Quality Through Specialization." *Marketing Science*, *27* (2): 168–184.

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

**Kreps, David M. and Jose A. Scheinkman.** 1983. "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes." *The Bell Journal of Economics*, *14* (2): 326–337.

**Levin, Dan, James Peck, and Lixin Ye.** 2009. "Quality Disclosure and Competition." *Journal of Industrial Economics*, *57* (1): 167–196.

**Lubensky, Dmitry.** "A Model of Recommended Retail Prices." Working Paper, Indiana University, Kelley School of Business, Department of Business Economics and Public Policy 2011.

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

**Milgrom, Paul and John Roberts.** 1986. "Price and Advertising Signals of Product Quality." *Journal of Political Economy*, *94* (4): 796–821.

**Milgrom, Paul R.** 1981. "Good news and bad news: Representation theorems and applications." *The Bell Journal of Economics*, pp. 380–391.

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

**Nelson, Phillip.** 1970. "Information and consumer behavior." *The Journal of Political Economy*, *78* (2): 311–329.

**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.

**Puppe, Clemens and Stephanie Rosenkranz.** 2011. "Why Suggest Non-Binding Retail Prices?" *Economica*, *78* (310): 317–329.

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

**Sadanand, Asha B. and Venkatraman Sadanand.** 1995. "Equlibria in Non-Cooperative Games II: Deviations Based Refinements of Nash Equilibrium." *Bulletin of Economic Research*, *47* (2): 93–113.

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

**Singh, Nirvikar and Xavier Vives.** 1984. "Price and Quantity Competition in a Differentiated Duopoly." *The RAND Journal of Economics*, *15*, 546–554.

**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.

**Spengler, Joseph J.** 1950. "Vertical Integration and Antitrust Policy." *Journal of Political Economy*, *58* (4): 347–352.

**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.

**Viscusi, W. Kip.** 1978. "A Note on "Lemons" Markets with Quality Certification." *The Bell Journal of Economics*, *9* (1): 277–279.

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