Three Essays in Industrial Organization

Inauguraldissertation

zur Erlangung des akademischen Grades

doctor rerum politicarum

eingereicht an der

Wirtschaftswissenschaftlichen Fakultät der Heinrich-Heine-Universität Düsseldorf

von

Irina Hasnas, M.Sc.

geb. am 26. April 1987 in Nisporeni.

Rektorin der Heinrich-Heine-Universität Düsseldorf:
Univ.-Prof. Dr. Anja Steinbeck

Dekan der Wirtschaftswissenschaftlichen Fakultät:
Univ.-Prof. Dr. Justus Haucap

Gutachter:
1. Univ.-Prof. Dr. Christian Wey
2. Jun.-Prof. Dr. Irina Baye

To my parents and my daughter.

Acknowledgments

I would like to thank my supervisors, Prof. Dr. Christian Wey and Jun.-Prof. Dr. Irina Baye, to whom I feel deep gratitude, for their immense support and guidance. I would like to thank Alexander Rasch and Gerhard Riener for their helpful comments and insights. I would also like to thank the whole Düsseldorf Institute for Competition Economics (DICE) team.

Table of Contents

Chapter 1

Introduction

This dissertation presents three essays in Industrial organization. In the first two essays I analyze, how the availability of data on consumer preferences influences strategic decisions of the firms, such as where to locate and whether to collude. In the last essay I analyze, whether full or partial collusion is more stable when firms are heterogeneous in the quality of their products.

Nowadays, in the era of big data, firms have access to advanced information technologies, which allow them to collect, store and analyze large amounts of customer data. Customer data provides an opportunity to the firms to segment consumers along different dimensions of their preferences. Depending on data quality and the resulting segmentation, firms can practice price discrimination by targeting specific groups of consumers with personalized offers. In Chapters 2 and 3, we distinguish between two types of customer data (on the two characteristics of consumer preferences) and analyze firms' location choices and their incentives to collude depending on data quality.

In Chapter 2, which is published in *Journal of Economics*, we analyze firms' location choices in a Hotelling model with two-dimensional consumer heterogeneity, along addresses and transport cost parameters (flexibility). Firms can price discriminate based on perfect data on consumer addresses and (possibly) imperfect data on consumer flexibility. We show that firms' location choices depend on how strongly consumers differ in flexibility. Precisely, when consumers are relatively homogeneous, equilibrium locations are socially optimal regardless of the quality of customer flexibility data. However, when consumers are relatively differentiated, firms make socially optimal location choices only when customer flexibility data becomes perfect. These results are driven by the optimal strategy of a firm on its turf, monopolization or market-sharing, which in turn depends on consumer heterogeneity in flexibility. Our analysis is motivated by the availability of customer data, which allows firms to practice third-degree price discrimination based on both consumer characteristics relevant in spatial competition, addresses and transport cost parameters.

Chapter 3 discusses the sustainability of collusion in a game of repeated interaction where firms can price discriminate among consumers based on the same two types of customer data described in Chapter 2: firms have perfect data on consumer addresses, data on their flexibility is imperfect. Three collusive schemes are considered to analyze the impact of the improvement in the quality of customer flexibility data on firms' incentives to collude. This work is related to Liu and Serfes (2007) and Sapi and Suleymanova (2013). In contrast to Liu and Serfes in this model it is the customer flexibility data which is imperfect and not the data on consumer addresses. However, the results in this chapter support their findings that with the improvement in data quality it is more difficult to sustain collusion.

Unlike the previous two chapters, Chapter 4 describes three firms heterogeneous in their quality located on a Salop circle. The quality gap among these firms emphasizes the quality difference between branded manufacturer products and private label products. In this chapter, we analyze the incentives to collude when brand manufacturers compete with a private label producer of inferior quality. Full collusion is easier to sustain than partial collusion from the brands' perspective when horizontal differentiation is large and vertical differentiation is small. The private label firm is better off under full collusion than under partial collusion if goods are sufficiently homogenous (horizontal and/or vertical). Partial collusion could be preferred by the private label exactly when full collusion is easier to sustain. Improving the private label's quality makes full collusion more likely, either because it relaxes the brand producers' incentive constraint or because it shifts the preference of the private label firm from partial collusion to full collusion. Fully collusive behavior reveals itself through a nonnegative price effect on the brands' side caused by a quality increase of the private label good.

Chapter 2

Consumer Flexibility, Data Quality and Location Choice

Co-authored by Irina Baye

2.1 Introduction

The widespread use of modern information technologies allows firms to collect, store and analyze customer data in many industries. For example, loyalty programs and consumers' online activity are very important sources of customer data in the retailing industry.[1,2] Collected data allows firms to conclude on both consumer characteristics relevant in spatial

[1] Jeff Berry, senior director of Knowledge Development and Application at LoyaltyOne Inc., global provider of loyalty solutions in different industries including retailing, said that "*For most organizations today, the loyalty program actually becomes the core of the ability to capture consumer data ...*" (Linkhorn, 2013).

[2] Electronic Privacy Information Center (EPIC), public interest research group with a focus on privacy protection, notes that "Online tracking is no longer limited to the installation of the traditional "cookies" that record websites a user visits. Now, new tools can track in real time the data people are accessing or browsing on a web page and combine that with data about that user's location, income, hobbies, and even medical problems. These new tools include flash cookies and beacons. Flash cookies can be used to re-install cookies that a user has deleted, and beacons can track everything a user does on a web page including what the user types and where the mouse is being moved." ("Online Tracking and Behavioral Profiling" at http://epic.org/privacy/consumer/online_tracking_and_behavioral.html).

competition, consumers' addresses and transport cost parameters (flexibility).[3] Customer location data is one of the first information items provided by consumers while signing up for a loyalty program and can be deducted from the IP address of a computer during the online purchase. This data is easily accessible and can be considered as (almost) perfect. In contrast, firms can estimate consumer flexibility only with less than perfect accuracy. Similarly, following a boom of location-based marketing enabled by the emergence of smartphones with built-in GPS devices, firms can make personalized offers to consumers depending on their precise physical locations delivered by smartphone signals.[4] However, physical location is not the only piece of customer data used by mobile advertisers and is usually combined with other types of data such as on previous sales, demographics, social networks and app usage.[5] Additional customer data can be used to estimate consumer flexibility. Gained customer insights are used by firms in two ways. First, customer data allows targeted advertising and pricing where firms can practice third-degree price discrimination

[3] The term "flexibility" captures the intuition that depending on whether transport costs are high or low, consumers are less or more likely to buy from the farther firm, respectively. Consumers with high (low) transport costs can be referred to as less (more) flexible.

[4] According to the forecasts of a mobile advertiser eMarketer, mobile advertising expenditures will overcome print advertising and will account for 20% of total advertising spendings in the UK in 2015 (see eMarketer, 2015).

[5] For instance, mobile advertising companies, such as Factual and YP Marketing Solutions, design personalized advertisements based on the physical locations of consumers and other types of customer data (see https://www.factual.com/products/geopulse-audience and http://national.yp.com/, respectively).

based on both consumer locations and their flexibility.[6,7,8] Second, customer data is widely used to decide on the optimal store location.[9]

In this article we consider a Hotelling model, where consumers differ both in their locations and transport cost parameters. There are two firms which compete in prices and have access to perfect data on consumers' locations. Additionally, firms hold data on consumer flexibility of an exogenously given quality, which allows them to distinguish between

[6] CEO of Safeway Inc., second-largest supermarket chain in the U.S., Steve Burd, said that "*There's going to come a point where our shelf pricing is pretty irrelevant because we can be so personalized in what we offer people.*" (Ross, 2013). Similarly, the spokesman of Rosetta Stone, which sells software for computer-based language learning said that "*We are increasingly focused on segmentation and targeting. Every customer is different.*" (Valentino-Devries, 2012).

[7] EPIC notes that "Advertisers are no longer limited to buying an ad on a targeted website because they instead pay companies to follow people around on the internet wherever they go. Companies then use this information to decide what credit-card offers or product pricing to show people, potentially leading to price discrimination." ("Online Tracking and Behavioral Profiling" at http://epic.org/privacy/consumer/online_tracking_and_behavioral.html).

[8] In electronic commerce there is evidence of both discrimination based on consumer locations and flexibility. Mikians et al. (2012) find that some sellers returned different prices to consumers depending on whether a consumer accessed a seller's website directly or through price aggregators and discount sites (like nextag.com). Those price differences can be explained through differences in price sensitivity (flexibility) of the two types of consumers. Consumers accessing a seller's website through price aggregators are likely to be more price-sensitive. Similarly, vice president of corporate affairs at Orbitz Worldwide Inc., which operates a website for travel booking, said that "*Many hotels have proven willing to provide discounts for mobile sites.*" (Valentino-Devries et al., 2012). The latter can also be explained as price discrimination based on consumer flexibility since smartphone users can be considered as more price-sensitive due to the availability of different mobile applications, which collect special offers depending on a user's location. The evidence of price discrimination based on consumer locations is provided in Valentino-Devries et al. (2012) who find the strongest correlation between the differences in online prices and the distance to a rival's store from the center of a ZIP Code of a buyer.

[9] To mention just one of many examples, Waitrose, a British supermarket chain, used services of data analytics company BeyondAnalysis to analyze data on their customers' Visa card transactions to decide on new store locations (see Ferguson, 2013).

different flexibility segments and attribute every consumer to one of them. We consider two versions of our model depending on how strongly consumers differ in flexibility, with relatively homogeneous and differentiated consumers. We analyze firms' location choices in the two versions of our model depending on the quality of customer flexibility data.

This article contributes to the literature on spatial competition in Hotelling-type models where firms first choose locations and then compete in prices given the ability to practice perfect third-degree price discrimination based on consumer addresses. The famous result in Lederer and Hurter (1986) states that in the latter case in equilibrium every firm chooses its location so as to minimize social costs equal to the minimal costs of serving a consumer at each address. In a standard Hotelling model (with a uniform distribution of consumers along a line segment) this result implies socially optimal equilibrium locations. Hamilton and Thisse (1992) introduce a vertical dimension of consumer heterogeneity along which firms can practice first-degree price discrimination. They get the same optimality result as in Lederer and Hurter and conclude that "*...we see that the Hurter-Lederer efficient location result relies on perfectly inelastic consumer demands. For firms to locate efficiently when demands are price-sensitive, they need more flexibility in pricing...*" (Hamilton and Thisse, 1992, p. 184) On the one hand, our results support this conclusion, as we show that when the quality of customer flexibility data improves (and firms can identify more flexibility segments) equilibrium locations become closer to the socially optimal ones. However, this happens only when consumers are relatively differentiated in flexibility. With relatively homogeneous consumers in equilibrium firms choose socially optimal locations regardless of their ability to discriminate based on consumer flexibility. Our results imply that in a model with price-sensitive demands at each address socially optimal locations can be an equilibrium under weaker requirements on the quality of customer data available to the firms than perfect data.

Valletti (2002) is another article, which introduces heterogeneity along the vertical dimension of consumer preferences and assumes that consumers can be of two types depending on their valuation for quality. While firms can practice perfect third-degree price discrimination based on consumer addresses, they do not observe their types and, hence, have to rely on second-degree price discrimination at each address. Valletti shows that firms' location choices influence their discriminating ability. Different from Valletti, in our model firms' ability to discriminate is given exogenously and depends on the quality of customer flexibility data, such that a firm's location choice influences only its market share (along the horizontal dimension of consumer preferences) and the profit on a given location. Our results show that for the same data quality firms make different location choices depending on how strongly consumers differ in flexibility.

Overall, our article contributes to Hamilton and Thisse (1992) and Valletti (2002) by introducing third-degree price discrimination along the vertical dimension of consumer preferences enabled by customer data, while the former assume first-degree and the latter considers second-degree price discrimination.

This article is also related to Anderson and de Palma (1988) who assume that products are heterogeneous not only in the spatial dimension, but also in the characteristic space, which also leads to price-sensitive demands at each location. Similar to Anderson and de Palma we show that socially optimal prices and locations are not always an equilibrium, in contrast to models where only spatial dimension of heterogeneity is considered. However, different from Anderson and de Palma, we show that socially optimal prices and locations can also be an equilibrium in a model with price-sensitive demands. This happens in two cases. First, if consumers are relatively homogeneous in transport cost parameters. Second, if firms have perfect data on consumer flexibility, and hence, can perfectly discriminate along that dimension. The intuition for our results is as follows. When consumers are relatively homogeneous, in equilibrium every firm serves all consumers on its turf, even when firms do not hold data on consumer flexibility. This happens because if a firm targets at some address its most loyal customer (with the highest transport cost parameter), it suffices to decrease the price slightly to gain even the least loyal customer (with the lowest transport cost parameter). As a result, similar to Lederer and Hurter (1986), every firm chooses its location so as to minimize social (transport) costs, which implies socially optimal locations.

However, when consumers are relatively differentiated, in equilibrium on any address on its turf a firm serves only the more loyal consumers and loses the less loyal ones to the rival. To mitigate competition firms deviate from the socially optimal locations, and the interfirm distance is larger in equilibrium compared to both the first-best and the second-best. With the improvement in the quality of customer flexibility data distortions in firms' equilibrium locations become smaller, because every firm can better target consumers on its turf, which weakens the rival's ability to attract its loyal consumers. When flexibility data becomes perfect, firms make socially optimal location choices with relatively differentiated consumers too.

Tabuchi (1994) and Irmen and Thisse (1998) assume that products are differentiated along different dimensions and analyze firms' locations in each dimension. The result of Tabuchi that firms choose maximal differentiation in one dimension and minimal differentiation in the other is generalized by Irmen and Thisse in a model of spatial competition in a multi-characteristic space who show that the Nash equilibrium implies maximal differentiation only in the dominant product characteristic. While in our model products differ only in one (horizontal) dimension, we introduce consumer heterogeneity in the strength of their preferences along that dimension. In contrast, in Tabuchi and Irmen and Thisse it is assumed that the transport cost parameters related to each product characteristic are same among all consumers. We show that firms' location choices depend on how strongly consumers differ in transport cost parameters and firms' ability to discriminate along that dimension of consumer preferences.

Finally, our article is related to Jenzsch, Sapi and Suleymanova (2013) and Sapi and Suleymanova (2013). Both articles assume that consumers differ in the strength of their brand preferences. In the former article the authors analyze firms' incentives to share different types of customer data depending on how strongly consumers differ in flexibility and consider only data of perfect quality. In the latter paper the authors analyze firms' incentives to acquire customer flexibility data depending on its quality and consumer heterogeneity along that dimension. The focus of our article is the analysis of firms' location choices depending on the quality of customer flexibility data and on how strongly consumers differ in flexibility. In addition, in one of our extensions we endogenize firms' decisions to acquire customer flexibility data and get results, which differ from those of Sapi and Suleymanova. We show that this difference is due to the location effect, which reduces firms' incentives to acquire flexibility data, and is absent in Sapi and Suleymanova, where firms do not choose their locations. Precisely, we show that firms do not necessarily end up in the prisoner's dilemma when flexibility data becomes precise and consumers are relatively differentiated. In the other extension where we assume that firms hold flexibility data of asymmetric quality, we show that with the improvement in data quality efficiency in location choices can be restored only when consumers are relatively homogeneous.

This chapter is organized as follows. In the next section we present the model. In Section 2.3 we provide the equilibrium analysis, state our results and compare them with Lederer and Hurter (1986) and Anderson and de Palma (1988). In Section 2.4 we consider two extensions. In the first one we analyze firms' location choices under the assumption that they hold flexibility data of asymmetric quality and in the second one we analyze firms' incentives to acquire flexibility data. Finally, in Section 2.5 we conclude.

2.2 The Model

We analyze Bertrand competition between two firms, A and B, located at $0 \leq d_A \leq 1$ and $0 \leq d_B \leq 1$ on a unit-length Hotelling line, respectively. Firms produce the same product of two different brands, A and B, respectively. Production costs are set to zero. There is a unit mass of consumers, each buying at most one unit of the product. We follow Jenzsch, Sapi and Suleymanova (2013) and assume that consumers are heterogeneous not only in locations, but also in transport cost parameters (flexibility). Each consumer is uniquely characterized by a pair (x, t), where $x \in [0, 1]$ denotes a consumer's address and $t \in \left[\underline{t}, \overline{t}\right]$ her transport cost parameter, with $\overline{t} > \underline{t} \geq 0$. We assume that x and t are uniformly and independently distributed with density functions $f_x(x) = 1$ and $f_t(t) = 1/\left(\overline{t} - \underline{t}\right)$, to which we will refer as f_x and f_t, respectively. If a consumer does not buy at her location, she has to incur linear transport costs proportional to the distance to the firm. The utility of a consumer (x,t) buying from firm $i = A, B$ at price p_i is

$$U_i(x,t) = v - p_i - t\,|x - d_i|\,,$$

where $v > 0$ is the basic utility, which is assumed to be high enough such that all consumers buy in equilibrium. A consumer buys from a firm, which delivers her a higher utility. In case of equal utilities we assume that a consumer buys from a closer firm.[10]

Firms know perfectly the location of each consumer in the market and can discriminate among consumers respectively. Firms also hold flexibility data, which is imperfect. To model imperfect customer data we follow Liu and Serfes (2004) and Sapi and Suleymanova (2013) and assume that data quality is characterized by the exogenously given parameter $k = 0, 1, 2, ..., \infty$. This data allows a firm to identify 2^k flexibility segments and allocate each consumer to one of them. Segment $m = 1, 2, ..., 2^k$ consists of consumers with transport cost parameters $t \in \left[\underline{t}^m(k); \overline{t}^m(k)\right]$, where $\underline{t}^m(k) = \underline{t} + \left(\overline{t} - \underline{t}\right)(m-1)/2^k$ and $\overline{t}^m(k) = \underline{t} + \left(\overline{t} - \underline{t}\right) m/2^k$ denote the most and the least flexible consumers on segment m, respectively. With the improvement in data quality (k becomes larger), firms are able to allocate consumers to finer flexibility segments. If $k \to \infty$, a firm knows perfectly the location and transport cost parameter of each consumer in the market and can charge individual prices. Otherwise, a firm has to charge group prices (to consumers with the same address and on the same flexibility segment). With $p_{im}(x)$, $i = A, B$, we will denote the price of firm i on address x on segment m.

Following Jenzsch, Sapi and Suleymanova (2013) and Sapi and Suleymanova (2013) we will consider two extreme versions of our model with respect to consumer heterogeneity in flexibility, measured by the ratio $l := \overline{t}/\underline{t}$. In the version with *relatively homogeneous* consumers we assume that $\underline{t} > 0$ and $l \leq 2$. In the version with *relatively differentiated*

[10] This is a standard assumption in Hotelling-type models, where firms can practice perfect discrimination based on consumer addresses, and allows to avoid relying on ε-equilibrium concepts (see Lederer and Hurter, 1986).

consumers we assume that $\underline{t} = 0$, in which case $\lim_{\underline{t} \to 0} l = \infty$. In a similar way we can distinguish between flexibility segments. Precisely, we will say that consumers on segment m are relatively homogeneous if $\underline{t}^m(k) > 0$ and $\overline{t}^m(k)/\underline{t}^m(k) \leq 2$. We will say that consumers are relatively differentiated there if $\underline{t}^m(k) = 0$. It is straightforward to show that in the version of our model with relatively homogeneous consumers for any data quality consumers on all flexibility segments are relatively homogeneous. In the version of our model with relatively differentiated consumers, for any data quality consumers are relatively differentiated only on segment $m = 1$ and are relatively homogeneous on all other segments.

We consider a standard sequence of firms' moves where firms first choose their locations and then make pricing decisions (see, for example, Lederer and Hurter, 1986; Anderson and de Palma, 1988; Irmen and Thisse, 1998). Hence, the game unfolds as follows:

Stage 1 (Location choices). Independently from each other firms A and B choose locations d_A and d_B, respectively.

Stage 2 (Prices). Independently from each other firms choose prices to different consumer groups.

2.3 Equilibrium Analysis

We solve for a subgame-perfect Nash equilibrium and start from the second stage, where we derive firms' optimal prices given their location choices in the first stage.[11] Without loss of generality we will assume that the firm which is located closer to $x = 0$ is firm A, such that $d_A \leq d_B$.

[11] Note that if in the second stage simultaneously with their pricing decisions both firms were deciding whether to acquire flexibility data of an exogenously given quality k, both of them would do that in equilibrium. Following the argument of Liu and Serfes (2007), we can say that by refraining from data acquisition a firm cannot influence the decision of the rival to acquire customer data and only decreases its degrees of freedom in pricing. In the Section "Extensions" we endogenize firms' decisions to acquire customer flexibility data and consider a different timing, where firms make their location choices after the decision to hold flexibility data.

Stage 2: Prices. As firms know consumer addresses, they can charge different prices on each location. It is useful to consider separately four intervals of the unit line: *i)* interval $x \leq d_A$, which constitutes the hinterland of firm A, *ii)* interval $d_A < x \leq (d_A + d_B)/2$ with consumers between the two firms, which are closer to firm A, *iii)* interval $(d_A + d_B)/2 < x \leq d_B$ with consumers between the two firms, which are closer to firm B, *iv)* interval $x > d_B$, which constitutes the hinterland of firm B. In the following we will refer to the intervals *i)* and *ii)* as "the turf of firm A" and to consumers there as "loyal consumers of firm A." Symmetrically, we will refer to the intervals *iii)* and *iv)* as "the turf of firm B" and to consumers there as "loyal consumers of firm B."

Consider the turf of firm A. Under moderate prices only consumers with relatively small transport cost parameters switch to firm B, because buying from the farther firm is not very costly for them. On interval $x \leq d_A$ on some segment m these are consumers with

$$t < \widetilde{t}\,(d_A, d_B, p_{Am}(x), p_{Bm}(x); x|\, x \leq d_A) := \tfrac{p_{Am}(x) - p_{Bm}(x)}{d_B - d_A},$$

provided $\widetilde{t}\,(\cdot|\, x \leq d_A) \in \left[\underline{t}^m(k), \overline{t}^m(k)\right]$, where $p_{im}(x)$ is the price of firm $i = A, B$ on address x on segment m. The transport cost parameter of the indifferent consumer, $\widetilde{t}\,(\cdot|\, x \leq d_A)$, does not depend on her address directly, only (possibly) through firms' prices. The reason is that a consumer with address $x \leq d_A$ always has to travel the distance $d_A - x$ independently of whether she buys from firm A or from firm B. In the latter case compared to the former she has to travel additionally the distance $d_B - d_A$. Hence, the difference in utility from buying at the two firms does not depend on a consumer's address.

If $d_A < x \leq (d_A + d_B)/2$, then on segment m the transport cost parameter of the indifferent consumer is

$$\widetilde{t}\left(d_A, d_B, p_{Am}(x), p_{Bm}(x); x|\, d_A < x \leq \tfrac{d_A + d_B}{2}\right) := \tfrac{p_{Am}(x) - p_{Bm}(x)}{d_A + d_B - 2x},$$

provided $\widetilde{t}(\cdot|\, d_A < x \leq (d_A + d_B)/2) \in \left[\underline{t}^m(k), \overline{t}^m(k)\right]$. Those consumers buy from firm A, who have relatively high transport cost parameters: $t \geq \widetilde{t}(\cdot|\, d_A < x \leq (d_A + d_B)/2)$. Different from $\widetilde{t}\,(\cdot|\, x \leq d_A)$, $\widetilde{t}(\cdot|\, d_A < x \leq (d_A + d_B)/2)$ depends on the address of the indifferent consumer. Precisely, when x increases, for given firms' prices the transport cost parameter of the indifferent consumer becomes larger. If a consumer is located close to firm B, she may find it optimal to buy from firm B even if she has a relatively high transport cost parameter. Using symmetry, we can also derive consumer demand on the turf of firm B.

Each firm maximizes its profit separately on each address x and each segment m. For example, on some $x \leq d_A$ and some m firm A solves the optimization problem

$$\begin{aligned} \max_{p_{Am}(x)} \Pi_{Am}\left(d_A, d_B, p_{Am}(x), p_{Bm}(x); x, k \middle| x \leq d_A\right) &= f_t\left(\bar{t}^m(k) - \widetilde{t}\left(\cdot \middle| x \leq d_A\right)\right) p_{Am}(x), \\ \text{s.t. } \widetilde{t}\left(\cdot \middle| x \leq d_A\right) &\in \left[\underline{t}^m(k), \bar{t}^m(k)\right], \end{aligned}$$

where $\Pi_{im}\left(\cdot \middle| x \leq d_A\right)$ denotes the profit of firm $i = A, B$ on address $x \leq d_A$ on segment m. The optimization problem of firm B is

$$\begin{aligned} \max_{p_{Bm}(x)} \Pi_{Bm}\left(d_A, d_B, p_{Am}(x), p_{Bm}(x); x, k \middle| x \leq d_A\right) &= f_t\left(\widetilde{t}\left(\cdot \middle| x \leq d_A\right) - \underline{t}^m(k)\right) p_{Bm}(x), \\ \text{s.t. } \widetilde{t}\left(\cdot \middle| x \leq d_A\right) &\in \left[\underline{t}^m(k), \bar{t}^m(k)\right]. \end{aligned}$$

In the following lemma we state firms' equilibrium prices, demand regions and profits depending on their location choices in the first stage of the game in the version of our model with relatively differentiated consumers. We will use the subscripts "d" and "h" to denote the equilibrium values in the versions of our model with relatively differentiated and homogeneous consumers, respectively.

Lemma 2.1 *(Stage 2: optimal prices. Relatively differentiated consumers). Assume that consumers are relatively differentiated in flexibility. Equilibrium prices and demand regions depend on consumer's address, flexibility segment and the quality of customer flexibility data.*

i) Consider some x in the hinterland of firm $i = A, B$. On $m = 1$ firms charge prices $p^d_{i1}(d_A, d_B; x, k) = 2\bar{t}\,(d_B - d_A) / \left(3 \times 2^k\right)$ and $p^d_{j1}(d_A, d_B; x, k) = \bar{t}\,(d_B - d_A) / \left(3 \times 2^k\right)$, where firm i serves consumers with $t \geq \bar{t} / \left(3 \times 2^k\right)$, $j = A, B$ and $i \neq j$. On $m \geq 2$ firms charge prices $p^d_{im}(d_A, d_B; x, k) = \underline{t}^m(k)\,(d_B - d_A)$ and $p^d_{jm}(d_A, d_B; x, k) = 0$, where firm i serves all consumers.

ii) Consider some $x \in [d_A, d_B]$ *on the turf of firm* $i = A, B$. *On* $m = 1$ *firms charge prices* $p^d_{i1}(d_A, d_B; x, k) = 2\bar{t}\,|d_A + d_B - 2x| / \left(3 \times 2^k\right)$ *and* $p^d_{j1}(d_A, d_B; x, k) = \bar{t}\,|d_A + d_B - 2x| / \left(3 \times 2^k\right)$, *where firm* i *serves consumers with* $t \geq \bar{t}/\left(3 \times 2^k\right)$, $j = A, B$ *and* $i \neq j$. *Firms' prices on* $m \geq 2$ *are* $p^d_{im}(d_A, d_B; x, k) = \underline{t}^m(k)\,|d_A + d_B - 2x|$ *and* $p^d_{jm}(d_A, d_B; x, k) = 0$, *where firm* i *serves all consumers.*

Firms realize profits

$$\begin{aligned}
\Pi^d_A(d_A, d_B; k) &= \frac{\bar{t}(d_B - d_A)\left(9 \times 2^k\left(2^k - 1\right)(3d_A + d_B) + 2(11d_A + d_B) + 8\right)}{9 \times 2^{2k+3}} \text{ and} \\
\Pi^d_B(d_A, d_B; k) &= \frac{\bar{t}(d_B - d_A)\left(32 - 9 \times 2^k(2^k - 1)(3d_B + d_A - 4) - 2(11d_B + d_A)\right)}{9 \times 2^{2k+3}}.
\end{aligned}$$

Proof. See Appendix.

For the intuition behind Lemma 2.1 we will consider the turf of firm A. On any address and any segment there firm A charges in equilibrium a higher price than the rival, because in case of buying at firm A consumers have to bear smaller transport costs. While all equilibrium prices of firm A are positive, firm B charges positive prices only on segment $m = 1$. Also, firm B serves consumers only on that segment. The differences in firms' equilibrium prices on segments $m = 1$ and $m \geq 2$ are driven by the differences in the equilibrium strategy of firm A on its turf, as shown in Sapi and Suleymanova (2013). Consider, for example, the interval $x \leq d_A$. On segment $m = 1$, where consumers are relatively differentiated, firm A follows a so-called market-sharing strategy, such that its best-response function takes the form

$$p_{A1}(d_A, d_B, p_{B1}(x); x, k|\, x \leq d_A) = \begin{cases} \frac{\bar{t}^1(k)(d_B - d_A) + p_{B1}(x)}{2} & \text{if } \quad p_{B1}(x) < \bar{t}^1(k)\,(d_B - d_A) \\ p_{B1}(x) & \text{if } \quad p_{B1}(x) \geq \bar{t}^1(k)\,(d_B - d_A)\,. \end{cases} \tag{2.1}$$

To monopolize segment $m = 1$ firm A has to charge a price equal to that of the rival, because the most flexible consumer (with $\underline{t}^1 = 0$) can switch brands costlessly. The best-response function (2.1) shows that firm A finds it optimal to monopolize segment $m = 1$ only if the rival's price is relatively high: $p_{B1}(x, k) \geq \bar{t}^1(k)\,(d_B - d_A)$. Otherwise, if the rival's price is relatively low ($p_{B1}(x, k) < \bar{t}^1(k)\,(d_B - d_A)$), firm A optimally charges a higher price and loses the more flexible consumers. To attract the loyal consumers of the rival firm B charges in equilibrium a low price ($p_{B1}(x) < \bar{t}^1(k)\,(d_B - d_A)$), which makes the market-sharing outcome optimal for firm A.

In contrast, on segments $m \geq 2$, where consumers are relatively homogeneous, firm A follows a so-called monopolization strategy, such that its best-response function takes the form

$$p_{Am}(\, d_A, d_B, p_{Bm}(x); x, k|\, x \leq d_A) = p_{Bm}(x) + \underline{t}^m(k)\,(d_B - d_A)\,,\text{ for any } p_{Bm}(x). \quad (2.2)$$

$p_{Am} = p_{Bm}(x) + \underline{t}^m(k)\,(d_B - d_A)$ is the highest price, which allows firm A to monopolize segment m on some address $x \leq d_A$ on its turf for a given price of the rival, $p_{Bm}(x)$. As the best-response function (2.2) shows, regardless of the rival's price firm A prefers to charge a relatively low price to serve all consumers on segment m. As a result, in equilibrium firm B cannot do better than charging the price of zero. Firm A serves all consumers on segment m although it charges a positive price there.

The type of the equilibrium strategy of a firm on some flexibility segment on its turf, market-sharing or monopolization, depends on how strongly consumers differ there in flexibility. When consumers are relatively homogeneous on some segment, it suffices for a firm to decrease slightly the price targeted at the least flexible consumer to serve all consumers there, such that regardless of the rival's price a firm finds it optimal to monopolize the segment. In contrast, when consumers are relatively differentiated on a given segment, serving all consumers there requires a substantial reduction in the price targeted at the least flexible consumer (because the most flexible consumer can switch brands costlessly), which makes the monopolization outcome optimal only when the rival's price (which serves as an anchor for a firm's price) is high enough.

It is also worth noting that the equilibrium distribution of consumers between the firms depends only on which firm's turf and on which flexibility segment (with relatively homogeneous or differentiated consumers) they are located. Precisely, in equilibrium all consumers on segments $m \geq 2$ buy from their preferred firms and on segment $m = 1$ one third of the more flexible consumers switches to the less preferred firm.[12] However, the equilibrium prices of a firm on its turf depend also on whether a consumer is located between the two firms or in its hinterland. Precisely, to consumers on the same segment on its turf a firm charges a higher price if they are located in its hinterland because switching to the other firm is more costly for them. In the next lemma we characterize the equilibrium of the second stage of the game in the version of our model with relatively homogeneous consumers.

Lemma 2.2 *(Stage 2: optimal prices. Relatively homogeneous consumers). Assume that consumers are relatively homogeneous in flexibility. Equilibrium prices and demand regions depend on consumer's address, flexibility segment and the quality of customer flexibility data.*

i) Consider some x in the hinterland of firm $i = A, B$. On $m \geq 1$ firms charge prices $p^h_{im}(d_A, d_B; x, k) = \underline{t}^m(k)(d_B - d_A)$ and $p^h_{jm}(d_A, d_B; x, k) = 0$, where firm i serves all consumers, $j = A, B$ and $i \neq j$.

ii) Consider some $x \in [d_A, d_B]$ on the turf of firm $i = A, B$. On $m \geq 1$ firms charge prices $p^h_{im}(d_A, d_B; x, k) = \underline{t}^m(k)|d_A + d_B - 2x|$ and $p^h_{jm}(d_A, d_B; x, k) = 0$, where firm i serves all consumers, $j = A, B$ and $i \neq j$.

Firms realize profits

$$\begin{aligned}\Pi^h_A(d_A, d_B; k) &= \frac{\underline{t}\left(\left(2^k+1\right)+l\left(2^k-1\right)\right)(d_B+3d_A)(d_B-d_A)}{2^{k+3}} \text{ and}\\ \Pi^h_B(d_A, d_B; k) &= \frac{\underline{t}\left(\left(2^k+1\right)+l\left(2^k-1\right)\right)(4-d_A-3d_B)(d_B-d_A)}{2^{k+3}}\end{aligned}$$

Proof. See Appendix.

In the version of our model with relatively homogeneous consumers, consumers are relatively homogeneous on any flexibility segment for any quality of customer data. As we showed above, in that case every firm follows a monopolization strategy on any segment on its turf. As a result, the rival charges the prices of zero on a firm's turf and serves no

[12] Note that $\lim_{k\to\infty} \overline{t}^m(k) = 0$, such that when firms can perfectly discriminate based on consumer flexibility, every firm serves in equilibrium all consumers on its turf.

consumers there. We next analyze firms' location choices given their optimal prices in the second stage of the game.

Stage 1: Location choices. We first derive socially optimal locations. Following Anderson and de Palma (1988) we will distinguish between first-best and second-best locations. In the former case social planner determines both prices and locations. In the latter case social planer only determines locations, while firms choose non-cooperatively prices to maximize their profits under the prescribed locations. In the following lemma we state first-best and second-best locations in both versions of our model.

Lemma 2.3 *(Socially optimal locations). In both versions of our model first-best prices satisfy*

$$p_{im}^{FB}(x) \leq t|x - x_j| - t|x - x_i| + p_{jm}^{FB}(x) \text{ if } |x - x_i| \leq |x - x_j|,$$

and first-best locations are $d_A^{FB} = 1/4$ *and* $d_B^{FB} = 3/4$. *Second-best locations depend on consumer heterogeneity in flexibility.*

i) If consumers are relatively homogeneous, then second-best locations coincide with first-best locations: $d_A^{SB,h} = 1/4$ *and* $d_B^{SB,h} = 3/4$.

ii) If consumers are relatively differentiated, then second-best locations are

$$d_A^{SB,d}(k) = \frac{9 \times 2^{2k}}{36 \times 2^{2k} - 4} > d_A^{FB} \text{ and } d_B^{SB,d}(k) = \frac{27 \times 2^{2k} - 4}{36 \times 2^{2k} - 4} < d_B^{FB}, \text{ for any } k \geq 0,$$

such that $\lim_{k\to\infty} d_A^{SB,d}(k) = 1/4$ *and* $\lim_{k\to\infty} d_B^{SB,d}(k) = 3/4$.

Proof. See Appendix.

First-best prices must induce an allocation of consumers between the firms where every consumer buys from the closer firm. Then first-best locations which minimize transport costs are symmetric, and every firm is located in the middle of its turf, which yields $d_A^{FB} = 1/4$ and $d_B^{FB} = 3/4$. Note that first-best locations do not depend on how strongly consumers differ in flexibility.

For given locations firms' equilibrium prices and the resulting distribution of consumers between the firms differ in the two versions of our model, such that we also get different second-best locations. As stated in Lemma 2.1, with relatively differentiated consumers every firm loses on its turf the more flexible consumers. Compared to the first-best, in the second-best firms are located closer to each other, which minimizes the transport costs of those consumers. In contrast, with relatively homogeneous consumers first-best and second-best locations coincide, because as stated in Lemma 2.2, in equilibrium every consumer buys from its preferred firm under any firms' locations. In the next proposition we characterize firms' equilibrium locations and compare them with the socially optimal ones.

Proposition 2.1 *(Stage 1: location choices). Equilibrium locations depend on consumer heterogeneity in flexibility.*

i) If consumers are relatively homogeneous, then for any data quality $k \geq 0$ in equilibrium firms choose locations:

$$d_A^h = 1 - d_B^h = \frac{1}{4},$$

which coincide with both first-best and second-best locations. Firms realize profits

$$\Pi_i^h(k) = \frac{3\underline{t}\left(2^k+1+l\left(2^k-1\right)\right)}{2^{k+5}},\ i = A, B.$$

ii) If consumers are relatively differentiated, then in equilibrium firms choose locations:

$$d_A^d(k) = 1 - d_B^d(k) = \frac{9\times 2^{2k} - 9\times 2^k + 6}{36\times 2^{2k} - 36\times 2^k + 32}, \tag{2.3}$$

where $\partial d_A^d(k)/\partial k > 0$ and $\partial d_B^d(k)/\partial k < 0$, such that with the improvement in data quality firms locate closer to each other. It holds that $d_A^d(k) < d_A^{FB} < d_A^{SB,d}(k)$ for any $k \geq 0$, while the inequality for firm B is symmetric. Moreover, $\lim_{k\to\infty} d_A^d(k) = d_A^{FB}$ and $\lim_{k\to\infty} d_B^d(k) = d_B^{FB}$. Firms realize profits

$$\Pi_i^d(k) = \frac{\overline{t}\left(9\times 2^{2k} - 9\times 2^k + 10\right)^2\left(27\times 2^{2k} - 27\times 2^k + 22\right)}{9\times 2^{2k+5}\left(9\times 2^{2k} - 9\times 2^k + 8\right)^2},\ i = A, B.$$

Proof. See Appendix.

In the following we will explain and provide intuition for our results in each version of our model using the approach of Lederer and Hurter (1986, in the following: LH). When consumers are relatively homogeneous, firms make socially optimal location choices, such that the equilibrium locations coincide with both first-best and second-best locations. This result is driven by the fact that every firm follows a monopolization strategy on any address on its turf and serves in equilibrium all consumers there. It can be shown that when firm i chooses the optimal location, it solves the optimization problem[13]

$$\min_{d_i} \left\{ \frac{(\bar{t}+\underline{t})}{2} \int_0^1 \min \left\{ D_j \left(d_j ; x \right), D_i \left(d_i ; x \right) \right\} dx \right\}, \tag{2.4}$$

where $D_i \left(d_i ; x \right) := |x - d_i|$ and $\left(\bar{t} + \underline{t} \right) \left(\int_0^1 \min \left\{ D_j \left(d_j ; x \right), D_i \left(d_i ; x \right) \right\} dx \right) /2$ can be, following LH, defined as *social transport costs*, which are the total transport costs incurred by consumers when they are served by firms in a cooperative manner minimizing transport costs. The latter implies that every consumer buys from the closer firm. It follows from (2.4) that the location choice of firm i minimizes social transport costs given the location of the rival, d_j, yielding first-best locations in equilibrium.[14,15] Equilibrium locations also coincide with the second-best locations. The latter minimize transport costs given the equilibrium allocation of consumers and, hence, also minimize social transport costs. Indeed,

[13] All the derivations of the formulas presented in this section are provided in the Appendix.

[14] To be more precise, in our case the equilibrium location of a firm minimizes *directly* the total distance travelled by consumers. In LH the equilibrium location of a firm minimizes directly social transport costs (if production costs are zero). This difference is related to the fact that in our model firms do not know the transport cost parameter of an individual consumer unless $k \to \infty$. Then in equilibrium in the version with relatively homogeneous consumers every consumer pays a price equal to the difference in the distances between the consumer and the two firms multiplied by the transport cost parameter of the most flexible consumer on the segment to which consumer belongs, and not consumer's own transport cost parameter. However, this difference between LH and our model does not change the main result that with relatively homogeneous consumers firms make socially optimal location choices.

[15] LH analyze firms' location choices in a two-dimensional market region. LH show that in that case equilibrium locations do not necessarily minimize social (transport) costs globally. However, this is always the case in our model (in a version with relatively homogeneous consumers), where firms choose locations on a unit-length Hotelling line over which consumers are distributed uniformly.

they solve the optimization problem

$$\min_{d_A,d_B}\left\{\frac{(\bar{t}+\underline{t})}{2}\int_0^1 \min\{D_A(d_A;x),D_B(d_B;x)\}\,dx\right\}. \tag{2.5}$$

When consumers are relatively differentiated, equilibrium locations differ both from the first-best and second-best locations (provided flexibility data is not perfect). Following LH and using the results of Lemma 2.1 we can show that when firm i chooses location d_i, it solves the optimization problem

$$\min_{d_i}\left[(1+\alpha(k))\frac{\bar{t}}{2}\int_0^1 \min\{D_j(d_j;x),D_i(d_i;x)\}\,dx-\alpha(k)\frac{\bar{t}}{2}\int_0^1 D_i(d_i;x)\,dx\right], \tag{2.6}$$

where $\alpha(k)=2/\left[9\times 2^k\left(2^k-1\right)+8\right]$. Different from the optimization problem with relatively homogeneous consumers (2.4), in the optimization problem (2.6) firm i minimizes the weighted difference between social transport costs and transport costs of buying at firm i given by the first and the second terms in (2.6), respectively. The latter term is new and is driven by the incentive of a firm to locate further apart from the rival to mitigate competition under the imperfect ability of a firm to protect market shares on its turf. Consider, for example, firm A. For a given location of the rival, the location choice which minimizes social transport costs is $d_A(d_B)=d_B/3$. And the location choice, which maximizes the transport costs of buying at firm A is $d_A=0$. Then depending on $\alpha(k)$, $d_A(d_B;k)$ which solves (2.6), takes some value between $d_A(d_B)=d_B/3$ and $d_A=0$. As first-best locations solve the optimization problem (2.5) with $\underline{t}=0$, it is straightforward that compared to them, equilibrium locations are closer to the end points of the unit interval and the inter-firm distance is larger in equilibrium than in the first-best. Second-best locations minimize the transport costs

$$\min_{d_A,d_B}\left\{\frac{\bar{t}\,[1-\beta(k)]}{2}\int_0^1 \min\{D_A(\cdot),D_B(\cdot)\}\,dx+\frac{\bar{t}\beta(k)}{2}\int_0^1 \max\{D_A(\cdot),D_B(\cdot)\}\,dx\right\}, \tag{2.7}$$

where $\beta(k)=1/\left[9\times 2^{2k}\right]$. Different from first-best locations, which minimize social transport costs, second-best locations minimize the weighted sum of the social transport costs and the maximal transport cost given by the first and the second terms in (2.7), respectively. The first term in (2.7) is the transport costs of consumers who buy from their preferred firms, and the second term in (2.7) is the transport costs of consumers who buy from the farther firms. The latter costs are minimized under minimal differentiation when both firms are located at the middle of the unit interval.[16] Then second-best locations are closer to the middle of the unit interval compared to first-best locations, and the interfirm distance is larger in equilibrium than in the second-best, too. Note finally that $\lim_{k\to\infty} \alpha(k) = 0$ and $\lim_{k\to\infty} \beta(k) = 0$, such that the optimization problem (2.6) is equivalent to (2.4) and the optimization problem (2.7) is equivalent to (2.5) with $\underline{t} = 0$, and the equilibrium locations coincide with both first-best and second-best locations when flexibility data becomes perfect. This happens because in that case every firm serves consumers only on its own turf.

Combining our results in both versions of our model we make the following conclusions on firms' location choices in a Hotelling model with two-dimensional consumer heterogeneity, where firms can practice perfect third-degree price discrimination based on consumer addresses and (possibly) imperfect one based on consumer flexibility. *First*, firms choose socially optimal locations in two cases. If either consumers are relatively homogeneous in flexibility or if firms have perfect customer flexibility data and thus can practice perfect third-degree price discrimination along that dimension, too. In both cases every firm serves all consumers on its turf. We conclude that the optimality result of LH may also hold when customer data is imperfect. *Second*, when consumers are relatively differentiated in flexibility and customer flexibility data is imperfect, firms make socially suboptimal location choices. However, with the improvement in the quality of customer data equilibrium locations become closer to the socially optimal ones. This result supports the intuition of Hamilton and Thisse (1992, p. 184) that more flexibility in pricing leads to more efficient location choices when demands at each location are price-sensitive. However, as our first conclusion shows, flexibility in pricing (based on consumer transport cost parameters) is

[16] Note that $\int_0^1 \max\{D_A(d_A;x), D_B(d_B;x)\}\,dx = d_A - 1/2 - (d_A + d_B)^2/4$. It is straightforward to show that the values $d_A = d_B = 1/2$ solve the following constrained optimization problem: $\max_{d_A,d_B}\left[d_A - 1/2 - (d_A + d_B)^2/4\right]$, s.t. $d_A - d_B \leq 0$, $-d_A \leq 0$ and $d_B \leq 1$.

not a necessary condition for socially optimal locations.

In the following we compare our results with the other closely related article of Anderson and de Palma (1988, in the following: AP). AP assume that products are heterogeneous not only in the spatial dimension, but also in the characteristic space, which leads to price-sensitive demands at each location. While both versions of our model imply price-sensitive demands, our results are similar to those of AP only in the version with relatively differentiated consumers, where firms' markets overlap in equilibrium and most importantly, in equilibrium firms do not choose optimal locations (apart from the case where firms can perfectly discriminate based on consumer flexibility). As we showed above, when consumers are relatively homogeneous, in equilibrium every firm serves all consumers on its turf, such that firms' markets do not overlap, which leads to socially optimal equilibrium locations.[17]

In AP (in the logit model) the equilibrium locations depend on parameter $\mu \geq 0$, which is interpreted as a measure of consumer/product heterogeneity. In our case it makes sense to define μ as $\mu_1(k) := \bar{t}(d_B - d_A)/2^k$, being inversely related to the quality of flexibility data. Similar to AP we can draw a graph, which represents the equilibrium location of firm A depending on the ratio $\mu(k) := \mu_1(k)/\bar{t}(d_B - d_A) = 1/2^k$, where $\mu(k) \in (0, 1]$ (see Figure 2.1). As Figure 2.1 shows, our results correspond to those in AP where μ is relatively small. Precisely, the first-best location of firm A is constant, the second-best location of firm A increases in $\mu(k)$ and its equilibrium location decreases in $\mu(k)$. To explain the behavior of the equilibrium locations, AP identify two effects. With an increase in μ from $\mu = 0$ in AP products become heterogeneous (at each location) and a firm loses the monopoly power over its turf. As a result, firms move further apart to mitigate competition (*first effect*). At the same time higher μ implies the increased ability of each firm to gain consumers on the rival's turf. With an increase in the size of the latter group firms tend to locate closer

[17] In AP the monopolization outcome is not possible, because regardless of the difference in firms' prices on a given address, some consumers choose the other firm than the majority of consumers. In contrast, in our model one firm gains all consumers on a given address and flexibility segment if firms' prices there are very different.

to the center to minimize the transport costs of serving those consumers (*second effect*). At some point the second effect starts to dominate, and the interfirm distance decreases in equilibrium. When μ increases from $\mu = 0$ in our model, a firm loses the perfect targeting ability on its turf, which allows the rival to gain the less loyal consumers of a firm. To mitigate competition firms move further apart according to the first effect in AP. On the other hand, the weakened ability of the rival to target consumers on its own turf allows a firm to gain more consumers there, which creates an incentive to move closer to the rival according to the second effect in AP. However, different from AP the second effect never dominates in our model, as a firm gains at most only one third of consumers on the rival's turf.[18]

Compared to AP, our analysis highlights the importance of the ability to discriminate along the vertical dimension of consumer preferences for firms' location choices. When data on consumer flexibility improves (μ decreases), firms choose locations as if their products became more homogeneous (at each location), which mitigates competition and pushes the equilibrium locations in the direction of the socially optimal ones.

[18] This result depends on the assumption of the uniform distribution of consumer transport cost parameters on each location. For example, if there were two large consumer groups with relatively high and low transport cost parameters, it could be optimal for a firm to serve only the former group on its turf, while the latter would switch to the rival. In that case every firm could serve in equilibrium more loyal consumers of the rival than the own loyal consumers.

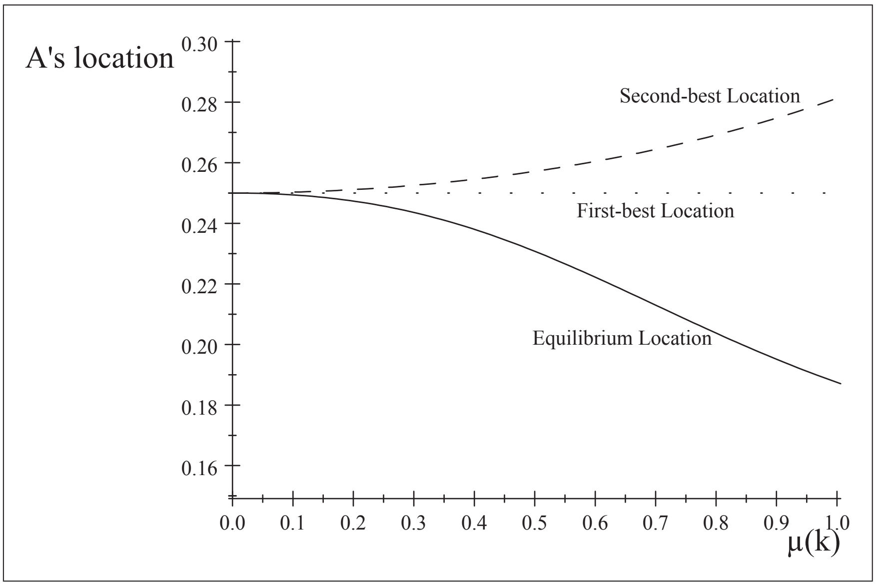

Figure 2.1. Equilibrium and optimal locations of firm A depending on $\mu(k)$

2.4 Extensions

In this section we consider two extensions. In the first one we allow firms to hold customer flexibility data of asymmetric quality. In the second one we analyze firms' incentives to acquire customer flexibility data.

Extension 1: Asymmetric information. In this extension we assume that firms hold customer flexibility data of asymmetric quality. Precisely, the quality of firm B's data is $k \geq 0$ and the quality of firm A's data is $k + n$, where $n \geq 1$. Hence, we assume that firm A always holds some data on consumer flexibility, while firm B may have not such data. Moreover, the quality of firm B's data is always worse than that of firm A and n measures the quality advantage of firm A's data. In reality we often observe that firms differ in their ability to collect and process customer data and to use it for price discrimination. One of the many examples is the UK's retail industry, where Tesco became a leading supermarket chain after the introduction of a loyalty program, which allowed it to collect data on consumer preferences and design individual discounts based on that data (see Winterman, 2013). The following proposition summarizes our results on the equilibrium location choices of the firms when they hold customer data of asymmetric quality and compares them with the first-best locations for the case of relatively homogeneous consumers. With the subscript "As" we refer to the case of asymmetric information.[19,20]

Proposition 2.2 *(Asymmetric data quality. Relatively homogeneous consumers). Assume that consumers are relatively homogeneous in flexibility. Firms' equilibrium location choices depend on the quality of their customer data and the ratio of consumer flexibility.*

i) If either $k \geq 1$ or if $k = 0$ and $l \leq 3/2$, then the equilibrium locations coincide with the first-best locations:

$$d_A^{h,As}(k,n,l) = 1 - d_B^{h,As}(k,n,l) = \tfrac{1}{4},$$

ii) If $k = 0$ and $l > 3/2$, then the equilibrium locations depend on the quality advantage of firm A's data in the following way.

a) If $n < \log_2((l-1)/(l-3/2))$, then in equilibrium firms choose locations:

$$\begin{aligned} d_A^{h,As}(0,n,l) &= \tfrac{-8l-12\times 2^n l+4l^2+13\times 2^n+4}{-32l-12\times 2^n l-12\times 2^n l^2+16l^2+25\times 2^n+16} \text{ and} \\ d_B^{h,As}(0,n,l) &= \tfrac{-24l-12\times 2^n l-8\times 2^n l^2+12l^2+21\times 2^n+12}{-32l-12\times 2^n l-12\times 2^n l^2+16l^2+25\times 2^n+16}. \end{aligned}$$

b) If $n \geq \log_2((l-1)/(l-3/2))$, then in equilibrium firms choose locations:

$$\begin{aligned} d_A^{h,As}(0,n,l) &= \tfrac{-l+2^{2n}-2^n l+2^{2n} l+2^n+1}{-l+4\times 2^{2n}-4\times 2^n l+4\times 2^{2n} l+4\times 2^n+1} \text{ and} \\ d_B^{h,As}(0,n,l) &= \tfrac{3\times 2^{2n}-l-3\times 2^n l+3\times 2^{2n} l+3\times 2^n+1}{4\times 2^{2n}-l-4\times 2^n l+4\times 2^{2n} l+4\times 2^n+1}. \end{aligned}$$

[19] We follow the literature and assume that in the case of asymmetric information (where one firm has more (accurate) data than the other), firms move sequentially. Precisely, the firm with less (accurate) data moves first and the other firm follows. This assumption can be justified by the observation that prices (discounts) designed for finer consumer groups can be changed easier than prices targeted at larger consumer groups. Moreover, under simultaneous moves a Nash equilibrium may not exist. For articles, which use the same assumption see, for instance, Thisse and Vives (1988), Shaffer and Zhang (1995, 2002) and Liu and Serfes (2004, 2005).

[20] To concentrate on the most important results, in Proposition 2.2 we omit the equilibrium prices and the allocation of consumers among the firms for any given location choices d_A and d_B, which can be found in the proof of Proposition 2.2 in the Appendix.

It holds that $\lim_{n\to\infty} d_A^{h,As}(\cdot) = d_A^{FB}$ *and* $\lim_{n\to\infty} d_B^{h,As}(\cdot) = d_B^{FB}$.

In both cases a) and b) it holds that $d_A^{h,As}(\cdot) < d_A^{FB}$ *and* $d_B^{h,As}(\cdot) < d_B^{FB}$. *With the improvement in the quality of firm A's data both firms move to the right and locate closer to each other:* $\partial d_i^{h,As}(\cdot)/\partial n > 0$, $i = A, B$ *and* $\partial\left(d_B^{h,As}(\cdot) - d_A^{h,As}(\cdot)\right)/\partial n < 0$. *When consumers become more differentiated in flexibility, then both firms move to the left and locate farther away from each other:* $\partial d_i^{h,As}(\cdot)/\partial l < 0$ *and* $\partial\left(d_B^{h,As}(\cdot) - d_A^{h,As}(\cdot)\right)/\partial l > 0$.

Proof. See Appendix.

Different from the symmetric case where both firms hold flexibility data of the same quality, with customer data of asymmetric quality firms' location choices are not always optimal even if consumers are relatively homogeneous. Precisely, the equilibrium locations coincide with the first-best locations if either consumers are strongly homogeneous in flexibility ($l \leq 3/2$) or if firm B holds some flexibility data ($k \geq 1$). In both cases firm B serves all consumers on its turf for any firms' location choices. As all consumers buy from their more preferred firms, firms choose optimally the locations, which minimize social transport costs, being the first-best locations, as we showed in our main analysis. However, if consumers are not strongly homogeneous in flexibility ($l > 3/2$) and firm B does not hold any customer data ($k = 0$), then firm B loses the more flexible consumers on its turf (those with the smaller transport cost parameters on segment 1 identified by firm A) for any firms' location choices. The reason is that with a customer data advantage firm A can better target consumers on the turf of firm B, such that protecting market shares becomes more costly for firm B (compared to the case when it holds some flexibility data, $k \geq 1$). As a result, firm B finds it optimal to target the more loyal consumers on its turf, while the more flexible consumers switch to the rival.

The inefficient distribution of consumers between the firms leads to the equilibrium locations, which do not minimize social transport costs. As a result, customer data asymmetry distorts both the equilibrium location of firm A and that of firm B and in equilibrium both firms locate to the left from the first-best locations. However, when flexibility data of firm A becomes more precise, firms move closer to the first-best locations, such that under perfect data equilibrium locations coincide with the first-best locations. The reason for this result is that when firm A can better target consumers on the turf of firm B (n gets larger), the latter is forced to decrease its price for the own loyal consumers (if $n \geq \log_2 ((l-1)/(l-3/2))$), such that it gains market shares among them.[21] When the data of firm A becomes almost perfect, firm B charges a price, which makes the most flexible consumer (with the transport cost parameter $\underline{t}$) indifferent between the two firms, and serves all of its loyal consumers. The efficient allocation of consumers between the firms leads to the equilibrium locations, which coincide with the first-best locations. It is also straightforward why the equilibrium locations become closer to the first-best ones when consumers become more homogeneous (l decreases): $\partial d_i^{h,As}(\cdot)/\partial l < 0$. The reason is that in that case less consumers buy from their less preferred firm A.

We can now conclude. With relatively homogeneous consumers, asymmetry in the available flexibility data distorts firs-best equilibrium locations only if firm B does not have any flexibility data ($k = 0$) and if consumers are not strongly homogeneous ($l > 3/2$). However, even in that case efficiency is restored when the quality of firm A's data becomes perfect. In the next proposition we summarize our results on the firms' equilibrium location choices when consumers are relatively differentiated in flexibility.

Proposition 2.3 *(Asymmetric data quality. Relatively differentiated consumers). Assume that consumers are relatively differentiated in flexibility. Firms' equilibrium location choices depend on the quality of their customer data.*
i) If $k = 0$*, then equilibrium locations are*

$$d_A^{d,As}(0,n) = \frac{8\times 2^{2n}+32\times 2^{3n}+16\times 2^{4n}-16\times 2^n+1}{8\times 2^{2n}+128\times 2^{3n}+112\times 2^{4n}-5} \text{ and}$$
$$d_B^{d,As}(0,n) = \frac{24\times 2^{2n}+96\times 2^{3n}+80\times 2^{4n}-7}{8\times 2^{2n}+128\times 2^{3n}+112\times 2^{4n}-5}.$$

With the improvement in firm A' data quality, both firms move to the left: $\partial d_i^{d,As}(\cdot)/\partial n < 0$, $i = A, B$*. The distance between the firms becomes (weakly) smaller if* $n \leq 2$ *and larger*

[21] As is shown in the Proof of Proposition 2.2 in the Appendix, in the case $n \geq \log_2 ((l-1)/(l-3/2))$ on its own turf firm B charges the price $p_B(d_A, d_B; x, n) = \left(\underline{t} + \left(\bar{t} - \underline{t}\right)/2^n\right)\left(|d_A - x| - |d_B - x|\right)$.

otherwise. Firms generally locate to the left from the first-best locations: for any $n \geq 1$ *it holds that* $d_A^{d,As}(\cdot) < d_A^{FB}$ *and for any* $n \geq 2$ *it holds that* $d_B^{d,As}(\cdot) < d_B^{FB}$*. Finally,*

$$\lim_{n\to\infty} d_A^{d,As}(\cdot) = \tfrac{1}{7} < d_A^{FB} \text{ and } \lim_{n\to\infty} d_B^{d,As}(\cdot) = \tfrac{5}{7} < d_B^{FB}.$$

ii) If $k \geq 1$*, then equilibrium locations are* $d_A^{d,As}(k,n)$ *and* $d_B^{d,As}(k,n)$*, such that*

$$\lim_{n\to\infty,k\to\infty} d_A^{d,As}(\cdot) = d_A^{FB} \text{ and } \lim_{n\to\infty,k\to\infty} d_B^{d,As}(\cdot) = d_B^{FB}.$$

Proof. See Appendix.

With relatively differentiated consumers, different from the case of relatively homogeneous consumers, perfect flexibility data held by firm A does not lead to first-best locations. Precisely, if firm B does not have any customer data ($k = 0$), then both firms choose locations to the left from the respective first-best locations and the efficiency is not achieved. The reason is that in equilibrium both firms lose consumers on their turfs, as we showed in the main analysis. Firm A loses consumers on its turf, because in the presence of a very flexible consumer (with the transport cost parameter of zero), it is too costly to serve all the loyal consumers. On the top of that, firm B loses consumers also due to a superior ability of firm A to target consumers on its turf. With the improvement in the quality of its data, firm A's ability to protect its market shares gets larger, such that with the perfect flexibility data firm A serves (almost) all of its loyal consumers. In contrast, firm B always loses some consumers on its turf. This is different in the case of relatively homogeneous consumers, where firm B serves (almost) all of its loyal consumers, when firm A holds data of perfect quality, while firm B does not have any flexibility data.

When both firms have customer flexibility data ($k \geq 1$), then with data of perfect quality firms choose first-best locations in equilibrium. We get the same result as in the main analysis with relatively differentiated consumers, where both firms hold flexibility data of the same quality. This similarity in the optimal locations is straightforward, because if $k \to \infty$, then both firms have flexibility data of a perfect quality and we are basically back in the symmetric case.

Combining our results from the two versions of our model we can make the following conclusions. When consumers are relatively homogeneous, then even in the asymmetric case where firm B does not hold any flexibility data, firms make socially optimal location choices provided that firm A holds perfect flexibility data. In contrast, when consumers are relatively differentiated in flexibility, then perfect data held by firm A does not lead to the optimal location choices if firm B does not have any flexibility data. Hence, if firms differ in the quality of their customer data, then whether or not perfect data held by firm A is sufficient to guarantee the optimal location choices of the firms, depends on how strongly consumers differ in flexibility.

Extension 2: Customer data acquisition. In the main analysis we assumed that both firms hold flexibility data. In contrast, in this extension we allow firms to choose whether they want to acquire flexibility data (which is costless). We assume here that firms make their location choices after their data acquisition decisions. Such a timing implies that data acquisition decision is a longer run decision than the location choice, which can be justified by the fact that firms often use the available customer data to find the optimal location.[22] Precisely, we consider the following sequences of firms' moves:

Stage 1 (Flexibility data acquisition). Independently from each other firms A and B decide whether to acquire customer flexibility data of an exogenously given quality $k \geq 1$.

Stage 2 (Location choice). Independently from each other firms A and B choose locations d_A and d_B, respectively.

Stage 3 (Prices). Independently from each other firms set prices to different consumer groups.

The following proposition summarizes firms' incentives to acquire customer flexibility data.

Proposition 2.4 *(Customer data acquisition). Firms' incentives to acquire customer flex-*

---

[22] Pitney Bowes Software, a company which provides among others Location Intelligence, mentions in one of its reports that "(t)o identify a profitable store location, acquiring customer and target market data is essential" (see Deakin University Australia Worldly and Pitney Bowes Software, 2012).

ibility data depend on consumer heterogeneity in flexibility.

i) If consumers are relatively homogeneous in flexibility, then a unique Nash equilibrium exists, where both firms acquire customer data of any quality $k \geq 1$*. In equilibrium firms are better-off compared to the case where none of them holds flexibility data.*

ii) If consumers are relatively differentiated in flexibility, then a Nash equilibrium exists, where none of the firms acquires customer data of any quality $k \geq 1$*. If* $k \geq 2$*, then the other Nash equilibrium exists, where both firms acquire flexibility data. In this equilibrium firms are worse-off compared to the case, where none of them holds flexibility data.*

Proof. See Appendix.

Our results show that firms' decisions to acquire customer data depend on consumer heterogeneity in flexibility and the quality of customer data. Precisely, with relatively homogeneous consumers in equilibrium firms acquire flexibility data of any quality. When consumers are relatively differentiated and customer data is very imprecise ($k = 1$), then in equilibrium none of the firms acquires flexibility data. With the improvement in data quality ($k \geq 2$) two equilibria coexist, where either both firms hold flexibility data or both of them refrain from that. It is instructive to compare our results with those in Sapi and Suleymanova (2013, in the following: SS), who analyze firms' incentives to acquire customer flexibility data in a similar setup, but in contrast to our article assume that firms' locations are exogenously given by $d_A = 0$ and $d_B = 1$. The main difference is that in SS in the case of relatively differentiated consumers firms necessarily end up in the prisoner's dilemma when data quality is sufficiently high ($k \geq 2$), because every firm has a unilateral incentive to acquire customer data making in the end both of them worse-off. While such an equilibrium exists also in our case, where firms can choose their locations, it is not unique and firms may end up in the equilibrium, where none of them holds flexibility data.

This difference is due to the fact that compared to SS in our analysis a firm never has a unilateral incentive to acquire flexibility data of any quality $k \geq 1$. In our case a firm acquires flexibility data in equilibrium only if the rival also does that. SS explain firms' data acquisition incentives through the interplay between the rent-extraction and competition effects. When data quality is low, competition effect dominates: using the new available data each firm targets aggressively the loyal consumers of the rival, which creates a downward pressure on the prices of the latter, such that both firms' profits decrease. With the improvement in data quality, however, the rent-extraction effect starts to dominate because each firm is able to extract more rents from its loyal consumers when it holds data of a better quality on their preferences. In our model where firms choose not only prices, but also their locations, the third effect plays a role, which is the location effect.

We explain now how the location effect influences the unilateral incentives of a firm to acquire flexibility data, which are responsible for the difference between our results and those of SS. Consider the turf of firm A and its incentives to gain flexibility data when the rival does not hold it. In equilibrium, on any address on a firms' turf both firms' prices are proportional either to the difference in firms' locations or to their sum (reduced by a consumer's address). Figure 2.2 depicts the difference between the equilibrium locations of firm B and firm A, when only firm A holds flexibility data of quality $k \geq 1$, and the difference between firms' locations when none of them holds customer data.[23] Figure 2.2 shows that although with the improvement in the quality of firm A's data ($k \geq 2$), the distance between firms' locations starts to increase, it never reaches the level of the case where both firms do not hold customer data.[24] This implies that other things being equal, equilibrium prices on any address on firm A's turf are smaller when it unilaterally acquires flexibility data compared to the case when it does not. Hence, the location effect contributes to the negative competition effect and reduces the unilateral incentives of a firm to acquire flexibility data. Intuitively, firms choose locations to protect their market shares, which intensifies competition compared to the case where locations are given exogenously.

[23] To calculate the former we used the equilibrium values $d_A^{d,As}(n)$ and $d_B^{d,As}(n)$ from case *i)* of Proposition 2.3. To calculate the latter we used the equilibrium values $d_A^d(k)$ and $d_B^d(k)$ from Proposition 2.1 and evaluated them at $k = 0$.

[24] Quite a similar result applies to the sum of firms' locations.

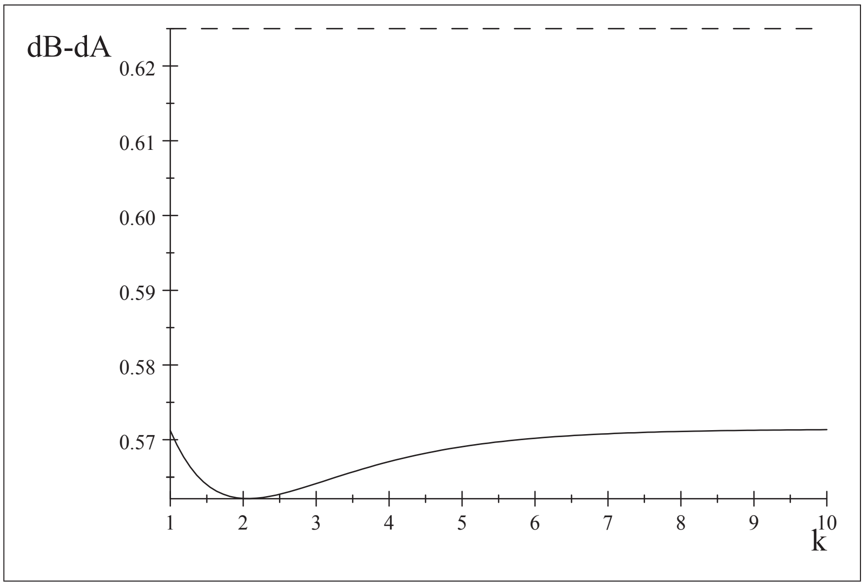

Figure 2.2. Difference in firms' locations when both firms do not hold flexibility data (dashed line) and when only firm A has data (solid line)

How do these new results change our conclusions on the equilibrium locations of the firms? We stated in Proposition 2.1 that when data on consumer flexibility becomes more precise, firms choose locations closer to the first-best ones, if consumers are relatively differentiated. However, as Proposition 2.4 shows, firms do not necessarily acquire flexibility data in that case and may end up in the equilibrium where none of them holds flexibility data, in which case equilibrium locations substantially depart from the first-best ones.

2.5 Conclusion

In this article we analyzed firms' location choices in a Hotelling model with two-dimensional consumer heterogeneity, along addresses and transport cost parameters (flexibility). We assumed that firms have perfect data on consumer locations, while the quality of customer flexibility data can be imperfect. Our results show that the optimality result of Lederer and Hurter (1986) holds even when firms' ability to practice third-degree price discrimination based on consumer transport cost parameters is imperfect, provided consumers are relatively homogeneous along that dimension. In that case under any location choices in equilibrium every firm serves all consumers on its turf, as in the case where firms have perfect data on consumer flexibility. In contrast, when consumers are relatively differentiated in flexibility, firms make socially suboptimal locations choices (unless the quality of customer flexibility data is perfect). However, with the improvement in the quality of customer flexibility data firms' location choices become closer to the socially optimal ones. This result supports the intuition of Hamilton and Thisse (1992) that to make socially optimal location choices firms need more flexibility in pricing. We also find that firms' ability to choose locations is crucial for their incentives to acquire flexibility data. Precisely, it reduces them through the location effect when consumers are relatively differentiated. Finally, we show that when firms hold flexibility data of asymmetric quality, efficiency in firms' location choices can be restored with the improvement in data quality only when consumers are relatively homogeneous. Our analysis is motivated by the availability of customer data, which allows firms to practice third-degree price discrimination based on both consumer characteristics relevant in spatial competition, addresses and transport cost parameters. It highlights the importance of consumer heterogeneity in flexibility and the quality of customer flexibility data for firms' location choices.

2.6 Appendix

In this Appendix we first provide the derivations of the formulas stated in the Section "Equilibrium Analysis" and then we present the proofs of the Lemmas and Propositions omitted in the text.

Derivations of the formulas. When consumers are relatively homogeneous, in equilibrium every firm charges on any address on its turf the highest price, which allows to monopolize a given flexibility segment. This price is proportional to the difference in the distances between the consumer and each of the firms. Following Lederer and Hurter (1986, in the following: LH) and using the results of Lemma 2.2 we can state the equilibrium prices of firm $i = A, B$ as

$$\begin{aligned} p_{im}^h(d_i, d_j; x, k) &= \underline{t}^m(k)\left(D_j(d_j; x) - D_i(d_i; x)\right) \text{ if } D_j(d_j; x) > D_i(d_i; x) \text{ and} \\ p_{im}^h(d_i, d_j; x, k) &= 0 \text{ if } D_j(d_j; x) \leq D_i(d_i; x), \end{aligned}$$

where $D_i(d_i; x) := |x - d_i|$. Then the equilibrium profit of firm i for given locations d_i and d_j can be expressed as

$$\begin{aligned} \Pi_i^h(d_i, d_j; k) &= \frac{1}{2^k}\sum_1^{2^k}\underline{t}^m(k)\int_{D_j(d_j;x)>D_i(d_i;x)}\left[D_j(d_j; x) - D_i(d_i; x)\right]dx \\ &= \frac{\underline{t}\left[\left(2^k+1\right)+l\left(2^k-1\right)\right]}{2^{k+1}}\int_{D_j(d_j;x)>D_i(d_i;x)}\left[D_j(d_j; x) - D_i(d_i; x)\right]dx, \end{aligned}$$

where $i \neq j$ and $j = A, B$. Similar to Lemma 4 in LH we can rewrite $\Pi_i^h(d_A, d_B; k)$ as

$$\begin{aligned} \frac{\Pi_i^h(d_i, d_j; k) \times 2^{k+1}}{\underline{t}\left[(2^k+1)+l\left(2^k-1\right)\right]} &= \int_{D_j(d_j;x)>D_i(d_i;x)} D_j(d_j; x)\,dx - \int_{D_j(d_j;x)>D_i(d_i;x)} D_i(d_i; x)\,dx \\ &= \int_{D_j(d_j;x)>D_i(d_i;x)} D_j(d_j; x)\,dx + \int_{D_j(d_j;x)\leq D_i(d_i;x)} D_j(d_j; x)\,dx \\ &\quad - \int_{D_j(d_j;x)\leq D_i(d_i;x)} D_j(d_j; x)\,dx - \int_{D_j(d_j;x)>D_i(d_i;x)} D_i(d_i; x)\,dx \\ &= \int_0^1 D_j(d_j; x)\,dx - \int_0^1 \min\left\{D_j(d_j; x), D_i(d_i; x)\right\}dx. \end{aligned}$$

Second-best locations solve the optimization problem

$$\min_{d_A, d_B}\left\{\frac{\left(\bar{t}+\underline{t}\right)}{2}\int_{D_A(d_A;x)<D_B(d_B;x)} D_A(d_A; x)\,dx + \frac{\left(\bar{t}+\underline{t}\right)}{2}\int_{D_B(d_B;x)\leq D_A(d_A;x)} D_B(d_B; x)\,dx\right\},$$

which is equivalent to the problem

$$\min_{d_A,d_B}\left\{\frac{\left(\bar{t}+\underline{t}\right)}{2}\int_0^1 \min\left\{D_A\left(d_A;x\right),D_B\left(d_B;x\right)\right\}dx\right\}.$$

Consider now the case of relatively differentiated consumers. As shown in Lemma 2.1, given any locations every firm serves on its own turf the more loyal consumers and the less loyal consumers on the rival's turf. In a similar way as above, following LH and using the results of Lemma 2.1 we can state the equilibrium prices of firm $i=A,B$ as

$$\begin{aligned}
p_{im}^d(d_i,d_j;x,k) &= \underline{t}^m\left(k\right)\left[D_j\left(d_j;x\right)-D_i\left(d_i;x\right)\right] \text{ if } D_j\left(d_j;x\right)>D_i\left(d_i;x\right) \text{ and } m\geq 2,\\
p_{im}^d(d_i,d_j;x,k) &= 2\bar{t}^m\left(k\right)\left[D_j\left(d_j;x\right)-D_i\left(d_i;x\right)\right]/3 \text{ if } D_j\left(d_j;x\right)>D_i\left(d_i;x\right) \text{ and } m=1,\\
p_{im}^d(d_i,d_j;x,k) &= 0 \text{ if } D_j\left(d_j;x\right)\leq D_i\left(d_i;x\right) \text{ and } m\geq 2,\\
p_{im}^d(d_i,d_j;x,k) &= \bar{t}^m\left(k\right)\left[D_i\left(d_j;x\right)-D_j\left(d_i;x\right)\right]/3 \text{ if } D_j\left(d_j;x\right)\leq D_i\left(d_i;x\right) \text{ and } m=1.
\end{aligned}$$

Then the profit of firm i for given locations d_i and d_j can be expressed as

$$\begin{aligned}
\Pi_i^d(d_i,d_j;k) &= \left[\frac{1}{2^k}\sum_2^{2^k}\underline{t}^m\left(k\right)+\frac{4\bar{t}}{9\times 2^{2k}}\right]\int_{D_j(d_j;x)>D_i(d_i;x)}\left[D_j\left(d_j;x\right)-D_i\left(d_i;x\right)\right]dx\\
&\quad+\frac{\bar{t}}{9\times 2^{2k}}\int_{D_j(d_j;x)\leq D_i(d_i;x)}\left[D_i\left(d_i;x\right)-D_j\left(d_j;x\right)\right]dx\\
&= \left[\frac{\bar{t}\left(2^k-1\right)}{2^{k+1}}+\frac{4\bar{t}}{9\times 2^{2k}}\right]\int_{D_j(d_j;x)>D_i(d_i;x)}\left[D_j\left(d_j;x\right)-D_i\left(d_i;x\right)\right]dx\\
&\quad+\frac{\bar{t}}{9\times 2^{2k}}\int_{D_j(d_j;x)\leq D_i(d_i;x)}\left[D_i\left(d_i;x\right)-D_j\left(d_j;x\right)\right]dx.
\end{aligned}$$

Similar to Lemma 4 in LH we can rewrite $\Pi_i^d(d_A,d_B;k)$ as

$$\begin{aligned}
\frac{\Pi_i^d(d_i,d_j;k)}{\left[\frac{\bar{t}(2^k-1)}{2^{k+1}}+\frac{4\bar{t}}{9\times 2^{2k}}\right]} &= \int_0^1 D_j\left(d_j;x\right)dx+\frac{2}{9\times 2^k\left(2^k-1\right)+8}\int_0^1 D_i\left(d_i;x\right)dx\\
&\quad-\left[1+\frac{2}{9\times 2^k\left(2^k-1\right)+8}\right]\int_0^1 \min\left\{D_j\left(d_j;x\right),D_i\left(d_i;x\right)\right\}dx.
\end{aligned}$$

Hence, when firm i chooses location d_i, its optimization problem is equivalent to

$$\min_{d_i}\left[\left(1+\alpha\left(k\right)\right)\frac{\bar{t}}{2}\int_0^1 \min\left\{D_j\left(d_j;x\right),D_i\left(d_i;x\right)\right\}dx-\alpha\left(k\right)\frac{\bar{t}}{2}\int_0^1 D_i\left(d_i;x\right)dx\right],$$

where $\alpha(k) = 2/\left[9 \times 2^k \left(2^k - 1\right) + 8\right]$.

Second-best locations minimize the transport costs

$$\left[\int\limits_{D_A(d_A;x)<D_B(d_B;x)} D_A(d_A;x)\,dx + \int\limits_{D_B(d_B;x)<D_A(d_A;x)} D_B(d_B;x)\,dx\right] \int\limits_{\frac{\bar{t}}{3\times 2^k}}^{\bar{t}} f_t t dt$$

$$+ \left[\int\limits_{D_A(d_A;x)>D_B(d_B;x)} D_A(d_A;x)\,dx + \int\limits_{D_B(d_B;x)>D_A(d_A;x)} D_B(d_B;x)\,dx\right] \int\limits_{0}^{\frac{\bar{t}}{3\times 2^k}} f_t t dt$$

$$= \frac{\bar{t}\left[1-\beta(k)\right]}{2}\int\limits_0^1 \min\left\{D_A(\cdot), D_B(\cdot)\right\} dx + \frac{\bar{t}\beta(k)}{2}\int\limits_0^1 \max\left\{D_A(\cdot), D_B(\cdot)\right\} dx,$$

where $\beta(k) = 1/\left[9 \times 2^{2k}\right]$.

We now derive the formulas necessary for the comparison of Anderson and de Palma (1988, in the following: AP) and the version of our model with relatively differentiated consumers. In that case the equilibrium prices on segment $m = 1$ (with relatively differentiated consumers) can be derived from Proposition 1 in AP. Consider, for example, the interval $x \leq d_A$.[25] We need to set $c_A = c_B = 0$ and replace F_1 with the demand of firm A on segment $m = 1$:

$$d_{A1}\left(p_{A1}(x) - p_{B1}(x), d_B - d_A, \bar{t}^1(k) \middle| x \leq d_A\right) := 1 - \frac{p_{A1}(x) - p_{B1}(x)}{\bar{t}^1(k)(d_B - d_A)}. \quad (2.8)$$

In AP (in the logit model) equilibrium locations depend on parameter $\mu \geq 0$, which is interpreted as a measure of consumer/product heterogeneity. In our case it makes sense to define μ as $\mu_1(k) := \bar{t}(d_B - d_A)/2^k$. Then from (2.8) we get that $|\partial d_{i1}(\cdot)/\partial p_{j1}(x)| = 1/\mu_1(k)$ $(i, j = A, B)$, such that with an increase in $\mu_1(k)$ firms' prices become less important in determining their demands. Parameter $\mu_1(k)$ is inversely related to the quality of customer flexibility data. If $k = 0$, then for any x, $\mu_1(k)$ gets its highest value of $\mu_1(0) = \bar{t}(d_B - d_A)$.

Proof of Lemma 2.1. As firms are symmetric, we will derive equilibrium on the two intervals on the turf of firm A.

i) Interval 1: $x \leq d_A$. Consider some $x \leq d_A$ and some m. The transport cost parameter of the indifferent consumer is

$$\widetilde{t}\left(d_A, d_B, p_{Am}(x), p_{Bm}(x); x \middle| x \leq d_A\right) := \frac{p_A(x) - p_B(x)}{d_B - d_A}, \text{ provided } \widetilde{t}(\cdot) \in \left[\underline{t}^m(k), \bar{t}^m(k)\right], \quad (2.9)$$

[25] On the other intervals one should proceed in a similar way.

such that consumers with $t \geq \widetilde{t}(\cdot)$ buy at firm A. Consider first $m = 1$. Maximization of firms' expected profits yields the best-response functions

$$p_{A1}\left(d_A, d_B, p_{B1}(x); x, k \middle| \, x \leq d_A\right) = \begin{cases} \frac{\bar{t}^1(k)(d_B - d_A) + p_{B1}(x)}{2} & \text{if} \quad p_{B1}(x) < \bar{t}^1(k)\left(d_B - d_A\right) \\ p_{B1}(x) & \text{if} \quad p_{B1}(x) \geq \bar{t}^1(k)\left(d_B - d_A\right) \end{cases} \tag{2.10}$$

and

$$p_{B1}\left(d_A, d_B, p_{A1}(x); x, k \middle| \, x \leq d_A\right) =$$
$$\begin{cases} \frac{p_{A1}(x)}{2} & \text{if} \quad p_{A1}(x) \leq 2\bar{t}^1(k)\left(d_B - d_A\right) \\ p_{A1}(x) - \bar{t}^1(k)\left(d_B - d_A\right) & \text{if} \quad p_{A1}(x) > 2\bar{t}^1(k)\left(d_B - d_A\right). \end{cases} \tag{2.11}$$

Given the best-response functions (2.10) and (2.11) we conclude that two types of equilibria are possible, where either firm A monopolizes segment m or where both firms serve consumers. Only the latter equilibrium exists where firms charge prices $p^d_{A1}(d_A, d_B; x, k) = 2\bar{t}\left(d_B - d_A\right) / \left(3 \times 2^k\right)$ and $p^d_{B1}(d_A, d_B; x, k) = \bar{t}\left(d_B - d_A\right) / \left(3 \times 2^k\right)$. Firm A serves consumers with $t \geq \bar{t}^1(k)/3$.

Consider now segments $m \geq 2$, where the best-response function of firm A is

$$p_{Am}(d_A, d_B, p_{Bm}(x); x, k | \, x \leq d_A) = p_{Bm}(x) + \underline{t}^m(k)\left(d_B - d_A\right). \tag{2.12}$$

As $\bar{t}^m(k) - 2\underline{t}^m(k) \leq 0$ for any $m \geq 2$, there is no $p_{Bm}(x) \geq 0$ under which it is optimal for firm A to share the market with firm B. (2.12) yields $p^d_{Bm}(d_A, d_B; x, k) = 0$. Indeed, assume that in equilibrium $p^d_{Bm}(d_A, d_B; x, k) > 0$ holds. Firm B gets in equilibrium the profit of zero, because (2.12) implies that firm B serves no consumers. But firm B can increase its profit through slightly decreasing its price. Hence, $p^d_{Bm}(d_A, d_B; x, k) = 0$ must hold. Firm A charges the price $p^d_{Am}(d_A, d_B; x, k) = \underline{t}^m(k)\left(d_B - d_A\right)$ and serves all consumers on segment m on address x.

On the interval $x \leq d_A$ firms realize profits

$$\begin{aligned} \Pi_A\left(d_A, d_B; k \middle| \, x \leq d_A\right) &= \int_0^{d_A} f_t \left(\left(\frac{2\bar{t}}{3 \times 2^k}\right)^2 \left(d_B - d_A\right) + \frac{\bar{t}(d_B - d_A)}{2^k} \sum_2^{2^k} \underline{t}^m(k) \right) dx \\ &= \frac{\bar{t}(d_B - d_A) d_A \left(9 \times 2^{2k} - 9 \times 2^k + 8\right)}{9 \times 2^{2k+1}} \end{aligned}$$

and

$$\Pi_B(d_A, d_B; k|\, x \leq d_A) = \tfrac{\bar{t}(d_B-d_A)d_A}{9\times 2^{2k}}.$$

Using symmetry we conclude on firms' profits on the interval $x \geq d_B$:

$$\begin{aligned}\Pi_A(d_A, d_B; k|\, x &\geq d_B) = \tfrac{\bar{t}(d_B-d_A)(1-d_B)}{9\times 2^{2k}} \text{ and}\\ \Pi_B(d_A, d_B; k|\, x &\geq d_B) = \tfrac{\bar{t}(d_B-d_A)(1-d_B)\left(9\times 2^{2k}-9\times 2^{k}+8\right)}{9\times 2^{2k+1}}\end{aligned}$$

ii) Interval 2: $d_A < x \leq (d_A + d_B)/2$. Consider some $d_A < x \leq (d_A + d_B)/2$ and segment m. The transport cost parameter of the indifferent consumer is

$$\widetilde{t}\left(d_A, d_B, p_{Am}(x), p_{Bm}(x); x|\, d_A < x \leq \tfrac{d_A+d_B}{2}\right) := \tfrac{p_A(x)-p_B(x)}{d_A+d_B-2x},$$

provided $\widetilde{t}(\cdot) \in \left[\underline{t}^m(k), \overline{t}^m(k)\right]$. Firm A serves consumers with $t \geq \widetilde{t}(\cdot)$. Consider first $m = 1$. Maximization of firms' profits yields the best-response functions

$$p_{A1}\left(d_A, d_B, p_{B1}(x); x, k|\, d_A < x \leq \tfrac{d_A+d_B}{2}\right) =$$

$$\begin{cases} \frac{\bar{t}^1(k)(d_A+d_B-2x)+p_{B1}(x)}{2} & \text{if} \quad p_{B1}(x) < \bar{t}^1(k)\left(d_A + d_B - 2x\right) \\ p_{B1}(x) & \text{if} \quad p_{B1}(x) \geq \bar{t}^1(k)\left(d_A + d_B - 2x\right) \end{cases} \tag{2.13}$$

and

$$p_{B1}\left(d_A, d_B, p_{A1}(x); x, k|\, d_A < x \leq \tfrac{d_A+d_B}{2}\right) =$$

$$\begin{cases} \frac{p_{A1}(x)}{2} & \text{if} \quad p_{A1}(x) \leq 2\bar{t}^1(k)\left(d_A + d_B - 2x\right) \\ p_{A1}(x) - \bar{t}^1(k)\left(d_A + d_B - 2x\right) & \text{if} \quad p_{A1}(x) > 2\bar{t}^1(k)\left(d_A + d_B - 2x\right). \end{cases} \tag{2.14}$$

Given the best-response functions (2.13) and (2.14) we conclude that two types of equilibria are possible where either firm A serves all consumers on segment m or shares it with the rival. Only the latter equilibrium exists, where firms charge prices

$$\begin{aligned}p^d_{A1}(d_A, d_B; x, k) &= 2\bar{t}\left(d_A + d_B - 2x\right)/\left(3\times 2^k\right) \text{ and}\\ p^d_{B1}(d_A, d_B; x, k) &= \bar{t}\left(d_A + d_B - 2x\right)/\left(3\times 2^k\right).\end{aligned}$$

On $m = 1$ firm A serves consumers with $t \geq \bar{t}/\left(3\times 2^k\right)$.

We now consider $m \geq 2$, where the best-response function of firm A takes the form

$$p_{Am}\left(d_A, d_B, p_{Bm}(x); x, k | \, d_A < x \leq \frac{d_A + d_B}{2}\right) = p_{Bm}(x,k) + \underline{t}^m(k)\,(d_A + d_B - 2x)\,,$$

such that firm A never finds it optimal to share segment m with firm B. Applying the logic described in part *i)* of the proof, we conclude that $p^d_{Bm}(d_A, d_B; x, k) = 0$ and $p^d_{Am}(d_A, d_B; x, k) = \underline{t}^m(k)\,(d_A + d_B - 2x)$. Firm A serves all consumers on any segment $m \geq 2$ on address x.

On the interval $d_A < x \leq (d_A + d_B)/2$ firms realize profits

$$\begin{aligned} &\Pi_A\left(d_A, d_B; k | \, d_A < x \leq \tfrac{d_A+d_B}{2}\right) \\ = &\int_{d_A}^{\frac{d_A+d_B}{2}} f_t \left[\left(\tfrac{2\bar{t}}{3\times 2^k}\right)^2 (d_A + d_B - 2x) + \tfrac{\bar{t}(d_A+d_B-2x)}{2^k} \sum_2^{2^k} \underline{t}^m(k)\right] dx \\ = &\tfrac{\bar{t}(d_B-d_A)^2\left(9\times 2^{2k} - 9\times 2^k + 8\right)}{9\times 2^{2k+3}} \text{ and} \end{aligned}$$

$$\Pi_B\left(d_A, d_B; k | \, d_A < x \leq \tfrac{d_A+d_B}{2}\right) = \tfrac{\bar{t}(d_B-d_A)^2}{9\times 2^{2k+2}}.$$

Using symmetry, we can conclude on firms' profits on the interval $(d_A + d_B)/2 < x \leq d_B$:

$$\begin{aligned} \Pi_A\left(d_A, d_B; k | \, \tfrac{d_A+d_B}{2} < x \leq d_B\right) &= \tfrac{\bar{t}(d_B-d_A)^2}{9\times 2^{2k+2}} \text{ and} \\ \Pi_B\left(d_A, d_B; k | \, \tfrac{d_A+d_B}{2} < x \leq d_B\right) &= \tfrac{\bar{t}(d_B-d_A)^2\left(9\times 2^{2k} - 9\times 2^k + 8\right)}{9\times 2^{2k+3}}. \end{aligned}$$

Summing up firms' profits on all the four intervals we get

$$\begin{aligned} \Pi^d_A(d_A, d_B; k) &= \tfrac{\bar{t}(d_B-d_A)\left(9\times 2^k\left(2^k-1\right)(3d_A+d_B)+2(11d_A+d_B)+8\right)}{9\times 2^{2k+3}} \text{ and} \\ \Pi^d_B(d_A, d_B; k) &= \tfrac{\bar{t}(d_B-d_A)\left(32-9\times 2^k(2^k-1)(3d_B+d_A-4)-2(11d_B+d_A)\right)}{9\times 2^{2k+3}}. \end{aligned}$$

Q.E.D.

Proof of Lemma 2.2. As firms are symmetric, we will only derive equilibrium on the two intervals on the turf of firm A and then conclude on the equilibrium on the turf of firm B.

i) Interval 1: $x \leq d_A$. Consider some $x \leq d_A$ and some $m \geq 1$. The transport cost parameter of the indifferent consumer is

$$\widetilde{t}\,(d_A, d_B, p_{Am}(x), p_{Bm}(x); x | \, x \leq d_A) := \tfrac{p_A(x)-p_B(x)}{d_B-d_A}, \text{ provided } \widetilde{t}(\cdot) \in \left[\underline{t}^m(k), \overline{t}^m(k)\right].$$

The best-response function of firm A is

$$p_{Am}(d_A, d_B, p_{Bm}(x); x, k|\, x \leq d_A) = p_{Bm}(x) + \underline{t}^m(k)\,(d_B - d_A)\,, \tag{2.15}$$

such that for any price of the rival firm A monopolizes segment m on address x. As $\bar{t}^m(k) - 2\underline{t}^m(k) \leq 0$ for any $m \geq 2$, there is no $p_{Bm}(x) \geq 0$ under which it is optimal for firm A to share the market with firm B. (2.15) yields $p^h_{Bm}(d_A, d_B; x, k) = 0$. Indeed, assume that in equilibrium $p^h_{Bm}(d_A, d_B; x, k) > 0$ holds. Firm B gets in equilibrium the profit of zero, because (2.15) implies that firm B serves no consumers. But firm B can increase its profit through slightly decreasing its price. Hence, $p^h_{Bm}(d_A, d_B; x, k) = 0$ must hold. Firm A charges the price $p^h_{Am}(d_A, d_B; x, k) = \underline{t}^m(k)\,(d_B - d_A)$ and serves all consumers on segment m on address x. Hence, $\Pi_B(d_A, d_B; k|\, x \leq d_A) = 0$ and the profit of firm A is computed as

$$\Pi_A\left(d_A, d_B; k|\, x \leq d_A\right) = \sum_1^{2^k} \underline{t}^m(k) \int_0^{d_A} \left(\tfrac{(d_B - d_A)}{2^k}\right) dx = \tfrac{d_A (d_B - d_A)\left(\bar{t}\left(2^k - 1\right) + \underline{t}\left(2^k + 1\right)\right)}{2^{k+1}}. \tag{2.16}$$

ii) Interval 2: $d_A < x \leq (d_A + d_B)/2$. Consider some $d_A < x \leq (d_A + d_B)/2$ and some $m \geq 1$. The transport cost parameter of the indifferent consumer is

$$\widetilde{t}\left(d_A, d_B, p_{Am}(x), p_{Bm}(x); x|\, d_A < x \leq \tfrac{d_A + d_B}{2}\right) := \tfrac{p_A(x) - p_B(x)}{d_A + d_B - 2x},$$

provided $\widetilde{t}(\cdot) \in \left[\underline{t}^m(k), \bar{t}^m(k)\right]$. Firm A serves consumers with $t \geq \widetilde{t}(\cdot)$. The best-response function of firm A is

$$p_{Am}\left(d_A, d_B, p_{Bm}(x); x, k|\, d_A < x \leq \tfrac{d_A + d_B}{2}\right) = p_{Bm}(x) + \underline{t}^m(k)\,(d_A + d_B - 2x)\,.$$

Following the logic applied in part *i)* of the proof we conclude that

$$\begin{aligned} p^h_{Bm}(x, k) &= 0 \text{ and} \\ p^h_{Am}(d_A, d_B; x, k) &= \underline{t}^m(k)\,(d_A + d_B - 2x)\,. \end{aligned}$$

Hence, $\Pi_B(d_A, d_B; k|\, d_A < x \leq (d_A + d_B)/2) = 0$ and the profit of firm A is computed as

$$\begin{aligned} \Pi_A\left(d_A, d_B; k|\, d_A < x \leq \tfrac{d_A + d_B}{2}\right) &= \sum_1^{2^k} \underline{t}^m(k) \int_{d_A}^{(d_A + d_B)/2} \left(\tfrac{d_A + d_B - 2x}{2^k}\right) dx \\ &= \tfrac{(d_B - d_A)^2 \left(\bar{t}\left(2^k - 1\right) + \underline{t}\left(2^k + 1\right)\right)}{2^{k+3}}. \end{aligned} \tag{2.17}$$

Summing up the profits (2.16) and (2.17) we get the profits of firm A as stated in the lemma. The profits of firm B are derived using symmetry. *Q.E.D.*

Proof of Lemma 2.3. We first derive first-best locations and prices. We will proceed in two steps. We will first derive first-best prices for any given locations and then will find first-best locations. Assume that firms are located at $d_A \leq d_B$. Social welfare is maximized when every consumer buys from the closer firm. Prices, which yield such a distribution of consumers between the firms are $p_{im}^{FB}(x), p_{jm}^{FB}(x) \geq 0$ such that

$$p_{im}^{FB}(x) \leq t|x - x_j| - t|x - x_i| + p_{jm}^{FB}(x) \text{ if } |x - x_i| \leq |x - x_j|. \tag{2.18}$$

Given (2.18), the first-best locations, d_A^{FB} and d_B^{FB}, have to minimize the transport costs

$$\begin{aligned} & TC^{FB}(d_A, d_B; k) \qquad (2.19) \\ = \; & \frac{(\bar{t}+t)}{2}\left(\int_0^{d_A}(d_A - x)\,dx + \int_{d_A}^{(d_A+d_B)/2}(x - d_A)dx + \int_{(d_A+d_B)/2}^{d_B}(d_B - x)\,dx + \int_{d_B}^{1}(x - d_B)\,dx\right), \end{aligned}$$

which yields the locations $d_A^{FB} = 1/4$ and $d_B^{FB} = 3/4$. Note that SOCs are fulfilled.

We now turn to the second-best locations. Here we have to distinguish between the cases of relatively homogeneous and differentiated consumers, because for any locations firms charge different prices in equilibrium depending on the case. We start with the case of relatively homogeneous consumers. Note that the equilibrium prices stated in Lemma 2.2 satisfy (2.18). Indeed, in equilibrium every consumer buys from the closer firm. Then second-best locations should also minimize (2.19), which yields $d_A^{SB,h} = 1/4$ and $d_B^{SB,h} = 3/4$.

We now consider the case of relatively differentiated consumers. According to Lemma 2.1, on its own turf each firm serves consumers with $t \geq \bar{t}/\left(3 \times 2^k\right)$, while consumers with $t < \bar{t}/\left(3 \times 2^k\right)$ switch to the rival. Then second-best locations have to minimize the transport costs

$$\begin{aligned}
& TC^{SB,d}\left(d_A,d_B;k\right) \\
= & \int_{\bar{t}/\left(3\times 2^k\right)}^{\bar{t}} \left(t f_t \int_0^{d_A} \left(d_A - x\right) dx \right) dt + \int_{\bar{t}/\left(3\times 2^k\right)}^{\bar{t}} \left(t f_t \int_{d_A}^{(d_A+d_B)/2} \left(x - d_A\right) dx \right) dt \\
& + \int_0^{\bar{t}/\left(3\times 2^k\right)} \left(t f_t \int_{(d_A+d_B)/2}^{d_B} \left(x - d_A\right) dx \right) dt + \int_0^{\bar{t}/\left(3\times 2^k\right)} \left(t f_t \int_{d_B}^{1} \left(x - d_A\right) dx \right) dt \\
& + \int_{\bar{t}/\left(3\times 2^k\right)}^{\bar{t}} \left(t f_t \int_{d_B}^{1} \left(x - d_B\right) dx \right) dt + \int_{\bar{t}/\left(3\times 2^k\right)}^{\bar{t}} \left(t f_t \int_{(d_A+d_B)/2}^{d_B} \left(d_B - x\right) dx \right) dt \\
& + \int_0^{\bar{t}/\left(3\times 2^k\right)} \left(t f_t \int_{d_A}^{(d_A+d_B)/2} \left(d_B - x\right) dx \right) dt + \int_0^{\bar{t}/\left(3\times 2^k\right)} \left(t f_t \int_0^{d_A} \left(d_B - x\right) dx \right) dt \\
= & -\tfrac{\bar{t}}{2^{2k}} \left(\tfrac{(d_A-d_B)^2}{36} + \tfrac{(d_A-d_B)}{18} - 3 \times 2^{2k-3} \left(\left(d_A\right)^2 + \left(d_B\right)^2 \right) \right) \\
& -\tfrac{\bar{t}}{2^{2k}} \left(-2^{2k-2} + 2^{2k-1} d_B + 2^{2k-2} d_A d_B \right),
\end{aligned}$$

which yields

$$d_A^{SB}\left(k\right) = \tfrac{9\times 2^{2k}}{36\times 2^{2k}-4} \text{ and } d_B^{SB}\left(k\right) = \tfrac{27\times 2^{2k}-4}{36\times 2^{2k}-4}.$$

SOCs are fulfilled. Note finally that $\lim_{k\to\infty} d_A^{SB}\left(k\right) = 1/4$ and $\lim_{k\to\infty} d_B^{SB}\left(k\right) = 3/4$. *Q.E.D.*

Proof of Proposition 2.1. *i)* Maximization of the profits

$$\begin{aligned}
\Pi_A^h(d_A,d_B;k) &= \tfrac{\underline{t}\left(\left(2^k+1\right)+l\left(2^k-1\right)\right)(d_B+3d_A)(d_B-d_A)}{2^{k+3}} \text{ and} \\
\Pi_B^h(d_A,d_B;k) &= \tfrac{\underline{t}\left(\left(2^k+1\right)+l\left(2^k-1\right)\right)(4-d_A-3d_B)(d_B-d_A)}{2^{k+3}}
\end{aligned}$$

with respect to d_A and d_B yields the FOCs: $d_B - 3d_A = 0$ and $d_A - 3d_B + 2 = 0$, respectively. Solving them simultaneously we get $d_A^h = 1/4$ and $d_B^h = 3/4$, such that the profit of firm $i = A, B$ is

$$\Pi_i^h\left(k\right) = \tfrac{3\underline{t}\left(\left(2^k+1\right)+l\left(2^k-1\right)\right)}{2^{k+5}}. \tag{2.20}$$

Note that SOCs are fulfilled. To prove that these locations constitute indeed the equilibrium, we have to prove that firm A does not have an incentive to choose a location $d_A \geq d_B^h = 3/4$. Due to symmetry, this would also imply that firm B does not have an incentive to choose

$d_B \leq d_A^h = 1/4$. If firm A locates at $d_A \geq d_B^h$, then according to Lemma 2.2 it realizes the profit

$$\Pi_B\left(d_B, d_A; k\right) = \frac{t\left(\left(2^k+1\right)+l\left(2^k-1\right)\right)(4-d_B-3d_A)(d_A-d_B)}{2^{k+3}}. \tag{2.21}$$

Maximizing (2.21) with respect to d_A yields the FOC: $d_A\left(d_B^h\right) = \left(d_B^h+2\right)/3 = 11/12$. Firm A realizes the profit

$$\Pi_B\left(\tfrac{3}{4}, \tfrac{11}{12}; k\right) = \frac{t\left(\left(2^k+1\right)+l\left(2^k-1\right)\right)(4-d_B-3d_A)(d_A-d_B)}{2^{k+3}} = \frac{t\left(\left(2^k+1\right)+l\left(2^k-1\right)\right)}{3\times 2^{k+5}}. \tag{2.22}$$

Comparing the profits (2.20) and (2.22) we conclude that

$$\Pi_i^h\left(k\right) - \Pi_B\left(\tfrac{3}{4}, \tfrac{11}{12}; k\right) = \frac{t\left(\left(2^k+1\right)+l\left(2^k-1\right)\right)}{3\times 2^{k+2}} > 0 \text{ for any } k \geq 0,$$

hence, firm A does not have an incentive to deviate to $d_A \geq d_B^h$. We conclude that the locations $d_A^h = 1/4$ and $d_B^h = 3/4$ constitute indeed the equilibrium.

ii) Maximization of the profits

$$\begin{aligned} \Pi_A^d(d_A, d_B; k) &= \frac{\bar{t}(d_B-d_A)\left(9\times 2^k\left(2^k-1\right)(3d_A+d_B)+2(11d_A+d_B)+8\right)}{9\times 2^{2k+3}} \text{ and} \\ \Pi_B^d(d_A, d_B; k) &= \frac{\bar{t}(d_B-d_A)\left(-9\times 2^k(2^k-1)(3d_B+d_A-4)-2(11d_B+d_A)+32\right)}{9\times 2^{2k+3}} \end{aligned}$$

with respect to d_A and d_B yields the FOCs

$$\begin{aligned} d_A\left(d_B; k\right) &= \frac{9\times 2^k d_B\left(2^k-1\right)+10d_B-4}{27\times 2^{2k}-27\times 2^k+22} \text{ and} \\ d_B\left(d_A; k\right) &= \frac{9\times 2^k\left(2^k-1\right)(d_A+2)+10d_A+16}{27\times 2^{2k}-27\times 2^k+22}, \end{aligned}$$

respectively. Solving FOCs simultaneously we get the locations

$$\begin{aligned} d_A^d\left(k\right) &= \frac{9\times 2^{2k}-9\times 2^k+6}{36\times 2^{2k}-36\times 2^k+32} \text{ and} \\ d_B^d\left(k\right) &= \frac{27\times 2^{2k}-27\times 2^k+26}{36\times 2^{2k}-36\times 2^k+32}, \end{aligned} \tag{2.23}$$

such that firm $i = A, B$ realizes the profit

$$\Pi_i^d(k) = \frac{\bar{t}\left(9\times 2^{2k}-9\times 2^k+10\right)^2\left(27\times 2^{2k}-27\times 2^k+22\right)}{9\times 2^{2k+5}\left(9\times 2^{2k}-9\times 2^k+8\right)^2}.$$

Note that the SOCs are also fulfilled. To prove that the locations $d_A^d\left(k\right)$ and $d_B^d\left(k\right)$ constitute indeed the equilibrium, we have to show that firm A does not have an incentive to locate at $d_A \geq d_B^d(k)$. As firms are symmetric, firm B then does not have an incentive to locate at $d_B \leq d_A^d(k)$ either. If firm A chooses $d_A \geq d_B^d(k)$, then it realizes the profit

$$\begin{aligned}&\Pi_B^d\left(d_B^d(k), d_A; k\right) \\ =\ &\frac{\bar{t}\left(d_A - d_B^d(k)\right)\left(-9\times 2^k(2^k-1)\left(3d_A + d_B^d(k) - 4\right) - 2\left(11 d_A + d_B^d(k)\right) + 32\right)}{9\times 2^{2k+3}}.\end{aligned} \tag{2.24}$$

Maximization of (2.24) with respect to d_A yields

$$\begin{aligned} d_A\left(d_B^d(k); k\right) &= \frac{9\times 2^k\left(2^k-1\right)\left(d_B^d(k)+2\right)+10 d_B^d(k)+16}{27\times 2^{2k}-27\times 2^k+22} \\ &= \frac{2547\times 2^{2k}-1782\times 2^{3k}+891\times 2^{4k}-1656\times 2^k+772}{(36\times 2^{2k}-36\times 2^k+32)(27\times 2^{2k}-27\times 2^k+22)}, \end{aligned}$$

such that firm A realizes the profit

$$\Pi_B\left(d_B^d(k), d_A\left(d_B^d(k); k\right); k\right) = \frac{\bar{t}\left(9\times 2^{2k}-9\times 2^k+10\right)^4}{9\times 2^{2k+5}(9\times 2^{2k}-9\times 2^k+8)^2(27\times 2^{2k}-27\times 2^k+22)}.$$

Comparison of the profits $\Pi_i^d(k)$ and $\Pi_B\left(d_B^d(k), d_A\left(d_B^d(k); k\right); k\right)$ yields

$$\begin{aligned}&\Pi_i^d(k) - \Pi_B\left(d_B^d(k), d_A\left(d_B^d(k); k\right); k\right) \\ =\ &\frac{\left(3\times 2^{2k}-3\times 2^k+2\right)\left(9\times 2^{2k}-9\times 2^k+10\right)^2}{3\times 2^{2k+2}\left(27\times 2^{2k}-27\times 2^k+22\right)(9\times 2^{2k}-9\times 2^k+8)} > 0 \text{ for any } k,\end{aligned}$$

hence, firm A does not have an incentive to locate at $d_A \geq d_B^d(k)$. We conclude that the locations $d_A^d(k)$ and $d_B^d(k)$ in (2.23) constitute indeed the equilibrium. *Q.E.D.*

Proof of Proposition 2.2. This proof consists of the two parts. In the first part we consider the second stage of the game and derive the equilibrium prices for any location choices. In the second part we consider the first stage of the game and derive the equilibrium locations of the firms. We also derive first-best locations and compare them with the equilibrium locations.

Part 1. To prove the proposition we may rely on the results of Proposition 2 in Sapi and Suleymanova (2013, in the following: SS), which states the equilibrium prices and consumer demands in the case when consumers are relatively homogeneous, one firm (A) has customer data on consumer flexibility of a given quality, the other does not, while both hold data on consumer brand preferences. Our problem can also be described in a similar way, because on each flexibility segment identified by firm B, firm A can identify further 2^n flexibility segments, so that there firm A holds flexibility data of quality n, while firm B does not hold any. Moreover, any segment m identified by firm B is a segment with relatively homogeneous consumers for any $k \geq 0$. Finally, as in SS and in our model both firms have perfect data on consumer brand preferences, we can apply their results by making only the necessary corrections related to the difference in firms' locations, which are always set at $d_A = 0$ and $d_B = 1$ in SS. Precisely, the equilibrium prices, which are always proportional to the module of the difference in the distances of a consumer to the two firms, have to be changed respectively.

SS show that on any address on its turf firm A optimally charges the price, at which it gains all consumers, independently of the price charged by firm B. As a result, in equilibrium firm B cannot do better than charging the price of zero. The equilibrium price of firm A on any address x on its turf ($x \leq (d_A + d_B)/2$) and any flexibility segment m identified by firm A is $p_{Am}(d_A, d_B; x, k) = \underline{t}^m (k) (|d_B - x| - |d_A - x|)$. Integrating over all addresses on the turf of firm A and summing up over all flexibility segments it can identify we arrive at the profits of firm A on its turf:

$$\Pi_A(d_A, d_B; k, n|\, x \leq (d_A+d_B)/2) = \frac{\underline{t}\left[\left(2^{k+n}+1\right)+l\left(2^{k+n}-1\right)\right](d_B+3d_A)(d_B-d_A)}{2^{k+n+3}}, \tag{2.25}$$

while firm B realizes a profit of zero on firm A's turf.

We now turn to the turf of firm B. Consider some segment m identified by firm B. SS show that the equilibrium there depends on whether the ratio $\bar{t}^m/\underline{t}^m$ is larger or smaller than 3/2. We next have to distinguish between the cases $k = 0$ and $k \geq 1$. Note that for any $k \geq 1$ and any segment $m = 1, 2, ..., 2^k$ we have that $\bar{t}^m/\underline{t}^m \leq 3/2$. Indeed, we have by definition that $\bar{t}^m/\underline{t}^m = \left[\underline{t} + \left(\bar{t} - \underline{t}\right) m/2^k\right] / \left[\underline{t} + \left(\bar{t} - \underline{t}\right)(m-1)/2^k\right]$. Then $\bar{t}^m/\underline{t}^m \leq 3/2$ holds if and only if

$$(l-1)(3-m) \leq 2^k \tag{2.26}$$

holds. If $k = 1$, then m can take only two values: 1 and 2. In both cases (2.26) holds (note that $l - 1 \leq 1$). If $k = 2$, then m can take four values: 1, 2, 3 and 4. Again (2.26) holds in all the cases. We can conclude that (2.26) holds for all $k \geq 3$ too. It is straightforward that (2.26) holds for any $m \geq 3$ as $(l-1)(3-m) \leq 0$ then and $2^k > 0$. And if $m \leq 2$, then (2.26) holds for any $k \geq 3$ just because it holds for $k \leq 2$ (as we have just showed) and 2^k increases in k.

We can now turn to the results of SS, which imply that on any flexibility segment with $\bar{t}^m/\underline{t}^m < 3/2$ firm B charges a price at which it gains all consumers there: $p_{Bm}(d_A, d_B; x, k+n) = \underline{t}^m (k+n)(|d_A - x| - |d_B - x|)$. Then firm A gets the profit of zero on the turf of firm B, while the profit of firm B is

$$\Pi_B(\, d_A, d_B; k, n|\, x > (d_A{+}d_B)/2) = \frac{\underline{t}\left(2^k+1+l\left(2^k-1\right)\right)(4-d_A-3d_B)(d_B-d_A)}{2^{k+3}}, \tag{2.27}$$

which we get by integrating the profit realized on one address and one flexibility segment over all addressees on the turf of firm B and summing up over all flexibility segments. Note that (2.27) is the total profit of firm B and (2.25) is the total profit of firm A if $k \geq 1$.

We finally consider the case $k = 0$, where the only flexibility segment identified by firm B includes all $t \in \left[\underline{t}, \bar{t}\right]$. If $l \leq 2/3$, then we get the same equilibrium as described above, such that the formula (2.27) again describes the profit of firm B on its turf and firms' total profits are given by the same functions as in the case $k \geq 1$. If $l > 2/3$, then as SS show, the profit depends on the quality of firm A's data, n (as $k = 0$). By changing the equilibrium prices in SS to account for varying locations of the firms, we can state the equilibrium for our case. If $n < \log_2((l-1)/(l-3/2))$, then firm B charges the price $p_B(d_A, d_B; x, n) = \left(2\bar{t} - \underline{t}\right)(|d_A - x| - |d_B - x|)/2$ on any address on its turf and serves consumers with $t \geq \left(2\bar{t} + \underline{t}\right)/4$. Firm A charges a positive price only on $m = 1$: $p_{A1}(d_A, d_B; x, n) = \left(2\bar{t} - 3\underline{t}\right)(|d_A - x| - |d_B - x|)/4$. We can then calculate firms' profits on the turf of firm B as

$$\Pi_A(\, d_A, d_B; 0, n|\, x \;\;>\;\; (d_A{+}d_B)/2) = \frac{\underline{t}(2l-3)^2(d_B-d_A)(4-d_A-3d_B)}{64(l-1)} \text{ and}$$
$$\Pi_B(\, d_A, d_B; 0, n|\, x \;\;>\;\; (d_A{+}d_B)/2) = \frac{\underline{t}(2l-1)^2(d_B-d_A)(4-d_A-3d_B)}{32(l-1)}.$$

Summing up $\Pi_A(\cdot|\, x > (d_A + d_B)/2)$ with (2.25), we get the total profit of firm A on both turfs together, while the profit of firm B is generated only on its own turf:

$$\begin{aligned} \frac{\Pi_A(\cdot)}{\underline{t}} &= \frac{(2l-3)^2(d_B-d_A)(4-d_A-3d_B)}{64(l-1)} + \frac{(2^n+1+l(2^n-1))(d_B+3d_A)(d_B-d_A)}{2^{n+3}} \\ \frac{\Pi_B(\cdot)}{\underline{t}} &= \frac{(2l-1)^2(d_B-d_A)(4-d_A-3d_B)}{32(l-1)}. \end{aligned} \tag{2.28}$$

If in contrast, $n \geq \log_2 [(l-1)(l-3/2)]$, then on any address on its turf firm B charges the price $p_B(d_A, d_B; x, n) = \left(\underline{t} + (\bar{t} - \underline{t})/2^n\right)(|d_A - x| - |d_B - x|)$ and serves consumers with $t \geq \underline{t} + (\bar{t} - \underline{t})/2^{n+1}$. Firm A charges a positive price only on segment $m = 1$: $p_{A1}(d_A, d_B; x, n) = (\bar{t} - \underline{t})(|d_A - x| - |d_B - x|)/2^{n+1}$. We can then calculate firms' profits on the turf of firm B as

$$\begin{aligned} \Pi_A(\, d_A, d_B; 0, n| \, x &> (d_A+d_B)/2) = \frac{\underline{t}(l-1)(d_B-d_A)(4-d_A-3d_B)}{2^{2n+4}} \\ \Pi_B(\, d_A, d_B; 0, n| \, x &> (d_A+d_B)/2) = \left(1-\frac{1}{2^{n+1}}\right)\left(1+\frac{l-1}{2^n}\right)\frac{\underline{t}(d_B-d_A)(4-d_A-3d_B)}{4}. \end{aligned}$$

Summing up these profits with the profits realized on the turf of firm A, we get firms' total profits:

$$\begin{aligned} \Pi_A(\cdot) &= \frac{\underline{t}(l-1)(d_B-d_A)(4-d_A-3d_B)}{2^{2n+4}} + \frac{\underline{t}((2^n+1)+l(2^n-1))(d_B+3d_A)(d_B-d_A)}{2^{n+3}} \\ \Pi_B(\cdot) &= \left(1-\frac{1}{2^{n+1}}\right)\left(1+\frac{l-1}{2^n}\right)\frac{\underline{t}(d_B-d_A)(4-d_A-3d_B)}{4} \end{aligned} \tag{2.29}$$

Part 2. We now turn to the first stage of the game and derive the equilibrium locations depending on parameters k, n and l. Assume first that $k \geq 1$. Maximizing the profits (2.25) and (2.27) with respect to d_A and d_B, respectively, and solving the two FOCs simultaneously we get the equilibrium locations: $d_A^{h,As}(k,n,l) = 1/4$ and $d_B^{h,As}(k,n,l) = 3/4$. Note that these are also the equilibrium locations in the case $k = 0$ and $l \leq 2/3$. Note finally that SOCs are fulfilled and none of the firms has an incentive to change its location such that $d_A \geq d_B$.

a) Assume now that $k = 0$, $l > 2/3$ and $n < \log_2((l-1)/(l-3/2))$. Maximizing the profits of firm A and firm B in (2.28) with respect to d_A and d_B, respectively, and solving the two FOCs simultaneously we get the equilibrium locations:

$$\begin{aligned} d_A^{h,As}(0,n,l) &= \frac{-8l-12\times 2^n l+4l^2+13\times 2^n+4}{-32l-12\times 2^n l-12\times 2^n l^2+16l^2+25\times 2^n+16} \text{ and} \\ d_B^{h,As}(0,n,l) &= \frac{-24l-12\times 2^n l-8\times 2^n l^2+12l^2+21\times 2^n+12}{-32l-12\times 2^n l-12\times 2^n l^2+16l^2+25\times 2^n+16}. \end{aligned} \tag{2.30}$$

Note that the SOCs are satisfied and any of the firms does not have an incentive to change its location such that $d_A \geq d_B$. We now analyze how the equilibrium locations change with the improvement in the quality of firm A's data. Taking the derivatives of $d_A^{h,As}(\cdot)$ and $d_B^{h,As}(\cdot)$ in (2.30) with respect to n we get

$$\begin{aligned}\frac{\partial d_A^{h,As}(\cdot)}{\partial n} &= \frac{3\times 2^{n+2}(\ln 2)\left(2l^2-5l+3\right)^2}{\left(16l^2-12\times 2^n l-12\times 2^n l^2-32l+25\times 2^n+16\right)^2} > 0, \text{ if } n \geq 1 \text{ and } l > 2/3,\\ \frac{\partial d_B^{h,As}(\cdot)}{\partial n} &= \frac{2^{n+2}(\ln 2)\left(2l^2-5l+3\right)^2}{\left(16l^2-12\times 2^n l-12\times 2^n l^2-32l+25\times 2^n+16\right)^2} > 0, \text{ if } n \geq 1 \text{ and } l > 2/3,\end{aligned}$$

which implies that with the improvement in data quality of firm A both firms move to the right. Consider now the difference $d_B^{h,As}(\cdot) - d_A^{h,As}(\cdot)$. We take the derivative of the latter difference with respect to n for any $n \geq 1$ and any $l > 2/3$,

$$\frac{\partial\left(d_B^{h,As}(\cdot)-d_A^{h,As}(\cdot)\right)}{\partial n} = -\frac{2^{n+3}(\ln 2)\left(2l^2-5l+3\right)^2}{\left(16l^2-12\times 2^n l-12\times 2^n l^2-32l+25\times 2^n+16\right)^2} < 0,$$

such that when the quality of firm A's data becomes better, firms locate closer to each other. We now analyze, which values the equilibrium locations take when $n \to \log_2\left((l-1)/(l-3/2)\right)$. We get that

$$\lim_{n\to\log_2((l-1)/(l-3/2))} d_A^{h,As}(\cdot) = \tfrac{4l^2-22l+19}{4l^2-52l+49}. \tag{2.31}$$

Let $f_1(l)$ denote the RHS of (2.31). Note that on the interval $l \in (3/2, 2]$, $f_1(l)$ is a monotonically decreasing function with $\lim_{l\to 3/2} f_1(l) = 1/4$ and $f_1(2) = 3/13 \approx 0.23$. Since $d_A^{h,As}(\cdot)$ monotonically increases in n and since for any $l \in (3/2, 2]$ it holds that $f_1(l) < d_A^{FB} = 1/4$, we conclude that for any $n < \log_2\left((l-1)/(l-3/2)\right)$ and any $l \in (3/2, 2]$ we have that $d_A^{h,As}(\cdot) < d_A^{FB}$. We next turn to the location of firm B. We get that

$$\lim_{n\to\log_2((l-1)/(l-3/2))} d_B^{h,As}(\cdot) = \tfrac{4l^2-42l+39}{4l^2-52l+49}. \tag{2.32}$$

Let $f_2(l)$ denote the RHS of (2.32). Note that on the interval $l \in (3/2, 2]$, $f_2(l)$ is a monotonically decreasing function with $\lim_{l\to 3/2} f_2(l) = 3/4$ and $f_2(2) = 29/39 \approx 0.74$. Since $d_B^{h,As}(\cdot)$ monotonically increases in n and since for any $l \in (3/2, 2]$ it holds that $f_2(l) < d_B^{FB} = 3/4$, we conclude that for any $n < \log_2\left((l-1)/(l-3/2)\right)$ and any $l \in (3/2, 2]$ we have that $d_B^{h,As}(\cdot) < d_B^{FB}$.

We next show that when consumers become more differentiated (l increases), the location of firm A moves to the left. Taking the derivative of $d_A^{h,As}(\cdot)$ with respect to l we get:

$$\frac{\partial d_A^{h,As}(\cdot)}{\partial l} = \frac{3\times 2^{n+3}(2l-3)\left(l-3\times 2^n l+2^{n+1}-1\right)}{(16l^2-12\times 2^n l-12\times 2^n l^2-32l+25\times 2^n+16)^2} < 0. \tag{2.33}$$

Note that $2l-3>0$ for any $l\in(3/2,2]$. To determine the sign of the expression $g_1(l) := l-3\times 2^n l+2^{n+1}-1$, note that for any $n\geq 1$ it holds that $2^n\geq 2$ and for all $l\in(3/2,2]$ it holds that $(l-1)/(3l-2)<1$, which implies that $(l-1)/(3l-2)<2^n$, such that $g_1(l)<0$ yielding $\partial d_A^{h,As}(\cdot)/\partial l<0$ as stated in (2.33). Consider now the address of firm B. We get that

$$\frac{\partial d_B^{h,As}(\cdot)}{\partial l} = \frac{2^{n+3}(2l-3)\left(l-3\times 2^n l+2^{n+1}-1\right)}{(16l^2-12\times 2^n l-12\times 2^n l^2-32l+25\times 2^n+16)^2} < 0,$$

where the sign of the derivative follows from the same considerations as above. We conclude that when consumers become more differentiated, firm B moves to the left. We finally analyze how the the difference $d_B^{h,As}(\cdot)-d_A^{h,As}(\cdot)$ responds to a change in l. We get that

$$\frac{\partial\left(d_B^{h,As}(\cdot)-d_A^{h,As}(\cdot)\right)}{\partial l} = -\frac{2^{n+4}(2l-3)\left(l-3\times 2^n l+2^{n+1}-1\right)}{(16l^2-12\times 2^n l-12\times 2^n l^2-32l+25\times 2^n+16)^2} > 0,$$

such that firms locate farther apart from each other when consumers become more differentiated.

b) Assume now that $k=0$, $l>2/3$ and $n\geq\log_2\left((l-1)(l-3/2)\right)$. Taking the derivative of $\Pi_A(\cdot)$ and $\Pi_B(\cdot)$ in (2.29) with respect to d_A and d_B, respectively, and solving simultaneously the two FOCs we get the equilibrium locations:

$$\begin{aligned} d_A^{h,As}(0,n,l) &= \frac{-l+2^{2n}-2^n l+2^{2n}l+2^n+1}{-l+4\times 2^{2n}-4\times 2^n l+4\times 2^{2n}l+4\times 2^n+1} \text{ and} \\ d_B^{h,As}(0,n,l) &= \frac{-l+3\times 2^{2n}-3\times 2^n l+3\times 2^{2n}l+3\times 2^n+1}{-l+4\times 2^{2n}-4\times 2^n l+4\times 2^{2n}l+4\times 2^n+1}. \end{aligned}$$

Note that SOCs are satisfied and any of the firms does not have an incentive to change its location such that $d_A\geq d_B$. All the dependencies of the optimal locations, which we derived in a) hold also in the case when $n\geq\log_2\left((l-1)(l-3/2)\right)$, which can be shown in a similar way as above. We only have to mention that for all $l\in(3/2,2]$ it holds that $\lim_{n\to\infty} d_A^{h,As}(\cdot)=1/4$ and $\lim_{n\to\infty} d_B^{h,As}(\cdot)=3/4$. *Q.E.D.*

Proof of Proposition 2.3. This proof consists of the two parts. In the first part we consider the second stage of the game and derive the equilibrium prices for any location choices. In the second part we consider the first stage of the game and derive the equilibrium locations of the firms. We also derive first-best locations and compare them with the equilibrium locations. In this proof we will sometimes denote a flexibility segment with two numbers, m, s, where m is the number of the segment identified by firm B and s is the number of the segment identified by firm A on segment m. In most of the cases (where it does not cause any misunderstanding) we will denote a segment with only one number.

Part 1. We start with the turf of firm A by considering some address there. It makes sense to consider separately each segment m identified by firm B. Note that on any such segment firm B holds data on consumer flexibility of quality zero, while the quality of firm A's data is n. Note next that any $m \geq 2$ is a segment with relatively homogeneous consumers. Then we can refer to the proof of Proposition 2.2 where we showed that when consumers are relatively homogeneous, firm A serves all consumers on any address on its turf. On each such segment firm A identifies further 2^n flexibility segments, where it charges a price, which makes the most flexible consumer (with the lowest transport cost parameter) indifferent between the two firms. Consider now segment $m = 1$, which is a segment with relatively differentiated consumers (as $\underline{t}^1 = 0$). To get the equilibrium prices on that segment we can refer to the results stated in Proposition 3 of Sapi and Suleymanova (2013, in the following: SS), which describe the optimal prices of the firms when consumers are relatively differentiated, firm A holds flexibility data of a given quality, while firm B does not. These results can be applied for $m = 1$, because firm A can identify there 2^n segments, while firm B none. Also, all the equilibrium prices on a given consumer address are proportional to the module of the difference in the distances between a consumer and each of the firms. Hence, in our case the prices only have to be corrected to take into account different firm locations (as in SS firm locations are fixed at $d_A = 0$ and $d_B = 1$). SS show that on any segment identified by firm A, starting from the second segment, it charges a price, which makes the mots flexible consumer (with the lowest transport cost parameter) indifferent between buying at the two firms and, hence, serves all consumers there. On the first segment (identified by firm A), firms charge the prices, such that firm A serves consumers with $t \geq \bar{t}/2^{k+n+2}$. Combining the equilibrium results on all the segments identified by firm B, we conclude that our problem yields the same results on the turf of firm A as if firm B had no customer data at all, while firm A had data of quality $k+n$. But then we can use the profit formulas derived in SS (replacing k with $k+n$ and correcting for the equilibrium prices in a way described above). As a result, we get the following profits on the turf of firm A[26]:

$$\begin{aligned} \Pi_A\left(\cdot|\, x\leq\tfrac{d_A+d_B}{2}\right) &= \bar{t}\,(d_B-d_A)\,(3d_A+d_B)\left(\tfrac{2^{k+n}-1}{2^{k+n+3}}+\tfrac{2^{k+n+3}+1}{2^{2(k+n)+6}}\right), \\ \Pi_B\left(\cdot|\, x\leq\tfrac{d_A+d_B}{2}\right) &= \tfrac{\bar{t}(d_B-d_A)(3d_A+d_B)}{2^{2(k+n)+5}}. \end{aligned} \tag{2.34}$$

We now turn to the turf of firm B. Consider first segments $m \geq 3$ identified by firm B (provided $k \geq 2$) with $\underline{t}^m(k) = \bar{t}(m-1)/2^k$ and $\bar{t}^m(k) = \bar{t}m/2^k$. Note that for any $k \geq 2$ and any $m \geq 3$ it holds that $\bar{t}^m(k)/\underline{t}^m(k) \leq 3/2$. Then we can apply the results of Proposition 2.2, which show that in that case firm B on any such segment charges the price, which makes the most flexible consumer on the segment indifferent between the two firms and gains all consumers on m. In contrast, firm A charges a price of zero on any $m \geq 3$ and realizes no profits there. Then the profits of firm B on segment $m \geq 3$ are given by

$$\Pi_{Bm\geq 3}(\,d_A, d_B; k, n|\, x > (d_A+d_B)/2) = \frac{\bar{t}\left(2^k+1\right)\left(2^k-2\right)(d_B-d_A)(4-d_A-3d_B)}{2^{2k+3}}. \tag{2.35}$$

Consider now segment $m = 2$ identified by firm B (provided $k \geq 1$) with $\underline{t}^2(k) = \bar{t}/2^k$ and $\bar{t}^2(k) = 2\bar{t}/2^k$. Note that for any $k \geq 1$ it holds that $\bar{t}^2(k)/\underline{t}^2(k) = 2$, such that $m = 2$ is a segment with relatively homogeneous consumers. On this segment firm A can identify further 2^k segments, while firm B none. Hence, we can apply the results of Proposition 2.2 for the case $l > 2/3$ (and only replace $\underline{t}$ with $\underline{t}^2(k)$ and $\bar{t}$ with $\bar{t}^2(k)$ in the formulas).

[26] Precisely, the profit formulas in SS have to be first divided by $\int_0^{1/2}(1-2x)\,dx$, which is the integral over the difference in the distance between a consumer with address x on firm A's turf and firm B and the distance between that consumer and firm A. In the second step the profits have to be multiplied with the sum of the two integrals, $\int_0^{d_A}(d_B-d_A)\,dx$ and $\int_{d_A}^{(d_A+d_B)/2}(d_A+d_B-2x)\,dx$, which is equivalent to the above integral when the locations of firm A and firm B are d_A and d_B instead of 0 and 1, respectively.

They state that the equilibrium depends on the quality of firm A's data, n. Note that the critical value of firm A's quality, $\log_2[(l-1)(l-3/2)]$, in our case takes the value $\log_2 2 = 1$. As we assume that $n \geq 1$, then only the case of Proposition 2.2 with $n \geq \log_2[(l-1)(l-3/2)]$ applies. In that case firm B charges the price $p_{B2}(d_A, d_B; x, k, n) = \bar{t}(1+1/2^n)(|d_A - x| - |d_B - x|)/2^k$ and serves consumers with $t \geq \bar{t}\left(1+1/2^{n+1}\right)/2^k$ on $m=2$. Firm A charges a positive price only on the first segment it can identify on $m=2$: $p_{A2,1}(d_A, d_B; x, k, n) = \bar{t}(|d_A - x| - |d_B - x|)/2^{n+k+1}$. Then on $m=2$ firms realize profits:

$$\begin{aligned} \frac{\Pi_{A2}(d_A,d_B;k,n|x>(d_A+d_B)/2)}{(d_B-d_A)(4-d_A-3d_B)} &= \frac{\bar{t}}{2^{2(k+n+2)}} \text{ and} \\ \frac{\Pi_{B2}(d_A,d_B;k,n|x>(d_A+d_B)/2)}{(d_B-d_A)(4-d_A-3d_B)} &= \frac{\bar{t}}{2^{2(k+1)}}\left(1+\frac{1}{2^n}\right)\left(1-\frac{1}{2^{n+1}}\right). \end{aligned} \tag{2.36}$$

Consider finally segment $m=1$ (identified by firm B) with $\underline{t}^1(k) = 0$ and $\bar{t}^1(k) = \bar{t}/2^k$, which is a segment with relatively differentiated consumers where firm A holds data of quality $n \geq 1$, while firm B does not hold any flexibility data. To derive the equilibrium profits on this segment we can again refer to the results stated in Proposition 3 of SS. We can use their profit formulas, in which we have to replace $\bar{t}$ with $\bar{t}^1 = \bar{t}/2^k$, divide with $\int_{1/2}^{1}(2x-1)\,dx$ and multiply with the sum of $\int_{(d_A+d_B)/2}^{d_B}(2x-d_A-d_B)\,dx$ and $\int_{d_B}^{1}(d_B-d_A)\,dx$, similarly as we did in the case of firm A's turf. As a result, we get

$$\begin{aligned} \frac{\Pi_{A1}(d_A,d_B;k,n|x>(d_A+d_B)/2)}{(d_B-d_A)(4-d_A-3d_B)} &= \frac{\bar{t}\left(2^{2n+2}-2^{n+2}+5\right)}{2^{2n+k+7}} \text{ and} \\ \frac{\Pi_{B1}(d_A,d_B;k,n|x>(d_A+d_B)/2)}{(d_B-d_A)(4-d_A-3d_B)} &= \frac{\bar{t}}{2^{k+4}}\left(1+\frac{1}{2^n}+\frac{1}{2^{2n+2}}\right). \end{aligned} \tag{2.37}$$

We can now summarize our results to compute firms' profits on the turf of firm B depending on data quality. If $k \geq 1$, then to calculate the profits of firm B on its turf, we have to sum up $\Pi_{Bm\geq 3}(\cdot|\, x > (d_A+d_B)/2)$ in (2.35), $\Pi_{B2}(\cdot|\, x > (d_A+d_B)/2)$ in (2.36) and $\Pi_{B1}(\cdot|\, x > (d_A+d_B)/2)$ in (2.37), which yields[27]

$$\frac{\Pi_B(d_A,d_B;k,n|x>(d_A+d_B)/2)}{(d_B-d_A)(4-d_A-3d_B)} = \frac{\bar{t}}{2^{2n}}\left(\frac{2^{2n}}{16}+\frac{3\times 2^n}{16}-\frac{7}{64}\right). \tag{2.38}$$

Summing up $\Pi_B(\cdot|\, x \leq (d_A+d_B)/2)$ in (2.34) and $\Pi_B(\cdot|\, x > (d_A+d_B)/2)$ in (2.38) we

[27] Note that if $k=1$, then $\Pi_{Bm\geq 3}(d_A, d_B; 1, n|\, x > (d_A+d_B)/2) = 0$.

get the total profits of firm B for any $k \geq 1$:

$$\frac{\Pi_B(d_A,d_B;k,n)}{\bar{t}(d_B-d_A)}=\frac{3d_A+d_B}{2^{2(k+n)+5}}+\frac{(4-d_A-3d_B)}{2^{2n}}\left(\frac{2^{2n}}{16}+\frac{3\times 2^n}{16}-\frac{7}{64}\right). \tag{2.39}$$

We calculate now the profits of firm A for $k \geq 1$. The profits of firm A on the turf of firm B can be calculated as the sum of $\Pi_{A2}(\cdot|\, x > (d_A+d_B)/2)$ in (2.36) and $\Pi_{A1}(\cdot|\, x > (d_A+d_B)/2)$ in (2.37), which yields

$$\frac{\Pi_A(\cdot|x>(d_A+d_B)/2)}{\bar{t}(d_B-d_A)(4-d_A-3d_B)}=\frac{1}{2^{2(k+n+2)}}+\frac{2^{2n+2}-2^{n+2}+5}{2^{2n+k+7}}. \tag{2.40}$$

Summing up the profits of firm A on its own turf and that of the rival, $\Pi_A(\cdot|\, x \leq (d_A+d_B)/2)$ in (2.34) and $\Pi_A(\cdot|\, x > (d_A+d_B)/2)$ in (2.40), respectively, we get the total profits of firm A for any $k \geq 1$:

$$\frac{\Pi_A(d_A,d_B;k,n)}{\bar{t}(d_B-d_A)}=$$

$$(4-d_A-3d_B)\left(\frac{1}{2^{2(k+n+2)}}+\frac{2^{2n+2}-2^{n+2}+5}{2^{2n+k+7}}\right)+(3d_A+d_B)\left(\frac{2^{k+n}-1}{2^{k+n+3}}+\frac{2^{k+n+3}+1}{2^{2(k+n)+6}}\right). \tag{2.41}$$

We finally consider the case $k=0$. On its own turf firm B realizes the profit $\Pi_{B1}(\cdot|\, x > (d_A+d_B)/2)$ stated in (2.37). Summing up this profit with $\Pi_B(\cdot|\, x \leq (d_A+d_B)/2)$ in (2.34) we get total profits of firm B when $k=0$:

$$\Pi_B(d_A,d_B;0,n)=\bar{t}\,(d_B-d_A)\left(\frac{(4-d_A-3d_B)}{2^4}\left(1+\frac{1}{2^n}+\frac{1}{2^{2n+2}}\right)+\frac{3d_A+d_B}{2^{2n+5}}\right). \tag{2.42}$$

On the turf of firm B firm A realizes the profit $\Pi_{A1}(\cdot|\, x > (d_A+d_B)/2)$ stated in (2.37) while on its own turf the profit is given by $\Pi_A(\cdot|\, x \leq (d_A+d_B)/2)$ in (2.34). Summing up those profits we get the total profits of firm A for $k=0$:

$$\frac{\Pi_A(\cdot)}{\bar{t}(d_B-d_A)}=(3d_A+d_B)\left(\frac{2^n-1}{2^{n+3}}+\frac{2^{n+3}+1}{2^{2(n)+6}}\right)+\frac{\left(2^{2n+2}-2^{n+2}+5\right)}{2^{2n+7}}(4-d_A-3d_B). \tag{2.43}$$

Part 2. We now turn to the first stage of the game, where firms simultaneously choose their locations. Consider first the case $k \geq 1$. Maximizing the profits $\Pi_B(\cdot)$ in (2.39) and $\Pi_A(\cdot)$ in (2.41) with respect to d_B and d_A, respectively, we get the equilibrium locations when $k \geq 1$:

$$d_A^{d,As}(k,n) = -\frac{g(k,n)}{f(k,n)} \text{ and} \tag{2.44}$$

$$d_B^{d,As}(k,n) = \frac{p(k,n)}{f(k,n)},$$

where

$$
\begin{aligned}
g(k,n) &= -84\times 2^{2k}2^{n}+4\times 2^{2n}2^{k}-88\times 2^{3k}2^{n}+49\times 2^{2k}+35\times 2^{3k}\\
&\quad -28\times 2^{2k}2^{2n}+56\times 2^{3k}2^{2n}-32\times 2^{3k}2^{3n}-56\times 2^{4k}2^{2n}\\
&\quad -16\times 2^{3k}2^{4n}+96\times 2^{4k}2^{3n}+32\times 2^{4k}2^{4n}-4\times 2^{k}2^{n}+5\times 2^{k}+8,
\end{aligned}
$$

$$
\begin{aligned}
f(k,n) &= 48\times 2^{2k}2^{n}+88\times 2^{3k}2^{n}-28\times 2^{2k}-35\times 2^{3k}+48\times 2^{2k}2^{2n}\\
&\quad -56\times 2^{3k}2^{2n}+32\times 2^{3k}2^{3n}+224\times 2^{4k}2^{2n}+16\times 2^{3k}2^{4n}\\
&\quad -384\times 2^{4k}2^{3n}-128\times 2^{4k}2^{4n}+4,
\end{aligned}
$$

$$
\begin{aligned}
p(k,n) &= 60\times 2^{2k}2^{n}+4\times 2^{2n}2^{k}+88\times 2^{3k}2^{n}-35\times 2^{2k}-35\times 2^{3k}\\
&\quad +20\times 2^{2k}2^{2n}-56\times 2^{3k}2^{2n}+32\times 2^{3k}2^{3n}+168\times 2^{4k}2^{2n}\\
&\quad +16\times 2^{3k}2^{4n}-288\times 2^{4k}2^{3n}-96\times 2^{4k}2^{4n}-4\times 2^{k}2^{n}+5\times 2^{k}+8.
\end{aligned}
$$

While the equilibrium locations in (2.44) are difficult to analyze, we may conclude that

$$\lim_{n\to\infty,k\to\infty} d_A^{d,As}(\cdot)=d_A^{FB} \text{ and } \lim_{n\to\infty,k\to\infty} d_B^{d,As}(\cdot)=d_B^{FB}.$$

Consider now $k=0$. Taking the derivative of the profits (2.43) and (2.42) with respect to d_A and d_B, respectively, we get the equilibrium locations when $k=0$:

$$
\begin{aligned}
d_A^{d,As}(0,n) &= \tfrac{8\times 2^{2n}+32\times 2^{3n}+16\times 2^{4n}-16\times 2^{n}+1}{8\times 2^{2n}+128\times 2^{3n}+112\times 2^{4n}-5} \text{ and}\\
d_B^{d,As}(0,n) &= \tfrac{24\times 2^{2n}+96\times 2^{3n}+80\times 2^{4n}-7}{8\times 2^{2n}+128\times 2^{3n}+112\times 2^{4n}-5}.
\end{aligned}
\tag{2.45}
$$

Note that the SOCs are fulfilled and none of the firms has an incentive to deviate in a way such that $d_A \geq d_B$ holds.

We next analyze how the equilibrium locations change with the improvement in the quality of firms' flexibility data. Taking the derivatives of $d_A^{d,As}(\cdot)$ and $d_B^{d,As}(\cdot)$ with respect to $n\geq 1$ we get

$$
\begin{aligned}
\tfrac{\partial d_A^{d,As}(\cdot)}{\partial n} &= -\tfrac{2^{n+4}\left(46\times 2^{2n}-208\times 2^{3n}-288\times 2^{4n}+96\times 2^{5n}+96\times 2^{6n}+6\times 2^{n}-5\right)\ln 2}{\left(2^{2n+3}+7\times 2^{4n+4}+2^{3n+7}-5\right)^2}<0,\\
\tfrac{\partial d_B^{d,As}(\cdot)}{\partial n} &= -\tfrac{2^{2n+5}\left(9\times 2^{3n+3}-3\times 2^{2n+4}+2^{5n+4}+2^{4n+7}-39\times 2^{n}+4\right)\ln 2}{\left(2^{2n+3}+7\times 2^{4n+4}+2^{3n+7}-5\right)^2}<0,
\end{aligned}
$$

such that with the improvement in the quality of firm A's data, both firms move to the left. We next analyze how the difference $d_B^{d,As}(\cdot) - d_A^{d,As}(\cdot)$ changes with the improvement in data quality. We find that

$$\frac{\partial d_B^{d,As}(\cdot)}{\partial n} - \frac{\partial d_A^{d,As}(\cdot)}{\partial n} = -\frac{\left(112\times 2^{3n} - 124\times 2^{2n} + 432\times 2^{4n} + 160\times 2^{5n} - 64\times 2^{6n} + 2\times 2^{n} + 5\right)2^{n+4}\ln 2}{\left(2^{2n+3} + 7\times 2^{4n+4} + 2^{3n+7} - 5\right)^2},$$

which is negative if $n < 2$ and positive if $n > 2$. We also find that

$$\lim_{n\to\infty} d_A^{d,As}(\cdot) = \tfrac{1}{7} \text{ and } \lim_{n\to\infty} d_B^{d,As}(\cdot) = \tfrac{5}{7}. \tag{2.46}$$

We calculate $d_A^{d,As}(0,1) = 513/2843 < 1/4$ and from the fact that $d_A^{d,As}(\cdot)$ is a strictly decreasing function in n, we conclude that for any $n \geq 1$ it holds that $d_A^{d,As}(\cdot) < d_A^{FB}$. We calculate now $d_B^{d,As}(0,1) = 2137/2843 > 3/4$ and $d_B^{d,As}(0,2) = 27001/36987 < 3/4$. Combining the latter results with the fact that $d_B^{d,As}(\cdot)$ is a monotonically decreasing function, we conclude that $d_B^{d,As}(\cdot) < d_B^{FB}$ for any $n \geq 2$. *Q.E.D.*

Proof of Proposition 2.4. Depending on the firms' decisions to acquire flexibility data of an exogenously given quality $k \geq 1$, we can distinguish among the three subgames. In the first subgame none of the firms acquires flexibility data, where firms realize profits stated in cases *i)* and *ii)* of Proposition 2.1 when consumers are relatively homogeneous and relatively differentiated and data quality is set to $k = 0$, respectively:

$$\Pi_i^{h,NN} = \tfrac{3\underline{t}}{16} \text{ and } \Pi_i^{d,NN} = \tfrac{275\bar{t}}{2304},\ i = \{A, B\},$$

where the superscript "NN" denotes the decision of both firms to acquire no data. In the second subgame both firms acquire customer data of an exogenously given quality $k \geq 1$, where they realize profits stated in cases *i)* and *ii)* of Proposition 2.1 when consumers are relatively homogeneous and relatively differentiated, respectively:

$$\begin{aligned}
\Pi_i^{h,DD}(k) &= \tfrac{3\underline{t}\left(2^k+1+l\left(2^k-1\right)\right)}{2^{k+5}} \text{ and} \\
\Pi_i^{d,DD}(k) &= \tfrac{\bar{t}\left(9\times 2^{2k}-9\times 2^k+10\right)^2\left(27\times 2^{2k}-27\times 2^k+22\right)}{9\times 2^{2k+5}\left(9\times 2^{2k}-9\times 2^k+8\right)^2},
\end{aligned}$$

where the superscript "DD" denotes the decision of both firms to acquire flexibility data. Finally, in the third subgame only one of the firms (firm A) acquires data of quality $k \geq 1$.

In this subgame firms realize profits, which can be calculated using the results on the firms' equilibrium location choices stated in Propositions 2.2 and 2.3 for the cases of relatively homogeneous and differentiated consumers, respectively. We get that if consumers are relatively homogeneous, then depending on the ratio of consumer heterogeneity, l, and the quality of customer data, k, firm A realizes the profits:

$$\text{if } l \leq 3/2 \text{, then } \Pi_A^{h,DN}(k) = \frac{3\underline{t}\left(2^k(l+1)+1-l\right)}{2^{k+5}},$$

$$\text{if } l > 3/2 \text{ and } k < \log_2\left((l-1)/(l-3/2)\right) \text{, then}$$

$$\Pi_A^{h,DN}(k) = -\frac{\underline{t}(l-1)\left(2^k l-l+2^k+1\right)^2\left(24l^2-12\times 2^k l-20\times 2^k l^2-48l+33\times 2^k+24\right)}{2^k\left(16l^2-12\times 2^k l-12\times 2^k l^2-32l+25\times 2^k+16\right)^2} \text{ and}$$

$$\text{if } l > 3/2 \text{ and } k \geq \log_2\left((l-1)/(l-3/2)\right) \text{, then}$$

$$\Pi_A^{h,DN}(k) = \frac{\underline{t}\left(2^k l-l+2^k+1\right)^2\left(6\times 2^{2k}-l-6\times 2^k l+6\times 2^{2k} l+6\times 2^k+1\right)}{4\left(4\times 2^{2k}-l-4\times 2^k l+4\times 2^{2k} l+4\times 2^k+1\right)^2},$$

where the subscript "DN" denotes the case, where firm A acquires flexibility data, while firm B does not. The profits of firm B are given in that case by

$$\text{if } l \leq 3/2 \text{, then } \Pi_B^{h,DN}(k) = \frac{3\underline{t}}{16},$$

$$\text{if } l > 3/2 \text{ and } k < \log_2\left((l-1)/(l-3/2)\right) \text{, then}$$

$$\Pi_B^{h,DN}(k) = \frac{6\underline{t}(2l-1)^2(l-1)\left(2^k l-l+2^k+1\right)^2}{\left(16l^2-12\times 2^k l-12\times 2^k l^2-32l+25\times 2^k+16\right)^2} \text{ and}$$

$$\text{if } l > 3/2 \text{ and } k \geq \log_2\left((l-1)/(l-3/2)\right) \text{, then}$$

$$\Pi_B^{h,DN}(k) = -\left(\frac{1}{2^{k+1}}-1\right)\frac{3\underline{t}\times 2^k\left(l+2^k-1\right)\left(2^k l-l+2^k+1\right)^2}{\left(4\times 2^{2k}-l-4\times 2^k l+4\times 2^{2k} l+4\times 2^k+1\right)^2}.$$

When consumers are relatively differentiated, we get the following profits in the subgame, where only firm A acquires customer data:

$$\Pi_A^{d,DN}(k) = \frac{\bar{t}\left(44\times 2^{2k}+4\times 2^k+1\right)\left(2^{k+1}+2^{2k+1}+2^{3k+3}+2^{4k+3}-1\right)^2}{f(k)}, \text{ where}$$

$$f(k) = 50\times 2^{2k}-160\times 2^{4k}-2560\times 2^{5k}-2112\times 2^{6k}$$
$$+4096\times 2^{7k}+36\,352\times 2^{8k}+57\,344\times 2^{9k}+25\,088\times 2^{10k} \text{ and}$$

$$\Pi_B^{d,DN}(k) = \frac{\bar{t}\left(12\times 2^{2k}+12\times 2^k+1\right)\left(2^{k+1}+2^{2k+1}+2^{3k+3}+2^{4k+3}-1\right)^2}{2^{2k}\left(2^{2k+3}+7\times 2^{4k+4}+2^{3k+7}-5\right)^2}.$$

We can next analyze firms' incentives to acquire customer data. We start with the case of the relatively homogeneous consumers. To conclude on the incentives of a firm to acquire customer data when the rival does not hold it, we have to consider the difference $\Pi_A^{h,DN}(k) - \Pi_i^{h,NN}$. Three cases depending on l and k are possible. Assume first that $l \leq 3/2$, in which case

$$\Pi_A^{h,DN}(k) - \Pi_i^{h,NN} = \frac{3\underline{t}\left(2^k-1\right)(l-1)}{2^{k+5}} > 0, \text{ for any } k \geq 1 \text{ and any } l \leq 3/2, \tag{2.47}$$

such that a firm has a unilateral incentive to acquire customer data of any quality $k \geq 1$ if $l \leq 3/2$. Assume next that $l > 3/2$ and $k < \log_2((l-1)/(l-3/2))$. As the direct comparison between $\Pi_A^{h,DN}(k)$ and $\Pi_i^{h,NN}$ can not be solved analytically, we consider instead the following comparison:

$$\begin{aligned} & \Pi_A^{h,DN}(k)\Big|_{l>3/2,\ k<\log_2((l-1)/(l-3/2))} - \Pi_A^{h,DN}(k)\Big|_{l\leq 3/2} \\ = & -\frac{\underline{t}(2l-3)^2\left(2^k l-l+2^k+1\right)\left(91\times 2^k-36\times 2^k l-52\times 2^k l^2-128l+64l^2+64\right)}{32\left(25\times 2^k-12\times 2^k l-12\times 2^k l^2-32l+16l^2+16\right)^2} > 0. \end{aligned} \tag{2.48}$$

To prove the sign of the inequality (2.48) note first that for any $l \leq 2$ it holds that $(l-1)/(l+1) < 1$, while $2^k > 1$ for any $k \geq 1$, which implies that $2^k l - l + 2^k + 1 > 0$. In a similar way we can conclude that $91 \times 2^k - 36 \times 2^k l - 52 \times 2^k l^2 - 128l + 64l^2 + 64 < 0$, which explains the sign of the inequality in (2.48). Note finally that $\Pi_A^{h,DN}(k)\Big|_{l\leq 3/2}$ increases in l. Hence, combining the inequalities (2.47) and (2.48) we conclude that

$$\Pi_A^{h,DN}(k) - \Pi_i^{h,NN} > 0, \text{ for any } l > 3/2 \text{ and } k < \log_2((l-1)/(l-3/2)).$$

Assume next that $l > 3/2$ and $k \geq \log_2((l-1)/(l-3/2))$. We proceed in a similar way as above and consider the difference

$$\begin{aligned} & \Pi_A^{h,DN}(k)\Big|_{l>3/2,\ k\geq\log_2((l-1)/(l-3/2))} - \Pi_A^{h,DN}(k)\Big|_{l\leq 3/2} \\ = & \frac{\underline{t}(l-1)\left(2^k l-l+2^k+1\right)\left(2^{2k+4}l-3l+2^{k+4}+2^{2k+4}-2^{k+4}l+3\right)}{32\times 2^k\left(4\times 2^{2k}-l-4\times 2^k l+4\times 2^{2k}l+4\times 2^k+1\right)^2} > 0. \end{aligned} \tag{2.49}$$

We prove now the sign of the inequality (2.49). We showed above that $2^k l - l + 2^k + 1 > 0$ holds for any $l \leq 2$ and $k \geq 1$. Note finally that

$$\begin{aligned} & \frac{2^{2k+4}l-3l+2^{k+4}+2^{2k+4}-2^{k+4}l+3}{(16l+16)} \\ = & \left(2^k - \frac{-(16-16l)-\sqrt{448l^2-512l+64}}{2(16l+16)}\right)\left(2^k - \frac{-(16-16l)+\sqrt{448l^2-512l+64}}{2(16l+16)}\right) > 0, \end{aligned}$$

which follows from the fact that $2^k > 1$ for any $k \geq 1$ and for any $l \leq 2$ it holds that

$$\frac{-(16-16l)-\sqrt{448l^2-512l+64}}{2(16l+16)} < \frac{-(16-16l)+\sqrt{448l^2-512l+64}}{2(16l+16)} < 1,$$

which explains the sign of the inequality in (2.49). We conclude that

$$\Pi_A^{h,DN}(k) - \Pi_i^{h,NN} > 0, \text{ for any } l > 3/2 \text{ and } k \geq \log_2\left((l-1)/(l-3/2)\right).$$

We can summarize that when consumers are relatively homogeneous, a firm always has a unilateral incentive to acquire customer data, such that an equilibrium where none of the firms acquires flexibility data does not exist. We next turn to the incentives of a firm to acquire customer data when the rival also holds it. Assume first that $l \leq 3/2$:

$$\Pi_i^{h,DD}(k) - \Pi_B^{h,DN}(k) = \frac{3\underline{t}\left(2^k-1\right)(l-1)}{32\times 2^k} > 0, \text{ for any } l \leq 3/2 \text{ and } k \geq 1.$$

Assume next that $l > 3/2$ and $k < \log_2\left((l-1)/(l-3/2)\right)$. We get that

$$\Pi_i^{h,DD}(k) - \Pi_B^{h,DN}(k) = -\frac{3\underline{t}\left(2^k l-l+2^k+1\right)\left(\frac{33}{4}l^2-17l^3+\frac{7}{2}l^4-\frac{689}{32}+\frac{107}{4}l\right)\left(2^k-\alpha\right)\left(2^k-\beta\right)}{2^k\left(25\times 2^k-12\times 2^k l-12\times 2^k l^2-32l+16l^2+16\right)^2}, \text{ where}$$

$$\begin{aligned}
\alpha &= \frac{-\left(74l-63l^2+12l^3+4l^4-27\right)+2(2l-1)(l-1)^2\sqrt{2(4l^2-8l+5)}}{2\left(\frac{33}{4}l^2-17l^3+\frac{7}{2}l^4-\frac{689}{32}+\frac{107}{4}l\right)} \text{ and} \\
\beta &= \frac{-\left(74l-63l^2+12l^3+4l^4-27\right)-2(2l-1)(l-1)^2\sqrt{2(4l^2-8l+5)}}{2\left(\frac{33}{4}l^2-17l^3+\frac{7}{2}l^4-\frac{689}{32}+\frac{107}{4}l\right)}.
\end{aligned}$$

Note that for any $3/2 < l \leq 2$ it holds that $\alpha < 1$ and $\beta < 2$, while for any $k \geq 1$ it holds that $2^k \geq 2$, which implies that $\left(2^k-\alpha\right)\left(2^k-\beta\right) > 0$. For all $3/2 < l \leq 2$ it holds that

$$\frac{33}{4}l^2-17l^3+\frac{7}{2}l^4-\frac{689}{32}+\frac{107}{4}l < 0.$$

Hence, we can conclude that if $l > 3/2$ and $k < \log_2\left((l-1)/(l-3/2)\right)$, then $\Pi_i^{h,DD}(k) - \Pi_B^{h,DN}(k) > 0$. Assume finally that $l > 3/2$ and $k \geq \log_2\left((l-1)/(l-3/2)\right)$, in which case

$$\Pi_i^{h,DD}(k) - \Pi_B^{h,DN}(k) = \frac{3\left(2^k l-l+2^k+1\right)f(l,k)}{2^k\left(4\times 2^{2k}-l-4\times 2^k l+4\times 2^{2k} l+4\times 2^k+1\right)^2}, \text{ where}$$

$$f(l,k) = (l-1)^2\left(\frac{7\times 2^{2k}}{4}-\frac{2^k}{4}+\frac{1}{32}\right)+2^{3k}\left(-2l^2+\frac{3}{2}+\frac{3l}{2}\right)+\frac{2^{4k}}{2}(l-1)(l+1).$$

Note that the following inequalities hold

$$\begin{aligned}
\frac{7\times 2^{2k}}{4}-\frac{2^k}{4}+\frac{1}{32} &> 0, \text{ for any } k \geq 1 \text{ and} \\
\frac{(l-1)(l+1)}{2}-\left(-2l^2+\frac{3}{2}+\frac{3l}{2}\right) &> 0, \text{ for any } l > 3/2,
\end{aligned}$$

which implies that $f(l,k) > 0$ and $\Pi_i^{h,DD}(k) - \Pi_B^{h,DN}(k) > 0$. We conclude that when consumers are relatively homogeneous, then a firm always has an incentive to acquire customer data when the rival holds it. Hence, we can conclude that a unique equilibrium exists where both firms acquire customer data.

We now turn to the case of relatively differentiated consumers. To conclude on the incentives of a firm to acquire flexibility data when the rival does not hold it, we have to consider the difference:

$$\Pi_A^{d,DN}(k) - \Pi_i^{d,NN} < 0, \text{ for any } k \geq 1. \tag{2.50}$$

We now prove the sign of the inequality (2.50). Note first that the function $\Pi_A^{d,DN}(k)$ decreases for any $1 \leq k \leq \widehat{k}$ (where $\widehat{k} \approx 2.55$) and increases for any $k > \widehat{k}$. Note next that $\Pi_A^{d,DN}(1) = 7623665\bar{t}/64661192 < \Pi_i^{d,NN}$. Note finally that $\lim_{k\to\infty} \Pi_A^{d,DN}(k) = 11/98 < \Pi_i^{d,NN}$. Hence, we conclude that inequality (2.50) indeed holds, which implies that a firm never has a unilateral incentive to acquire customer data of any quality $k \geq 1$. It follows then that firms' decisions not to acquire flexibility data constitute the equilibrium. We next analyze whether an equilibrium exists, where both firms acquire flexibility data. This equilibrium exists if a firm has an incentive to acquire data given that the rival also holds it. We have to consider the difference:

$$f(k) := \Pi_i^{d,DD}(k) - \Pi_B^{d,DN}(k).$$

Note that $f(k)$ is a monotonically increasing function, such that $f(1) < 0$ and $f(2) > 0$. Hence, $f(k) > 0$ for any $k \geq 2$. We conclude that if $k \geq 2$, then a Nash equilibrium exists, where both firms acquire customer flexibility data. Summarizing our results for the case of relatively differentiated consumers, we conclude that if $k = 1$, then a unique Nash equilibrium exists, where none of the firms acquires flexibility data. If $k \geq 2$, then two Nash equilibrium exist where either both firms do not acquire flexibility data or both of them do.

Finally, we analyze how firms' profits change in the equilibrium where both of them acquire customer data compared to the case, where none of the firms holds flexibility data.

We start with the case of relatively homogeneous consumers. The comparison shows that

$$\Pi_i^{h,DD}(k) - \Pi_i^{h,NN} = \frac{3\underline{t}\left(2^k-1\right)(l-1)}{32\times 2^k} > 0, \text{ for all } k \geq 1 \text{ and } l \leq 2,$$

such that both firms are better-off. We now turn to the case of relatively differentiated consumers. The comparison shows that for any $k \geq 2$:

$$\begin{aligned} & \Pi_i^{d,DD}(k) - \Pi_i^{d,NN} \\ = & \frac{\left(84\,168\times 2^{3k} - 88\,816\times 2^{2k} - 43\,749\times 2^{4k} + 7938\times 2^{5k} + 4779\times 2^{6k} + 53\,280\times 2^k - 17\,600\right)\bar{t}}{-147\,456\times 2^{2k} + 331\,776\times 2^{3k} - 518\,400\times 2^{4k} + 373\,248\times 2^{5k} - 186\,624\times 2^{6k}} < 0, \end{aligned}$$

such that both firms are worse-off. *Q.E.D.*

Declaration of Contribution according to § 6 (1) PO as of 25.11.2013:

Hereby I, Irina Hasnas, declare that the chapter “Consumer Flexibility, Data Quality and Location Choice” is co-authored by Irina Baye. It has been pre-published as a DICE Discussion paper (Baye and Hasnas, 2014).

My contributions with regard to content and methods are the following:

- I have searched for the relevant literature and examples.
- I have contributed to the set-up and the specifications of the model.
- I have contributed to the calculation of the two-stage game.
- I have contributed to the proofs.
- I have contributed to Conclusion.

Signatures:

____________________	____________________
Irina Hasnas, M.Sc.	Jun.-Prof. Dr. Irina Baye

Chapter 3

Consumer Flexibility, Data Quality and Collusion

3.1 Introduction

Advances in information technologies allow firms to collect, store and analyze various types of customer data including demographics (address, gender, age, income), data on previous purchases etc.[1] This data may give insights into consumers' brand preferences and the strength thereof (flexibility), allowing firms to price discriminate respectively. Ever since Thisse and Vives (1988) it is known that competitive price discrimination may intensify competition and decrease firms' profits, as a result firms could collude not to acquire customer data and/or share the market.[2,3]

[1] For example, IBM provides data processing platforms and Business analytics software which help firms to store, process, forecast and statistically analyze various data (http://www-01.ibm.com/software/data/bigdata/platform/product.html).

[2] In telecommunications market, firms collect large amounts of customer data such as name, gender, physical address and calling history. In 2005 Conseil de la Concurrence fined three biggest French mobile operators for engaging in anticompetitive agreements. These companies were accused of sharing customer data and sharing the market. http://www.autoritedelaconcurrence.fr/user/standard.php?id_rub=160&id_article=502

[3] In the airline industry firms collect customer data through Frequent Flyer Program (FFP) and use it for third degree price discrimination. Customers can opt in and receive discounts based on the total amount of miles they fly. In 1999 two Scandinavian airlines SAS and Maersk Air notified the European

In this chapter I analyze how firms' incentives to collude depend on the quality of customer data. Following Sapi and Suleymanova (2013) I introduce two-dimensional consumer heterogeneity and assume that consumers differ both with respect to their brand preferences and flexibility. Respectively, two types of customer data are available to the firms. Furthermore, I assume that data on consumer addresses is perfect, while data on their flexibility is not. For example, in location-based marketing firms know the precise location of each consumer, while consumer sensitivity to different marketing activities (like price reductions and advertising) can be estimated only with less-than-perfect accuracy.[4] I follow Liu and Serfes (2007) to model imperfect customer data. I assume that firms are able to correctly identify different flexibility segments and can allocate any consumer to one of them.[5] With the improvemnt in the quality of customer data consumer segmentation becomes finer.

The articles most closely related to this work are Liu and Serfes (2007) and Sapi and Suleymanova (2013). Liu and Serfes analyze firms' incentives to collude depending on the quality of data on consumer brand preferences when consumers are differentiated only along that dimension. Sapi and Suleymanova analyze firms' incentives to acquire imperfect customer flexibility data when data on consumer addresses is perfect and both firms hold it. Following Sapi and Suleymanova (2013) I introduce two-dimensional consumer heterogeneity, on brand preferences and flexibility (transport cost parameters) and analyze firms' incentives to collude depending on the quality of the latter characteristic of consumer

Commission about a cooperative agreement that included code-sharing on a number of routes and FFP extension that allowed Maersk customers to earn points when flying with SAS and *vice versa*. However, in 2001 the European Commission fined the airlines for market-sharing agreement. (Sun-Air versus SAS and Maersk Air, 2001)

[4] Epling (2002) uses data from long-distance telephony to show that information on customer's location and income allows firms to better price-discriminate among consumers.

[5] Angwin (2010) describes how various internet companies collect personal information (location, age, gender, income, education, marital status, etc.) about websites' users and sell it to marketers and advertisers. The data can be of any quality: "We can segment it all the way down to one person" says Eric Porres, Lotame's chief executive officer.

preferences.

This chapter is organized as follows. Section 3.2 presents the model. In Section 3.3 I provide the equilibrium analysis and consider firms' incentives to collude. I distinguish between three collusive schemes: in the first scheme firms collude both in prices and their data acquisition decisions; in the second, they collude only in prices; in the third, they compete in prices and collude in data acquisition decisions. Section 3.4 compares the three schemes. I conclude in Section 3.5.

3.2 The Model

There are two firms, A and B, each situated at the end of a unit interval. Firm A is located at $x_A = 0$, and Firm B at $x_B = 1$. Each firm produces a brand of the same good. I normalize the marginal cost of production of such good to zero. There is a mass of consumers normalized to unity. I follow Sapi and Suleymanova (2013) and assume that consumers are differentiated both with respect to their addresses and transport cost parameters. Therefore, every consumer is uniquely described by a pair of parameters (x, t), where $x \in [0, 1]$ represents consumer's address and $t \in [\underline{t}, \overline{t}]$ is her transport cost parameter (flexibility), where $\underline{t} \geq 0$ and $\overline{t} > \underline{t}$. I assume that x and t are uniformly and independently distributed; i.e., $f_t = 1/(\overline{t} - \underline{t})$, $f_x = 1$ and $f_{t,x} = 1/(\overline{t} - \underline{t})$.

I assume that both firms hold perfect information on consumer locations. Firms can also acquire customer flexibility data at zero cost. Data on consumer transport costs is imperfect and is characterized by the exogenously given quality parameter $k = 0, 1, 2, ..., \infty$. For any k firms are able to divide the interval $[\underline{t}, \overline{t}]$ into $n := 2^k$ segments and allocate each consumer to one of them. Every segment $m = 1, 2, ..., 2^k$ is characterized by the transport cost parameters $t^m \in [\underline{t}^m, \overline{t}^m]$, where $\underline{t}^m = \underline{t} + (\overline{t} - \underline{t})(m - 1)/n$ and $\overline{t}^m = \underline{t} + (\overline{t} - \underline{t})m/n$. Higher k implies customer data with a finer segmentation of the interval $[\underline{t}, \overline{t}]$ and, hence, better quality. If $k \to \infty$, firms have perfect information on consumers' transport cost parameters.

I follow Sapi and Suleymanova (2013) and consider two versions of the model, depending on consumer heterogeneity in flexibility measured by the ratio $l(\underline{t}, \overline{t}) := \overline{t}/\underline{t}$. In the first version consumers are *relatively heterogeneous* and $\underline{t} = 0$, such that $\lim_{\underline{t}\to 0} l(\underline{t}, \bar{t}) = \infty$. In the second version consumers are *relatively homogeneous* with $\underline{t} > 0$ and $l(\underline{t}, \bar{t}) \leq 2$.

If a firm acquires flexibility data, it charges different prices depending on a consumer's address, flexibility segment and the quality of customer data: $p_i(x, m; n)$ with $i = A, B$. Therefore, two consumers at the same location that belong to different flexibility segments can be charged different prices. The utility of a consumer (x, t) in case of buying from Firm i is

$$U_i(p_i(x, m; n), t, x) = V - t|x - x_i| - p_i(x, m; n),$$

where $V > 0$ is the basic valuation of a product. In order to ensure that the market is always covered in equilibrium, I assume that V is large enough. A consumer always buys from a firm offering the highest utility. The indifferent consumer buys from the nearest firm.

I consider an infinitely repeated game. In a stage game firms decide simultaneously and independently whether to acquire customer flexibility data and which prices to charge. These decisions become a common knowledge at the end of each stage game and firms have a perfect memory of all past actions.

Firms may collude using trigger strategies, such that in period t a firm plays cooperatively if in period $t - 1$ the rival played cooperatively. In other words, firms stick to the collusion agreement as long as nobody has deviated. However, if deviation takes place in period t, firms will play Nash Equilibrium from period $t + 1$ to infinity. I denote the one-shot collusive profit of Firm i by $\pi_i^C(n)$, the deviation profit by $\pi_i^D(n)$, and non-cooperative profit by $\pi_i^N(n)$. Collusion can only be sustained in the infinitely repeated game if the discount factor is sufficiently high:

$$\delta \geq \bar{\delta}(n) \equiv \frac{\pi_i^D(n) - \pi_i^C(n)}{\pi_i^D(n) - \pi_i^N(n)}, \; i = A, B.$$

3.3 Equilibrium Analysis

Non-cooperative profits. Note that in this model firms make data acquisition decisions simultaneously with their price choices in a stage game, and therefore, in the non-cooperative equilibrium firms always acquire customer data. This is different from Sapi and Suleymanova (2013), where these two decisions are made sequentially. Hence, the equilibrium prices in my analysis are the same as in Sapi and Suleymanova in the subgame where both firms acquire customer data. The interval $[0, 1/2)$ represents *Firm A's turf* and the interval $(1/2, 1]$ is *Firm B's turf*. Given prices $p_A(x, m; n)$ and $p_B(x, m; n)$, the transport cost parameter of the indifferent consumer on the segment m with address x is

$$\widetilde{t}(x, m; n) = \frac{p_A(x, m; n) - p_B(x, m; n)}{1 - 2x}, \text{ where } \widetilde{t}(x, m; n) \in [\underline{t}^m, \overline{t}^m].$$

On any segment m, Firm A serves its most loyal consumers with high transport cost parameters $t \geq \widetilde{t}(x, m; n)$, and Firm B serves the least loyal consumers of Firm A with low transport cost parameter, $t < \widetilde{t}(x, m; n)$. For any address $x \in [0, 1/2)$ and on any segment m, Firm A maximizes the expected profit

$$E[\pi_A(x, m; n)|x < 1/2] = p_A(x, m; n) \Pr\{t \geq \widetilde{t}(x, m; n)\},$$

while Firm B maximizes the expected profit

$$E[\pi_B(x, m; n)|x < 1/2] = p_B(x, m; n) \Pr\{t < \widetilde{t}(x, m; n)\}.$$

In my version of the model, the equilibrium prices and profits in the non-cooperative case are identical to those stated in Proposition 1 in Sapi and Suleymanova (2013) and depend on data quality and consumer heterogeneity in flexibility.

Proposition 3.1. (From Sapi and Suleymanova, 2013)

i) Assume that consumers are relatively differentiated. In equilibrium on Firm i's turf on the segment $m = 1$, *firms charge prices* $p_i^*(x, 1; n) = 2\overline{t}\,|1 - 2x|\,/(3n)$ *and* $p_j^*(x, 1; n) = \overline{t}\,|1 - 2x|\,/(3n)$. *Firm i serves consumers with* $t \geq \overline{t}/\,(3n)$. *On the segments* $2 \leq m \leq n$ *equilibrium prices are* $p_i^*(x, m; n) = \overline{t}\,(m - 1)\,|1 - 2x|\,/n$ *and* $p_j^*(x, m; n) = 0$, *where Firm i serves all consumers. Equilibrium profits are* $\Pi_i^*(n) = 5\overline{t}/(36n^2) + \overline{t}/8\,(1 - 1/n)$.

ii) Assume that consumers are relatively homogeneous. In equilibrium on Firm i's turf firms charge prices $p_i^*(x, m; n) = \left[\underline{t} + \left(\bar{t} - \underline{t}\right)(m-1)/n\right]|1-2x|$ *and* $p_j^*(x, m; n) = 0$. *Firm i serves all consumers on its turf. Firms realize profits* $\Pi_i^*(n) = \underline{t}/4 + (\bar{t} - \underline{t})/8\,[1 - 1/n]$.

The following two graphs show how firms' non-cooperative profits change with the improvement in the quality of customer flexibility data. I use the values $\underline{t} = 1$ and $\bar{t} = 2$ for the cases of relatively homogeneous and relatively differentiated consumers, respectively.

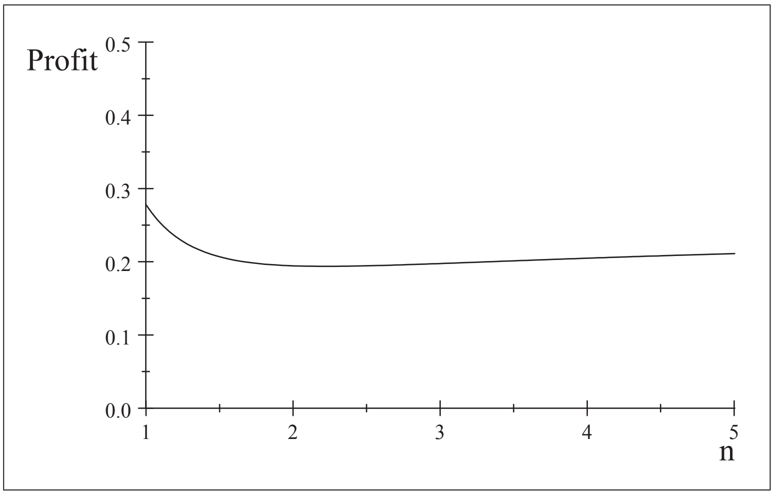

Figure 3.1. Non-cooperative profit with relatively differentiated consumers

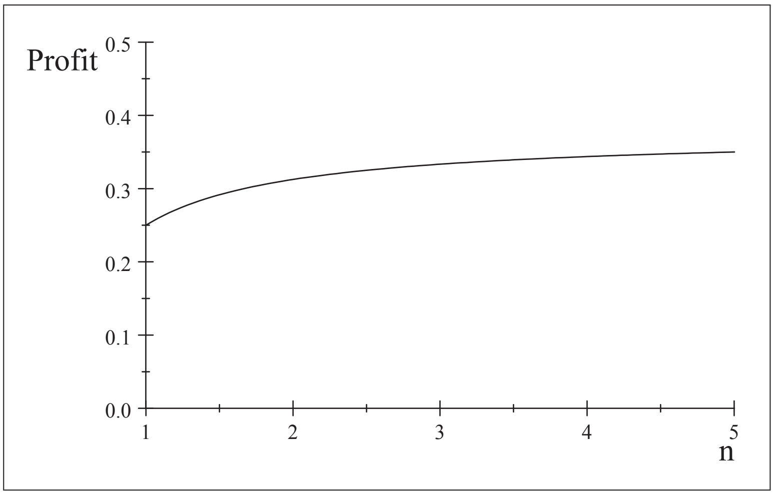

Figure 3.2. Non-cooperative profit with relatively homogeneous consumers

The two versions of the model (with relatively homogeneous and differentiated consumers) yield two different equilibria which are driven by the type of the best-response function of a firm on its turf. Precisely, as is shown in Sapi and Suleymanova (2013) when consumers are heterogeneous (homogeneous) a firm follows a *market-sharing* (*monopolization strategy*) on its turf in the absence of data on consumer flexibility. In the former case a firm optimally serves all consumers on its turf only if the rival's price is sufficiently high. Otherwise, a firm shares consumers on its turf with the rival. The reason is that it is costly for a firm to serve all consumers on its turf, because the most flexible consumer can switch brands costlessly. In equilibrium the rival charges indeed a relatively low price and targets the least loyal consumers of a firm some of whom switch. As a result, in equilibrium firms serve consumers on both turfs. The acquisition of data of quality $k = 1$ intensifies competition as on both new segments the rival charges lower prices. As a result, profits decrease. However, with a further improvement in data quality profits start to increase since a firm can better target consumers, and the rent-extraction effect dominates.

When consumers are relatively homogeneous, for any price of the rival it suffice for a firm to decrease a little the price targeted at the least flexible consumers to attract all consumers with a given address. This makes it optimal for a firm to follow a monopolization strategy on its turf, such that for any price of the rival a firm optimally serves all consumers on its turf. Then in equilibrium the rival charges the price of zero on a firm's turf and competition is very intense. As the rival cannot decrease its price below zero, the acquisition of additional data gives rise only to the rent-extraction effect, and a firm's profits monotonically increase in data quality. In the case of relatively differentiated consumers the behavior of non-cooperative profits depending on data quality is similar to the one in Liu and Serfes (2007), because in both cases there is a consumer who can switch brands costlessly. This is the consumer with $x = 1/2$ in Liu and Serfes (2007) and the consumer with $t = 0$ in my case.

Collusive profits. Firms may collude along two dimensions: customer data acquisition decisions and pricing decisions. I follow Liu and Serfes (2007) and consider three collusive schemes. In the first scheme firms acquire customer flexibility data and charge monopoly discriminatory prices. Each firm acts as a monopolist on its own turf. Provided the basic valuation is high enough, all consumers are served under collusion, and every consumer buys from her most preferred firm at a price, which makes her indifferent between buying at that firm and not buying. This type of collusion leads to both firms using monopolization strategies regardless of consumers' heterogeneity in flexibility. If a firm deviates, it gains all consumers on the rival's turf. The following proposition states the collusive and deviation prices and profits for the first scheme.

Proposition 3.2. *Consider the collusive scheme under which firms acquire customer flexibility data and charge collusive prices. Assume that the basic utility is relatively large:* $V > \max\left\{\underline{t}/2 + (\bar{t} - \underline{t})(m+1)/(2n), \underline{t} + (m+1)(\bar{t} - \underline{t})/n\right\}$ *for any* m, n.

i) Under collusion, on its own turf on the segment m *and address* x, *Firm* $i = A, B$ *charges the price* $p_i^C(x, m; n) = V - \left(\underline{t} + (\bar{t} - \underline{t})m/n\right)|x - x_i|$ *and serves all consumers there. The collusive profit of Firm* i *is* $\pi_i^C(n) = V/2 - \underline{t}/8 - (\bar{t} - \underline{t})(1+n)/(16n)$.

ii) If Firm j *deviates on the turf of Firm* i, *it charges the price* $p_j^D(x, m; n) = V - \left(\underline{t} + (\bar{t} - \underline{t})m/n\right)|x_j - x|$ *and serves all consumers. The price on its own turf does not change. The deviation profit of Firm* j *is* $\pi_j^D(n) = V - \underline{t}/2 - (\bar{t} - \underline{t})(1+n)/(4n)$.

Proof. See Appendix.

Let's now turn to the second collusive scheme. Under this scheme, firms collude by deciding not to acquire data on consumer flexibility and charge monopoly prices independently of consumer's flexibility. The following proposition states the collusive and deviation prices and profits for the second scheme.

Proposition 3.3. *Consider the collusive scheme under which firms do not acquire customer flexibility data and collude in prices. Assume that the basic utility is relatively large:* $V > \max\left\{\bar{t} - \underline{t}/2, \underline{t} + (m+1)/n(\bar{t} - \underline{t})\right\}$ *for any* m *and* n.

i) Under collusion, Firm $i = A, B$ *charges the price* $p_i^C(x) = V - \bar{t}\,|x - x_i|$ *to consumers with address* x. *The collusive profit of Firm* i *is* $\pi_i^C = V/2 - \bar{t}/8$.

ii) If Firm i *deviates, it acquires consumer flexibility data. On its own turf it charges the price* $p_i^D(x, m; n) = V - \left(\underline{t} + (\bar{t} - \underline{t})m/n\right)|x - x_i|$, and $p_i^D(x, m; n) = V - \bar{t}x - \left(\underline{t} + (\bar{t} - \underline{t})m/n\right)|2x - x_i|$ *on the rival's turf. The deviation profit of Firm* i *is* $\pi_i^D(n) = V - (\bar{t} + 3\underline{t})/8 - 3(\bar{t} - \underline{t})(1 + n)/(16n)$.

Proof. See Appendix.

When a firm deviates under the second collusive scheme, it acquires data on consumer flexibility and discriminates consumers with respect to their address and flexibility on both turfs. If the basic consumer valuation is high enough, then firms charge prices under which all consumers buy both under collusion and deviation. Firms use monopolization strategies. This implies that under collusion each firm serves all consumers on its own turf and none of the firms wants to target the most flexible consumers of the rival. However, if a firm deviates, then it serves all consumers on both turfs.

The results from Propositions 3.2 and 3.3 do not depend on consumers' heterogeneity. Under these collusive schemes, both firms optimally share and monopolize the market. They extract the highest possible rent from consumers and set such collusive prices that every consumer buys from its nearest firm. Under deviation, a firm undercuts the rival's collusive price and serves all consumers.

Finally, I consider the third collusive scheme, where firms agree not to acquire customer flexibility data and set competitive prices. This scheme does not make sense with relatively homogeneous consumers as for any quality of customer flexibility data it yields profits which are (weakly) smaller than the non-cooperative profits.[6] The reason being that in the non-cooperative equilibrium firms discriminate consumers based on their address and flexibility. Since the data on consumer address is perfect, the non-cooperative profit depends on the quality of the customer flexibility data: Sapi and Suleymanova (2013) find that the non-cooperative profits increase monotonically in data quality when consumers are relatively homogeneous. When firms agree to not acquire customer flexibility data and compete in prices (collusive scheme three), they discriminate consumers based solely on their address. Therefore when consumers are relatively homogeneous, the profits in the third scheme correspond to the lowest non-cooperative profits; i.e., when the quality of data is zero.

Hence, I consider this scheme only for relatively differentiated consumers. The following

[6] Under third scheme the collusive profit in case of relatively homogeneous consumers is $\pi_i^C = \underline{t}/4$.

proposition states collusive and deviation profits under the third scheme.

Proposition 3.4. *Consider the collusive scheme under which firms do not acquire customer flexibility data and compete in prices.*

i) Under collusion, on Firm i's turf firms charge prices $p_i^C(x) = 2\bar{t}\,|1-2x|\,/3$ *and* $p_j^C(x) = \bar{t}\,|1-2x|\,/3$, *where* $i = A, B$ and $i \neq j$. *Firm i serves consumers with* $t \geq \bar{t}/3$. *Collusive profits are* $\pi_i^C = 5\bar{t}/36$.

ii) If Firm i deviates, it acquires customer flexibility data. Assume that $n > 2$ $(k > 1)$. *On its own turf Firm i charges the deviation price* $p_i^D(x, m; n) = \bar{t}/3(3(m-1)/n+1)\,|1-2x|$. *On the rival's turf it charges the price* $p_i^D(x, m; n) = \bar{t}/3(2-3m/n)\,|1-2x|$ *to all consumers if* $m < 2n/3-1$, *and* $p_i^D(x, m; n) = \bar{t}/6(2-3(m-1)/n)\,|1-2x|$ *to consumers with flexibility* $t \in [\underline{t}^m, \bar{t}/3 + \underline{t}^m/2]$ *if* $m > 2n/3-1$. *The deviation profit of Firm i is*

$$\pi_i^D(n) = \frac{(5n-3)\bar{t}}{24n} + \sum_{m=1}^{[\frac{2n}{3}-1]} \int_{\frac{x_i}{2}}^{\frac{1}{2}(1+x_i)} \int_{\underline{t}^m}^{\bar{t}^m} f_t \frac{\bar{t}}{3}(2-\frac{3m}{n})\,|1-2x|\,dtdx$$

$$+ \sum_{m=[\frac{2n}{3}]}^{n} \int_{\frac{x_i}{2}}^{\frac{1}{2}(1+x_i)} \int_{\underline{t}^m}^{\bar{t}(\frac{1}{3}+\frac{m-1}{2n})} f_t \frac{\bar{t}}{6}(2-3\frac{m-1}{n})\,|1-2x|\,dtdx$$

Assume now that $n = 2$ $(k = 1)$. *On its own turf Firm i charges the deviation price* $p_i^D(x, 1; 2) = (5\bar{t}/12)\,|1-2x|$ *if* $t \in [\bar{t}/12, \bar{t}/2]$, *and* $p_i^D(x, 2; 2) = (5\bar{t}/6)\,|1-2x|$ *if* $t \in [\bar{t}/2, \bar{t}]$. *On Firm j's turf Firm i charges* $p_i^D(x, 1; 2) = (\bar{t}/3)\,|1-2x|$ *if* $t \in [0, \bar{t}/3]$, *and* $p_i^D(x, 2; 2) = (\bar{t}/12)\,|1-2x|$ *if* $t \in [\bar{t}/2, 7\bar{t}/12]$. *The deviation profit of Firm i is* $\pi_i^D = (17\bar{t})/96$.

Proof. See Appendix.

When firms collude under this scheme and consumers are relatively differentiated, every firm follows a market-sharing strategy on its turf and loses the less loyal consumers to the rival. The reason is that this type of collusion does not allow for price discrimination with respect to consumer flexibility, firms must set uniform prices to consumers with the same address. With relatively differentiated consumers, it is optimal for a firm to charge a relatively high price and target the most loyal consumers on its turf. As a result, under this collusive scheme each firm serves consumers on both turfs.

If a firm deviates, it acquires customer flexibility data of quality: $n \geq 2$.[7] When $n = 2$, it adopts a different deviation strategy than when $n > 2$. Let's consider the optimal deviation strategy on Firm i's own turf. If $n = 2$, Firm i follows a market-sharing strategy on the first segment and a monopolization strategy on the second. However, if $n > 2$, it follows a monopolization strategy on all segments. Now I turn to the optimal deviation strategy on a rival's turf. If $n = 2$, Firm i adopts a market-sharing strategy on both segments. However, if $n > 2$, it follows also a monopolization strategy on some segments according to the rule described in Proposition 3.4. These results are driven by the fact that consumers are relatively differentiated. When the deviation takes place on a firm's own turf, the negative competition effect is very strong when $n = 2$, and becomes weaker when n increases. When a firm deviates on the rival's turf, the competition effect there is even stronger, therefore, even with a data of better quality the deviating firm still targets only the most flexible consumers of the rival. Hence, when a firm deviates under the third scheme, it does not serve all consumers in the market.

3.4 Comparison of the collusive schemes

Assumption 3.1. *Basic valuation is assumed to be sufficiently high, precisely* $V > 2\bar{t}$.[8]

To compare the three collusive schemes, I distinguish between two cases based on consumer heterogeneity. Consider first the case of relatively homogeneous consumers. Collusive scheme one can be sustained if the discount factor is relatively high:

$$\delta \geq \overline{\delta}_1(n) := \frac{\frac{V}{2} - \frac{3\underline{t}}{8} - \frac{(\bar{t}-\underline{t})}{16}(3 + \frac{3}{n})}{V - \frac{3\underline{t}}{4} - \frac{\bar{t}-\underline{t}}{8}(3 + \frac{1}{n})}.$$

The first order derivative of $\overline{\delta}_1(n)$ with respect to n is positive under Assumption 3.1. As

[7] When $n = 1$, it stands for no data acquisition or customer data of zero quality, therefore I do not consider this case.

[8] Assumption 3.1 represents the strictest condition on V from Propositions 3.2 and 3.3 and it is obtained when $n = 1$ and $\underline{t} = 0$.

$\overline{\delta}_1(n)$ is an increasing function of n, it implies that collusion becomes more difficult to sustain as the quality of customer flexibility data improves.

The second collusive scheme can be sustained if $\delta \geq \overline{\delta}_2(n)$, where

$$\overline{\delta}_2(n) := \frac{\frac{V}{2} - \frac{3\underline{t}}{8} - \frac{(\overline{t}-\underline{t})}{16}(3 + \frac{3}{n})}{V - \frac{\overline{t}+5\underline{t}}{8} - (\overline{t} - \underline{t})(\frac{5n+1}{16n})}.$$

In Figure 3.3 I present $\overline{\delta}_2(n)$ as a function of $n = 1, ..., 2^k$.[9]

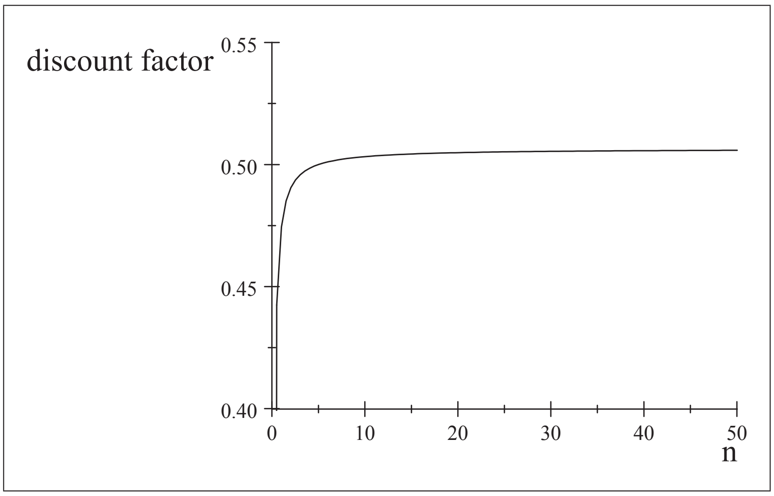

Figure 3.3. $\overline{\delta}_2(n)$ in case of relatively homogeneous consumers.

Again, the first-order derivative of $\overline{\delta}_2(n)$ with respect to n is positive. Similarly to the first collusive scheme, it becomes more difficult to sustain collusion when the quality of customer flexibility data improves. Moreover, it is easier to sustain the first collusive scheme: $\overline{\delta}_1(n) < \overline{\delta}_2(n)$ for any $n \geq 2$.[10] Liu and Serfes (2007) also get the latter result in a model where data on consumer addresses is imperfect.

I now turn to the case of relatively differentiated consumers, where

$$\overline{\delta}_1(n) := \frac{\frac{V}{2} - \frac{3\overline{t}(1+n)}{16n}}{V - \frac{5\overline{t}}{36n^2} - \frac{\overline{t}(3n+1)}{8n}}$$

---

[9] In Figures 3.3 and 3.4 I use the following values: $V = 6$, $\overline{t} = 2$ and $\underline{t} = 1$ (in the case of relatively homogeneous consumers).

[10] If $n = 1$ then $\overline{\delta}_1(n) = \overline{\delta}_2(n)$.

$$\overline{\delta}_2(n) := \frac{\frac{V}{2} - \frac{3\bar{t}(1+n)}{16n}}{V - \frac{5\bar{t}}{36n^2} - \frac{\bar{t}}{8} - \frac{\bar{t}(5n+1)}{16n}}$$

First-order derivatives of $\overline{\delta}_1(n)$ and $\overline{\delta}_2(n)$ with respect to n are positive. In Figure 3.4 I present $\overline{\delta}_2(n)$ in case of relatively differentiated consumers.[11]

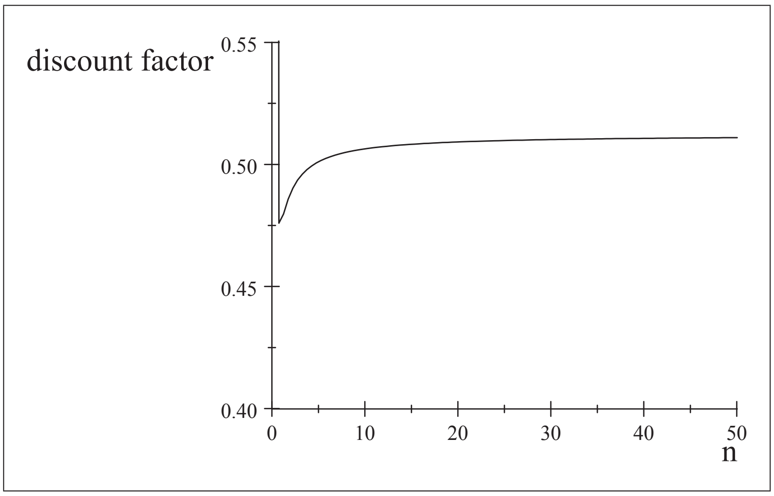

Figure 3.4. $\overline{\delta}_2(n)$ in case of relatively differentiated consumers.

In the third collusive scheme, $\overline{\delta}_3(n)$ cannot be derived analytically. I estimate $\overline{\delta}_3(n)$ for different values of the initial parameters in the model. The results show that $\overline{\delta}_3(n)$ is an increasing function of n.[12]

My results support the findings of Liu and Serfes (2007) that it becomes more difficult to sustain collusion when the quality of customer data improves. I conclude that the above result holds not only when the quality of data on consumer addresses improves, but also when the quality of data on consumer flexibility improves and firms hold perfect data on consumer locations. This result holds both when consumers are relatively homogeneous and relatively differentiated in flexibility, although the behavior of non-cooperative profits is different in the two cases. The intuition for this result is the following: As non-cooperative profits increase with the improvement in data quality when consumers are relatively homo-

[11] I do not show the graphics for $\overline{\delta}_1(n)$, because they are analogous to those already presented for $\overline{\delta}_2(n)$.

[12] For example, for $\bar{t} = 10$, I get $\overline{\delta}_3(2) = 0.477$; $\overline{\delta}_3(4) = 0.679$; $\overline{\delta}_3(8) = 0.790$; $\overline{\delta}_3(16) = 0.856$.

geneous, the punishment following deviation becomes less severe with better data quality, which makes deviation more attractive.

3.5 Conclusion

I analyze the sustainability of collusion in an infinitely repeated game depending on the quality of customer flexibility data. I follow Sapi and Suleymanova (2013) and assume that consumers are differentiated both with respect to their addresses and flexibility. Therefore, firms can price-discriminate among consumers using two types of customer data: consumer addresses and flexibility. I assume that data on consumer addresses is perfect, while data on consumer flexibility is not. In this way I depart from Liu and Serfes (2007) who assume that consumers are differentiated only in their addresses and this data is imperfect. I consider two cases with respect to consumer heterogeneity in flexibility: relatively differentiated consumers and relatively homogeneous consumers. In the former case the behavior of non-cooperative profits as a function of data quality is similar to the behavior of non-cooperative profits in Liu and Serfes: The profits first decrease and then increase. When consumers are relatively homogeneous, the non-cooperative profits increase monotonically with the improvement in data quality, which makes it more difficult to sustain collusion with the improvement in quality of customer data compared to the case of relatively differentiated consumers. My results support the findings of Liu and Serfes that collusion becomes more difficult to sustain with the improvement in data quality and show that they are also relevant for a different type of customer data, that on consumer flexibility.

3.6 Appendix

Proof Proposition 3.2. I characterize the collusive outcome under the first collusive scheme. Under collusion every firm acts as a monopolist on its own turf. If the basic consumer valuation is high enough, then every firm serves all consumers on its turf. The transport cost parameter of the consumer on the segment m indifferent between buying from Firm A and not buying is

$$V - t^m x - p_A^C(x, m; n) = 0 \Longrightarrow \widetilde{t}^m = \frac{V - p_A^C(x, m; n)}{x}, \text{ where } \widetilde{t}^m \in [\underline{t}^m, \overline{t}^m].$$

$p_A^C(x, m; n)$ is the collusive price set by Firm A for a consumer on segment m with address x, given the quality of data is k and there are $n := 2^k$ segments. Firm A serves consumers with $t \leq \widetilde{t}^m$ for any address $x < 1/2$. The reason is that in each segment consumers with relatively low transport cost parameters get positive utility when Firm A charges the collusive price, $p_A^C(x, m; n)$. As a monopolist, Firm A can extract a higher rent from them. Consumers with high transport cost, $t > \widetilde{t}^m$, choose to not buy at all, otherwise they get a negative utility. The expected profit is

$$E[\pi_A^C(x, m; n)|x < 1/2] = p_A^C(x, m; n) \Pr\{t \leq \widetilde{t}^m\} = p_A^C(x, m; n) f_t \left(\frac{V - p_A^C(x, m; n)}{x} - \underline{t}^m \right).$$

Solving the maximization problem of Firm A w.r.t. $p_A^C(x, m; n)$ yields the condition on V, which guarantees that Firm A serves all consumers for any $x < 1/2$ on any segment m:

$$\text{if } V > \underline{t}x + \frac{x}{n}(\bar{t} - \underline{t})(m + 1) \text{ then } p_A^C(x, m; n) = V - \left(\underline{t} + (\bar{t} - \underline{t})\frac{m}{n} \right) x \text{ and } \widetilde{t}^m = \overline{t}^m.$$

The strongest condition implies $x = 1/2$, yielding $V > \underline{t}/2 + (\bar{t} - \underline{t})(m + 1)/(2n)$. The collusive profit of Firm A is

$$\pi_A^C = \sum_{m=1}^{n} \int_0^{1/2} \int_{\underline{t}^m}^{\overline{t}^m} f_t p_A^C(x, m; n) dt dx = \frac{V}{2} - \frac{\underline{t}}{8} - \frac{(\bar{t} - \underline{t})(1 + n)}{16n}.$$

Since two firms are symmetric, $\pi_B^C = \pi_A^C$.

I now characterize the optimal deviation strategy. If a firm deviates, it charges a different price only on the rival's turf. Consider a deviation by Firm B. The indifferent consumer on the turf of Firm A is characterized by the equation

$$V - tx - p_A^C(x, m; n) = V - (1 - x)t - p_B^D(x, m; n),$$

where $p_B^D(x, m; n)$ is the deviation price of Firm B on Firm A's turf. The consumer indifferent between buying from Firm A charging the collusive price and Firm B charging the deviation price is

$$\widehat{t}^m = \frac{V - p_B^D(x, m; n) - \overline{t}^m x}{1 - 2x}, \text{ where } \widehat{t}^m \in [\underline{t}^m, \overline{t}^m] \text{ and } x < 1/2.$$

Firm B serves consumers with $t \leq \widehat{t}^m$ for any address $x < 1/2$. The reason is that under deviation Firm B offers a lower price than the rival and the most flexible consumers find it attractive to switch since they have a low transport cost. The expected deviation profit of Firm B is $E[\pi_B^D(x, m; n)|x < 1/2] = p_B^D(x, m; n) \Pr\{t \leq \widehat{t}^m\}$. Solving the maximization problem of Firm B I get the following condition, which guarantees that Firm B serves all consumers on Firm A's turf:

$$\begin{aligned} \text{if } V &> \underline{t}(1 - x) + \frac{(\overline{t} - \underline{t})}{n}(m(1 - x) + 1 - 2x), \\ \text{then } p_B^D(x, m; n) &= V - \left(\underline{t} + (\overline{t} - \underline{t})\frac{m}{n}\right)(1 - x) \text{ and } \widehat{t}^m = \overline{t}^m. \end{aligned}$$

The strongest condition implies $x = 0$, yielding $V > \underline{t} + (\overline{t} - \underline{t})(m + 1)/n$. The deviation profit of Firm B is

$$\begin{aligned} \pi_B^D &= \sum_{m=1}^{n} \left[\int_0^{1/2} \int_{\underline{t}^m}^{\overline{t}^m} f_t p_B^D(x, m; n) dt dx\right] + \pi_B^C(x > 1/2) \\ &= V - \frac{\underline{t}}{2} - \frac{(\overline{t} - \underline{t})(1 + n)}{4n}. \end{aligned}$$

Q.E.D.

Proof Proposition 3.3. I first characterize the collusive outcome under the second collusive scheme. Every firm acts as a monopolist on its turf and charges a monopoly price for any address on its turf which does not depend on the segment. Consider the turf of Firm A. For some $x < 1/2$ the transport cost parameter of the consumer indifferent between buying at Firm A and not buying is

$$V - tx - p_A^C(x) = 0 \Longrightarrow \widetilde{t} = \frac{V - p_A^C(x)}{x}, \text{ where } \widetilde{t} \in [\underline{t}, \overline{t}].$$

Firm A serves consumers with $t < \widetilde{t}$, because only consumers with relatively low transport costs get positive utility when Firm A charges the collusive price, $p_A^C(x)$. Its expected profit is

$$E[\pi_A^C(x)|x < 1/2] = p_A^C(x) \Pr\{t \leq \widetilde{t}\} = p_A^C(x) f_t \left(\frac{V - p_A^C(x)}{x} - \underline{t}\right).$$

I solve the maximization problem of Firm A w.r.t. $p_A^C(x)$. If the basic consumer valuation is large enough, Firm A serves all consumers on any segment on its turf:

$$\text{if } V > (2\bar{t} - \underline{t})x, \text{ then } p_A^C(x) = V - \bar{t}x \text{ and } \widetilde{t} = \bar{t}.$$

The strongest condition implies $x = 1/2$, yielding $V > \bar{t} - \underline{t}/2$. The collusive profit of Firm A is

$$\pi_A^C = \int_0^{1/2} \int_{\underline{t}}^{\bar{t}} f_t p_A^C(x) dt dx = \frac{V}{2} - \frac{t}{8}.$$

Since two firms are symmetric, $\pi_B^C = \pi_A^C$.

Every firm deviates through acquiring customer flexibility data of quality k and discriminates among flexibility segments on both turfs. Consider the deviation by Firm B. The consumer indifferent between buying from Firm A or Firm B with some $x < 1/2$ on segment m is given by the equation

$$V - tx - p_A^C(x) = V - (1 - x)t - p_B^D(x, m; n),$$

where $p_B^D(x, m; n)$ is the deviation price of Firm B on Firm A's turf. Plugging in $p_A^C(x)$ into the above equation yields

$$\widehat{t}^m = \frac{V - p_B^D(x, m; n) - \bar{t}x}{1 - 2x}, \text{ where } \widehat{t}^m \in [\underline{t}^m, \bar{t}^m].$$

Firm B serves consumers with $t \leq \widehat{t}^m$, because their utility is higher when they buy from Firm B than from Firm A. The expected deviation profit of Firm B is $E[\pi_B^D(x, m; n)|x < 1/2] = p_B^D(x, m; n) \Pr\{t \leq \widehat{t}^m\}$. I solve the maximization problem of Firm B w.r.t. $p_B^D(x, m; n)$ and get that if the basic consumer valuation is large enough, then Firm B serves all consumers for some $x < 1/2$ and some m:

$$\begin{aligned} \text{if } V &> \bar{t}x + (\underline{t} + (\bar{t} - \underline{t})\frac{m+1}{n})(1 - 2x), \\ \text{then } p_B^D(x, m; n) &= V - \bar{t}x - \left(\underline{t} + (\bar{t} - \underline{t})\frac{m}{n}\right)(1 - 2x) \text{ and } \widehat{t}^m = \bar{t}^m. \end{aligned}$$

The strongest condition implies $x = 0$, yielding $V > \underline{t} + (\bar{t} - \underline{t})(m + 1)/n$.

I now compute the deviation prices of Firm B on it own turf. The indifferent consumer on the segment m for some $x > 1/2$ is characterized by the equation: $V - t(1-x) - p_B^D(x, m; n) =$

0. The transport cost parameter of the consumer on the segment m indifferent between buying from Firm B and not buying is

$$\widehat{t}^m = \frac{V - p_B^D(x,m;n)}{1-x}, \text{ where } \widehat{t}^m \in [\underline{t}^m, \overline{t}^m].$$

Firm B stays a monopolist and serves consumers with $t \leq \widehat{t}^m$. The expected deviation profit of Firm B on its own turf is $E[\pi_B^D(x,m;n)|x > 1/2] = p_B^D(x,m;n)\Pr\{t \leq \widehat{t}^m\}$. I solve the maximization problem w.r.t. $p_B^D(x,m;n)$ and get that if the basic consumer valuation is large enough, then Firm B serves all consumers on some segment m and some address $x > 1/2$:

$$\begin{aligned} \text{if } V &> \left(\underline{t} + (\overline{t} - \underline{t})\frac{m+1}{n}\right)(1-x), \\ \text{then } p_B^D(x,m;n) &= V - \left(\underline{t} + (\overline{t} - \underline{t})\frac{m}{n}\right)(1-x). \end{aligned}$$

The strongest condition implies $x = 1/2$, yielding $V > \underline{t}/2 + (\overline{t} - \underline{t})(m+1)/(2n)$. The deviation profit of Firm B is

$$\begin{aligned} \pi_B^D &= \sum_{m=1}^{n} \int_0^{1/2} \int_{\underline{t}^m}^{\overline{t}^m} f_t p_B^D(x,m;n) dt dx + \sum_{m=1}^{n} \int_{1/2}^{1} \int_{\underline{t}^m}^{\overline{t}^m} f_t p_B^D(x,m;n) dt dx \\ &= V - \frac{\overline{t} + 3\underline{t}}{8} - \frac{3(\overline{t} - \underline{t})(1+n)}{16n}. \end{aligned}$$

Q.E.D.

Proof Proposition 3.4. Under the third collusive scheme firms agree not to acquire flexibility data and charge uniform competitive prices. I consider the case of the model with relatively differentiated consumers. Collusive prices and profits are a special case of non-cooperative equilibrium, with $k = 0$. In the equilibrium both firms follow a market-sharing strategy, where Firm A sets the price $p_A^C(x|x < 1/2) = 2\overline{t}(1-2x)/3$ on its own turf and serves consumers with $t \in [\overline{t}/3, \overline{t}]$. On the rival's turf it charges $p_A^C(x|x > 1/2) = \overline{t}(2x-1)/3$, and serves consumers with $t \in [0, \overline{t}/3]$. Firm B charges symmetric prices. The collusive profit is $\pi_i^C = 5\overline{t}/36$, $i = A, B$.

Now, I turn to the deviation prices and profits. Suppose that Firm B deviates on A's turf by acquiring customer data ($n \geq 2$) and charging discriminatory prices. For $x < 1/2$, the consumer indifferent between buying from Firm A or from Firm B is given by the condition:

$$V - tx - p_A^C(x|x < 1/2) = V - (1-x)t - p_B^D(x, m; n)$$

where $p_B^D(x, m; n)$ is the deviation price of Firm B on Firm A's turf. Plugging $p_A^C(x|x < 1/2) = 2\bar{t}(1-2x)/3$ into the above equation I obtain the transport cost parameter of the indifferent consumer:

$$\widehat{t^m} = \frac{2\bar{t}(1-2x) - 3p_B^D(x, m; n)}{3(1-2x)}, \text{ where } \widehat{t^m} \in [\underline{t}^m, \bar{t}^m].$$

On the rival's turf Firm B serves consumers with relatively low transport cost, $t \leq \widehat{t^m}$. Since consumers are relatively differentiated, there are always consumers who can switch brands costlessly. Therefore Firm B targets only the most flexible consumers of the rival. The expected deviation profit of Firm B is $E[\pi_B^D(x, m; n)|x < 1/2] = p_B^D(x, m; n) \Pr\{t \leq \widehat{t^m}\}$. I solve the maximization problem of Firm B w.r.t. $p_B^D(x, m; n)$ taking into account that $\widehat{t^m} \in [\underline{t}^m, \bar{t}^m]$. I obtain the following results:

i) If $m < 2n/3 - 1$, *then* $p_B^D(x, m; n) = (\bar{t}/3)(2 - 3m/n)(1-2x)$, *each firm follows a monopolization strategy on segment* m *and serves all consumers there.*

ii) If $m > 2n/3 - 1$, *then* $p_B^D(x, m; n) = (\bar{t}/6)(2 - 3(m-1)/n)(1-2x)$, *Firm* B *follows a market-sharing strategy on the segment* m *and serves only consumers with* $t \in [\underline{t}^m, \bar{t}/3 + \underline{t}^m/2]$.

Firm B optimally deviates on its own turf as well. For $x > 1/2$ the transport cost parameter of the indifferent consumer is derived from the following condition:

$$V - tx - p_A^C(x|x > 1/2) = V - (1-x)t - p_B^D(x, m; n)$$

where $p_B^D(x, m; n)$ is the deviation price of Firm B on its own turf. Plugging $p_A^C(x|x > 1/2) = \bar{t}(2x-1)/3$ into the above equation I obtain:

$$\widehat{t^m} = \frac{3p_B^D(x, m; n) - \bar{t}(2x-1)}{3(2x-1)}, \text{ where } \widehat{t^m} \in [\underline{t}^m, \bar{t}^m].$$

On its own turf Firm B serves consumers with relatively high transport cost, $t \geq \widehat{t^m}$. Since consumers are relatively differentiated, it prefers to lose the most flexible consumers to the rival because serving them is not profitable. The expected deviation profit of Firm B is $E[\pi_B^D(x,m;n)|x > 1/2] = p_B^D(x,m;n)\Pr\{t \geq \widehat{t}^m\}$. I solve the maximization problem of Firm B w.r.t. $p_B^D(x,m;n)$ taking into account that $\widehat{t}^m \in [\underline{t}^m, \overline{t}^m]$ and get that:

i) If $m < 2 - n/3$, *then* $p_B^D(x,m;n) = (\overline{t}/6)(3m/n+1)(2x-1)$, *Firm B follows a market-sharing strategy on the segment* m *and serves only consumers with* $t \in [\overline{t}^m/2 - \overline{t}/6, \overline{t}^m]$.

ii) If $m > 2 - n/3$, *then* $p_B^D(x,m;n) = (\overline{t}/3)(2x-1)(3(m-1)/n+1)$ *and Firm B follows a monopolization strategy on the segment* m.

When $n = 2$ ($k = 1$) Firm B adopts a different deviation strategy. Since consumers are relatively differentiated, it prefers to lose the most flexible consumers, whose transport cost parameter is $t \in [0, \overline{t}/12]$, to Firm A. Hence, if $m = 1$, Firm B adopts a market-sharing strategy and charges the price $p_B^D(x,1;2) = (5\overline{t}/12)(2x-1)$ to consumers with $t \in [\overline{t}/12, \overline{t}/2]$. If $m = 2$, Firm B adopts a monopolization strategy and charges the price $p_B^D(x,2;2) = (5\overline{t}/6)(2x-1)$ to consumers with $t \in [\overline{t}/2, \overline{t}]$. The total deviation profit, when $n = 2$ ($k = 1$), is

$$\pi_B^D = \int_0^{1/2}\int_0^{\overline{t}/3} f_t p_B^D(x,1;2)dtdx + \int_0^{1/2}\int_{\overline{t}/2}^{7\overline{t}/12} f_t p_B^D(x,2;2)dtdx$$

$$+\int_{1/2}^{1}\int_{\overline{t}/12}^{\overline{t}/2} f_t p_B^D(x,1;2)dtdx + \int_{1/2}^{1}\int_{\overline{t}/2}^{\overline{t}} f_t p_B^D(x,2;2)dtdx = \frac{17\overline{t}}{96}$$

If $n \geq 4$ ($k \geq 2$), then Firm B follows a monopolization strategy on every segment on its own turf. Thus, the total deviation profit is

$$\pi_B^D = \sum_{m=1}^{[\frac{2n}{3}-1]} \int_0^{1/2}\int_{\underline{t}^m}^{\overline{t}^m} f_t \frac{\overline{t}}{3}(2 - \frac{3m}{n})(1-2x)dtdx + \sum_{m=[\frac{2n}{3}]}^{n} \int_0^{1/2}\int_{\underline{t}^m}^{\overline{t}(\frac{1}{3}+\frac{m-1}{2n})} f_t \frac{\overline{t}}{6}(2-3\frac{m-1}{n})(1-2x)dtdx$$

$$+\sum_{m=1}^{n}\int_{1/2}^{1}\int_{\underline{t}^m}^{\overline{t}^m} f_t \frac{\overline{t}}{3}(3\frac{m-1}{n}+1)(2x-1)dtdx$$

Q.E.D.

Chapter 4

Full Versus Partial Collusion Among Brands and Private Label Producers

Co-authored by Christian Wey

4.1 Introduction

We present a Salop-circle model which captures market competition between branded products and a private label substitute.[1] All products are differentiated in the horizontal dimension. In addition, the private label good is assumed to be inferior in the vertical dimension. We use an infinitely repeated game approach to examine under which circumstances full collusion of heterogeneous firms is easier or harder to sustain than partial collusion. In the former case all firms (the brands and their private label substitute) collude, while in the latter case only the brands form a self-enforcing cartel.

Private label products (also called *store brands*) encompass all merchandise sold under a retailer's brand. Their market share has risen significantly and today private labels are an integral part of almost all retail markets. For instance, based on 2011 sales data of

[1] Another application of our analysis are pharmaceutical products where brands and their generic equivalents compete against each other after the end of patent protection (see Frank and Salkever, 1997).

Nielsen private labels accounted for 42 per cent in the United Kingdom.[2] Private labels were initially part of a low-price, low-quality strategy allowing retailers to compete for price-sensitive consumers (Hassan and Monier-Dilhan 2006). These budget private labels were often designed as *me too* products and were positioned at the lower end of the quality and price spectrum. Private labels were especially successful in markets where no strong national brands were present (European Commission, 2011). However, over recent years they have grown in the segment of added-value and *premium* products.[3] Based on GfK German retailing data, Inderst (2013, Figure 3, p. 14) reports that over the period 2007-2012 budget private labels' market share stayed put at around 25 per cent, while the market share of premium private labels increased from 9 per cent to 12.9 per cent over the same period.

The rise of private label products has been explained by cost savings, buyer power reasons and retailer differentiation.[4] The role private labels play within the context of collusion has been largely neglected so far.[5] This is surprising because collusion is an

[2] Market shares are calculated based on the turnover of fast-moving consumer goods, excluding fresh food. The market shares differ significantly across countries (for instance, Spain: 39%, Germany and Portugal 32%, while Greece has the lowest with 10%). Market shares have been increasing steadily. According to European Commission (2011), from 2003 to 2009 their share increased by 2-7 percentage points in Western and Southern Europe (except Spain) and by 10-26 percentage points in Spain and Central Europe. Inderst (2013) provides a survey of these developments.

[3] For instance, the German supermarket chain Real offers its own premium labels in many product categories.

[4] Hoch and Banerji (1993) have shown that cost savings can be so large that private labels may generate even higher profit margins than the respective national brands. Private labels can substantially enhance a retailer's bargaining position vis-à-vis brand manufacturers because it enhances their outside option (Mills, 1998, Bontems, Monier-Dilhan, and Requillart, 1999, and Steiner, 2004). Moreover, private labels can increase retailer differentiation as retailers would otherwise carry the same assortment of branded goods (Gabrielsen and Sörgard, 2007).

[5] An exception is Steiner (2004) who warns explicitly that the issue of collusion between private label goods and national brands may become more of an issue in the future. Interestingly, the issue of collusion

ongoing issue among food manufacturers. In Germany, for instance, three cartel cases with food manufacturers involved were decided recently. The coffee roasters' cartel was a partial cartel in which only brand coffee makers were found guilty of forming a cartel, while there was no evidence found that private label producers (mainly store brands of retailers Aldi and Lidl) have participated in the cartel (Bundeskartellamt, 2009). The coffee roasters' cartel is remarkable, because it lasted for many years even though it did not include the private labels which have a market share of 17 per cent in the German coffee market in 2011 (Bundeskartellamt, 2014a, p. 208). In the sausage cartel, to the opposite, private label producers participated in the cartel which included almost all branded sausage producers (see Bundeskartellamt, 2014b). The confectionery manufacturers' cartel only included six branded products (see Bundeskartellamt, 2013).[6]

Behind this background our main research questions are the following: *First*, how do horizontal product differentiation and the private label's (vertical) quality affect the stability of full and partial collusion (the former being an all-encompassing cartel, while the latter only includes the branded goods)? And relatedly: When is a more homogenous cartel among branded manufacturers more likely to form than a heterogeneous cartel which also includes private label substitutes? *Second*, how is the private label producer's incentive to close the quality gap towards the branded goods affected by market conduct which can be competitive, partially collusive or fully collusive?

We analyze a Salop circle model with three firms that differ in their (vertical) quality parameter. Two out of three firms are high-quality brand producers, while the third firm is a private label producer that has an inferior quality. We are interested to analyze whether

---

between private labels and branded goods does not play a major role in recent retailing sector inquires by competition authorities (see, e.g., Bundeskartellamt, 2014a, Competition Commission, 2008, European Commission, 1999).

[6] It should be noted that cartel cases are decided on explicit evidence of cartel formation. The question, therefore, whether or not private label producers participated in the cartel via tacit collusion was not decided in those cases where only brand manufacturers found guilty.

collusion among three heterogeneous firms is easier to sustain than partial collusion among the two brand producers. We use an infinitely repeated game approach to examine the stability of collusion. We first show that the brand producers' incentive constraint is critical to obtain full collusion over partial collusion, whenever nonparticipation of one firm leads to non-cooperative market conduct. In those instances, the private label firm always joins the brand producers for a full collusion outcome, given that full collusion is incentive compatible for the brand manufacturers. If, however, nonparticipation of the private label firm induces the brand manufacturers to form a partial cartel (which is always better than non-cooperative conduct), then the private label may prefer partial collusion over full collusion. This is, *ceteris paribus*, more likely to be the case, the lower the intensity of competition (i.e., the higher the horizontal product differentiation) and the larger the (vertical) quality gap between the private label and the branded goods. Thus, a private label firm is more likely to join the branded goods producers to form an all-encompassing cartel the higher the quality of the private label good and the more intense competition are.

We also show that the incentives to increase the private label's quality are the largest under full collusion with partial collusion and non-cooperative behavior following in that order. The incentives are further enhanced by the prospect of making full collusion feasible in the first place. There are two reasons why a quality upgrade of the private label good can trigger full collusion. *First*, it relaxes the incentive constraint for the brand producers, and *second*, it makes it more likely that the private label firm prefers full collusion over partial collusion.

Our paper contributes to the collusion literature that deals with cartel stability when firms' are heterogeneous.[7] Häckner (1994) shows that an all-inclusive cartel is harder to

[7] Selten (1973) analyzes cartel stability as a coalition formation process (i.e., without referring to an infinitely repeated game context). He assumes homogeneous products and Cournot competition. Full cartelization is only possible when there are few firms. See Prokop (1999) for a related approach within a model of price competition, where it is also shown that the chance of full cartelization is very much limited. For a survey, see Bos and Harrington (2010).

sustain when products become more vertically differentiated.[8] Based on a spatial model of horizontal product differentiation Ross (1992) argues that increased product differentiation could enhance cartel stability (see also Chang, 1991).[9] Given those results, opposing forces are present in a model that combines variety-differentiated products with (vertical) quality differentiation. In addition, partial collusion has not been addressed in those works.[10] Closely related is Bos and Harrington (2010) who analyze the sustainability of collusion (full and partial) within an infinitely repeated game framework. Their focus is on capacity asymmetries among firms, while firms' products are homogenous. Overall, they show that full collusion is harder to sustain, when firms become more asymmetric (with regard to their capacity). Moreover, smaller firms are more likely to stay out of the cartel giving rise to partial collusion among the largest firms in the market. We apply the same stability analysis as they do; namely, we suppose that nonparticipation of a firm in the all-encompassing cartel will keep collusion among the remaining firms if it is profitable for them.[11]

We also contribute to the economic analysis of private labels (for a survey, see Berges-Sennou, Bontems, and Requillart, 2004). Price effects and product positioning incentives were analyzed in Mills (1995) and Bontems, Monier-Dilhan, Requillart (1999), and Gabrielsen and Sörgard (2007).[12] Those works focused on the strategic effects within a vertical relations setting without considering the collusion problem. Empirical works have

---

[8] A related result is obtained in Rothschild (1992).

[9] Thomadsen and Rhee (2007) show that collusion is always harder to sustain the more differentiated the products if costs of forming the cartel are sufficiently large.

[10] Our model builds on Economides (1989, 1993) which are early models (of the Hotelling and Salop type, respectively) with both horizontal and vertical product differentiation.

[11] The debate about how to formalize a cartel's stability in case of heterogeneous firms is ongoing (see Bos and Harrington, 2010, for a survey). The impact of cost asymmetries in association with an indivisible cost of collusion is analyzed in Ganslandt, Persson, and Vasconcelos (2012).

[12] Choi and Coughlan (2006) show (disregarding the vertical relation problem) that a private label should position close to a strong (weak) national brand when its quality is high (low).

shown ambiguous price effects of private labels on branded substitutes. Quite interestingly, Putsis (1997) and Cotterill and Putsis (2000) have provided evidence that brands' prices decreased after the introduction of private label substitutes. In contrast, Ward et al. (2002) show a positive association of private labels' market shares and branded products' prices.[13] A similar relationship is uncovered in Frank and Salkever (1997) who investigated price responses of branded pharmaceuticals after patent protection expired and generic substitutes entered the market.

The chapter proceeds as follows. In Section 4.2 we present the model setup and Section 4.3 provides the equilibrium analysis. In Section 4.4 we analyze the private label producer's quality incentives and we show how collusion can be detected from market data. Section 4.5 provides further extensions and discussion. Finally, Section 4.6 concludes.

4.2 The Model

We specify a variant of the Salop circle model (Salop, 1979) which combines horizontal product differentiation (as a measure of overall competition intensity) and vertical product differentiation (which mirrors the inferior quality of the private label vis-à-vis the branded goods). Let there be three firms ($j = 0, 1, 2$) located equidistantly on the unit circle. Firms 1 and 2 produce two brands with (vertical) quality index s_i, where $i = 1, 2$.[14] Both firms are horizontally differentiated and they are located at $x_1 = 1/3$ and $x_2 = 2/3$, respectively. We refer to these goods as brands 1 and 2, respectively. Firm 0 is located at $x_0 = 0$ on the unit circle and produces a private label product which is a horizontally differentiated variant, but of a lower (vertical) quality $s_0 \leq s_i$ for $i = 1, 2$. We set production costs to zero.[15]

---

[13] Bontemps, Orozco, and Requillart (2008) provide related evidence for France.

[14] We use the index i to refer only to the brands $i = 1, 2$, whereas the index j is used to refer to all firms $j = 0, 1, 2$.

[15] A three-firms Salop model is also used in Rasch and Wambach (2009) to analyze the effect of a two-firm merger and internal-decision making rules on cartel stability. Yet, in their model, all products have the same

Consumers are distributed uniformly along the unit circle with mass of one. Each consumer buys at most one unit of the good. A consumer's position x on the unit circle represents her most preferred product variant in the horizontal dimension. The utility of a consumer with address $x \in [0, 1]$ buying from firm $i = 1, 2$ is given by

$$U_i^x = s_i - t\,|x_i - x| - p_i, \tag{4.1}$$

where p_i is firm i's price. According to (4.1), we consider a linear transport cost function, where $t > 0$ is the exogenously given transport cost parameter and s_i is the (vertical) quality index.[16] Correspondingly, the utility of a consumer with address x buying the private label product is given by

$$U_0^x = s_0 - \min\{tx; t(1-x)\} - p_0, \tag{4.2}$$

where $s_0 \leq 1$ and p_0 is the price charged by firm 0. We set $s_1 = s_2 = 1$ and define $s := s_0$. Thus both brands are assumed to be of the same quality and their quality is higher than the private label's quality. The quality gap between the brands and the private label is given by $1 - s \geq 0$. Consumers only buy a product if their utility is not negative.

We consider an infinitely repeated price competition game to study firms' collusion incentives. In the stage game all firms set their prices simultaneously. All firms have the same discount factor $\delta \in [0, 1]$. In the infinitely repeated game we focus on trigger strategies with *Nash reversal* in the punishment phase.[17]

We consider two types of collusion: *i)* full collusion (FC), where all three firms collude, and *ii)* partial collusion (PC) where only firms 1 and 2 collude, while firm 0 behaves

vertical quality.

[16] See Economides (1989) for a similar approach to combine both horizontal and vertical product differentiation within a Salop model.

[17] We use a grim strategy as in the seminal paper of Friedman (1971) to derive firms' collusion incentives. While this is standard practice in the tacit collusion literature (see Mas-Colell, Whinston, and Green, 1995, Chap. 12D), optimal punishments (so-called stick-and-carrot strategies) can be more effective in sustaining collusion (see Abreu, 1986, 1988, and Abreu, Pearce, and Stacchetti 1986). Both approaches can be expected to lead to the same qualitative results (see Häckner, 1996, and Chang, 1991).

non-cooperatively. In addition, we denote by N the case that all firms behave always non-cooperatively.

We analyze the stability of collusion under full and partial collusion. We denote by π_i^N the non-cooperative (stage game) profit of firm i, by π_i^C the collusive (stage game) profit of firm i and by $\pi_i^{D,C}$ the deviation profit of a colluding firm i, where the superscript C refers either to the partial collusion case or the full collusion case; i.e., $C = FC, PC$. Given trigger strategies with Nash-reversal, firm i has no incentive to deviate from the collusive behavior if and only if the discount factor is large enough; i.e., if

$$\delta \geq \delta_i^C := \frac{\pi_i^{D,C} - \pi_i^C}{\pi_i^{D,C} - \pi_i^N} \tag{4.3}$$

holds. We impose the following parameter restrictions which ensure that the market is always covered in equilibrium.

Assumption 4.1. *We restrict the analysis to all parameter pairs* (t, s) *which fulfill the conditions* $1 \geq s \geq \max\{1 - 5t/6, t/2, 13t/6 - 2\}$. *The minimal possible values of* t and s are $t_{\text{min}} = 6/11$ *and* $s_{\text{min}} = 3/8$, *while the maximal possible values are* $t_{\text{max}} = 18/13$ *and* $s_{\text{max}} = 1$.

Assumption 4.1 specifies the feasible set of parameters we are considering throughout the analysis (all conditions are derived in the Appendix). Specifically, restriction $s \geq 1 - 5t/6$ ensures that the equilibrium market share and price of the private label good are positive under non-cooperative behavior. Conditions $s \geq t/2$ and $s \geq 13t/6 - 2$ ensure that the market is covered under full and partial collusion, respectively. Specifically, condition $s \geq t/2$ implies that the deviation price of the private label firm under full collusion is lower than its collusive price. Finally, $t \geq t_{\text{min}} = 6/11$ makes sure that the deviating firm under (both partial and full) collusion realizes a market share which is less than 100 per cent.[18]

Nash Equilibrium of the Stage Game. Before we analyze the infinitely repeated game, we solve the stage game to derive π_j^N, for $j = 0, 1, 2$. We first derive the demand functions.

[18]The remaining maximal and minimal values of t and s follow from the restrictions that constrain s as stated in the first sentence of Assumption 4.1.

Note that firms are located equidistantly on the unit circle. Denote the indifferent consumer between firms 0 and 1 by x_0. Given prices p_0 and p_1 the indifferent consumer between firms 0 and 1 is given by

$$s - p_0 - tx_0 = 1 - p_1 - t\left(\frac{1}{3} - x_0\right)$$

which gives her location on the segment $x \in (0, 1/3)$:

$$x_0 = \frac{1}{2t}\left(s + \frac{t}{3} - p_0 + p_1 - 1\right) \text{ for } x_0 \in \left(0, \frac{1}{3}\right). \tag{4.4}$$

Similarly, the indifferent consumer between firms 1 and 2 (denoted by x_1) is obtained from

$$1 - p_1 - t\left(x_1 - \frac{1}{3}\right) = 1 - p_2 - t\left(\frac{2}{3} - x_1\right),$$

which gives her location on the segment $x \in (1/3, 2/3)$:

$$x_1 = \frac{t - p_1 + p_2}{2t} \text{ for } x_1 \in \left(\frac{1}{3}, \frac{2}{3}\right). \tag{4.5}$$

Finally, the indifferent consumer between firms 2 and 0 (denoted by x_2) is given by

$$1 - p_2 - t\left(x_2 - \frac{2}{3}\right) = s - p_0 - t(1 - x_2)$$

which gives her location on the segment $x \in (2/3, 1)$:

$$x_2 = \frac{1}{2t}\left(1 - s + \frac{5t}{3} + p_0 - p_2\right) \text{ for } x_2 \in \left(\frac{2}{3}, 1\right). \tag{4.6}$$

Using (4.4), (4.5), and (4.6) we can write firm i's demand D_i as

$$D_i = \begin{cases} x_0 + 1 - x_2, & \text{if } i = 0 \\ x_i - x_{i-1}, & \text{if } i = 1, 2. \end{cases} \tag{4.7}$$

Using (4.7) we can solve the firms' maximization problems $\max_{p_j \geq 0} \pi_j = p_j D_j$ simultaneously to obtain the equilibrium prices and firms' equilibrium profits under non-cooperative behavior.

Proposition 4.1. *Suppose that all firms behave non-cooperatively. We obtain the following equilibrium values:*

i) Prices: $p_i^N = (3(1-s) + 5t)/15$, *for* $i = 1, 2$, *and* $p_0^N = (6(s-1) + 5t)/15$. *Moreover,* $p_1^N = p_2^N \geq p_0^N$ *if* $s \leq 1$ *(with equality holding for* $s = 1$*).*

ii) Profits: $\pi_i^N = (5t + 3(1-s))^2/(225t)$, *for* $i = 1, 2$, *and* $\pi_0^N = (6(s-1) + 5t)^2/(225t)$; *moreover,* $\pi_1^N = \pi_2^N \geq \pi_0^N$ *(with equality holding for* $s = 1$*).*

iii) Locations of the indifferent consumers: $x_0^N = 1/6 - (1-s)/(5t)$, $x_1^N = 1/2$, *and* $x_2^N = 5/6 + (1-s)/(5t)$.

Proof. See Appendix.

Parts *i)* and *ii)* of Proposition 4.1 state firms' prices and profits, respectively. The prices and profits of the brand producers decrease when the quality of the private label increases, while the opposite holds for the private label producer. As long as the private label good is of a strictly lower quality than the branded goods ($s < 1$), the brand producers realize higher profits than the private label producer. Part *iii)* of Proposition 4.1 shows that the private label producer 0 serves the consumers with addresses $(0, x_0^N) \cup (x_2^N, 1)$, while the branded manufacturers 1 and 2 serve the consumers on the intervals $(x_0^N, 1/2)$ and $(1/2, x_2^N)$, respectively. If the quality of the private label good is inferior, $s < 1$, then firm 0 serves less consumers than firms 1 and 2, despite the fact that it charges the lowest price. The relatively low quality of the private label good reduces its equilibrium demand. This benefits the brands, because they can sell their products at a higher price and also enjoy a larger equilibrium demand. However, if firm 0's quality increases, firms 1 and 2 face stronger competition and reduce their prices. If $s = 1$, then all three firms are homogeneous in the vertical dimension and they share the market equally.

4.3 Equilibrium Analysis

We next analyze the infinitely repeated game for the cases of full collusion and partial collusion. We then compare our results and we relate them to the case where firms always behave non-cooperatively. Finally, we compare the stability of both types of collusion.

4.3.1 Full Collusion

Assume that all firms in the market collude (case FC). Then all firms maximize their joint profit $\pi^{FC} := \sum_{i=0}^{2} \pi_i$ and charge collusive prices.[19] The maximization problem is given by

$$\max_{p_0,\, p_1,\, p_2 \geq 0} \pi^{FC} = p_0 D_0 + p_1 D_1 + p_2 D_2.$$

We impose the following constraints: *First*, the market is always covered and each firm obtains a strictly positive market share. *Second*, all consumers realize a nonnegative utility when buying one of the offered products. In the Appendix we show that these constraints pin down the equilibrium under full collusion. The branded firms set the same price (they are both symmetric) such that the indifferent consumer located at $x = 1/2$ gets a utility of zero. The private label firm then sets a price such that the indifferent consumers located at $x = 1/6$ and $x = 5/6$ obtain also a utility of zero and are therefore indifferent between buying the private label good or the next branded good. In addition, we also derive the optimal deviation prices where we impose that the maximal market share of the deviating firm is less than 100 per cent. The following proposition states the fully collusive prices and profits as well as the deviation prices and deviation profits.

Proposition 4.2. *Consider collusion by all three firms. We obtain the following equilibrium values:*

i) Prices: $p_i^{FC} = 1 - t/6$, *for* $i = 1, 2$, *and* $p_0^{FC} = s - t/6$, *so that* $p_0^{FC} \leq p_i^{FC}$ *holds (with equality holding at* $s = 1$*).*

ii) Profits: $\pi_i^{FC} = 1/3 - t/18$, *for* $i = 1, 2$ *and* $\pi_0^{FC} = s/3 - t/18$.

iii) Demands: Firm 0 *serves consumers located at* $[0, 1/6) \cup (5/6, 1]$. *Firm* 1 *serves consumers located at* $(1/6, 1/2)$ *and firm* 2 *at* $(1/2, 5/6)$, *respectively.*

iv) Deviation by firm i, $i = 1, 2$*: The deviation prices and profits are* $p_i^{D,FC} = t/12 + 1/2$ *and* $\pi_i^{D,FC} = (t+6)^2 / (144t)$, *respectively.*

---

[19] We follow Donsimoni (1985) and Athey and Bagwell (2001) who select the collusive outcome which maximizes joint profits (see Bos and Harrington, 2010, and Thomadsen and Rhee, 2007, for discussions of this issue and for related literature).

v) Deviation by firm 0*: The deviation price and profits are* $p_0^{D,FC} = t/12 + s/2$ *and* $\pi_0^{D,FC} = (6s + t)^2/(144t)$*, respectively.*

Proof. See Appendix.

Parts *i)* and *ii)* of Proposition 4.2 state the collusive prices and profits when all three firms collude. The firms charge higher prices than in the non-cooperative case. According to part *iii)* of Proposition 4.2 each firm's market share is 1/3. It also implies that the demand for the private label good increases compared to the non-cooperative case, while the market share of the brands is reduced accordingly.

Parts *iv)* and *v)* give the deviation prices and profits of the firms. The deviating firm undercuts its rivals by setting a lower price; it then obtains a higher market share and a higher profit. The number of consumers served depends on the transport cost parameter. We assume that the transport cost is large enough, so that the deviating firm obtains a market share of less than 100 per cent (which is ensured by $t \geq 6/11$; see Assumption 4.1).

4.3.2 Partial Collusion

In case of partial collusion (PC), firms 1 and 2 collude, while firm 0 behaves non-cooperatively. Let p_i^{PC} denote the price of firm i, for $i = 1, 2$, and let p_0^{PC} be the non-cooperative price set by firm 0 under partial collusion. The colluding brands maximize their joint profit and their maximization problem is given by

$$\max_{p_1, p_2 \geq 0} \pi^{PC} = p_1 D_1 + p_2 D_2,$$

while the maximization problem of firm 0 is

$$\max_{p_0 \geq 0} \pi_0 = p_0 D_0.$$

Solving the maximization problems gives rise to a set of first-order conditions which determine the equilibrium outcome.

Proposition 4.3. *Consider partial collusion between firms* 1 *and* 2*, while firm* 0 *behaves non-cooperatively. We obtain the following equilibrium values:*

i) Prices: $p_i^{PC} = 5t/9 + (1-s)/3$, *for* $i = 1, 2$, *and* $p_0^{PC} = 4t/9 - (1-s)/3$, *so that* $p_0^{PC} < p_i^{PC}$ *holds always.*

ii) Profits: $\pi_0^{PC} = (4t - 3(1-s))^2/(81t)$ *and* $\pi_i^{PC} = (5t + 3(1-s))^2/(162t)$, *for* $i = 1, 2$.

iii) Demands: Firm 0 *serves consumers located at* $[0, x_0^{PC}) \cup (x_2^{PC}, 1]$, *where* $x_0^{PC} > x_0^N$ *and* $x_2^{PC} < x_2^N$. *Firm* 1 *serves consumers located at* $(x_0^{PC}, 1/2)$ *and firm* 2 *at* $(1/2, x_2^{PC})$, *respectively. Moreover,* $x_0^{PC} = (4t - 3(1-s))/(18t)$, $x_1^{PC} = 1/2$, and $x_2^{PC} = (14t + 3(1-s))/(18t)$.

iv) If firm i, $i = 1, 2$, *deviates from partial collusion, then its price and profits are* $p_i^{D,PC} = (1-s)/4 + 5t/12$ *and* $\pi_i^{D,PC} = ((1-s)/4 + 5t/12)^2/t$, *respectively.*

Proof. See Appendix.

Part *iii)* of Proposition 4.3 says when the two brands collude, they reduce their market shares and serve less consumers compared to the non-cooperative case. Part *iv)* of Proposition 4.3 states that a brand could deviate from the collusive agreement by charging a lower price than those set by the rival brand and the private label. Such a deviation increases its profits.

4.3.3 Comparison of Results

Comparing the results derived so far we can order firms' prices, demands and profits under the three different types of conduct (non-cooperative, partially collusive and fully collusive).

Corollary 4.1. *Comparing the equilibrium prices under non-cooperation, full and partial collusion, we get:* $p_j^{FC} > p_j^{PC} > p_j^N$, *for* $j = 0, 1, 2$.

Proof. Follows directly from comparing the equilibrium prices as stated in Propositions 4.1-4.3.

Corollary 4.1 states that prices are increasing when firms' conduct becomes more collusive. Prices are maximal when there is full collusion. They remain higher under partial collusion than under non-cooperative behavior. Combining the latter observation with the fact that the private label firm sets a lower price than the branded goods producers in case of a partial cartel (see Proposition 4.3), we get that the brands' prices serve as an *umbrella* such that the private label's price increases above the fully non-cooperative price under partial collusion (see Bos and Harrington, 2010, for related results).[20]

Corollary 4.2. *Comparing the equilibrium demands under non-cooperation, full and partial collusion, we get the following orderings:*

i) $D_0^{FC} > D_0^N$, $D_0^{PC} > D_0^N$ *and* $D_0^{PC} > D_0^{FC}$ *for* $s > 1 - t/3$.

ii) $D_i^N > D_i^{FC}$, $D_i^N > D_i^{PC}$ *and* $D_i^{FC} > D_i^{PC}$, *for* $s > 1 - t/3$ *with* $i = 1, 2$.

iii) $D_0^{FC} = D_1^{FC} = D_2^{FC} = 1/3$.

Proof. Follows directly from calculating firms' demands (4.7) by using the locations of the indifferent consumers as stated in Propositions 4.1-4.3.

By comparing the equilibrium demands, we notice that the brands serve the highest share of the market in the non-cooperative case. Each brand serves more than one third of the market. Under full collusion all firms share the demand equally. Under partial collusion, brands charge a lower price than under full collusion and serve less consumers; thus, their market shares become smaller than one third.

Corollary 4.3. *Comparing the equilibrium profits under non-cooperation, full and partial collusion, we get the following orderings:*

i) $\pi_i^{FC} > \pi_i^{PC} > \pi_i^N$, *for* $i = 1, 2$.

ii) $\pi_0^{FC} > \pi_0^N$ and $\pi_0^{PC} > \pi_0^N$ *hold always.*

iii) $\pi_0^{FC} > \pi_0^{PC}$ *if* $s > s^*(t) := \left[t - 3\sqrt{3}\sqrt{t\,(4 - 3t)}\right]/6 + 1$ *and* $\pi_0^{FC} < \pi_0^{PC}$ *if* $s < s^*(t)$ *(with equality holding at* $s = s^*(t)$*). Moreover,* $\partial s^*(t)/\partial t > 0$.

[20] In the EU, umbrella effects are potentially becoming more important for the assessment of the harm created by a cartel in private law suits. According to the new EU Damages Directive (see EU, 2014) members of a cartel can be held responsible for higher prices independently charged by firms competing with cartel members. Umbrella pricing then refers to a market outcome, where independent firms increased their prices in response to the cartel's price increases.

Proof. See Appendix.

Corollary 4.3 states that firms' profits are always higher under collusion (both partial and full collusion) compared to the profits under non-cooperative behavior. Full collusion always leads to higher profits than partial collusion for the brand producers, $i = 1, 2$, which is not the case for the private label firm. In fact, the private label firm can realize higher profits under partial collusion than under full collusion. This observation is important for the stability of full and partial collusion, respectively.[21] If $s < s^*(t)$, then it is optimal for the private label firm not to join the brands for a full collusion outcome, given that the brands keep their collusive conduct (i.e., partial collusion is realized).

The reason is that the market share of the private label firm always increases under partial collusion when compared with its market share under full collusion. In the former case, the private label's market share on the segment $x \in [0, 1/3]$ is $x_0^{FC} = 1/6$ and in the latter case the private label's market share is $x_0^{PC} = (4t - 3(1-s))/(18t)$. We then get $x_0^{PC} > x_0^{FC}$ if $s > 1 - t/3$ (see Corollary 4.2). Note also that $\partial(x_0^{PC} - x_0^{FC})/\partial t = (1-s)/(6t^2) > 0$ holds, so that the difference of the market shares is increasing in t. Accordingly, from part *iii)* of Corollary 4.3 we can infer that $\pi_0^{PC} > \pi_0^{FC}$ only becomes feasible when $t > 6/5$ holds. That means that for the profit of the private label to be larger under partial collusion than under full collusion, the increase in the market share must be large enough to compensate for the price decrease. In line with this observation, part *iii)* of Corollary 4.3 also states that the critical value $s^*(t)$ is increasing in t. This means that the range of the quality parameter s for which partial collusion is preferred by the private label firm increases in t. Thus, everything else equal, a reduced competitive intensity (high value of t) makes it more likely that the private label firm prefers partial collusion over full collusion.

Corollary 4.4. *Comparing the optimal deviation prices and profits under full and partial*

[21] This result is related to Donsimoni (1985) who analyzed cartel stability when firms differ in costs (but produce a homogenous good). He showed that the most efficient firms always join the cartel while the less efficient firms could stay outside.

collusion, we get the following orderings:

i) $\pi_j^{D,FC} > \pi_j^{FC}$ *and* $p_j^{D,FC} < p_j^{FC}$ *with* $j = 0, 1, 2.$

ii) $\pi_j^{D,PC} > \pi_j^{PC}$ *and* $p_j^{D,PC} < p_j^{PC}$ *with* $j = 0, 1, 2.$

Proof. Follows directly from comparing the respective values as stated in Propositions 4.2-4.3.

Corollary 4.4 states that firms always deviate by charging a lower price to earn a higher profit. This result also implies that firms' critical discount factor (4.3) is always in the range between zero and one.

4.4 Stability Analysis

Now we check the sustainability of the two collusive cases. Collusion is sustainable if the discount factor (4.3) is large enough; i.e., $\delta \geq \delta_j^C$ holds, where $C = FC, PC$ and $j = 0, 1, 2$. We find that $\delta_i^{FC}(s) \geq \delta_0^{FC}(s)$ for any $s \in (3/8, 1]$, where $i = 1, 2$. When $s \to 1$, firms become more homogeneous in quality, and the two incentive constraints converge, $\delta_0^{FC}(s = 1) = \delta_i^{FC}(s = 1)$. Thus the private label's incentive constraint holds, whenever it holds for the brand manufacturers.

The critical discount factor of firm i, $i = 1, 2$, under full collusion is given by

$$\delta_i^{FC} = \frac{\pi_i^{D,FC} - \pi_i^{FC}}{\pi_i^{D,FC} - \pi_i^{N}} = \frac{75\,(2-t)^2}{-48s^2 + 160st + 96s - 125t^2 - 60t + 252}, \tag{4.8}$$

where $\partial \delta_i^{FC}/\partial s < 0$ and $\partial \delta_i^{FC}/\partial t < 0$. Thus, the critical discount factor δ_i^{FC} is reduced (implying that full collusion is easier to sustain) when the transport cost parameter and/or the quality parameter of the private label are increasing. Intuitively, when s increases the profit in the non-cooperative stage game (see part *ii)* of Proposition 4.1) is reduced which lowers the expected profit of deviation. At the same time both the fully collusive profit and the profit in the deviation stage are independent of private label's quality. Thus a higher quality of the private label makes full collusion easier to sustain (lower value of δ_i^{FC}). Also, increasing horizontal product differentiation increases the likelihood of full collusion as in Ross (1992).

The critical discount factor of firm 0 under full collusion is given by

$$\delta_0^{FC} = \frac{\pi_0^{D,FC} - \pi_0^{FC}}{\pi_0^{D,FC} - \pi_0^{N}} = \frac{75\,(2s-t)^2}{108s^2 - 220st + 384s - 125t^2 + 320t - 192}, \tag{4.9}$$

where it can be shown that $\partial\delta_0^{FC}/\partial s > 0$ and $\partial\delta_0^{FC}/\partial t < 0$. The former derivative says that the private label firm's incentive to collude is reduced when s increases. This is due to the fact that in the punishment phase the profit of the private label good increases the smaller the quality gap, so that the expected profit in the punishment phase increases in s. As for the brands, the incentive constraint of the private label firm is more likely to be fulfilled when the transport cost parameter increases (i.e., horizontal product differentiation is high).

The critical discount factor of firm i, $i = 1, 2$, under partial collusion is given by

$$\delta_i^{PC} = \frac{\pi_i^{D,PC} - \pi_i^{PC}}{\pi_i^{D,PC} - \pi_i^{N}} = \frac{25}{81}, \tag{4.10}$$

so that the stability of partial collusion among the brand producers neither depends on the intensity of competition t nor on the private label's quality s. Thus partial collusion between the branded products is immune against changes of the quality of the private label good and changing intensities of competition.[22] We summarize the comparison and properties of the critical discount factors as follows.

Proposition 4.4. *The orderings of the critical discount factors are as follows:*

i) $\delta_0^{FC} < \delta_i^{FC}$ *holds always (equality holding at* $s = 1$*).*

ii) $\delta_i^{FC} < \delta_i^{PC}$ $(\delta_i^{FC} > \delta_i^{PC})$ *holds if* $s > \overline{s}(t)$ $(s < \overline{s}(t))$*, with equality holding at* $s = \overline{s}(t)$*, where* $\overline{s}(t) := 5t/3 - (\sqrt{201t - 44t^2 - 126})/3 + 1$*. Moreover,* $\partial\overline{s}(t)/\partial t < 0$ *and* $\partial^2\overline{s}(t)/\partial t^2 > 0$*.*

Furthermore, $\partial\delta_i^{FC}/\partial s < 0$*,* $\partial\delta_i^{FC}/\partial t < 0$*,* $\partial\delta_0^{FC}/\partial s > 0$*, and* $\partial\delta_0^{FC}/\partial t < 0$ *hold always. Finally,* $\partial\delta_i^{PC}/\partial s = \partial\delta_i^{PC}/\partial t = 0$*.*

Proof. See Appendix.

[22] This may explain the relative stability of recently detected partial cartels among branded manufacturers in Germany as mentioned in the Introduction.

Part *i)* of Proposition 4.4 states that the brand producers are critical for sustaining full collusion.[23] If the discount factor is high enough such that the branded goods collude, so does the private label firm. Part *ii)* shows that either full collusion or partial collusion is easier to sustain. There exists a critical value $\overline{s}(t)$ such that for $s > \overline{s}(t)$ full collusion is easier to sustain than partial collusion. If this condition is reversed, then the opposite holds with partial collusion being easier to sustain than full collusion. The function $\overline{s}(t)$ consists of all pairs (t, s) at which the critical discount factors under full and partial collusion are the same. This function is convex and negatively sloped in a (t, s) diagram over the feasible set. This means that relatively large values of s and t make it more likely that full collusion is easier to sustain than partial collusion. Or put differently, partial collusion is easier to sustain than full collusion when the values of s and t are relatively low.[24]

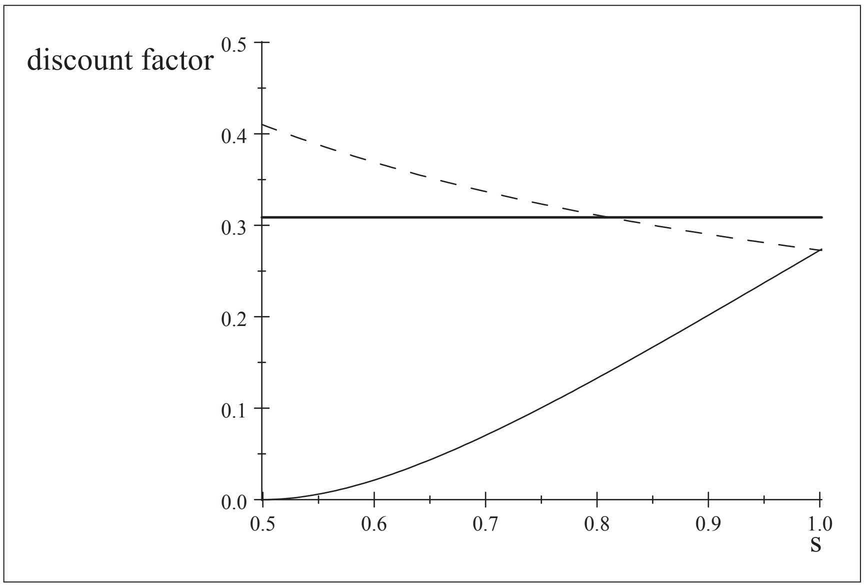

Figure 4.1. dashed line: δ_i^{FC}; thin line: δ_0^{FC}; bold line: δ_i^{PC}

[23] This result is also obtained in Häckner (1994) who shows that it is always the high quality firm which has the largest deviation incentives.

[24] If we allow for prices under full collusion which take care of the incentive constraint of the brand producers, then the critical discount factor can be reduced for the brands (see Bos and Harrington, 2010). This would mean that the private label had to increase its price to shift revenues to the branded firms. However, such an optimization problem is also constrained by the private label firm's participation constrained as given by part *iii)* of Corollary 4.3.

Figure 4.1 illustrates the critical discount factors as functions of s when we set the parameter value $t = 1$. The upward sloping curve represents the critical discount factor of the private label firm (thin line) which is never binding for $s < 1$. If the incentive constraint for full collusion is fulfilled for the branded firms, then it is also always fulfilled for the private label firm. Comparison of the critical discount factors of the brand producers under full collusion (dashed line) and under partial collusion (bold line) shows that full collusion is easier to sustain than partial collusion when s is larger than the threshold value $\overline{s}$, (which is reached at the intersection of both curves), while the opposite holds for lower values of s. Thus, full collusion is, *ceteris paribus*, easier to sustain than partial collusion if the quality gap of the private label good is not too large.

From Corollary 4.3 (which states firms' profit levels under the three types of conduct) we know that the brand producers always prefer full collusion over partial collusion, while both types of collusion are preferred over non-cooperative behavior. The private label firm also realizes the lowest profit level under non-cooperative behavior. In contrast to the brand producers, the private label firm, however, may prefer partial collusion over full collusion. Assuming that the firms select the type of conduct which maximizes their profits and taking into account the feasibility of collusion, we get the following market conduct depending on firms' discount factors (the same stability criterion is applied in Bos and Harrington, 2010).

Proposition 4.5. *Depending on the discount factor δ firms market conduct is as follows:*

i) If $\delta > \max\{\delta_i^{FC}, \delta_i^{PC}\}$, then market conduct is FC if $s > s^*$, whereas it is PC if $s < s^*$.

ii) If $\delta_i^{FC} > \delta > \delta_i^{PC}$, then market conduct is PC.

iii) If $\delta_i^{PC} > \delta > \delta_i^{FC}$, then market conduct is FC.

iv) If $\delta < \min\{\delta_i^{FC}, \delta_i^{PC}\}$, then market conduct is N.

Proof. Follows from combining the results of Proposition 4.5 with part *iii)* of Corollary 4.3.

Part *i)* of Proposition 4.5 refers to the case, where firms' discount factor is sufficiently large to make both full collusion and partial collusion stable. In this area, the private label firm's preference for either type of collusion determines the market conduct. From Corollary 4.3, part *iii)*, we know that the private label firm prefers full collusion if $s > s^*$, whereas the opposite holds for $s < s^*$. The type of conduct then follows immediately from the private label firm's preference. Interestingly, we notice that the private label firm prefers partial collusion over full collusion in the parameter area, where full collusion is easier to sustain than partial collusion from the brand producers' perspective. A prediction of the likely market conduct only based on the critical discount factors would be misleading in those instances. One has to consider the incentives of the private label firm to join the brands in their collusive conduct or whether to behave non-cooperatively. Inspection of the critical value s^* yields that full collusion is always the outcome for low values of t (precisely, $t < 6/5$), while in the remaining parameter area partial collusion is preferred by the private label firm only if $s < s^*(t)$. In that area it holds that intense competition (low value of t) makes, *ceteris paribus*, full collusion more likely (given that both types of collusion are feasible). In the same way, it follows from the shape of s^*, that low values of s tend to make partial collusion more likely, while for large values of s it becomes less likely.

Part *ii)* deals with cases where partial collusion is easier to sustain than full collusion. If firm's discount factor allows only for partial collusion, it will also be the market conduct chosen by the firms. Part *iii)* refers to the opposite case, where full collusion is easier to sustain than partial collusion, while only the former is feasible. Clearly, in this case full collusion is the type of conduct selected by the firms.

Finally, part *iv)* gives the case where neither type of collusion is incentive compatible, so that non-cooperative behavior is the market conduct. A non-cooperative outcome is more likely the larger the quality gap and/or the more intense competition. This follows directly from $\partial\delta_i^{FC}/\partial s < 0$ and $\partial\delta_i^{FC}/\partial t < 0$, so that an increasing quality of the private label good and reduced competition make full collusion easier to sustain. The former relation is in line with Steiner's (2004) observation that private labels' quality has been increasing, while signs of collusion between brands and private label goods have also emerged more recently.

Figure 4.2 illustrates our results presented in Proposition 4.5. The thin lines represent the constraints of the feasible set as specified in Assumption 4.1.[25] The upward sloping dashed curve is the locus of all pairs (t, s) such that the private label firm is indifferent between partial and full collusion. It represents the critical value $s^*(t)$ as specified in part *iii)* of Corollary 4.3. The private label firm gets a higher (lower) profit under full collusion than under partial collusion northeast (southwest) of the dashed curve. Inspection of the dashed curve $s^*(t)$ yields that t must pass a minimal value (namely, $\widetilde{t} = 6/5$ where $\overline{s}(\widetilde{t}) = s^*(\widetilde{t})$ holds), so that partial collusion can be more attractive than full collusion for the private label firm. Only if the intensity of competition is sufficiently low ($t > 6/5$) partial collusion can be preferred by the private label firm. In that range, however, the private label's quality must not surpass the critical value $s^*(t)$ (dashed curve), to get partial collusion instead of full collusion (while assuming that both are feasible).

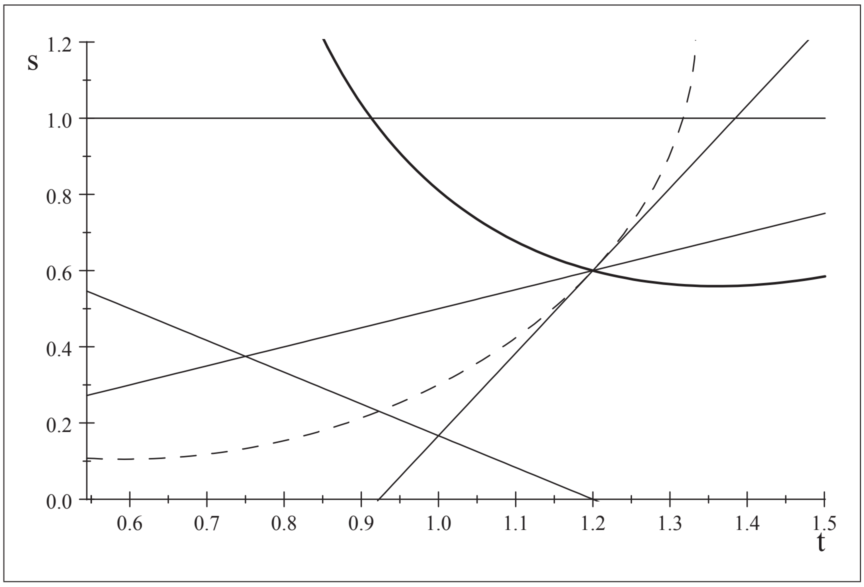

Figure 4.2. thin lines: feasible set; bold curve: $\overline{s}(t)$; dashed curve: $s^*(t)$

The downward sloping bold curve in Figure 4.2 depicts the critical value $\overline{s}(t)$ which is the locus of all (t, s) pairs where the critical discount factors of the brand producers are equal under full and partial collusion. Full collusion is easier (harder) to sustain than partial collusion for all (t, s) pairs northeast (southwest) of the bold curve. Interestingly,

[25] Of course, the following discussion of Figure 4.2 only refers to the range of parameters within the feasible set.

exactly when full collusion is easier to sustain than partial collusion (i.e., we are in the area northeast of the bold curve), then it can happen that the private label firm prefers partial collusion over full collusion.

The intensity of competition is measured by the parameter t. If the intensity of competition is large (low value of t), then full collusion becomes harder to sustain, so that partial collusion is more likely. In contrast to this observation (which relies on the brand producers' incentives to engage in full or partial collusion), the private label producer tends to prefer partial collusion over full collusion when the intensity of competition is reduced (high value of t). With regard to vertical quality differentiation (parameter s), both the incentives of the brand manufacturers and the private label producer are more in line. An increase of the quality of the private label makes it more likely that the incentive constraint of the brand producers is fulfilled and that the private label producer prefers full collusion over partial collusion.

4.5 Extensions and Discussion

In this section we analyze the private label firm's incentives to improve its quality. We also show how our results can be used to detect collusive conduct from market data. This argument is based on the markedly different market responses to an improvement of the private label's quality under the three types of market conduct (non-cooperative, partially or fully collusive).

Strategic choice of private label quality. The incentive of the private label producer to close the quality gap, $1-s$, between the private label's and the brands' qualities critically depends on the type of conduct. Assume an initial decision knot, where the private label's quality is set before the infinitely repeated market game starts. Suppose that the type of market conduct is fixed being either N, PC or FC. Taking the total derivative of the private label producer's expected flow of profits (around the equilibrium values p_0, p_1, and

p_2 under the three different conduct regimes) with respect to the quality parameter s yields

$$\frac{d\pi_0}{ds} = \frac{\partial \pi_0}{\partial s} + \frac{\partial \pi_0}{\partial p_0}\frac{dp_0}{ds} + \sum_{i=1}^{2} \frac{\partial \pi_0}{\partial p_i}\frac{dp_i}{ds}, \tag{4.11}$$

where we canceled out the factor $1/(1-\delta)$. The first term of the right-hand side of (4.11) is the direct effect of a change in s on the private label firm's profit which is positive but different depending on the type of market conduct. The second term on the right-hand side is zero because of the envelope theorem. The third term on the right-hand side represents the strategic effect of the quality investment. It is noteworthy that it is negative under N and PC (because both dp_i^N/ds and dp_i^{PC}/ds are strictly negative), but disappears under FC (because of $dp_i^{FC}/ds = 0$). In the parlance of Fudenberg and Tirole (1984), there exists an investment reducing strategic effect under N and PC (*puppy dog ploy*), while there is no such effect present under FC. Not surprisingly, marginal investment incentives are largest under FC, what can be derived from calculating the total derivatives of (4.11) under the three different types of market conduct.

Proposition 4.6. *The private label firm's marginal incentives to close the quality gap* $1-s$ *are largest under FC with PC and N following in that order; i.e.,* $d\pi_0^{FC}/ds > d\pi_0^{PC}/ds > d\pi_0^{N}/ds > 0$ *hold always.*

Proof. See Appendix.

From Proposition 4.4 and Corollary 4.3, we know that an increase of the quality of the private label makes full collusion more likely because of two reasons: *First*, a higher value of s makes it easier for brand producers to sustain full collusion; i.e., $\partial\delta_i^{FC}/\partial s < 0$. As $\partial\delta_i^{PC}/\partial s = 0$ holds, it then also follows that full collusion becomes relatively easier to sustain than partial collusion the larger s (see Proposition 4.4). *Second*, a higher value of s makes full collusion more attractive than partial collusion for the private label firm (part *iii)* of Corollary 4.3).

Suppose now a generic change in s, say from s_1 to s_2, with $s_1 < s_2$, such that this increase in the private label's quality changes market conduct from non-cooperative (N) or partially collusive (PC) to fully collusive (FC). Investment incentives, which are given by the profit differential $\pi_0^{FC}(s_2) - \pi_0^{C}(s_1)$ (with $C = FC, PC, N$) are then driven only by the *replacement effect* and the following result follows immediately.[26]

Proposition 4.7. *Suppose a generic increase of the private label's quality from s_1 to s_2 with $s_1 < s_2$ which changes market conduct from N or PC to FC. The private label firm's incentives to close the quality gap $1 - s$ are then ordered as follows:*

$$\pi_0^{FC}(s_2) - \pi_0^{N}(s_1) > \pi_0^{FC}(s_2) - \pi_0^{PC}(s_1) > \pi_0^{FC}(s_2) - \pi_0^{FC}(s_1) > 0.$$

Proof. Follows directly from $\pi_0^{FC} > \pi_0^{PC} > \pi_0^{N}$.

Proposition 4.7 follows from the ordering of the private label firm's profit under the three different types of market conduct. The private label producer has larger investment incentives the lower the degree of collusion, because the net gain from a quality increase is higher when competition is more intense initially.[27]

These observations have two implications: *First*, an increase in the quality of the private label may be strategic so as to obtain a full collusion outcome as the type of market conduct chosen by the firms. *Second*, an increase in the quality of the private label may have significant adverse effects for consumers, whenever it is used to trigger a full collusion outcome. In fact, incentives to close the quality gap between the private label good and the brands are maximal whenever non-cooperative behavior would prevail without any investment. Such a constellation is obtained if $\delta_i^{FC}(s_1) > \delta > \delta_i^{FC}(s_2)$ holds.

The pro-collusive effect of a higher quality of the private label good sheds new light on the possible market effects of so-called premium private label goods. While the trend of an increasing quality of private label goods has been generally interpreted as pro-competitive, our investigation highlights their role in stabilizing full collusion between private label and brand producers.

[26] Note that a change from N to PC is not possible through an increase of s as δ_i^{PC} does not depend on s.

[27] A qualitatively similar result is stated in Aubert, Rey, and Kovacic (2006) where a drastic innovation is considered.

Identifying collusive conduct. The following Table 1 describes the change in the prices of the brands and the private label depending on an improvement of the private label's quality s. Rows 2-4 relate to cases where the type of conduct is fixed as being non-cooperative, partially collusive or fully collusive, respectively. The fifth row refers to those instances where initially conduct is either non-cooperative or partially collusive, while a quality improvement of the private label induces a fully collusive outcome.

Conduct	Brands' prices	Private label's price
N	$\downarrow$ (by $-\frac{1}{5}$)	$\uparrow$ (by $\frac{2}{5}$)
PC	$\downarrow$ (by $-\frac{1}{3}$)	$\uparrow$ (by $\frac{1}{3}$)
FC	no change	$\uparrow$ (by 1)
switch to FC	$\uparrow$	$\uparrow$

Table 4.1: Effects of an increase in s on brands' and private label's prices

Two observations are noteworthy: *First*, given that the type of conduct does not change, the private label's price is the more quality-sensitive, the more collusive the market conduct (Table 4.1 states in the third column the sign of the own-price effect and the marginal effect is given in brackets). If the type of conduct changes through an increase in s to full collusion (see last row of Table 4.1), then the change in the private label's price is a discrete jump upward. *Second*, full collusion can be inferred from the absence of a negative sensitivity of the brands' prices with respect to the private label's quality. According to Table 4.1, the brands' prices stay put if full collusion prevails. The price effect is even positive for the branded goods if a higher private label quality leads to fully collusive conduct.[28]

Incidentally, Ward et al. (2002) showed in their empirical analysis of scanner data (obtained at cash registers) from US grocery stores that an increasing market share of

[28] See Gabrielsen and Sörgard (2007) for a vertical restraint theory which also shows that branded suppliers could increase their prices in the presence of a private label (particularly of poor perceived quality). See also Gilo (2008) for a survey of vertical restraints leading to a cartel outcome between retailer-controlled private labels and branded goods.

private labels tends to increase the brands' prices. While this can be explained by same static theories of product differentiation and vertical relations, our model suggests that such an outcome can also result from a combination of an increasing private label quality and collusive conduct.

The relations stated in Table 1 suggest two empirical strategies to identify collusion in markets where brands and private labels compete. From the first observation it follows that the private label's own price response to a quality improvement is the larger the more collusive industry conduct. The second observation suggests that fully collusive behavior can be inferred from a nonnegative price effect of the brands resulting from an increase of the private label's quality.

4.6 Conclusion

We have analyzed collusion between brands and a private label substitute. We focused on heterogeneity of firms due to product differentiation; both horizontal and vertical. We assumed that nonparticipation of the private label producer in a cartel does not necessarily lead to fully non-cooperative behavior, but may induce the brand manufacturers to form a partial cartel. In fact, forming such a partial cartel is always optimal for the brand producers when being incentive compatible. Given that both partial and full collusion are feasible, the private label firm is more likely to revert to non-cooperative behavior (with a partial cartel following), if horizontal and/or vertical differentiation is sufficiently large. Thus, focusing on the private label's incentive to participate in a full cartel, we get that this is more likely whenever product differentiation (vertical and/or horizontal) is low enough.

Interestingly, this picture is different when considering the brand producers' incentive conditions. Then a higher degree of horizontal product differentiation works in favor of full collusion, while the quality gap of the private label must not be too large. Thus, comparing the brand producers' incentive constraints gives the result, that full collusion is more likely when horizontal product differentiation is large but vertical differentiation is low. Taking both results together, we find that exactly in the parameter range, where full collusion is easier to sustain than partial collusion (from brands' perspective), it could happen that the private label producer opts for a partial collusion outcome by behaving non-cooperatively. For this to happen, horizontal and vertical differentiation must be sufficiently large.

The partial collusion case has two characteristic features which merit mentioning. *First*, the stability conditions of the brand manufacturers are independent of the degree of product differentiation (both horizontal and vertical). *Second*, the private label firm can increase its price under partial collusion above the equilibrium price under fully non-cooperative conduct. The first observation can be used to explain the remarkable stability of cartels among brand manufacturers even in an environment in which the degree of competition and the private label's quality change over time. The second result is potentially important for the assessment of the harm created by a cartel that only involves explicit collusion among the brand manufacturers. As private law suits also allow for damages created by a cartel's umbrella effects on outsiders' prices the question emerges whether or not the outsiders did join the explicit cartel by colluding tacitly (so that in fact a full cartel was in operation) or whether the outsiders behaved non-cooperatively (in which case the cartel was partial). In both instances, the outsiders increase their prices, however, by different amounts.

We have also analyzed the private label's incentive to increase its quality where we showed it is maximal under full collusion with partial collusion and non-cooperative behavior following in that order. In addition, a quality increase makes full collusion more likely because of two reasons: *First*, it relaxes the brand producers' incentive constraint for full collusion, and *second*, it makes it more likely that the private label firm prefers full collusion over partial collusion. The latter observation has also implications for the competitive assessment of private label goods. As long as private label goods were of the budget type, private label producers had only little incentives to join into a full cartel, while branded firms found it also easier to sustain a partial cartel. This may have changed as private labels' quality increased over time. As the quality gap becomes smaller, private label firms and branded producers should have found it more attractive to form an all-encompassing cartel.

We have also shown that the price responses associated with a quality improvement of the private label good can give important information for detecting cartelization. *First*, if a quality increase does not induce a negative effect on brand producers' prices (everything else equal), then either full collusion is present or the market is triggered into fully collusive conduct. *Second*, the larger the private label's own-price effect of a quality increase, the more likely it is that collusive conduct exists in the market.

4.7 Appendix

In this Appendix we present the missing proofs.

Proof of Proposition 4.1. Using firms' demand functions (4.7), we can write the firms' profits as

$$\pi_0 = D_0 p_0 = (x_0 + 1 - x_2)p_0, \tag{4.12}$$

$$\pi_1 = D_1 p_1 = (x_1 - x_0)p_1, \text{ and} \tag{4.13}$$

$$\pi_2 = D_2 p_2 = (x_2 - x_1)p_2. \tag{4.14}$$

Substituting the values of the indifferent consumers (4.4), (4.5), and (4.6) into these expressions, we get

$$\begin{aligned} \pi_0 &= \frac{p_0\,(6s + 2t - 6p_0 + 3p_1 + 3p_2 - 6)}{6t}, \\ \pi_1 &= \frac{p_1\,(2t - 3s + 3p_0 - 6p_1 + 3p_2 + 3)}{6t}, \text{ and} \\ \pi_2 &= \frac{p_2\,(2t - 3s + 3p_0 + 3p_1 - 6p_2 + 3)}{6t}. \end{aligned}$$

Maximization of each firm's profit gives the following system of first-order conditions:

$$\begin{aligned} \frac{6s + 2t - 12p_0 + 3p_1 + 3p_2 - 6}{6t} &= 0, \\ \frac{2t - 3s + 3p_0 - 12p_1 + 3p_2 + 3}{6t} &= 0, \text{ and} \\ \frac{2t - 3s + 3p_0 + 3p_1 - 12p_2 + 3}{6t} &= 0. \end{aligned}$$

All profit functions are strictly concave, so that this system of first-order conditions determines the unique equilibrium outcome. Solving for the prices we get the following equilibrium values as stated in part *i)* of the proposition:

$$p_0^N = \frac{6(s-1)+5t}{15} \text{ and}$$
$$p_i^N = \frac{3(1-s)+5t}{15} \text{ for } i=1,2.$$

Substituting the equilibrium prices into (4.4), (4.5), and (4.6), we get the equilibrium locations of the indifferent consumers (see part *iii)* of the proposition)

$$x_0^N = \frac{1}{6} - \frac{1-s}{5t}, \tag{4.15}$$
$$x_1^N = \frac{1}{2}, \text{ and} \tag{4.16}$$
$$x_2^N = \frac{5}{6} + \frac{1-s}{5t}. \tag{4.17}$$

Note that $x_0^N \in (0, 1/3)$ and $x_2^N \in (2/3, 1)$ must hold in an interior solution. This is true for $s > 1 - 5t/6$ which is implied by Assumption 4.1. Substituting the equilibrium prices and locations of the indifferent consumers into the profit functions (4.12), (4.13), and (4.14) yields

$$\pi_0^N = \frac{(6(s-1)+5t)^2}{225t} \text{ and}$$
$$\pi_1^N = \pi_2^N = \frac{(5t+3(1-s))^2}{225t}.$$

which are stated in part *ii)* of the proposition. We finally check whether the utilities of the indifferent consumers (4.15), (4.16), and (4.17) are nonnegative. We obtain

$$U_{x_0^N} = U_{x_2^N} = \frac{2s+3}{5} - \frac{t}{2} \text{ and}$$
$$U_{x_1^N} = \frac{s+4}{5} - \frac{t}{2}.$$

Setting $U_{x_0^N}, U_{x_1^N}, U_{x_2^N} > 0$, we get the condition $s > 5t/4 - 3/2$, which holds by Assumption 4.1.

Q.E.D.

Proof of Proposition 4.2. Consider collusion by all three firms. The three firms maximize their joint profit

$$\max_{p_0^{FC},\, p_1^{FC},\, p_2^{FC} \geq 0} \pi^{FC} = p_0 D_0 + p_1 D_1 + p_2 D_2$$

subject to consumers' reservation utilities (which must be nonnegative). Substituting the demand functions into π^{FC} and differentiating with respect to the prices, we get the following system of first-order conditions

$$\begin{aligned} \frac{3s + t - 6p_0 + 3p_1 + 3p_2 - 3}{3t} &= 0 \text{ and} \\ \frac{2t - 3s + 6p_0 - 12p_i + 6p_{i'} + 3}{6t} &= 0, \text{ for } i, i' = 1, 2,\ i \neq i'. \end{aligned}$$

This system of first-order conditions is only fulfilled at $t = 0$. Hence, there cannot exist an interior solution to the maximization problem. We next show that Assumption 4.1 ensures that the joint profit is maximized in the corner solution where all indifferent consumers get a utility of zero, while the branded goods prices p_1 and p_2 are the same and the market is fully covered. Suppose that all indifferent consumers get a utility of zero; i.e., $U_{x_0} = U_{x_1} = U_{x_2} = 0$ holds. Firms 1 and 2 are symmetric, hence, the location of the indifferent consumer is $x_1 = 1/2$ with $p_1 = p_2$. Substituting theses values into $U_{x_1=1/2} = 0$, we get the following equilibrium prices of the brands under full collusion

$$p_1^{FC} = p_2^{FC} = 1 - \frac{t}{6}. \tag{4.18}$$

We assume that the market is always covered. Therefore, the utility of the indifferent consumer on the segments $(0, 1/3)$ and $(2/3, 1)$ must be zero, $U_{x_0} = U_{x_2} = 0$ (given the brands' prices (4.18) the indifferent consumers are located at $x = 1/6$ and $x = 5/6$, respectively). By substituting (4.4) and the collusive prices of the brands (4.18) into the utility functions we get the price for the private label good:

$$p_0^{FC} = s - \frac{t}{6}. \tag{4.19}$$

Given the prices of the brands (4.18), we still have to check whether it is indeed optimal to set the price $p_0^{FC} = s - t/6$ which ensures that the market is fully covered. Note first that the joint profit can never increase with a lower price for the private label, because $p_0^{FC} < p_i^{FC}$ for $i = 1, 2$. However, we still have to ensure that there is no incentive to set a higher price for the private label than p_0^{FC}. If this were optimal, the market would not be covered. In other words, under a higher collusive price charged by the private label there is a new indifferent consumer whose address is $x_0 = (s - p_0)/t < 1/6$ on the segment $(0, 1/3)$ (correspondingly, the new location on the segment $(2/3, 1)$ is then $x_2 = 1-(s-p_0)/t > 5/6$). Hence, given the prices of the brands (4.18), the joint profit cannot be increased by a price of the private label which is higher than p_0^{FC} if

$$p_0^{FC} \geq \widehat{p}_0 \text{ with } \widehat{p}_0 := \arg\max_{p_0} D_0(p_0)p_0.$$

Solving the maximization problem

$$\max_{p_0} D_0(p_0)p_0 = \max_{p_0} \frac{1}{t} 2p_0 (s - p_0) \tag{4.20}$$

gives

$$\widehat{p}_0 = \frac{1}{2}s$$

which implies

$$\widehat{p}_0 \leq p_0^{FC} \text{ if and only if } s \geq \frac{1}{3}t.$$

The latter inequality is assumed in Assumption 4.1. Again, it ensures that the market is fully covered under full collusion. We have, therefore, proven part *i)* and part *iii)* of the proposition. The indifferent consumer between the private label and brand 1 (2) is located at $x_0 = 1/6$ $(x_2 = 5/6)$. Substituting the equilibrium prices and locations of the indifferent consumers into the profit functions (4.12), (4.13), and (4.14) yields the profits under full collusion (as stated in part *ii)* of the proposition)

$$\begin{aligned} \pi_0^{FC} &= \frac{s}{3} - \frac{t}{18} \text{ and} \\ \pi_1^{FC} &= \pi_2^{FC} = \frac{1}{3} - \frac{t}{18}. \end{aligned} \tag{4.21}$$

The sum of all firms' profits under full collusion is then given by

$$\pi^{FC} = \frac{s+2}{3} - \frac{t}{6}.$$

Derivation of the deviation profits. We first solve the deviation problem of one of the brand producers which are symmetric. Next we solve the deviation problem of the private label firm.

Case 1 (deviation by firm 1*):* If firm 1 deviates, it charges a deviation price $p_1^{D,FC}$. Substituting the collusive prices (4.18)-(4.19) into (4.4) and (4.5), we get the locations of the indifferent consumers depending on firm 1's deviation price:

$$x_0^{D,FC} = \frac{t + 2p_1^{D,FC} - 2}{4t} \text{ and} \tag{4.22}$$

$$x_1^{D,FC} = \frac{5t - 6p_1^{D,FC} + 6}{12t}. \tag{4.23}$$

Note that we must ensure that $x_1^{D,FC} - x_0^{D,FC} < 1$. Firm 1 maximizes its deviation profit

$$\max_{p_1^{D,FC} \geq 0} \pi_1^{D,FC} = p_1^{D,FC}(x_1^{D,FC} - x_0^{D,FC}).$$

By substituting the locations of the indifferent consumers (4.22) and (4.23) into the profit function and solving the maximization problem, we find the optimal deviation price of firm 1:

$$p_1^{D,FC} = \frac{t}{12} + \frac{1}{2}.$$

Note that the optimal deviation price of the brand $p_1^{D,FC}$ is smaller than the collusive price p_i^{FC} for all $t < 2$ which holds by Assumption 4.1. Substituting the optimal deviation price into (4.22) and (4.23), we get the locations of the indifferent consumers

$$x_0^{D,FC} = \frac{7t - 6}{24t} \text{ and}$$

$$x_1^{D,FC} = \frac{3t + 2}{8t}.$$

We must check whether the demand of the deviating firm is smaller than one, $x_1^{D,FC} - x_0^{D,FC} < 1$. This is true if $t > 6/11$ which is assumed in Assumption 4.1. The deviation profit is then given by

$$\pi_1^{D,FC} = \frac{(t + 6)^2}{144t}.$$

This proves part *iv)* of the proposition.

Case 2 (deviation by firm 0*):* Substituting the collusive prices into (4.4) and (4.6), we get the locations of the indifferent consumers depending on the private label firm's deviation price:

$$x_0^{D,FC} = \frac{6s + t - 6p_0^{D,FC}}{12t} \text{ and} \tag{4.24}$$

$$x_2^{D,FC} = \frac{11t - 6s + 6p_0^{D,FC}}{12t}. \tag{4.25}$$

Again, we must ensure that $x_0^{D,FC} + 1 - x_2^{D,FC} < 1$. Firm 0 maximizes its deviation profit

$$\max_{p_0^{D,FC} \geq 0} \pi_0^{D,FC} = p_0^{D,FC}(x_0^{D,FC} + 1 - x_2^{D,FC}).$$

This problem is only well defined for $p_0^{D,FC} \leq p_0^{FC}$ because in this case the market is fully covered. For $p_0^{D,FC} > p_0^{FC}$ the maximization problem (4.20) applies where we already showed that p_0^{FC} is optimal for $s \geq t/3$. By substituting the locations of the indifferent consumers (4.24) and (4.25) into the profit function and solving the maximization problem, we find the optimal deviation price of firm 0:

$$p_0^{D,FC} = \frac{t}{12} + \frac{s}{2}. \tag{4.26}$$

The optimal deviation price (4.26) is smaller than the collusive price of the private label (4.19) if $p_0^{FC} - p_0^{D,FC} \geq 0$ holds, which gives the condition

$$s \geq \frac{t}{2}. \tag{4.27}$$

Condition (4.27) is part of Assumption 4.1. Note that a deviation price of the private label larger than the collusive price p_0^{FC} cannot be optimal for $s \geq t/3$ as we have shown by solving the maximization problem (4.20). Thus, the optimal deviation price of the private label would be the collusive price p_0^{FC} for $t/2 > s \geq t/3$. By Assumption 4.1 this case is ruled out, so that the private label firm finds it optimal to deviate with a price which is smaller than the collusive price.

By substituting the optimal deviation price of the private label good into the locations of the indifferent consumers (4.24) and (4.25), we get

$$\begin{aligned} x_0^{D,FC} &= \frac{6s+t}{24t}, \text{ and} \\ x_2^{D,FC} &= \frac{23t-6s}{24t}. \end{aligned}$$

We must check that under deviation the demand of the private label does not exceed one, $x_0^{D,FC} + 1 - x_2^{D,FC} < 1$. This holds for $t > 6s/11$ which is implied by assuming $t > 6/11$ (see Assumption 4.1). The deviation profit of firm 0 is then given by

$$\pi_0^{D,FC} = \frac{(6s+t)^2}{144t}.$$

This proves part *v)* of the proposition.

Q.E.D.

Proof of Proposition 4.3. Consider collusion by firms 1 and 2, while firm 0 behaves non-cooperatively. Firms 1 and 2 maximize their joint profit

$$\max_{p_1,\, p_2 \geq 0} \pi^{PC} = p_1 D_1 + p_2 D_2,$$

while the maximization problem of firm 0 is

$$\max_{p_0 \geq 0} \pi_0 = p_0 D_0.$$

Substituting the demand functions into both problems and maximizing over the respective prices, we get the following system of first-order conditions:

$$\begin{aligned} \frac{2t - 3s + 3p_0 - 12p_i + 6p_{i'} + 3}{6t} &= 0, \text{ for } i, i' = 1, 2,\ i \neq i', \text{ and} \\ \frac{6s + 2t - 12p_0 + 3p_1 + 3p_2 - 6}{6t} &= 0. \end{aligned}$$

All maximization problems are strictly concave, so that the solution of this system of first-order conditions gives the equilibrium prices as stated in part *i)* of the proposition; namely,

$$\begin{aligned} p_0^{PC} &= \frac{4t}{9} - \frac{1-s}{3}, \\ p_1^{PC} &= p_2^{PC} = \frac{5t}{9} + \frac{1-s}{3}. \end{aligned}$$

All prices are strictly positive under Assumption 4.1. Substituting the collusive prices into (4.4), (4.5), and (4.6), we get the equilibrium locations of the indifferent consumers under partial collusion (see part *iii)* of the proposition)

$$\begin{aligned} x_0^{PC} &= \frac{4t - 3(1-s)}{18t}, \\ x_1^{PC} &= \frac{1}{2}, \text{ and} \\ x_2^{PC} &= \frac{14t + 3(1-s)}{18t}. \end{aligned}$$

Assumption 4.1 ensures that $x_0^{PC} \in (0, 1/3)$ and $x_2^{PC} \in (2/3, 1)$. Substituting the equilibrium prices and locations of the indifferent consumers into the profit functions (4.12), (4.13), and (4.14) yields (part *ii)* of the proposition)

$$\begin{aligned} \pi_0^{PC} &= \frac{(4t-3(1-s))^2}{81t}, \\ \pi_1^{PC} &= \pi_2^{PC} = \frac{(5t+3(1-s))^2}{162t}. \end{aligned} \tag{4.28}$$

We finally check whether the utilities of the indifferent consumers (4.15), (4.16), and (4.17) are nonnegative. We obtain

$$\begin{aligned} U_{x_0^{PC}} &= U_{x_2^{PC}} = \frac{1+s}{2} - \frac{2t}{3}, \text{ and} \\ U_{x_1^{PC}} &= \frac{s+2}{3} - \frac{13t}{18}. \end{aligned}$$

These utility levels are nonnegative if $s > \max\{13t/6 - 2, 4t/3 - 1\}$ which is assumed in Assumption 4.1.

We next derive the deviation price and profit of one of the brand producers (both are symmetric) as stated in part *iv)* of the proposition. Consider deviation by firm 1. Substituting the collusive prices p_0^{PC} and p_2^{PC} into (4.4) and (4.5), we get the locations of the indifferent consumers depending on the deviation price of firm 1:

$$x_0^{D,PC} = \frac{6s - t + 9p_1^{D,PC} - 6}{18t} \text{ and} \tag{4.29}$$

$$x_1^{D,PC} = \frac{14t - 9p_1^{D,PC} + 3 - 3s}{18t}. \tag{4.30}$$

Firm 1 maximizes its deviation profit

$$\max_{p_1^{D,PC} \geq 0} \pi_1^{D,PC} = p_1^{D,PC}(x_1^{D,PC} - x_0^{D,PC}).$$

By substituting the locations of the indifferent consumers (4.29) and (4.30) into the profit function and solving the maximization problem, we find the optimal deviation price of firm 1:

$$p_1^{D,PC} = \frac{1-s}{4} + \frac{5t}{12}.$$

Substituting the optimal deviation price into (4.29) and (4.30), we get the locations of the indifferent consumers under deviation

$$\begin{aligned} x_0^{D,PC} &= \frac{15s + 11t - 15}{72t} \text{ and} \\ x_1^{D,PC} &= \frac{41t - 3s + 3}{72t}. \end{aligned}$$

We can show that $x_1^{D,PC} - x_0^{D,PC} < 1$ if $s > 1 - 7t/3$ which holds by Assumption 4.1. The deviation profit of firm 1 is then given by

$$\pi_1^{D,PC} = \frac{1}{t}\left(\frac{5t}{12} + \frac{1-s}{4}\right)^2.$$

Q.E.D.

Proof of Corollary 4.3. The proof of parts *i)* and *ii)* of the corollary follows immediately from comparing the respective profit levels under non-cooperative behavior, full collusion and partial collusion. Part *iii)* compares the profit levels of the private label firm under full collusion (4.21), and under partial collusion (4.28). This comparison gives rise to the condition

$$\pi_0^{FC} - \pi_0^{PC} = \frac{-18s^2 + 6st + 36s - 41t^2 + 48t - 18}{162t} \geq 0$$

which holds if and only if

$$s \geq s^*(t) := \frac{1}{6}t - \frac{1}{2}\sqrt{3}\sqrt{t(4-3t)} + 1.$$

Moreover,

$$\frac{\partial s^*(t)}{\partial(t)} = \frac{\sqrt{t(4-3t)}\left(\sqrt{t(4-3t)} + 9\sqrt{3}t - 6\sqrt{3}\right)}{6t(4-3t)}$$

which obtains three zeros at $t \in \{0, (2/3) - (\sqrt{2}\sqrt{41})/123, 4/3\}$. It is easily checked that $\partial s^*(t)/\partial(t) > 0$ for all $t \in [(2/3) - (\sqrt{2}\sqrt{41})/123, 4/3]$. As $s^*(t)$ cuts through the feasible set (as specified in Assumption 4.1) over the interval $t \in [6/5, 54/41]$, it then follows that $\partial s^*(t)/\partial(t) > 0$ holds always.

Q.E.D.

Proof of Proposition 4.4. *Part i)* Inspecting the critical discount factors δ_0^{FC} and δ_i^{FC} (see (4.9) and (4.8), respectively), we get that $\delta_0^{FC} = \delta_i^{FC}$ holds at $s = 1$. To prove that $\delta_i^{FC} > \delta_0^{FC}$ holds for $s < 1$, we first show that δ_i^{FC} is monotonically decreasing in s over the relevant parameter range. Taking the derivative with respect to the parameter s, we get

$$\frac{\partial \delta_i^{FC}}{\partial s} = 75(2-t)^2 \frac{-160t + 96s - 96}{(48s^2 - 160st - 96s + 125t^2 + 60t - 252)^2},$$

so that the sign of the derivative depends on the sign of the term $-160t + 96s - 96$. This term is negative for all $s < 5t/3+1$ which is implied by Assumption 4.1. Thus, $\partial\delta_i^{FC}/\partial s < 0$ holds everywhere. We next show that δ_0^{FC} is monotonically increasing in s over the relevant parameter range. Taking the respective derivative, we get

$$\frac{\partial\delta_0^{FC}}{\partial s} = \frac{-28s^2t + 96s^2 - 76st^2 + 160st - 96s + 45t^3 - 104t^2 + 48t}{\left(108s^2 - 220st + 384s - 125t^2 + 320t - 192\right)^2/1200}, \tag{4.31}$$

so that the sign of the derivative depends on the sign of the numerator. The numerator has two potentially relevant roots at

$$s'(t) = \frac{1}{2}t \text{ and } s''(t) = \frac{-104t + 45t^2 + 48}{14t - 48} \text{ for } t \neq \frac{24}{7}.$$

Further inspection yields that $s > \max\{s'(t), s''(t)\}$ holds in the feasible area as specified in Assumption 4.1. This implies that the numerator of the right-hand side of (4.31) and thus $\partial\delta_0^{FC}/\partial s$ is strictly positive in the relevant parameter range. Combining these results concerning the slopes of both critical values with the fact that both values are equal at $s = 1$ gives the ordering stated in the proposition.

Part *ii)* Setting $\delta_i^{FC} = \delta_i^{PC}$ we can calculate the unique threshold value $\overline{s}(t) := 5t/3 - (\sqrt{201t - 44t^2 - 126})/3 + 1$ which cuts through the feasible set (the expression below the square root sign is always positive). The orderings stated in the proposition are then easily verified. Calculating the first and second derivative with respect to t we get

$$\begin{aligned}\frac{\partial\overline{s}(t)}{\partial t} &= \frac{1}{6}\frac{88t + 10\sqrt{-44t^2 + 201t - 126} - 201}{\sqrt{-44t^2 + 201t - 126}} < 0 \text{ and} \\ \frac{\partial^2\overline{s}(t)}{\partial t^2} &= \frac{6075}{4\left(-44t^2 + 201t - 126\right)^{\frac{3}{2}}} > 0,\end{aligned}$$

where the signs hold within the considered parameter range.

Q.E.D.

Proof of Proposition 4.6. We have to calculate the marginal profit changes of the private label producer under the three types of market conduct. This yields

$$\begin{aligned}\frac{d\pi_0^N}{ds} &= \frac{72s + 60t - 72}{225t}, \\ \frac{d\pi_0^{PC}}{ds} &= \frac{18s + 24t - 18}{81t}, \text{ and} \\ \frac{d\pi_0^N}{ds} &= \frac{1}{3}.\end{aligned}$$

Comparison of those values gives the ordering stated in the proposition. It remains to show that

$$\frac{\partial \delta_i^{FC}}{\partial t} = 1200\,(t-2)\,\frac{32s - 35t + 10st - 6s^2 + 24}{\left(48s^2 - 160st - 96s + 125t^2 + 60t - 252\right)^2} < 0,$$

which holds because the numerator of the right-hand side is positive if

$$s > 5t/6 - 5(\sqrt{t^2 - 2t + 16})/6 + 8/3.$$

The above condition is implied by Assumption 4.1. Finally,

$$\frac{\partial \delta_0^{FC}}{\partial t} = \frac{28s^3 + 76s^2t - 176s^2 - 45st^2 + 48st + 48s + 20t^2 - 24t}{\left(108s^2 - 220st + 384s - 125t^2 + 320t - 192\right)^2/1200} < 0,$$

which holds because the numerator of the right-hand side is negative if

$$s > \frac{22}{7} - \frac{5}{28}\sqrt{81t^2 - 272t + 256} - \frac{45}{28}t,$$

which is implied by Assumption 4.1, again.

Q.E.D.

Declaration of Contribution according to § 6 (1) PO as of 25.11.2013:

Hereby I, Irina Hasnas, declare that the chapter "Full Versus Partial Collusion Among Brands and Private Label Producers" is co-authored by Christian Wey. It has been pre-published as a DICE Discussion paper (Hasnas and Wey, 2015).

My contributions with regard to content and methods are the following:

- I have contributed for the positioning in relevant literature and the searching for the relevant examples.
- I have created the set-up and the specifications of the model.
- I have written the Equilibrium analysis.
- I have written the Stability analysis.
- I have contributed to a major part of the proofs.
- I have contributed to the Conclusion.

Signatures:

Irina Hasnas, M.Sc.	Prof. Dr. Christian Wey

References

Abreu, D., (1986), Extremal Equilibria of Oligopolistic Supergames, Journal of Economic Theory 39, 191-225.

Abreu, D., (1988), On the Theory of Infinitely Repeated Games with Discounting, Econometrica 56, 383-396.

Abreu, D., Pearce, D.G., Stacchetti, E., (1986), Optimal Cartel Equilibria with Imperfect Monitoring, Journal of Economic Theory 39, 251-269.

Anderson, S.P., de Palma, A., (1988), Spatial Price Discrimination with Heterogeneous Products, Review of Economic Studies 55, 573-592.

Angwin, J., (2010), The Web's New Golden Mine: Your Secrets, The Wall Street Journal, 30 July 2010 (available online: www.wsj.com).

Anticompetitive Agreements in the Mobile Telephony Market. Autorité de la Concurrence, Press Release: 1 December 2005, Paris (available online: www.autoritedelaconcurren ce.fr).

Athey, S.K., Bagwell, K., (2001), Optimal Collusion with Private Information, Rand Journal of Economics 32, 428-465.

Aubert, C., Rey, P., Kovacic, W.E., (2006), The Impact of Leniency and Whistle-Blowing Programs on Cartels, International Journal of Industrial Organization 24, 1241-1266.

Berges-Sennou, F., Bontems, P., Requillart, V., (2004), Economics of Private Labels: A Survey of Literature, Journal of Agricultural and Food Industrial Organization 2, Article 3.

Bontemps, C., Orozco, V., Requillart, V., (2008), Private Labels, National Brands and Food Prices, Review of Industrial Organization 33, 1-22.

Bontems, P., Monier-Dilhan, S., Requillart, V., (1999), Strategic Effects of Private Labels, European Review of Agricultural Economics 26, 147-165.

Bos, I., Harrington, J.E. Jr., (2010), Endogenous Cartel Formation with Heterogeneous Firms, Rand Journal of Economics 41, 92-117.

Bundeskartellamt (2009), Bundeskartellamt Imposes Fines on Coffee Roasters, Press

Release: 21 December 2009, Bonn (available online: www.bundeskartellamt.de).

Bundeskartellamt (2013), Bundeskartellamt Imposes Fine of More Than € 60 Million on Confectionery Manufacturers, Press Release: 31 January 2013, Bonn (available online: www.bundeskartellamt.de).

Bundeskartellamt (2014a), Sektoruntersuchung Lebensmitteleinzelhandel: Darstellung und Analyse der Strukturen und des Beschaffungsverhaltens auf den Märkten des Lebensmitteleinzelhandels in Deutschland, September 2014, Bonn.

Bundeskartellamt (2014b), Bundeskartellamt Imposes Fines on Sausage Manufacturers, Press Release: 15 July 2014, Bonn (available online: www.bundeskartellamt.de).

Chang, M.-H., (1991), The Effects of Product Differentiation on Collusive Pricing, International Journal of Industrial Organization 9, 453-469.

Choi, S.C., Coughlan, A.T., (2006), Private Label Positioning: Quality Versus Feature Differentiation from the National Brand, Journal of Retailing 82, 79-93.

Competition Commission (2008), The Supply of Groceries in the UK, Market Investigation, Final Report, Competition Commission (UK), London.

Cotterill, R.W., Putsis, W.P. Jr., (2000), Market Share and Price Setting Behavior for Private Labels and National Brands, Review of Industrial Organization 17, 17-39.

Deakin University Australia Worldly and Pitney Bowes Software (2012), New Location Perspectives in Retail: In the Zone, Retail White Paper (available online: http://www.pitneybowes.com.au/docs/Australia/PBS%20Retail%20Part%202-New%20Location%20Perspectives%202013_print.pdf).

Donsimoni, M.P., (1985), Stable Heterogenous Cartels, International Journal of Industrial Organization 3, 451-467.

Economides, N., (1993), Quality Variations in the Circular Model of Variety-differentiated Products, Regional Science and Urban Economics 23, 235-257.

Economides, N., (1989), Quality Variations and Maximal Variety Differentiation, Regional Science and Urban Economics 19, 21-29.

eMarketer (2015), Mobile Ad Spending to Overtake Print in the UK, 30 September 2015 (available online: www.emarketer.com).

Epling, N., (2002), Price Discrimination amid Heterogeneous Switching Costs: A Competitive Tactic of the Telephony Resale Fringe. Yale University, Mimeo.

EU (2014), Directive on Certain Rules Governing Actions for Damages, Directive 2014/1 04/EU of 26 November 2014, Official Journal of the European Union, 5.12.2015 (L349).

European Commission (1999), Buyer Power and Its Impact on Competition in the Food Retail Distribution Sector of the European Union, European Commission, Brussels.

European Commission (2011), The Impact of Private Labels on the Competitiveness of the European Food Supply Chain, European Commission, Brussels.

Ferguson, D., (2013), How Supermarkets Get Your Data- and What They Do with It, The Guardian, 8 June 2013 (available online: www.theguardian.com).

Frank, R.G., Salkever, D.S., (1997), Generic Entry and the Pricing of Pharmaceuticals, Journal of Economics and Management Strategy 6, 75-90.

Friedman, J., (1971), A Non-cooperative Equilibrium for Supergames, Review of Economic Studies 38, 1-12.

Fudenberg, D., Tirole, J., (1984), The Fat Cat Effect, the Puppy Dog Ploy, and the Lean and Hungry Look, American Economic Review Papers and Proceedings 74, 361-368.

Gabrielsen, T.M., Sörgard, L., (2007), Private Labels, Price Rivalry, and Public Policy, European Economic Review 51, 403-424.

Ganslandt, M., Persson, L., Vasconcelos, H., (2012), Endogenous Mergers and Collusion in Asymmetric Market Structures, Economica 79, 766-791.

Gilo, D., (2008), Private Labels, Dual Distribution, and Vertical Restraints -An Analysis of the Competitive Effects, in: Ezrachi, A. and Bernitz, U. (eds.), Private Labels, Brands and Competition Policy, The Changing Landscape of Retail Competition, Chapter 7, 2009.

Häckner, J., (1994), Collusive Pricing in Markets for Vertically Differentiated Products, International Journal of Industrial Organization 12, 155-177.

Häckner, J., (1996), Optimal Symmetric Punishments in a Bertrand Differentiated Products Duopoly, International Journal of Industrial Organization 14, 611-630.

Hamilton, J.H., Thisse, J.-F., (1992), Duopoly with Spatial and Quantity-Dependent Price Discrimination. Regional Science and Urban Economics 22, 175-185.

Hassan, D., Monier-Dilhan, S., (2006), National Brands and Store Brands: Competition Through Public Quality Labels, Agribusiness 22, 21-30.

Hoch, S.J., Banerji, S., (1993), When Do Private Labels Succeed?, Sloan Management Review 34, 57-67.

IBM: http://www-01.ibm.com/software/data/bigdata/platform/product.html

Inderst, R., (2013), Implications of Buyer Power and Private Labels on 'Vertical Competition' and Innovation, Report commissioned by Markenverband, Research and Consulting Services (October 2013), Goethe University Frankfurt/Main.

Irmen, A., Thisse, J.-F., (1998), Competition in Multi-characteristic Spaces: Hotelling Was Almost Right, Journal of Economic Theory 78, 76-102.

Jentzsch, N., Sapi, G., Suleymanova, I., (2013), Targeted Pricing and Customer Data Sharing Among Rivals, International Journal of Industrial Organization 31, 131-144.

Lederer, P.J., Hurter, A.P., (1986), Competition of Firms: Discriminatory Pricing and Location, Econometrica 54, 623-640.

Linkhorn, T., (2013), Retailers Use Variety of Ways to Track Consumer Habits Through Loyalty Programs, The Blade, 11 August 2013 (available online: www.toledoblade.com).

Liu, Q., Serfes, K., (2004), Quality of Information and Oligopolistic Price Discrimination. Journal of Economics & Management Strategy 13, 671-702.

Liu, Q., Serfes, K., (2005), Imperfect Price Discrimination in a Vertical Differentiation Model. International Journal of Industrial Organization 23, 341-354.

Liu, Q., Serfes, K., (2007), Market Segmentation and Collusive Behavior. International Journal of Industrial Organization 25, 355-378.

Mas-Colell, A., Whinston, M.D., Green, J.R., (1995), Microeconomic Theory, Oxford University Press, Oxford.

Mikians, J., Gyarmati, L., Erramilli, V., Laoutaris, N., (2012), Detecting Price and Search Discrimination on the Internet, Hotnets' 12, 29-30 October 2012, Seattle, WA, USA (available online: http://conferences.sigcomm.org/hotnets/2012/papers/hotnets12-final94.pdf).

Mills, D.E., (1995), Why Retailers Sell Private Labels, Journal of Economics and Management Strategy 4, 509-528.

Mills, D.E., (1998), Private Labels and Manufacturer Counterstrategies, European Review of Agricultural Economics 26, 125-145.

Prokop, J., (1999), Process of Dominant-cartel Formation, International Journal of Industrial Organization 17, 241-257.

Putsis, W.P., (1997), An Empirical Study of the Effect of Brand Proliferation on Private Label-National Brand Pricing Behavior, Review of Industrial Organization 12, 335-371.

Rasch, A., Wambach, A., (2009), Internal Decision-Making Rules and Collusion, Journal of Economic Behavior and Organization 72, 703-715.

Ross, B., (2013), Canned Goods–not Canned Prices With Personalized Pricing. How Grocers Can Implement a Personalized Pricing Strategy in Four Steps, Canadian Grocer, 24 July 2013 (available online: www.canadiangrocer.com).

Ross, T.W., (1992), Cartel Stability and Product Differentiation, International Journal of Industrial Organization 10, 1-13.

Rothschild, R., (1992), On the Sustainability of Collusion in Differentiated Duopolies, Economics Letters 40, 33-37.

Salop, S.C., (1979), Monopolistic Competition with Outside Goods, Bell Journal of Economics 10, 141-156.

Sapi, G., Suleymanova, I., (2013), Consumer Flexibility, Data Quality and Targeted Pricing, DICE Discussion Paper No 117.

Selten, R., (1973), A Simple Model of Imperfect Competition, Where 4 Are Few and 6 Are Many, International Journal of Game Theory 2, 141-201.

Shaffer, G., Zhang, Z. J., (1995), Competitive Coupon Targeting. Marketing Science 14, 395-416.

Shaffer, G., Zhang, Z. J., (2002), Competitive One-to-One Promotions. Management

Science 48, 1143-1160.

Steiner, R.L., (2004), The Nature and Benefits of National Brand/Private Label Competition, Review of Industrial Organization 24, 105-127.

Sun-Air versus SAS and Maersk Air, Commission decision 2001/716/EC [2001] OJ L 265/15, Case COMP/D.2 37.444, Official Journal of the European Communities, 2001.

Tabuchi, T., (1994), Two-stage Two-dimensional Spatial Competition Between Two Firms, Regional Science and Urban Economics 24, 207-227.

Thisse, J.-F., Vives, X., (1988), On the Strategic Choice of Spatial Price Policy, American Economic Review 78, 122-137.

Thomadsen, R., and Rhee, K.-E., (2007), Costly Collusion in Differentiated Industries, Marketing Science 26, 660-665.

Valentino-Devries, J., Singer-Vine, J., Soltani, A., (2012), Websites Vary Prices, Deals Based on Users' Information, The Wall Street Journal, 24 December 2012 (available online: www.wsj.com).

Valletti, T. M., (2002), Location Choice and Price Discrimination in a Duopoly, Regional Science and Urban Economics 32, 339-358.

Ward, M.B., Shimshack, J.P., Perloff, J.M., Harris, J.M., (2002), Effects of the Private-Label Invasion in Food Industries, American Journal of Agricultural Economics 84, 961-973.

Winterman, D., (2013), Tesco: How One Supermarket Came to Dominate. BBC News Magazine, 9 September 2013, (available online: www.bbc.co.uk).

Appendix

Ich erkläre hiermit an Eides Statt, daß ich die vorliegende Arbeit ohne Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Düsseldorf, 11. April 2016. Irina Hasnas