MSc Economics

Competition and Nonlinear Pricing in Telecommunications

A Master's Thesis submitted for the degree of
"Master of Science"

supervised by

Dr. Anna Hammerschmidt

Gábor Béla Uhrin

1026547

Vienna, 14 June 2012

MSc Economics

Affidavit

I, Gábor Béla Uhrin

hereby declare

that I am the sole author of the present Master's Thesis,

Competition and Nonlinear Pricing in Telecommunications

60 pages, bound, and that I have not used any source or tool other than those referenced or any other illicit aid or tool, and that I have not prior to this date submitted this Master's Thesis as an examination paper in any form in Austria or abroad.

Vienna, 14 June 2012

Signature

Contents

List of Tables

List of Figures

List of Appendices

Abstract

The aim of the present thesis is to develop a better understanding of the mechanisms underlying the imperfectly competitive environment of telecommunications industries. To this end, we first identify price schedules that are being commonly offered. These are: linear, two-part tariffs, three-part tariffs and flat-rates. As a second step, we systematically review linear and two-part tariff models of the two-way interconnection literature. We concentrate on equilibrium comparative statics with respect to the Hotelling transport cost parameter. Moving on to more complex price schedules, we extend a model of consumer overconfidence (denoted M). M is capable of explaining three-part tariffs. The extension (H) incorporates duopolistic price competition, and draws on the framework of competition in utility space. Finally, we propose an extension of H to H^+ that is motivated by the two-way interconnection literature: variable costs differ based on the network on which the call terminates.

Our results can be interpreted at two levels. At the general level, if the setting of a fixed fee is allowed, then if competition increases, the fixed fee is going to drop. Unit prices will be charged at marginal costs. If a fixed fee is not allowed to be set, then competition decreases the unit price both on- and off-net. Our specific contribution is that we are successful in extending M to H, and show that a symmetric equilibrium in the utility space exists. The extension does not alter the marginal price structure of M, which is a three-part tariff. The monopolistic and the prefect competitive results of M turn out to be flanking cases of the result from H. The transport cost parameter provides smooth transition between the two extremes, only affecting a fixed fee. We are not able to derive sharp conclusions from the extension H^+. Thus, we present a simple numerical example. The conclusions from this example are: we are able to find a symmetric equilibrium of the utility game, but this equilibrium might not exist in general; the price structure of H is not likely to H^+. Comparative static analysis, however, shows that our general level conclusions apply to this particularly simple setting.

Chapter 1

Introduction

"The telecommunications industry has been changing rapidly for years, but academic research is still lagging behind" – Laffont and Tirole (2000, p. XIII.)

The aim of the present thesis is to develop a better understanding of the mechanisms underlying the imperfectly competitive environment of telecommunications industries.

The telecommunications sector is a dynamic part of the economy in all OECD countries. After the liberalisation of the markets heavy price competition had arisen in many countries. This competitive pressure induced providers' pricing behaviour to change dynamically. As a result of this, various new pricing schedules had emerged and the application of more conventional price schedules had fallen out of common practice. As Laffont and Tirole point out, academic research is lagging behind this rapidly changing reality. For this reason, the present thesis identifies some real-life developments and investigates theoretical models from this new perspective.

Starting from models of linear pricing and concluding with models suitable for generating three-part tariffs[1] we give a systematic account on what the strategic variables are and how they react to changes in competition (suitably defined) within the respective model.

As mentioned above, the telecommunications industry is an important sector to analyse and understand. This is due to its inherent imperfections such as network externalities, oligopolistic environments, high sunk costs amongst others. As a result of these properties the sector is intensively monitored and heavily regulated in all OECD countries by the national regulatory authorities (NRAs). Therefore, it would not only be of academic interest if we were able to derive general conclusions

[1]By "generating" we mean that the optimal fully nonlinear price schedule will closely resemble a three-part tariff structure, and this result stems from the primitives of the model.

regarding the mechanisms underlying price setting and the effect of competition thereon.

To achieve this general understanding, we review existing models that are suit for the analysis of the telecommunications industry and concentrate on the equilibrium comparative statics. This review can easily be carried out in models using linear and two-part tariffs that are based on models of two-way interconnection—a model framework explicitly designed for the analysis of telecommunications industries (Armstrong and Wright, 2009; Laffont, Rey, and Tirole, 1998a,b).

Unfortunately, these models are only tuned towards simple price structures. If we want to examine models of more complex price schedules (namely three-part tariffs), then we have to turn to the monopolistic screening, nonlinear pricing literature.

Monopolistic models are, however, clearly unsatisfactory for our purposes. Therefore, the main specific contribution of this thesis is to extend an existing monopolistic nonlinear pricing model that is able to generate three-part tariffs (Grubb (2009) – henceforth called M) along two dimensions. First, a Hotelling oligopolistic environment (H) is introduced into M. Second, we propose an extension of H that utilizes the structure of a telecommunications industry better (henceforth H^+). In principle, H^+ is motivated by the two-way interconnection model framework.

Our results are positive. At the general level we are able to identify a pattern that is common to almost all the models under scrutiny: if fixed fees are allowed, then the intensified competition will drive this fixed fee down; if only linear pricing is available, then the unit price will be smaller the more intense the competition is. On the specific level, our main results are Proposition 3.3 and Corollary 3.4 which show that the main results of M will be flanking outcomes of H. Further, while getting to this result we demonstrate that the oligopolistic extension from M to H does not destroy the original price structure of M—which is a three-part tariff. From model H^+, however, we are not able to derive sharp results. Therefore we resort to an illustratory example that sheds light on the complications of solving for the equilibrium optimal nonlinear tariff in H^+. In particular, the properties of model H (and thus also M) are not likely to carry on to H^+.

The thesis proceeds as follows. Section 1.1 highlights some facts about the importance of the telecommunications sector and gives real-world examples for all the price schedules that we examine in detail. Chapter 2 gives a systematic review of the literature on two-way interconnection, thereby illustrating the main principles and themes in the models with linear and two-part tariffs. Chapter 3 consists of two loosely connected parts. First, Section 3.1 gives an introduction into some key notions of mechanism design and nonlinear pricing that are necessary to understand for the subsequent analysis. Second, Section 3.2 develops model H by extending model M. By discussing H, this section also highlights the important points in M. Section 3.3 illustrates the proposed extension from model H to H^+. Finally, Chapter 4 concludes.

1.1 Facts about the telecommunications sector

The OECD Communications Outlook 2011 (OECD, 2011) provides a detailed and up-to-date collection of facts and figures about the telecommunications industry based on various data sources. Highlighting some of these figures, and observing the patterns in the contractual tariffs among the Austrian mobile providers gives a well-grounded motivation for the subsequent discussion.

The importance of the telecommunications sector

An everyday observation might lead to the conclusion that mobile and Internet telecommunications play an important role in our life. This observation holds true for the macroeconomic and microeconomic levels. Table 1.1 shows the revenues from telecommunications as a percentage of the country's GDP for selected countries and the OECD as a whole.

Table 1.1: Telecom revenues in percentage of GDP

	Revenues in percentage of GDP							
Country	2002	2003	2004	2005	2006	2007	2008	2009
Austria	2.57	2.66	2.61	2.53	2.35	2.11	1.90	1.85
France	2.75	2.75	2.77	2.88	2.76	2.71	2.74	2.79
Germany	2.89	2.95	3.02	3.00	2.85	2.63	2.51	2.53
Hungary	5.79	5.58	4.68	4.63	4.44	4.19	3.74	3.62
Poland	3.48	3.53	3.80	3.77	3.76	3.43	3.38	3.22
United Kingdom	3.04	3.04	3.10	3.08	3.02	2.95	2.88	2.91
United States	3.21	3.07	2.93	2.89	2.78	2.71	2.72	2.71
OECD	3.14	3.08	3.00	2.97	2.89	2.82	2.76	2.81

Source: OECD (2011), `http://dx.doi.org/10.1787/888932397568` (29.05.2012)

Since the numbers in the table represent mobile telecommunication *revenues*, the price level of the sector affects the magnitude of the data. Thus, for example in the case of Austria, the large decline in the price level of this sector lowers the figures throughout the years (Lappöhn, Pohl, and Zucker, 2011). Nevertheless, even with the significant price decrease, this sector creates substantial revenues relative to GDP in each of the respective countries. These figures demonstrate that the telecommunications sector plays an important role in the macroeconomy of a country.

Table 1.2: Mobile phone subscriptions per 100 inhabitants

	Subscriptions per 100 inhabitants							
Country	2002	2003	2004	2005	2006	2007	2008	2009
Austria	83.33	87.39	97.75	101.65	111.74	118.52	127.2	136.7
France	62.64	67.22	71.33	76.55	81.75	87.06	90.4	95.3
Germany	71.69	78.53	90.08	96.04	103.99	118.10	130.6	132.2
Hungary	67.79	78.43	86.35	92.40	98.95	109.69	121.8	117.7
Poland	36.35	45.56	60.49	76.43	96.36	108.59	115.2	117.4
United Kingdom	83.52	88.65	99.76	108.72	115.15	120.99	125.0	129.9
United States	51.27	54.55	62.90	71.87	80.82	87.06	85.7	89.2
OECD	59.25	64.16	71.96	79.88	87.79	96.06	99.9	102.6

Source: OECD (2011), `http://dx.doi.org/10.1787/888932398005` (29.05.2012)

Table 1.2 displays the number of mobile phone subscriptions per 100 inhabitants for the same sample of countries. It shows more about the patterns of subscriptions: starting from the early 2000's, the mobile penetration rate grew heavily, now amounting to over 100 per cent.[2]

If we consider increasing penetration rates together with the fact that mobile telecommunication revenues constitute a significant part (above 60 per cent in most highlighted countries) of the total telecommunication revenues (OECD, 2011, Table 3.4, p. 113), then we can conclude that the telecommunications industry plays an important role both at the macro- and microeconomic levels. This conclusion leads to the need of understanding this industry, its patterns, and most importantly its mechanics. Most fruitfully, this analysis can be done in simple theoretical microeconomic (industrial organisation – IO) models.

Trends in pricing

If we would like to examine (or build) models of telecommunications, then we first have to decide on the strategic variable and look at its characteristics. This, in the case of the telecommunications, is unarguably the price or, in more complicated situations, the price schedule. Through casual observation we can identify four common pricing schedules: linear tariffs, two-part tariffs, three-part tariffs and flat-rate tariffs.

Table 1.3 illustrates that three-part tariffs do indeed exist and are common in practice by giving examples for contractual tariffs of the four Austrian main mobile phone providers: A1, Hutchinson 3G, Orange, T-Mobile.[3] The numbers make

[2] This implies that many subscribers have more subscriptions, perhaps some not heavily used. For example, in Hungary it is common to offer tariffs with two SIM-cards that give the opportunity of free calling between the two SIM-cards included. Besides this, however, the outbound tariffs might be worse than with other subscriptions.

[3] The table concentrates on voice telephony. Inclusive SMS-s are displayed to give an idea about

Table 1.3: Contractual tariffs in Austria

Provider	Tariff	Inclusive Minutes	Inclusive SMS (pcs)	Price (€) Fix	Price (€) Overage	Comment
A1*	Smart 1	1000	500	19.9	0.29	
	Smart 2	2000	1000	29.9	0.29	
	Smart 3	3000	1000	39.9	0.29	
	Smart Unlimited	∞	∞	59.9	0	Flat-rate
	Talk	500	0	14.9	0.29	On-net calls free
	Jugend	2000	1000	19.9	0.29	Youth tariff
H3G*	Light	250	0	5	0.35	
	L	1000	1000	20	0.35	
	XL	2000	1000	30	0.35	
	XXL	3000	1000	40	0.35	
	Comfort	1000	100	10	0.35	Limited data usage
	Comfort Plus	1000	100	15	0.35	
Orange	All in 10	1000	100	10	n.a.	1000 plus minutes for € 10
	All in 15	1000	1000	15	n.a.	1000 plus minutes for € 10
	All in 20	2000	1000	20	n.a.	1000 plus minutes for € 10
	All in 30	3000	1000	30	n.a.	1000 plus minutes for € 10
	All in 40	4000	1000	40	n.a.	1000 plus minutes for € 10
	Supernet 3000	1000	2000	15	n.a.	–"–, youth tariff
T-Mobile	Hit 1000	1000	0	8	0.29	No data volume included
	All Inclusive	1000	1000	15.92	0.29	

*: The contract includes a phone when first signed.

Sources: http://www.orange.at/Content.Node/tarife/; http://www.a1.net/handys-telefonie http://shop.t-mobile.at/002/1_1_1/10026/index.html; http://www.drei.at/webshop (all 15.05.2012).

obvious that firms usually offer so-called *three-part tariffs*, where a fixed fee contains inclusive minutes and over these inclusive minutes a per-minute price must be paid. In the case of provider A1's Smart Unlimited tariff we can see an example of a *flat-rate tariff*, i.e. a fixed fee with unlimited usage.

The pre-paid subscriptions, i.e. where the consumer buys a certain amount of minutes and uses them up can, by definition, be perceived as a *linear tariff*.

An example to the existence of pure *two-part tariffs*, i.e. where the payment of a fixed fee is followed by unit pricing from the first minute onwards, is harder to find. However, the inspection of some Hungarian contractual mobile tariffs shows that they are closer to being two-part than three-part, especially compared to Austria. Choosing one particular tariff, the Basic Tariff 1 (Alaptarifa 1) from Telenor Hungary we can make the following observations:[4] a) The monthly fee is € 5.5, b) The per-minute price is € 0.13, c) The monthly fee can be used up for payment, thus approximately 40 minutes are included.[5] Compared to the Austrian tariffs, it can be inferred, that this tariff is close to being two-part.[6]

why seemingly similar tariffs differ. Included data volumes are usually unlimited except for Orange. These latter two services are, however, not part of the subsequent analysis.

[4]Calculating with the rate 100 Hungarian Forints equal 34.5 Eurocents.

[5]Source: http://www.telenor.hu/szamlas-tarifa/alaptarifa/ (29.05.2012).

[6]A comprehensive analysis of the Hungarian tariffs would be more complex than it is for Austria. This is due to the greater complexity and variety of the tariffs being offered. Sticking to the example

Having given examples for all the tariff types mentioned above, the need of examining competition in all of these pricing schedules becomes evident. The rest of the pages documents this adventure. Welcome on board!

---

of Telenor Hungary, 16 different tariffs can be counted. The youth division of Telenor Hungary runs under a different name (djuice), and offers 7 different tariffs. Thus altogether, Telenor Hungary offers 25 tariffs. This great variety is also observable at the other two providers: T-Mobile Hungary and Vodafone Hungary.

Chapter 2

Competing in linear and two-part tariffs

Analysing the game of two (or more) service providers in telecommunications simply as a Bertrand price setting game would lead to the trivial result that (given symmetric providers and perfect information) the price of one minute call is equal to its marginal cost. This result, however, is hardly observed in the real world, hence a more sophisticated reasoning and a more satisfactory model is needed.

Accordingly, Laffont, Rey, and Tirole (1998a) argue, the telecommunications industry was in need of a theoretical framework in the caller party pays (CPP) principle[1] that allows regulators and academics to evaluate regulatory intervention, pricing behaviour, welfare and to identify the relationships and channels through which these measures act. To this end, seminal papers of Armstrong (1998), Laffont, Rey, and Tirole (1998a,b) created a simple framework that allowed academics to organise thoughts. Due to their simplicity, these models did not necessarily match even the most straightforward and intuitive real world observations, thus, much work has been done to accommodate these mismatches.

Since the seminal papers mentioned above, much progress has been made for the understanding of the economics of two-way interconnection. Starting from linear, non-discriminatory prices, the literature has incrementally built up a model that includes two-part tariffs and termination based price discrimination. This framework gives a consistent way of thinking about the—possibly nontrivial—interaction between two mobile or one fixed and one mobile provider.

Models of two-way interconnection are the presently the most advanced frameworks that allow us to think about telecommunications industries. Thus, in the following I present a simple model of such kind following Armstrong and Wright (2009)

[1] This is the most widespread principle that is applied all throughout Europe. In the US the receiver pays.

that—due to its simplicity—allows me to concentrate on the main points of pricing behaviour, and the effect of competition thereon. Their model is in the spirit of Gans and King (2001).

2.1 Different levels of model sophistication

The key observation of the models of two-way interconnection is that a call that is *originated* in network A and *terminates* in network B requires the two networks A and B to be interconnected. In case of telecommunications, the networks are obviously reciprocally interconnected, hence the term *two-way* interconnection.

If we assume that the retail market and the market for termination are different markets and let providers charge different prices, then this creates very interesting tradeoffs and channels that affect the pricing of these services. The reason for this is that a network operator acts as a *competitive bottleneck* for its own customers on the termination market thereby having quasi-monopoly power in this respect. Thus, a network operator has the incentives to get as many customers as it can under his belt for the prospect of the termination rents that these customers will generate. This results in a stronger retail market competition. However, if operator A drives out operator B from the market, there will be no termination rents at all. Understanding this interaction, therefore is not straightforward.

Different levels of model sophistication can be examined, all of which have their own channels of "exercising" competition and each have their own equilibrium properties. Since in the basic frameworks we have two networks, each operating on a retail and a termination market, each firm $i \in \{0,1\}$ has the option to choose the optimal value of the following instruments: p_i for price per minute of on-net calls, $\hat{p}_i$ for price per minute of off-net calls, F_i for the fixed (subscription) fee and a_i for termination charge. These instruments allow for a wide variety of model setups with the combination of the following properties:

1. Termination rates can be reciprocal ($a_i = a$ for $i = 0,1$) or unilaterally chosen.

2. Per minute prices can be uniform ($p_i = \hat{p}_i$ for $i = 0,1$) or discriminated based on termination.

3. The pricing schedule can be linear ($F_i = 0$ for $i = 0,1$), or can consist of two part tariffs.

At this first sight, we can already infer something about the possible channels of competition. As a generic result in two-part tariff pricing, the important instrument of earning a market share will be the fixed fee. If competition is in linear prices, all interaction will be picked up in the per minute prices. Termination rates may act as channels of competition if they are unilaterally chosen. This is less so, if we allow for termination based price discrimination. Discriminated prices also allow us to separate the effects of termination charges on off-net and on-net prices.

2.2 Baseline model

Customers

Suppose that there is a unit mass of customers distributed uniformly on the Hotelling line[2] where at the two endpoints 0 and 1 firm $A \equiv x_0$ and $B \equiv x_1$ can be found respectively. Transport cost (the exogenously given level of product differentiation) is denoted by τ. In the Hotelling model each consumer can be uniquely identified by his location on the line. Thus, a consumer located at $x \in [0,1]$ on the line, calling (on average) $q \in \mathbb{R}_+$[3] minutes, having income $y \in \mathbb{R}_+$ and joining network $i \in \{0,1\}$ obtains the utility of

$$v_0 + y - \tau|x - x_i| + u(q)\,, \tag{2.1}$$

where v_0 is the utility gain from subscribing to either one of the networks. For simplicity, v_0 is assumed to be big enough to ensure full participation. $u(q)$ is some well-behaved utility function, i.e. satisfying $u'(q) > 0$, $u''(q) < 0$ and $\lim_{q\to\infty} u'(q) = 0$, $\lim_{q\to 0} u'(q) = \infty$.[4]

The consumer's demand function for call minutes is

$$q(p) = \arg\max_q \{u(q) - pq\}\,, \tag{2.2}$$

and the variable net surplus is for this above given demand function is

$$v(p) = \max_q \{u(q(p)) - pq(p)\}\,. \tag{2.3}$$

Consequently, $v'(p) = -q(p)$.

Given the triple $(p_i, \hat{p}_i, F_i)$, i.e. on-net, off-net prices and fixed fee respectively, and assuming that the calling pattern is balanced the total net surplus from joining network i is

$$w_i = s_i v(p_i) + (1 - s_i) v(\hat{p}_i) - F_i\,, \tag{2.4}$$

where s_i is the market share of firm i. In the Hotelling specification consumer lo-

[2] See for example Shy (1995) or Tirole (1988).

[3] Here and henceforth $\mathbb{R}_+$ denotes the nonnegative reals (i.e. including zero), and $\mathbb{R}_{++}$ denotes the strictly positive reals.

[4] An example of this would be—as in Laffont, Rey, and Tirole (1998a)—the function $u(q) = \frac{q^{1-1/\eta}}{1-1/\eta}$.

cated at s_i is indifferent between joining the two networks if and only if

$$w_i - \tau s_i = w_j - \tau(1 - s_i)\,. \tag{2.5}$$

This implies that the market share of network i is made up from precisely those consumers that are located between x_i and s_i. Therefore, in the unit interval this results in the well known formula

$$s_i = \frac{1}{2} + \frac{w_i - w_j}{2\tau}\,. \tag{2.6}$$

Firms

Each network $i \in \{0,1\}$ incurs fixed cost f per subscriber. Originating a call results in cost c_O, while terminating a call costs c_T. Further, if network i's customer calls a customer in network j, then a unilateral or negotiated access charge $a_i \geq -c_O$[5] has to be paid by i to j. Given these specifications, the costs of providing an on-net call is $c_{\text{on}} = c_O + c_T$, whereas the costs of providing an off-net call is $c_{\text{off}} = c_O + a_i$.

With respect to pricing, for the moment, there are two different alternatives: linear pricing and two-part tariffs. Linear tariffs simply involve a pair $(p_i, \hat{p}_i)$ for on- and off-net calls per minute respectively, whereas two part tariffs constitute a triple $(p_i, \hat{p}_i, F_i)$, where F_i is a fixed fee paid by the customer e.g. monthly.[6] On the level of the model setup, allowing for the decision over the full triple is without loss of generality. When solving for different variations of the model, restrictions can be made appropriately.

Given market shares s_i, $s_j = 1 - s_i$, demand function $q(p)$ and triple $(p_i, \hat{p}_i, F_i)$, firm i's profit can be written as

$$\pi_i = s_i \cdot \left(\underbrace{s_i q(p_i)(p_i - c_{\text{on}}) + (1 - s_i) q(\hat{p}_i)(\hat{p}_i - c_{\text{off}}) + F_i - f}_{\text{revenues from own customers}} \right) + \underbrace{(1 - s_i) \cdot s_i (a_i - c_T) q(\hat{p}_j)}_{\text{termination revenues}}\,. \tag{2.7}$$

[5] This assumption rules out that the cost of outgoing calls is negative. This would induce the possibility of making arbitrarily large revenues by simply originating calls to the rival's network infinitely often.

[6] Note, that in this case, there are no quantity allowances for this fixed fee. The fixed fee must be paid and above that, the customer pays per minute. If the fixed fee also involved quantity allowances then we would talk about *three-part tariffs* and if the price per call was zero, then we would have *flat-rate tariffs*.

Timing

1. Firms negotiate cooperatively a reciprocal access charge *or* firms choose termination charge unilaterally,

2. Firms choose simultaneously, non-cooperatively the pair $(p_i, \hat{p}_i)$ *or* the triple $(p_i, \hat{p}_i, F_i)$.

3. Upon observing these, the customers make their subscription decision and their calling quantity decision.

Equilibrium

Berger (2004a) calls the equilibrium in the subscription strategies of the third subgame *consumer equilibrium*. The equilibrium is the subscription decision of the individual from which she has no incentive to deviate given that all other individuals play their equilibrium subscription strategies. Put differently and more concisely, the equilibrium subscription decisions simply result in equilibrium market share for firm i such that none of its consumers has the incentive to deviate to firm j and vice versa.

Notice that this consumer equilibrium is not unique. With the plausible assumptions of: 1. on-net prices being lower that off-net prices; 2. prices being the same for both networks; 3. goods being sufficiently homogeneous, it is an equilibrium that one of the two networks has market share $s_i = 1$, i.e. it corners the market Berger (2004b). This is true, because if a customer expects that all other customers subscribed to network i, then she has no incentive to subscribe to network j given that the on–off-net price differential is sufficiently large compared to the transport costs τ.[7]

Generically, however, we are not interested in such corner solutions, but rather in a shared equilibrium. Equation 2.6 and the arguments preceding it shows a formulation of this shared equilibrium in an implicit way (w_i contains s_i). Under appropriate conditions, the shared equilibrium exists and is stable.[8]

To simplify the game further, it is solved for a symmetric equilibrium, i.e. where $s_i = 1/2$ and $(p_i, \hat{p}_i, F_i) = (p_j, \hat{p}_j, F_j)$. This greatly simplifies the calculations and the analysis.

For LINEAR, NONDISCRIMINATORY PRICES AND NEGOTIATED ACCESS CHARGES Propositions

[7]As to which cornered solution arises, it is the usual question that arises in coordination games, like the "battle of sexes".

[8]Detailed treatment can be found in Laffont, Rey, and Tirole (1998a) and in even more detail in Berger (2004b).

1 and 2 of Laffont, Rey, and Tirole (1998a) show interesting results based on the utility function $u(q) = \frac{q^{1-1/\eta}}{1-1/\eta}$, with $\eta > 1$. According to Proposition 1 (p. 10), for small enough a and large τ, a unique symmetric equilibrium for prices exists.[9] However, for small τ and large a, there is no equilibrium. This observation is interesting, especially if we see it jointly with part ii. of Proposition 2 (p. 11). This states that if τ decreases then so does p^*, the equilibrium price and if τ goes to zero then p^* approaches the socially optimal price. Putting this result differently: if the networks become closer substitutes, competition drives prices down ultimately to the socially optimal level. Due to the nonexistence result mentioned above, in order to guarantee the existence of the equilibrium, a has to come close to its cost, c_T.[10]

In the case of linear, nondiscriminatory prices and UNILATERAL TERMINATION CHARGES, we might ask whether this termination charge can become an instrument of competition between the two networks. The answer is, on the one hand, positive: it does act as a strategic device. However, the direction of the termination charges in response to increased competition (lower τ) is ambiguous according to Proposition 4 (p. 14). Note, however, that the proposition gives a condition for the decrease of a if τ decreases with a small amount from the no substitution case: a decreases if $\pi_0 < (\eta - 1)f$, where π_0 is the equilibrium profit in the case of no substitution.

In LINEAR, DISCRIMINATORY PRICES, the results of Laffont, Rey, and Tirole (1998b), Berger (2004a) and Berger (2004b) make the analysis intuitive and more plastic. The equilibrium on- and off-net prices can be found by intersecting two curves: a linear, and a nonlinear one. The linear curve is the so-called proportionality rule

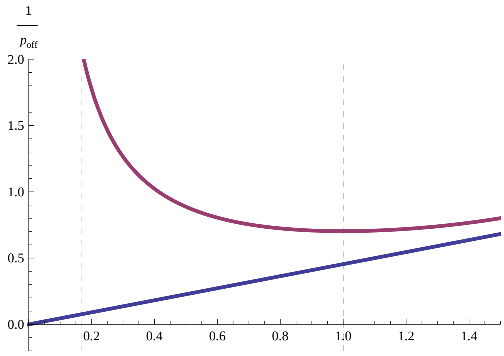
$$p_{\text{off}} = \left(1 + \frac{a - c_T}{c_O + c_T}\right) p_{\text{off}} \tag{2.8}$$

whereas the second curve is a non-linear curve giving the value of p_{off} dependent on a, η, τ, c_O, c_T and p_{on}. Following Berger (2004a), if we express these equations in reciprocals, it is easy to visualise in the $\left(\frac{1}{p_{\text{off}}}, \frac{1}{p_{\text{on}}}\right)$ coordinate space. This is done in Figure 2.2.

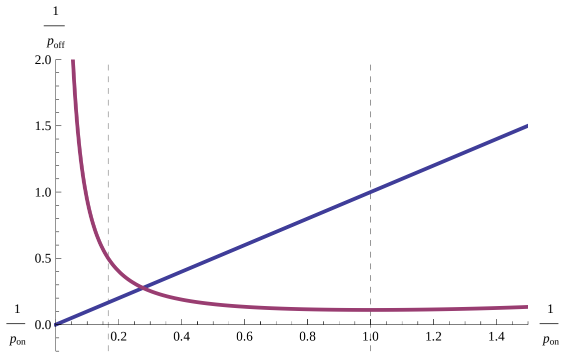
The relevant region where we have to pay attention to the behaviour of the curves is within the two dashed lines. The line closer to the y axis is the reciprocal of the monopolistic price $p^M = \eta(c_O + c_T)/(\eta - 1)$, which is calculated from the relation of the price–cost-margin and η, the elasticity of demand (cf. Tirole (1988)). The second line is the reciprocal of the total cost $c_O + c_T$. As we can see, in Subfigure (a),

[9] But the prices cannot be given explicitly, only implicitly.

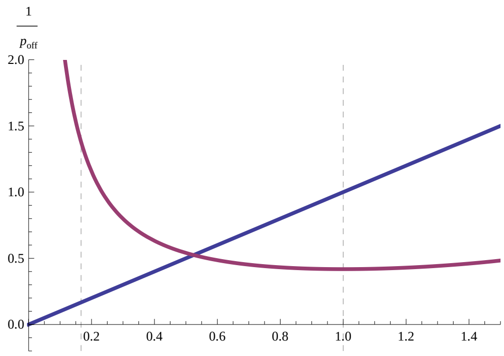
[10] Part i. of Proposition 2 raises the issue that termination charges can be instruments of tacit collusion insomuch as p^* is increasing with a. This view has rapidly been dispensed with even in Laffont, Rey, and Tirole (1998a) but also in most other follow-up works.



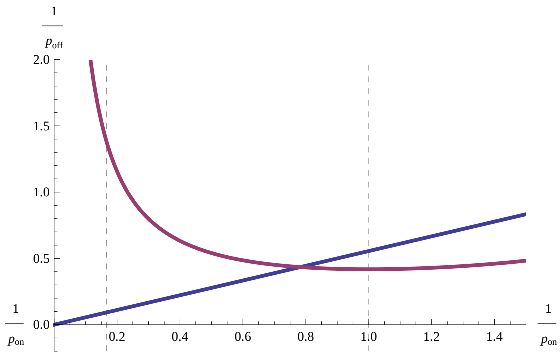
(a) Equilibrium may fail to exist ($a = 1.8 > c_T$, $\tau = 0.34$)



(b) High τ increases both prices ($a = c_T = 0.5$, $\tau = 1.78$)



(c) Low τ decreases both prices ($a = c_T = 0.5$, $\tau = 0.8$)



(d) High a decreases p_{on}, increases p_{off} ($a = 1.3 > c_T$, $\tau = 0.8$)

Figure 2.1: Comparative statics in the linear discriminatory price model

The equilibrium may fail to exist for small τ and large a; in Subfigures (b)-(c), both on- and off net prices increase with τ (note that the values are in reciprocals); in Subfigure (d) The increase of the access charge a decreases p_{on}, but increases p_{off}, regardless of the value of τ. These results are all intuitive, and analytically provable (Laffont, Rey, and Tirole, 1998a).

NONLINEAR, DISCRIMINATORY PRICING allows us to separate a different channel of competition: the fixed fee. To see this, consider the equilibrium results from the above, baseline model following Armstrong and Wright (2009)[11]:

$$p_i = c_{\text{on}}\ ;\ \hat{p}_i = c_{\text{off}} \tag{2.9}$$

$$F_i = f + \tau - v(c_{\text{on}}) + v(c_{\text{off}}) \tag{2.10}$$

$$\Pi'(a)|_{a=c_T} < 0 \tag{2.11}$$

First of all, the prices of calls are equal to their respective perceived marginal costs. Thus, the only factor that causes the difference between the on- and off-net

[11] See Appendix A for detailed calculations.

prices is their cost difference. This cost difference is only caused by the termination price.

Second, the fixed fee is set such that the firm extracts above cost (f) profits from the higher product differentiation (τ) and the difference in the consumers' indirect utility from calling in and out of the network: $v(c_{\text{on}}) - v(c_{\text{off}})$. If the termination charge is high, so that this "surplus" is positive, the firm optimally lowers its fixed fee in order to gather more customers under its belt therefore extracting higher termination rents. If a is lowered (or cut), this results in a higher fixed fee (besides lower per-usage fees).[12] Further, observe the result that firms compete de facto in the fixed fees. Since the optimal usage prices are cost based, τ, the product differentiation parameter only appears in the optimal fixed fee. Thus, firms compete for subscriptions through adjusting the fixed fee depending on the degree of product differentiation which can be thought of as a proxy of the intensity of competition. The closer the products are (lower τ), the more intense firms' interaction is.

Thirdly, even though the actual value of the termination charge depends on the specification of the demand function, under the fairly general assumption that $q(p) > 0$ for $p > 0$, we can establish that in the ε-vicinity of the termination marginal cost c_T firms have the incentive to charge *lower* termination charges. This result is in contrast with the first results of the literature. Those models, however, use simple linear pricing (without price discrimination).

Note the puzzling implication of this result with respect to retail pricing: since $c_{\text{off}} = c_O + a$ and $c_{\text{on}} = c_O + c_T$, if $a < c_T$ then $\hat{p} < p$ from Equation 2.9, which is clearly not the case in the real world. Why can it be optimal to set such a termination price that makes off-net calls less expensive than on-net calls? If operator A ceteris paribus charges below cost termination prices so that operator B will choose lower off-net prices the result will be that customers of B call more to A thereby generating higher termination revenues for A. On the subscription level observe, that lower a results in higher F, i.e. this termination pricing decision relaxes competition in the subscription market. This is also intuitive: if $\hat{p} < p$, then customers prefer joining the network with smaller market share.

As regards welfare, since the socially optimal prices are equal to their respective (exogenous) marginal costs $c_O + c_T$, the first relationship above also implies that the socially optimal termination charge equals its marginal cost c_T. Thus, the incentive to charge lower termination charges also results in that they are going to be below

[12] This observation is related to the waterbed effect. The *waterbed effect* is the phenomenon that a drop in rents from termination services (e.g. due to regulation) might result in an increase in the price of retail market services. Viewed from the opposite perspective it means that the profit per customer generated through termination will be to some extent competed away on the retail market (Jullien, Rey, and Sand-Zantman, 2010). Genakos and Valletti (2011a) and (2011b) show the existence and the magnitude of this effect in relation to the fixed-to-mobile termination.

the socially optimal.

The below cost termination charge result is only true for MTM termination. If we analyse the fixed-to-mobile termination rates between an incumbent fixed network operator and a mobile operator (assuming that fixed and mobile calls constitute two separate markets) we can easily intuitively conclude that the profit maximising FTM termination charge will be the monopolist one. To argue, why this is so, note that since the two markets are separate, a mobile operator does not compete with the fixed operator for the customers. Instead, they want to extract the maximal amount of profits from terminating FTM calls. Thus, the profit maximising FTM termination charge will generically be the monopolist charge.[13] This result is even more appealing if we note that the per-usage call prices charged by fixed operators are widely and tightly regulated. Thus, even if we expect that termination rates will raise FTM call charges thereby lowering call demand and consequently termination rents, most probably this cannot happen due to the regulated prices.

We still did not comment on NONLINEAR, NONDISCRIMINATORY PRICES. The reason for this is that the results are quite similar to that of the discriminatory case. In particular, the following results hold for a symmetric shared equilibrium (if exists):

$$p = c_{\text{on}} + \frac{a - c_T}{2} \tag{2.12}$$

$$F = f - \frac{(a - c_T)}{2} q(p) + \tau \,. \tag{2.13}$$

Here, the intuition is, that the uniform price will contain the expected costs of outbound and inbound calls. The symmetric equilibrium may fail to exist because of too low substitutability or too high termination charge.

2.3 Extensions

2.3.1 Call externalities

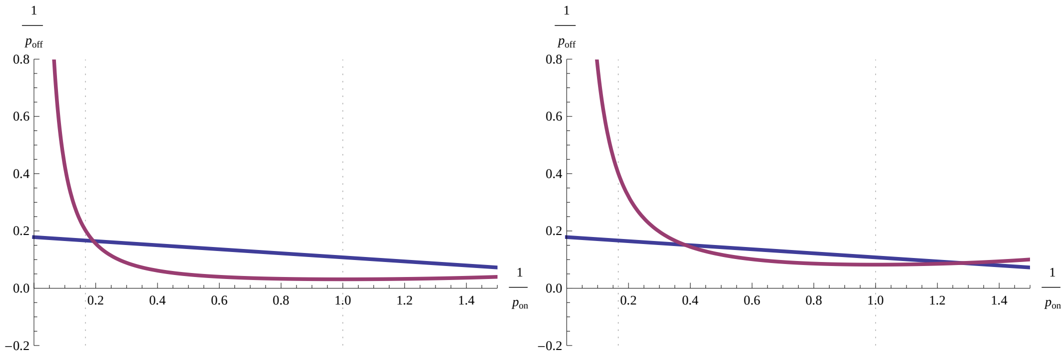
One side of the research agenda of Ulrich Berger in his economics doctoral dissertation (Berger, 2004b) was to incorporate demand side network externalities into the baseline framework models.[14] The justification for this seems natural: we do not

[13] A slightly more elaborate treatment can be found in Armstrong and Wright (2009). The intuition, however, is already described.

[14] The second theme of this research work was to dynamise market shares thereby in a way conducting a robustness check for the static models of two-way interconnection. This extension tackles an important question whose outcomes cannot be seen in a straightforward way. According to the

only care for calling, we do care about being called. In fact, as Berger (2004b) argues, by answering the phone we "reveal" that we gain at least as high utility from being called as from rejecting it. Call externalities are modeled by the condition $\bar{u}(q) = \beta u(q)$. In this condition $u(q)$ is the utility from outwards calling, $\bar{u}(q)$ is the utility from receiving calls and $\beta \geq 0$.

Berger (2004a) embeds call externalities into a linear, discriminatory pricing model, whereas Berger (2005) extends the discussion to a two-part tariff, discriminatory pricing model. In the former, the call externality parameter β creates an ambiguity with respect to the effect of τ on the off-net prices. This result is illustrated in figure 2.2 which is to be interpreted analogously to Figure 2.2.[15]



(a) $\beta = 0.96$, $\tau = 5.04$

(b) $\beta = 0.96$, $\tau = 4.23$

Figure 2.2: The effect of high call externalities on off- and on-net prices

As standard results show, the on-net price falls as τ falls. However, lowering τ combined with high β results in an upward pressure to the off net price because of the *negative* slope of the first (blue) curve. This is due to the intuition that if provider A charges high off-net prices, it restricts the utility that provider B can provide to its customers, since customers care about being called (the more, the higher β is), but the amount of off-net calls from A is decreasing if $\hat{p}_A$ increases. Further, if the differentiation is sufficiently low, consumers have an incentive to join A. This incentive is strengthened by the low on-net price that A is able to offer due to high off-net prices.

Berger (2005) reinforces the results that we have already seen in the framework of two-part tariffs with termination-based price discrimination. There is only one different observation: cost-based termination pricing can never be socially optimal. Thus, it can be the case that bill-and-keep agreements are welfare improving, in

results, for fixed price structures and at least three networks, the stable consumer equilibrium need not exist, except for the case where the off-net price is below the on-net price (Berger, 2004b, p. 64).

[15] Note, however, the different scales on the y axis. In Figure 2.2, the range runs from -0.2 to 2, whereas in Figure 2.2, the upper bound is 0.8. This is for the ease of display.

contrast to Gans and King (2001) who show that zero termination charges can lead to the softening of retail competition.

2.3.2 Low termination charges?

As we have seen above, one of the main outcome of the network competition model with two-part tariffs and price discrimination is that MTM termination charges will be set too low. This outcome is puzzling on two levels: 1. consequently, the off-net price will be below the on-net charge, and 2. it has never been the concern of any regulator that MTM charges will can be too low—even below the socially optimal level. This discrepancy between the model and the intuition have been tried to be treated recently by Armstrong and Wright (2009) and Jullien, Rey, and Sand-Zantman (2010).

Armstrong and Wright (2009) argue that the huge difference between MTM and FTM termination prices (remember, the FTM charge was at the monopolist level) allows for arbitrage possibilities. For example, the fixed provider can set up a device that routes all outgoing calls through either one of the MTM networks.[16] This leads to he natural conclusion that if we impose that $a = \tilde{a}$, i.e. the MTM and FTM termination charge is the same, then it can lead to the result that the uniform termination charge will be below the monopolist level, but above the perceived marginal cost if the networks choose their access charges *unilaterally* (Armstrong and Wright, 2009).

Coordination between the two mobile providers does not yield this result. This is due to the fact, that if we incorporate FTM termination charges into the above described model, $R(\tilde{a})$ the revenue from it enters both the profit function and the optimal fixed fee additively, with opposite signs, thus they cancel out each other.[17] Consequently, the industry profit is also independent of $\tilde{a}$, the FTM termination charge.

Jullien, Rey, and Sand-Zantman (2010) propose an interesting and plausible solution to this problem. In their model, which is a two-part tariff model in both nondiscriminatory and discriminatory prices, the demand side is made up of heavy and light usage customers. Heavy usage customers constitute a fix unit mass, whereas the subscribing amount of low usage customers ($\tilde{\alpha}_T$ in their notation) is determined endogenously depending on the fixed fee offered to them. To make their point, they assume that the light usage consumers do not call, they are only being called. How-

[16]Interestingly, this happened in France where, until recently, the MTM termination charges were zero (Genakos and Valletti, 2011a).

[17]To see this more clearly, in Appendix A. if we modify the profit function A.5 to $\tilde{\pi}_i = s_i \cdot [\tilde{F}_i - f + (1-s_i)(a-c_T)q(c_{\text{off}}) + R(\tilde{a})]$, and substitute $\tilde{F}_i = f + \tau - v(c_{\text{on}}) + v(c_{\text{off}}) - R(\tilde{a})$, then $R(\tilde{a})$ cancels out.

ever, they also demonstrate that their results are robust to releasing this assumption (as long as the light users call sufficiently less than the heavy users). In this setup they are able to generate MTM termination charges that are above their costs, *and* above the socially optimal level, giving ground to conventional regulatory wisdom.

2.4 Two part tariffs and beyond

The analysis of interconnecting networks in linear tariffs is clearly an oversimplification of the observed telecommunications industries. Nevertheless, the close interaction between unit pricing, access pricing and market share building creates and interesting intellectual exercise that we have tried to investigate above.

The possibility of charging a two-part tariff in the presence of fixed fees has two merits in my view: 1. it allows us to disentangle unit pricing from market share building, thereby simplifying the analysis of competition to basically one strategic variable (fixed fee) and 2. it is plausible that a two-part tariff is the optimal fully non-linear tariff. This latter statement is true for homogeneous consumers, who do not differ in, say, calling demands only with respect to subscription demand.

In the case of heterogeneous consumers, pricing has an additional role besides acting as strategic devices towards the competitor: they act as strategic devices towards the customers insomuch as they allow to separate different types through (possibly complex) type-dependent mechanisms. As Armstrong and Vickers (2001), Rochet and Stole (2002) and Armstrong and Vickers (2010) show, two-part nonlinear tariffs still can be optimal in a general competitive environment with heterogeneous consumers in the presence of fixed cost plus constant marginal cost. These results (sometimes interpreted as "no screening results") show that a large amount of intuition can be used to replace mechanism design in possibly complex market situations.

However, as Dessein (2003) and Hahn (2004) show, this result disappears if the termination charges are away from their costs. Hahn (2004) characterises the fully nonlinear separating equilibrium price schedule for a continuum of consumer types $\theta \sim F(\theta)$ with support $[\underline{\theta}, \bar{\theta}]$ which is assumed to be independent from the "location" type $x \in [0,1]$ (distributed uniformly on its support). In this scenario, the boundary types pay according to the marginal cost, but the interior types pay unit prices larger or smaller prices than marginal costs if $a > c_T$ or $a < c_T$ respectively (Hahn, 2004, Proposition 2.).

This result shows that finishing our investigation in a two-part tariff model is not satisfying. In fact, per call prices can be lower than their costs, therefore the possibility of zero per-call prices (either for all quantities or for some quantities) is possible.[18] This leads us to the fuller exploration of the possibility of flat-rate and, especially, *three-part tariffs*.[19]

[18]But not for all types. The boundary types pay according to marginal costs.

[19]Further, Jensen (2006) argues that in a duopolistic environment it may even not be feasible to implement a two-part tariff. However, a three-part tariff can be implementable even if the two-part is not feasible.

Chapter 3

Competing in three-part tariffs

Generally, a k*-part tariff* consists of a fixed fee (the first part) and $k-1$ marginal prices for different levels of consumption. A *three-part tariff* in our usage will be a tariff, where the marginal price for the first Q units is zero and strictly positive afterwards.[1] Put differently, in return for the fixed fee there is a quantity allowance and usage based price above that.

As we can see, a three-part tariff is a quite complex pricing schedule which includes three variables for one firm to decide upon: 1. the fixed fee, 2. the cutoff quantity below which the marginal price is zero and 3. per-usage price above the cutoff quantity.

The investigation of these tariffs thus falls into the investigation of non-linear pricing. What is important and interesting to see is that standard economic theory with standard simplifying assumptions cannot easily give a rationale for such complex price structures. According to standard results (cf. Wilson (1993)) optimal pricing schedule for a monopolist should be usage based. Indeed, in the view of the discussion of competition in telecommunications in the preceding chapter, even though two-part tariffs need not be optimal, there will always be types, who pay cost-based prices (Hahn, 2004).

Hence, to create a model environment in which three-part tariffs are set in optimum is a challenging intellectual exercise. There are two ways to argue for the optimality of such tariffs: I. make usage-based pricing costly so that it becomes (at least weakly) optimal to incorporates flat parts in the tariffs[2] and II. approach from the demand side. In the present discussion we stick to the second alternative. In setting a proper demand side environment, there are two ways to proceed: 1. set up an environment in which consumers have *taste* for such tariffs or 2. create a model where

[1] Generally, we can call a tariff three-part if the marginal price of he first Q minutes is not zero, but in the context of telecommunications industries the zero marginal price case is more common.

[2] This approach has been taken by Sundararajan (2004).

firms find it optimal to *exploit* their customers' cognitive biases through these tariffs. The former approach is taken by Herweg and Mierendorff (2012) in the context of flat-rate tariffs, however, this idea has only partially been examined in the context of three-part tariffs. Nevertheless, in the subsequent discussion we examine the competitive extension of a model based on the latter approach, namely, exploiting *consumer overconfidence*.[3]

Seeing these general nonlinear pricing results one might ask the question, why we investigate them in the context of telecommunications industries.[4] There are two answers to that: 1. as will be seen both analytically and intuitively, all of the results that will be elaborated below hold only if (marginal) costs of providing one more unit of consumption is sufficiently low (even zero) which is a natural property of the telecommunications industries, and 2. the structure of decision making from the consumers' side fits intuitively the way how consumers make their subscription and calling decisions.

Note, that the existing results are tuned towards monopolistic nonlinear pricing (screening). One of our main aim is to examine the extension of these results to an oligopolistic setting. This way we will be able to investigate whether the established price structures remain the same and through what channels do competition affects the strategic variables.

The next section (3.1) covers methodological notions that will be necessary to investigate the completely nonlinear pricing problems in the sections to follow. The discussion of these methodological notions, however, at points exceeds the textbook-level treatment (as can be found e.g. in Fudenberg and Tirole (1991) or Börgers (2010)) in that we incorporate results from the most recent literature.

Section 3.2 presents the model of consumer overconfidence in a generalised setting and discusses results concerning three-part tariffs. Section 3.3 presents a possible extension of the consumer overconfidence model that is tuned towards telecommunications industries.

3.1 Methodological odds and ends

The theory of nonlinear pricing and screening partly belongs to the theory of *mechanism design*. The difference between the classical mechanism design problems and the nonlinear pricing, monopolistic screening is that in these latter there is one

[3]As a general, short survey of these kinds of models see Grubb (2012).

[4]We also have to give up from the complexity of the two-way access interconnection models. However, by examining a more simple structure we are able to generate more complex price schedules. In the end of this chapter, however, we take steps towards extending the models into a setting that resembles the structure of the telecommunications industry more.

principal (uninformed actor) and a single agent (informed actor), thus there is no strategic interaction between the agents. In our settings, the principal is going to be the seller and the single agent is going to be the buyer. It is without loss of generality to say that the single agent represents a *unit mass of consumers* each possessing a *type* that is known by the agent. The principal, however, only has prior probabilistic beliefs that are characterised by the distribution of the type..

3.1.1 Direct mechanisms and the revelation principle

The crucial problem in our settings is for the seller to induce the buyers to reveal their type correctly when making their purchase decision. This she does by implementing a *mechanism*. We are particularly interested in the so called *direct mechanisms*, since they are the contracts that we are looking for.

Let the type of the agent be a pair $(x,\theta) \in X \times \Theta$, and, for the ease of notation denote $\Xi \equiv X \times \Theta$. A contract is a pair $\{q(x,\theta), P(x,\theta)\}$ of quantity and payment schedule respectively for each type that is offered by a monopolist.

Definition 3.1. A *direct mechanism* is the pair of functions $q : X \times \Theta \to \mathbb{R}_+$ and $P : X \times \Theta \to \mathbb{R}$. Put differently, in a direct mechanism the type that reports (x,θ) is offered $q(x,\theta)$ quantities for $P(x,\theta)$ payment.

The *revelation principle* allows us to concentrate on direct mechanisms from the class of mechanisms in general. Recall that $\Xi := X \times \Theta$. Let $\Sigma \subseteq \Xi$. A strategy $\sigma : \Xi \to \Sigma$ is a function that gives a type $\xi' \in \Sigma$ for each input $\xi \in \Xi$.

Proposition 3.1 (Revelation principle (Börgers, 2010))**.** *For every mechanism M and every optimal buyer strategy σ in M there is a direct mechanism M^* and an optimal buyer strategy σ^* in M^* such that:*

1. *$\sigma^*(\xi) = \xi$ for all $\xi \in \Xi$ (thus, also $\Sigma = \Xi$),*
2. *For every $\xi \in \Xi$, the purchases $q(\xi)$ and payments $P(\xi)$ under M^* are equal to those under M if the buyer plays her optimal strategy σ.*

This means that if a strategy is generally optimal, then: 1. there is a mechanism that induces truthtelling and, 2. this mechanism is a direct mechanism.

3.1.2 Sequential screening

The general nonlinear pricing, contracting problem can be demonstrated by the *sequential screening* problem. Moreover, the intuitive idea behind this modeling setup suits our purposes well in studying telecommunications.[5]

A very plausible assumption regarding mobile service subscriptions is that at the time of signing a contract, we do not exactly know how much we are going to call during the following month, but have beliefs about this. In the meantime, we "learn" this private information from our preferences and make our calling quantity decision accordingly. This setup makes sense especially in the case of monthly contracts, where this "learning" and "buying" can simply be interpreted as the amount of calls made in that given monthly period and not in an instant.

This idea can be studied in a model of sequential screening as proposed for example by Courty and Hao (2000).[6] Since the original, "canonical" model displays a binary decision over consumption (or production) and the valuation of the commodity is linear in the type[7], it is necessary to generalise the formulation to serve our illustrational purposes in two ways: 1. by adding a type dependent utility function and 2. by allowing for a quantity decision.

Let a continuum of consumers have two types: 1. x, distributed according to $F(x)$ on $X := [\underline{x}, \bar{x}]$, and 2. $\theta \in [\underline{\theta}, \bar{\theta}] =: \Theta$ that is distributed according to the conditional distribution $G(\theta \mid x)$. Thus, each consumer is described by the pair (x, θ). Suppose, that x is known to the consumer at the time of contracting, whereas θ will be revealed only in the second period, after contracting. One simplifying additional assumption that Courty and Hao (2000) make is that $G(\theta \mid x)$ have the same support for all x.

A contract is a pair $\{q(x, \theta), P(x, \theta)\}$ of quantity and payment schedule respectively for each type that is offered by a monopolist. Supplying a unit q costs k and the subscription of one customer costs K. Consumers obtain a general type-dependent utility, $u(x, \theta, q(x, \theta))$ from consumption. The payoff functions (profit and utility) and the beliefs about types are common knowledge.

Given the above specifications, the optimal design problem is the following:

$$\max_{\{q(x,\theta), P(x,\theta)\}} \int_{\underline{x}}^{\bar{x}} \int_{\underline{\theta}}^{\bar{\theta}} [P(x,\theta) - kq(x,\theta) - K]\, dG(\theta \mid x)\, dF(x)$$

subject to

$$u(x, \theta, q(x, \theta)) - P(x, \theta) \geq u(x, \theta', q(x, \theta')) - P(x, \theta')\ \forall x, \forall \theta, \theta' , \tag{IC$_2$}$$

[5]Note, that the setup in this section is general and we do not intend to solve the model at this stage in any way. The concrete application will always be simpler, and will have more structure so that we can derive results.

[6]This inspection will also give hints on how to possibly formulate an oligopolistic model in which the firms are differentiated à la Hotelling.

[7]See also Krähmer and Strausz (2011).

$$\int_{\underline{\theta}}^{\bar{\theta}} [u(x,\theta,q(x,\theta)) - P(x,\theta)]\, dG(\theta \mid x) \geq \int_{\underline{\theta}}^{\bar{\theta}} [u(x,\theta,q(x',\theta)) - P(x',\theta)]\, dG(\theta \mid x')\ \forall x,x'\,, \tag{IC$_1$}$$

$$\int_{\underline{\theta}}^{\bar{\theta}} [u(x,\theta,q(x,\theta)) - P(x,\theta)]\, dG(\theta \mid x) \geq 0\ \forall x\,. \tag{IR}$$

The constraints can be interpreted as follows. IC$_2$ (incentive compatibility) requires the firm to set up a contract that forces type (x,θ) customer to report her type θ in period 2 correctly for every type x. IC$_1$ represents the same with respect to type x taking into account that customers have probabilistic beliefs about their type in the θ dimension. IR ensures that the customer is individually rational, i.e. she does not accept a contract in the first period that yields negative expected surplus.[8]

This is a very general program for contracts with general nonlinear tariffs ($P(x,\theta)$) that—imposing certain assumptions on the parameters and functions—might or might not yield the flat-rate or three-part tariffs that we are aiming at.

3.1.3 Local and global incentive compatibility

If we observe IC$_2$, we can see the quantor $\forall\theta,\theta'$, i.e. the constraint must hold for all combinations of types. This results in a great number of constraints that must hold, and we have to check whether all of them are satisfied simultaneously.

A simplifying technique that is very common in the literature is to observe the incentive compatibility property in the small vicinity of the true type θ and impose constraints in this vicinity. Then it is necessary to find assumptions under which the local incentive compatibility implies global IC. As is demonstrated by Carroll (2012), the sufficiency of local IC to imply global IC generally depends on: 1. the properties and shape of the type space, 2. the utility function. In the case of mechanisms without transfers very general results can be established regarding when local IC constraints are sufficient to imply global IC. In contrast, in the present case of mechanism with transfers (payments), Carroll (2012) (in its extended web-appendix) suggests that the utility function being linear in own type is an indispensable assumption (besides convex type spaces) for the implication local IC $\rightarrow$ global IC to hold

[8]Krähmer and Strausz (2011) also allow for an outside option in the second period, thus, a second IR constraint is necessary in that setting. Generally, this is a plausible extension, since e.g. in job contracts it is possible to terminate the contract during the second period, or it is possible to bring back a good to the shop and get full refund. In a telecommunications setting, however, sticking to only ex ante (i.e. contractual) outside options and only period one IR constraint is a reasonable simplification.

generally.[9]

In the light of this discussion we see that the implication "local IC → global IC" may not hold generally in our settings. Therefore, we have to check, for which types of problem formulations this implication actually holds. These are standard assumptions in the literature that can be found e.g. in Fudenberg and Tirole (1991) or Börgers (2010). We will investigate these properties in the actual problems.[10]

3.1.4 Competition in utility space

Observe, that we assumed above, that the firm (principal) is a monopolist. Thus, intuitively, the individual rationality constraint (participation constraint) binds, since the monopolist aims at extracting the highest surplus that is possible. What happens if we would like to incorporate an imperfect competitive environment (assuming symmetric firms)? The IR constraint will possibly not bind at 0, because the extractable surplus is bounded from below by the competitive environment. Thus, a way to proceed is to set the IR formula equal to a certain level of v, expected consumer surplus, and find the optimal contract for that particular v. Then, the optimal profit function $\pi(v)$ is going to depend on v. Further, as is suggested by Armstrong and Vickers (2001), in a second optimisation problem the firms A and B (now in the extended model) are going to play a game where they choose strategies $v^i \in \mathbb{R}_+$. Firm A is going to have payoff:

$$m(v^A, v^B)\pi(v^A)\,, \tag{3.1}$$

where $m(\cdot,\cdot)$ is the market share function that is increasing in its first and decreasing in its second argument, and $\pi(v^A)$ is the maximal profit as a function of v^A.[11] Thus, given the strategy of firm B, firm A maximises should maximise own market shares times profits and vice versa. The equilibrium will give optimal v^A and v^B, which we can substitute back into the profit function to find the optimal profits of each firm.

[9] However, the extent of this problem is unclear.

[10] Of course, the formulations should also ensure that IC holds locally. This is also a conclusion of the usual approaches.

[11] In this discussion a confusion might arise. Strictly speaking Armstrong and Vickers (2001), while suggesting this formulation do not propose a model. They only discuss about functions $m(\cdot,\cdot)$ and $\pi(\cdot)$ satisfying certain, general assumptions. Besides these assumptions, however, no additional structure is imposed on the problem e.g. by specifying cost structure. This way, they are able to illustrate the general idea of competition in utility space and derive generally applicable results.

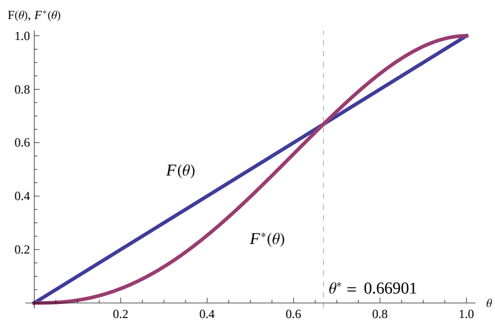
3.2 Competing in a model of consumer overconfidence

Grubb (2009) investigates the effects of consumer overconfidence to pricing in a monopolistic and perfectly competitive screening model. The timing of decisions in this model are analogous to the sequential screening model described above. The crucial assumption that drives his results is that consumers underpredict the actual variance of their future consumption, i.e. they are overconfident about predicting their future preferences and consumption.

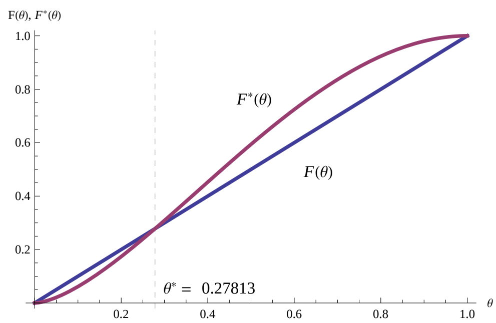
This assumption can be easily formalised. Let the taste parameter θ be distributed according to $F(\theta)$ with support $\Theta := [\underline{\theta}, \bar{\theta}]$. This F is the true distribution of θ. Denote consumers' (wrong) priors about this distribution with $F^*(\theta)$ with the same support. Then overconfidence can be modeled by assuming that

Assumption 3.1 (Grubb (2009))**.** $F^*(\theta)$ crosses $F(\theta)$ only once from below at θ^*.

This means that for every $\underline{\theta} < \theta < \theta^*$, $F^*(\theta) < F(\theta)$ and for every $\bar{\theta} > \theta > \theta^*$, $F^*(\theta) > F(\theta)$[12], i.e. more mass is concentrated around θ^* in the distribution $F^*(\theta)$. Figure 3.1 illustrates, with μ^* and μ denoting the mean of F^* and F respectively.[13] Note also, that if θ^* is the expectation of both F and F^*, then this assumption can be interpreted such that F is the mean preserving spread of F^*.



(a) $\mu^* > \mu$



(b) $\mu^* < \mu$

Figure 3.1: $F(\theta)$ and $F^*(\theta)$ satisfying Assumption 3.1

In the following we discuss Grubb's model in a generalised setting by extending it to a Hotelling oligopolistic environment. Our main result is that this extension does not alter the results in Grubb (2009) about the optimal contracts in a crucial way, but the comparative statics with respect to τ, the differentiation parameter can be investigated more thoroughly.

[12]Since the two distributions $F(\theta)$ and $F^*(\theta)$ have common support, at $\underline{\theta}$ and $\bar{\theta}$ the functional values must be the same too.

[13]In both plots F is the uniform distribution on $[0,1]$. F^* is the beta distribution with parameters (a) $\alpha = 2.5$, $\beta = 2$, (b) $\alpha = 1.6$, $\beta = 2$. Plot (b) illustrates that Assumption 3.1 is not equivalent to F^* second order stochastically dominating (SOSD) F, only if $\mu^* \geq \mu$, as in plot (a).

There are two simplifying assumptions that make the analysis more transparent by allowing us to rely on Armstrong and Vickers (2001). These are: 1. the Hotelling location parameter x is stochastically independent from the consumption type parameter θ, and 2. at the stage of contracting consumers are homogeneous and decide according to their expectations of their future demand.[14]

3.2.1 Model setup

Firms

Suppose that there are two firms offering contracts specified by the pair $\{q(\theta), P(\theta)\}$ of quantities and payment schedules for each consumer type θ. Firms have the correct priors about the distribution $F(\theta)$. The firms are symmetric and the profits are given by $\pi_i = P_i(\theta) - C(q(\theta))$, where C is a cost function that is assumed to be increasing and convex in q. The two firms are spatially differentiated and are located at the two ends of the interval $[0,1]$.

Note that this firm specification implicitly assumes that that the location parameter x and the consumption type parameter θ are independent, since the contract offered does not depend on the location type, i.e. it is the same for all types x.

Consumers

Consumers are heterogeneous with respect to a pair of types (x,θ). The two types are stochastically independent. Consumers' priors on θ are represented by $F^*(\theta)$. The true distribution $F(\theta)$ is related to $F^*(\theta)$ according to Assumption 3.1. Type x is distributed uniformly on $[0,1]$. Consumers know their location x at the time of contracting, but the realisation of type θ will only be revealed to the consumers after the contracting stage.

Preferences are represented by the type dependent utility function $u(q,\theta)$. For each type there is a satiation point $q^s(\theta)$ above which quantities are *freely disposable*. The utility function satisfies the following assumptions:

Assumption 3.2.

1. For all θ the satiation point $q^s(\theta)$ is finite, and given by $\min\{q : u_q(q,\theta) = 0\}$ for all θ. This satiation point is increasing in θ.

[14] The framework of Rochet and Stole (2002) is of the same spirit, however they model firms as selling qualities, rather than quantities. The mechanics of an oligopolistic nonlinear pricing model can be well understood from Stole (1995). Stole's modeling approach also allows for interacting (not independent) types.

2. $u_q(q,\theta) > 0$ for $q < q^s(\theta)$; and $u_q(q,\theta) = 0$ for all $q \geq q^s(\theta)$, $u_{qq}(q,\theta) < 0$; $u_{q\theta} > 0$ for all θ.

3. $u(0,\theta) = 0$ for all θ.

4. $u(q,\theta)$ is three times continuously differentiable with bounded derivatives.

In the next step we can verify the following claim that is going to be important in the subsequent analysis.

Claim 3.2. *Given Assumption 3.2, $u(q,\theta)$ has weakly increasing differences, i.e. for all $\theta' > \theta$ and $q' > q$*

$$u(q',\theta') - u(q',\theta) \geq u(q,\theta') - u(q,\theta).$$

The inequality is strict if and only if $q < q^s(\theta')$.

Proof. Suppose Assumption 3.2 holds. Take any $q' > q$ and $\theta' > \theta$, and show that

1. If $q \geq q^s(\theta')$, then the inequality holds with equality.

   Since for any θ, $u_q(q,\theta)$ is zero for all $q \geq q^s(\theta)$, then $u(q',\theta') = u(q,\theta')$. Since $q^s(\theta)$ is increasing in θ, $q^s(\theta') \geq q^s(\theta)$, and, in particular $q' > q \geq q^s(\theta)$. Thus, also $u(q',\theta) = u(q,\theta)$. Hence, $u(q',\theta') - u(q',\theta) = u(q,\theta') - u(q,\theta) = 0$.

2. If $q < q^s(\theta')$, then the inequality is strict.

   First of all, let $q < q^s(\theta)$ (θ without a prime). Since $u_{q\theta}(q,\theta) > 0$ for all θ, and $u_q(q,\theta) > 0$ for all q in this region, thus $u_q(q,\theta') - u_q(q,\theta) > 0$. If $q' < q^s(\theta)$, then this implies that $u(q',\theta') - u(q,\theta') > u(q',\theta) - u(q,\theta)$, what we wished for. If $q^s(\theta) \leq q' < q^s(\theta')$, then $u_q(q',\theta') > 0$ and $u_q(q',\theta) = 0$. In particular then $u_q(q',\theta') > u_q(q',\theta)$. Thus, for all $q < q'$, $u(q',\theta') - u(q,\theta') > u(q',\theta) - u(q,\theta)$. If $q' \geq q^s(\theta')$, then $u(q',\theta') > u(q',\theta)$ (from $u_{q\theta} > 0$), and the functions are maximised, thus for all $q < q^s(\theta')$, $u(q',\theta') > u(q,\theta')$, and $u(q',\theta) \geq u(q,\theta)$. These two inequalities together give a strict inequality that we wished for.

   If $q^s(\theta) \leq q < q^s(\theta')$, then $u(q,\theta) = 0$. But, since $u_q(q,\theta')$ is still positive, $u_q(q,\theta') - u_q(q,\theta) > 0$, and $u(q',\theta') - u(q,\theta') \geq 0$, with equality only if $q \geq q^s(\theta')$, which condition is a contradiction. Thus the inequality is strict.

3. $q < q^s(\theta')$ is necessary for the strict inequality. Suppose the contrary, i.e. $q \geq q^s(\theta')$. This is by the first point a contradiction.

□

Since for each consumer type there is a unique satiation point, beyond which quantities are freely disposable, every type θ reporting type θ' is going to consume $\min[q(\theta'), q^s(\theta)]$. The variable surplus of type θ reporting θ' is denoted by $V(\theta,\theta') \equiv u(\min[q(\theta'), q^s(\theta)], \theta) - P(\theta')$. The surplus from joining provider $i \in \{0,1\}$ is generally given by

$$W_i(\theta,\theta',x) \equiv V(\theta,\theta') - \tau|i-x|\,. \tag{3.2}$$

Note, that in the contracting first period consumers only have expectations about their future surplus. Thus, there is no place for not telling the truth about the location type x. The contracts will be incentive compatible with respect to x if they are with respect to θ. Hence, the above expression simplifies to $W_i(\theta,x)$, or $\mathbb{E}^*[W_i(\theta,x)]$.

Timing

1. Firms simultaneously offer contracts to consumers.
2. Consumers choose firm subscription based on their prior beliefs by accepting a particular firm's contract.
3. The real types are revealed, consumers make consumption (calling) decision accordingly and pay an amount that the contract prescribes.

3.2.2 Solving for the optimal contract $\{q(\theta), P(\theta)\}$

In order to solve this problem in a transparent way, we use the approach of Armstrong and Vickers (2001) and model the two firms as competing in the utility space. This we do by solving the problem in two steps: 1. Finding the optimal tariff for a given level $\bar{V}$ of utility supplied to consumers and 2. Choosing the optimal utility supplied to consumers given the other firm's analogous decision by playing a game. Besides this, the arguments closely follow Grubb (2009).[15]

Since both firms are symmetric, we can formulate the optimal contract problem without a firm index. This problem is the following:

$$\max_{\{q(\theta),P(\theta)\}} \mathbb{E}[P(\theta) - C(q(\theta))]$$

subject to

$$V(\theta,\theta) \geq V(\theta,\theta')\ \forall \theta,\theta' \in \Theta \tag{IC}$$

$$\mathbb{E}^*[V(\theta,\theta)] = \bar{V} \tag{IR}$$

[15] All statements phrased as claims or propositions so far and henceforth are stated and proven by me, except where the contrary is clearly indicated.

$$\bar{V} \geq 0;\ q(\theta) \geq 0\, \forall \theta$$

This program specifies the incentive compatible optimal contract for a given level of expected consumer surplus. At this point, we differentiate between quantities offered by the firm $q(\theta)$ and the quantities actually consumed $q^c(\theta,\theta') \equiv \min[q(\theta'), q^s(\theta)]$ denoting $q^c(\theta,\theta) = q^c(\theta)$.

The derivative of V is:

$$\frac{\partial V(\theta,\theta')}{\partial \theta} = u_q(q^c(\theta,\theta'),\theta) \cdot \frac{\partial q^c(\theta,\theta')}{\partial \theta} + u_\theta(q^c(\theta,\theta'),\theta).$$

Recall, that q^c is a minimum function, therefore we have to split for the two cases $q(\theta') < q^s(\theta)$ and $q(\theta') \geq q^s(\theta)$. In the former case $q^c(\theta,\theta') = q(\theta')$, thus the derivative according to θ is zero. In the latter case, $q^c(\theta,\theta') = q^s(\theta)$, in which case $dq^s(\theta)/d\theta$ is zero by assumption. Thus

$$\frac{\partial V(\theta,\theta')}{\partial \theta} = u_\theta(q^c(\theta,\theta'),\theta)\ a.e.,$$

where a.e. denotes "almost everywhere", i.e. except for a set whose (Lebesgue) measure is zero. In our case, this is the point where the minimum function is not differentiable.

As Grubb (2009, Lemma 1., p. 1778.) shows, if the contract $\{\hat{q}(\theta), \hat{P}(\theta)\}$ is optimal, then so is $\{\min[\hat{q}(\theta), q^s(\theta)], \hat{P}(\theta)\}$, and that if costs are strictly increasing then $\hat{q}(\theta) \leq q^s(\theta)$ (a.e.). Hence, we can invoke the strict increasing property from claim 3.2, combine it with the property that $q^c(\theta)$ is nondecreasing and refer to Fudenberg and Tirole (1991, Theorem 7.3, p. 261.) to conclude that the optimal contract is globally incentive compatible. If the optimal contract is globally incentive compatible, then we can refer to Milgrom and Segal (2002, Theorem 2, p. 586), to conclude that the following envelope formula holds

$$V(\theta) - V(\underline{\theta}) = \int_{\underline{\theta}}^{\theta} V_\theta(s)\, ds$$

$$V(\theta) = V(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} u_\theta(q^c(s), s)\, ds\,. \tag{3.3}$$

From this this equation, we can express $\mathbb{E}[V(\theta)]$ and $\mathbb{E}^*[V(\theta)]$ as

$$\mathbb{E}^\circ[V(\theta)] = V(\underline{\theta}) + \mathbb{E}\left[u_\theta(q^c(\theta),\theta)\frac{1-F^\circ(\theta)}{f(\theta)}\right], \tag{3.4}$$

where $\circ$ either stands for $*$ or nothing.

Next, we have to find $P(\theta)$ by expressing it as $P(\theta) = V(\theta) - u(q^c(\theta),\theta)$. From

Equation 3.3, we get that

$$P(\theta) = V(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} u_\theta(q^c(s), s)\, ds - u(q^c(\theta), \theta)\,. \tag{3.5}$$

Up until now, we have only used the IC constraint for the derivation. Let us turn to the IR constraint and combine it with 3.4. This way we can express $V(\underline{\theta})$ and substitute, thus,

$$\hat{P}(\theta, \bar{V}) = \underbrace{\bar{V} - \mathbb{E}\left[u_\theta(q^c(\theta), \theta)\frac{1 - F^*(\theta)}{f(\theta)}\right]}_{V(\underline{\theta})} + \int_{\underline{\theta}}^{\theta} u_\theta(q^c(s), s)\, ds - u(q^c(\theta), \theta)\,, \tag{3.6}$$

which is essentially the same as the result of Grubb (2009), except for the constant term $\bar{V}$. In the next stage of optimisation, we have to pin down $\bar{V}$.

Note already, however, one interesting aspect of this pricing schedule: the marginal price will be independent of the particular utility level $\bar{V}$ that the firm offers. This means, that the marginal price cannot be the channel through which the two firms compete. Indeed, $\bar{V}$ is part of the fixed fee, hence this fixed fee will be one of the channels of competition.

What about the quantity offered for a particular type θ consumer? Notice, that in the derivation of the above pricing schedule we used both the IC and IR constraints. Thus $P(\theta)$, which depends on the truthtelling $q^c(\theta)$ identifies the optimal (incentive compatible and individually rational) contract price for any $q(\theta)$ and $\bar{V}$. What remains then is to maximise profits again along the $q(\theta)$ dimension subject to only two constraints: 1. $0 \leq q(\theta) \leq q^s(\theta)$ and 2. $q(\theta)$ nondecreasing. This way, we can find the optimal quantity schedule $\hat{q}(\theta)$.[16]

As we have already noted, the second constraint is a necessary condition for a contract in the q dimension to be incentive compatible (Börgers, 2010; Fudenberg and Tirole, 1991). Whenever this constraint is binding, an "ironing" procedure (cf. Fudenberg and Tirole (1991, Chapter 7., Appendix.)) is applied which results in that the optimal quantity offered to the types where the constraint is binding will be the same (pooling).

Note, that this maximisation problem immediately implies that $\bar{V}$ (being a constant) does not enter the quantity decision either, so the fixed fee is the only channel through which firms actually compete.

[16]For the purposes of the present discussion, the actual characterisation and properties of this optimal quantity schedule is redundant. For the details please refer to Grubb (2009, Section D.).

Now we turn to the second phase: a noncooperative game for consumers. As we have already seen in the Hotelling specification, in the case of full consumer participation the expected market share function for firm i will be

$$m(\bar{V}^i, \bar{V}^{1-i}) \equiv \begin{cases} 0, & \text{if } (\bar{V}^i - \bar{V}^{1-i}) \leq -\tau \\ \frac{1}{2} + \frac{\bar{V}^i - \bar{V}^{1-i}}{2\tau}, & \text{if } (\bar{V}^i - \bar{V}^{1-i}) \in (-\tau, \tau) \\ 1, & \text{if } (\bar{V}^i - \bar{V}^{1-i}) \geq \tau \end{cases} . \tag{3.7}$$

Note, that because of the constraint is related to $\bar{V}$, this market share function is based on the expectation of the *consumers'* priors. This is clearly the case, since consumers subscribe to either one of the providers based on their own beliefs.

Recall, that $\hat{q}(\theta)$ is the optimal quantity schedule offered in the contract. Bearing this in mind, define

$$\hat{\Pi}(\theta, \bar{V}^i) \equiv \hat{P}(\theta, \bar{V}^i) - C(\hat{q}(\theta)). \tag{3.8}$$

Further, for the ease of notation, define

$$\hat{\Omega}(\theta) \equiv \mathbb{E}\left[u_\theta(\hat{q}(\theta), \theta)\frac{1 - F^*(\theta)}{f(\theta)}\right] - \int_{\underline{\theta}}^{\theta} u_\theta(\hat{q}(s), s)\, ds + u(\hat{q}(\theta), \theta),$$

i.e. the optimal price schedule without $\bar{V}$. Given the two symmetric firms, the optimal contracts, the already calculated profit functions and the market share function, we can make the following claim

Proposition 3.3. *Define a two player normal form game as: firm* $i \in \{0, 1\}$ *can choose actions* $\bar{V}^i \in \mathbb{R}_+$ *and has the payoff function*

$$m(\bar{V}^i, \bar{V}^{1-i}) \cdot \mathbb{E}\left[\hat{\Pi}(\theta, \bar{V}^i)\right], \tag{3.9}$$

where $m(\cdot, \cdot)$ *and* $\hat{\Pi}(\cdot, \cdot)$ *are given by Equations 3.7 and 3.8 respectively.*

This game has a symmetric pure strategy Nash-equilibrium $\hat{\bar{V}}^0 = \hat{\bar{V}}^1 = \hat{\bar{V}}$ *that is given by*

$$\hat{\bar{V}} = \mathbb{E}\left[\hat{\Omega}(\theta) - C(\hat{q}(\theta))\right] - \tau, \tag{3.10}$$

if $\mathbb{E}\left[\hat{\Omega}(\theta) - C(\hat{q}(\theta))\right] - \tau \geq 0$, *or by* $\hat{\bar{V}} = 0$ *else.*

Proof. The optimal contracts, thus the variable profits are the same for both firms, therefore we can drop the subscripts for $\hat{\Omega}$ and C.

1. $\mathbb{E}\left[\hat{\Omega}(\theta) - C(\hat{q}(\theta))\right] - \tau \geq 0.$

   If both firms offer $\hat{\bar{V}} = \mathbb{E}\left[\hat{\Omega}(\theta) - C(\hat{q}(\theta))\right] - \tau$, then the equilibrium payoff for firm i is $\frac{1}{2}\tau$, since $m(\hat{\hat{V}}, \hat{\hat{V}}) = 1/2$, and $\mathbb{E}\left[\hat{\Pi}(\theta, \hat{\hat{V}})\right] = \mathbb{E}\left[\hat{\Omega}(\theta) - C(\hat{q}(\theta))\right] - \hat{\hat{V}}$. Now suppose that that firm $1-i$ offers $\hat{\hat{V}}$, but firm i deviates by offering $\bar{V}'$.

First, consider the deviation, where $D \equiv (\bar{V}' - \hat{\hat{V}}) \in (-\tau, \tau)$. Depending on the sign of the deviation, the increase (or decrease) in market share $m(\bar{V}', \hat{\hat{V}})$ for firm i is given by $D/2\tau$. The change in expected profit is simply given by $-D$. Thus, the payoff change is $\frac{-(D)^2}{2\tau}$, which (as $\tau > 0$) is negative regardless of the sign of the deviation.

Second, consider a deviation where $D \equiv (\bar{V}' - \hat{\hat{V}}) \geq \tau$. Then the market share is going to be 1, and the payoff will be

$$\tau - \bar{V}' \leq \bar{V}' - \hat{\hat{V}} - \bar{V}' = -\hat{\hat{V}} \leq 0,$$

since $\hat{\hat{V}} \geq 0$. This is smaller than $1/2\tau$ for all τ.

Last, consider a deviation such that $D \equiv (\bar{V}' - \hat{\hat{V}}) \leq -\tau$. Then the payoff is going to be zero, because of the zero market shares. Zero is smaller than the equilibrium payoff.

2. $\mathbb{E}\left[\hat{\Omega}(\theta) - C(\hat{q}(\theta))\right] - \tau < 0$, and $\hat{\hat{V}} = 0$.

   The equilibrium payoff is $\frac{1}{2}\mathbb{E}\left[\hat{\Omega}(\theta) - C(\hat{q}(\theta))\right]$.

   Offering a strictly negative deviation is not a feasible action. If firm i offers a strictly positive deviation $\bar{V}' < \tau$, then the change in the payoff for firm i is, again, going to be $\frac{-(\bar{V}')^2}{2\tau}$, which is negative. If $\bar{V}' \geq \tau$, then the payoff is going to be

$$\mathbb{E}\left[\hat{\Omega}(\theta) - C(\hat{q}(\theta))\right] - \bar{V}' \leq \mathbb{E}\left[\hat{\Omega}(\theta) - C(\hat{q}(\theta))\right] - \tau < 0.$$

Reversing the role of the firms and repeating the argument concludes the proof.

□

Plugging back the equilibrium $\hat{\hat{V}}$ into the price schedule $\hat{P}$ yields:

$$\begin{aligned}\hat{P}^H(\theta, \tau) = u(\hat{q}(\theta), \theta) - \int_{\underline{\theta}}^{\theta} u_\theta(\hat{q}(s), s)\, ds + \mathbb{E}\left[u_\theta(\hat{q}(\theta), \theta)\frac{1 - F^*(\theta)}{f(\theta)}\right] \\ - \mathbb{E}\left[u_\theta(\hat{q}(\theta), \theta)\frac{1 - F^*(\theta)}{f(\theta)}\right] + \mathbb{E}\left[u_\theta(\hat{q}(\theta), \theta)\frac{1 - F(\theta)}{f(\theta)}\right] \\ - \mathbb{E}\left[\underbrace{u(\hat{q}(\theta), \theta) - C(\hat{q}(\theta))}_{S(\hat{q}(\theta))}\right] + \tau, \quad (3.11)\end{aligned}$$

where $S(\hat{q}(\theta))$ is the total surplus.

Note the following two observations:

1. the mere existence of this oligopolistic environment cancels the extra surplus $\mathbb{E}\left[u_\theta(\hat{q}(\theta),\theta)\frac{1-F^*(\theta)}{f(\theta)}\right]$ that is based on the *wrong priors*, and replaces it with the extra surplus $\mathbb{E}\left[u_\theta(\hat{q}(\theta),\theta)\frac{1-F(\theta)}{f(\theta)}\right]$ that is based on the *correct priors*,
2. τ, the transport cost enters into the formula only additively, without any multipliers.

The following immediate corollary provides a direct connection between our results and the results of Grubb (2009).

Corollary 3.4.

1. *If τ is large enough, then $\hat{\hat{V}} = 0$ and*

$$\hat{P}^H(\theta,\tau) = u(q^c(\theta),\theta) - \int_{\underline{\theta}}^{\theta} u_\theta(q^c(s),s)\,ds + \mathbb{E}\left[u_\theta(q^c(\theta),\theta)\frac{1-F^*(\theta)}{f(\theta)}\right]. \tag{3.12}$$

2. *If $\tau \to 0$, then*

$$\hat{P}^H(\theta,\tau) \to u(\hat{q}(\theta),\theta) - \int_{\underline{\theta}}^{\theta} u_\theta(\hat{q}(s),s)\,ds$$
$$+\mathbb{E}\left[u_\theta(\hat{q}(\theta),\theta)\frac{1-F(\theta)}{f(\theta)}\right] - \mathbb{E}\left[\underbrace{u(\hat{q}(\theta),\theta) - C(\hat{q}(\theta))}_{S(\hat{q}(\theta))}\right]. \tag{3.13}$$

These are the exact pricing formulas that Grubb (2009) derives in Proposition 1. (p. 1779.) for the monopolist firm and the perfectly competitive firm respectively. This is not surprising: if the transport cost is very high, each firm acts like a local monopolist, and if the transport cost is near zero, then a Bertrand-type intuition suggests that the outcome will be the competitive one in the limit.

This is exactly the case here, and the Hotelling-specification provides a smooth transition between the two extremes with the transport cost parameter that enters additively into the pricing schedule. In particular, if we denote the right-hand-side of Equation 3.13 as $\hat{P}^C(\theta)$, then, if τ is not too high, we have the following relationship:

$$\hat{P}^H(\theta,\tau) = \hat{P}^C(\theta) + \tau. \tag{3.14}$$

If we compare Equation 3.12 (denoting it $\hat{P}^M(\theta)$) with $\hat{P}^C(\theta)$, then we can see that the main difference between a monopolist, and the (perfect) competitive industry is a difference in the fixed fee. If we consider the Hotelling setting, then we can see that the "comparative statics" with respect to τ are very simple: the smaller the transportation cost, the lower the fixed part of the price is.

The difference between the monopolist pricing in this overconfident setup and the usual monopolistic screening result (as is described for example in Fudenberg and Tirole (1991)) is the term with positive sign $\mathbb{E}\left[u_\theta(\hat{q}(\theta),\theta)\frac{1-F^*(\theta)}{f(\theta)}\right]$. I.e. the monopolist sets up a generally *higher* price schedule, and this extra revenue is *fixed for all types*, since it is an expectation.

As noted earlier, the effect of the oligopolistic environment is that its existence sweeps away the extra revenue *based on customers misperception of the expectation* and replaces it with an expected virtual surplus based on *the real distribution*. In addition, the difference (with negative sign towards the competitive outcome) is going to be a fixed fee that is exactly the expected total surplus.

If the transport costs are close to zero, and the firms are playing a simple Bertrand-type game, the oligopolistic price schedule will be precisely *the same as the competitive pricing.*

After highlighting these aspects of the optimal contract result of Grubb (2009) and our results in a Hotelling duopolistic setup, we can conclude this section with the following summary:

Result. *The quantity schedule specified by the optimal contract of the above program is the same for the three specifications: Monopoly (M), Perfect competition (PC), Hotelling duopoly (H). The optimal price schedule differs across the three specifications by only a fixed fee. In the case of the H, this fixed fee is precisely τ in addition to the price schedule of PC. Thus, in the case of $\tau = 0$, the outcome of H and PC is precisely the same. If, however, τ is sufficiently large, then the provided utility in a symmetric equilibrium will be zero and the result of H will be the same as M.*

3.2.3 Further pricing results

The importance of demonstrating (showing) that the results from M, PC and H differ only by a fixed fee in the price schedule is that this leaves the results of Grubb (2009) regarding the marginal price identically valid. Thus, also in the case of H, we do not lose the structure of three part tariffs that is discovered by Grubb for the model of overconfidence.

If we link the optimal price schedule to the quantity consumed by denoting $\hat{P}(q) = \hat{P}(\hat{\theta}(q))$, with $\hat{\theta}(q) \equiv \inf\{\theta : \hat{q}(\theta) = q\}$, then for those quantities where we do not have to care about the monotonicity constraint, the following marginal pricing result is true (Grubb, 2009, Proposition 2., p.1781):

$$\frac{d\hat{P}(q)}{q} = \max\left\{ C_q(q) + u_{q\theta}(q, \hat{\theta}(q)) \frac{F^*(\hat{\theta}(q)) - F(\hat{\theta}(q))}{f(\hat{\theta}(q))}, 0 \right\}. \tag{3.15}$$

From this result we can immediately see that if the real and perceived priors of the customers were the same ($F(\theta) = F^*(\theta)$ for all θ), then the optimal marginal price would be marginal cost pricing for all θ-s—a result already familiar from the case of two-part tariffs.

Further, recall that $F(\theta)$ and $F^*(\theta)$ coincide in each elements of the set $\{\underline{\theta}, \theta^*, \bar{\theta}\}$. For $\theta < \theta^*$, $F^*(\theta) < F(\theta)$, and for $\theta > \theta^*$ the relationship is reverse. Thus, if we denote $\underline{q} \equiv \hat{q}(\underline{\theta})$, $q^* \equiv \hat{q}(\theta^*)$, $\bar{q} \equiv \hat{q}(\bar{\theta})$, then we can conclude[17] that whenever the marginal cost is zero, and the monotonicity constraint does not bind in the derivation of optimal quantity schedule,

1. Marginal price is zero for all quantities below q^* and for $\bar{q}$,
2. Marginal price is positive for all quantities in the complement of the above set, i.e. strictly between q^* and $\bar{q}$.

In fact, for quantities above q^* the marginal price is generally *higher than marginal cost* $C_q(q)$, which in the case of the above result is assumed to be zero. If the marginal cost is strictly positive for all quantities (without the monotonicity constraint taken into account), then for quantities below q^*, marginal price is lower than marginal cost, for quantities above q^*, marginal price is higher than marginal cost and for the points of coincidence $(\underline{q}, q^*, \bar{q})$ marginal price equals marginal cost.

Recall, that this result is independent of the market structure that we are investigating, as payments differ only by a fixed fee between market structures.

Since in the case of telecommunications, marginal costs can safely be assumed to be close to zero, we have found the payment schedule structure that we have aimed at: a three-part tariff. Thus we can phrase the following summarising result.

Result. *With a model of overconfidence in the spirit of Grubb (2009), assuming zero marginal costs we can generate a pricing structure that is qualitatively similar to a three-part tariff. In this structure we can identify and show the following properties concerning the three-part tariff:*

1. *The fixed fee is determined by the competitiveness of the market structure in which we are investigating. The more competitive the industry, the lower the fixed fee is.*

[17] This is also stated in Grubb in Corollary 1 (p.1781).

2. *The quantity allowance is given by the quantity consumed by a particular intermediate demand type θ^*, whose anticipated and correct priors (F^* and F resp.) coincide.*

3. *The marginal price is independent of the market structure and is higher than marginal cost after the quantity allowance is used up. The steepness of this above-allowance marginal price is bigger if the difference between the anticipated and correct priors is bigger, or if the probability density (given by the pdf f) is smaller.*

3.3 Towards telecommunications – A possible extension

In the light of the discussion of Chapter 2, assuming that the costs of making inbound calls is the same as that of making outbound calls might be with loss of generality if we investigate the telecommunications industry. The negotiated termination charge makes off-net (outbound) calls more costly for both firms.

If we observe todays' tariffs in Austria for example (cf. Section 1.1), then it is usual that there is no termination based price discrimination in the three part tariff contract offered. Thus, building on the above extension of Grubb (2009) a different model can be formulated that resembles the telecommunications industry more. In this model costs differ based on the network where a call terminates with $c_{\text{on}} < c_{\text{off}}$ for both firms, but the payment schedule $P(\theta)$ is uniform for the two network terminations. As we will see, this extension makes the model more complicated than the simple oligopolistic extension of Grubb (2009).

3.3.1 Model setup

In this model, consumers are the same as in the model of section 3.2, i.e. have the same utility function as described by assumption 3.2, and their incorrect priors are, again, described by assumption 3.1, the satiation assumption applies. Motivated by the models of Chapter 2, we make the simplifying assumption of balanced calling pattern in this section. This means, that consumers do not intentionally choose the network to which they call, rather, their call termination depends on the relative market shares of the two firms.

The timing of the game is what is described in section 3.2.

Firms offer type dependent contracts $\{q(\theta), P(\theta)$ and are located at the two extremes of the Hotelling line. This section's extension is, that firm $i \in \{0, 1\}$ has variable cost c_{on} for calls terminating in i's network, and c_{off} for calls terminating in the rival's network.[18] The relationship is $c_{\text{on}} < c_{\text{off}}$. There is a fixed cost K associated with each subscribing consumer. Writing $m_i \equiv m(\bar{V}^i, \bar{V}^{1-i})$ analogously to Equation 3.7, and assuming (as natural in the Hotelling specification) that $m_{1-i} = 1 - m_i$, firm i's profit is from selling contract $\{q(\theta), P(\theta)\}$ to consumer of type θ, given opponent's $\bar{V}^{1-i}$ and own $\bar{V}^i$:

$$\Pi_i(\theta, \bar{V}^i, \bar{V}^{1-i}) \equiv P(\theta) - m_i c_{\text{on}} q(\theta) - (1 - m_i) c_{\text{off}} q(\theta) - K. \tag{3.16}$$

As costs did not appear in the derivation of the optimal pricing schedule given $q^c(\theta)$, θ and $\bar{V}^i$, in this setting the optimal (monopolist) price schedule will, again, be given by Equation 3.6 which we rewrite here:

$$\hat{P}(\theta, \bar{V}) = u(q^c(\theta), \theta) - \int_{\underline{\theta}}^{\theta} u_\theta(q^c(s), s)\, ds + \mathbb{E}\left[u_\theta(q^c(\theta), \theta)\frac{1 - F^*(\theta)}{f(\theta)}\right] - \bar{V}^i. \tag{3.17}$$

Thus, since the structure is the same, we could hope, that the game in utility levels will be as simple, as in the previous section.

However, this is not the case here, because $\hat{q}(\theta)$, the optimal quantity offered will depend on the market share function, thus on both $\bar{V}^i$ and $\bar{V}^{1-i}$. As a consequence, the payment function $P(\theta)$ in the optimal contract will also depend on $\bar{V}^i$ and $\bar{V}^{1-i}$ through $\hat{q}(\theta, \bar{V}^i, \bar{V}^{1-i})$.

To characterise $\hat{q}$, we have to maximise the expected profits given the optimal price schedule $\hat{P}(\theta)$, imposing the monotonicity and the satiation constraint. Ignoring for the moment these constraints, the maximizand objective for firm $i \in \{0, 1\}$ therefore is:

$$\mathbb{E}\left[u(q, \theta) - \int_{\underline{\theta}}^{\theta} u_\theta(q, s)\, ds + \mathbb{E}\left[u_\theta(q, \theta)\frac{1 - F^*(\theta)}{f(\theta)}\right] \right.$$
$$\left. + [m_i(c_{\text{off}} - c_{\text{on}}) - c_{\text{off}}]q - \bar{V}^i - K\right]. \tag{3.18}$$

Which, after integrating by parts and using that the expectation of the expectation (w.r.t. the same measure) is simply the expectation, simplifies to

$$\mathbb{E}\left[u(q, \theta) + u_\theta(q, \theta)\frac{F(\theta) - F^*(\theta)}{f(\theta)} + [m_i(c_{\text{off}} - c_{\text{on}}) - c_{\text{off}}]q\right] - \bar{V}^i - K. \tag{3.19}$$

If we maximise this expression with respect to q for all θ without the expectation,

[18] Note that incoming calls do not generate costs. This is a simplification, because if it was not so, the opponent's optimal contract (or, rather the quantity offered therein) would also appear in the profit function, as $q(p_j)$ does in equation 2.7.

then the expected value is also going to be maximised.

Denote

$$\Psi(q,\theta,\bar{V}^i,\bar{V}^{1-i}) \equiv u(q,\theta)+u_\theta(q,\theta)\frac{F(\theta)-F^*(\theta)}{f(\theta)}+[m_i(c_{\text{off}}-c_{\text{on}})-c_{\text{off}}]q-\bar{V}^i-K\,,$$

and make the following technical assumption:

Assumption 3.3.

$$u_{\theta qq}=\begin{cases} \geq 0, & \text{if } F(\theta)-F^*(\theta)<0 \\ \leq 0, & \text{if } F(\theta)-F^*(\theta)>0 \end{cases}. \tag{3.20}$$

Recall, further, that $u_{qq}<0$ for $q\in[0,q^s(\theta)]$ for all θ. These assumptions, together with the recognition that if $F(\theta)=F^*(\theta)$, then $u_{\theta qq}(q,\theta)(F(\theta)-F^*(\theta))/f(\theta)=0$, make Ψ strictly concave ($\Psi_{qq}<0$). Hence, maximising Ψ with respect to q is a concave optimisation problem on $[0,q^s(\theta)]$, for all θ, where for each θ the optimum (in the interior of $[0,q^s(\theta)]$) is characterised by the FOC

$$u_q(q,\theta)+u_{\theta q}(q,\theta)\frac{F(\theta)-F^*(\theta)}{f(\theta)}=c_{\text{off}}-m_i(c_{\text{off}}-c_{\text{on}}). \tag{3.21}$$

Observe, that $m_i\equiv m(\bar{V}^i,\bar{V}^{1-i})$ enters this equation, thus, given $\bar{V}^{1-i}$, by changing $\bar{V}^i$ also the optimal quantity $\hat{q}(\theta)$ is going to change in contrast to the problem in the previous section, where $\bar{V}^i$ entered the price schedule only additively, and had no bearing on the optimal contract besides that.

While from Equation 3.21 in its generality we cannot say much about the explicit dependence of $\hat{q}$ on $\bar{V}^i$, we can invoke the implicit function theorem to make the following statement:

Proposition 3.5. *Under Assumptions 3.1, 3.2 and 3.3, if the monotonicity, satiation and non-negativity constraints are* not *binding, then $\hat{q}(\theta,\bar{V}^i,\bar{V}^{1-i})$ is increasing in $\bar{V}^i$.*

Proof. If neither of the three constraints (monotonicity, non-negativity, satiation) are binding, then $\hat{q}(\theta,\bar{V}^i,\bar{V}^{1-i})$ is characterised by Equation 3.21. This FOC is given by the Equation $\Psi_q(q,\theta,\bar{V}^i,\bar{V}^{1-i})=0$, and $\hat{q}\in(0,q^s(\theta))$ open interval for all θ.

Since $\Psi_{qq}<0$ on $(0,q^s(\theta))$, the implicit function theorem implies that

$$\frac{d}{d\bar{V}^i}\hat{q}(\cdot,\bar{V}^i)=\frac{\Psi_{q\bar{V}^i}}{\Psi_{qq}}\,,$$

and thus the sign of $\Psi_{q\bar{V}^i}$ tells how $\hat{q}$ changes in response to a bigger $\bar{V}^i$.

Since $\partial m(\bar{V}^i, \bar{V}^{1-i})/\partial \bar{V}^i = 1/2\tau$, thus

$$\Psi_{q\bar{V}^i} = \frac{1}{2\tau}(c_{\text{off}} - c_{\text{on}}).$$

Since $c_{\text{off}} > c_{\text{on}}$ and $\tau > 0$ by assumption, this expression is strictly greater than zero. Hence, $\frac{d}{d\bar{V}^i}\hat{q}(\cdot, \bar{V}^i) > 0$. □

Interpreting this result is intuitive. Suppose firm i ceteris paribus supplies less utility. Then m_i decreases, and, in response, $1 - m_i$ increases, thus own consumers are calling off-net with greater probability, which calls are more costly than on-net calls. Hence, in order to suffer less losses, firm i reduces on the quantity it offers in the contract. The magnitude of this effect is driven by the transport cost parameter and the on–off-net cost difference in exactly the way we would intuitively anticipate it. If the networks are more differentiated (τ is bigger), or the cost difference is small, this effect is going to be smaller.

Along the same line of argument, it is evident that under the assumptions of Proposition 3.5, the opponent firm can exercise competitive pressure on firm i through its quantity schedule. In particular, $\hat{q}$ changes downwards if $\bar{V}^{1-i}$ rises.

3.3.2 A simple example

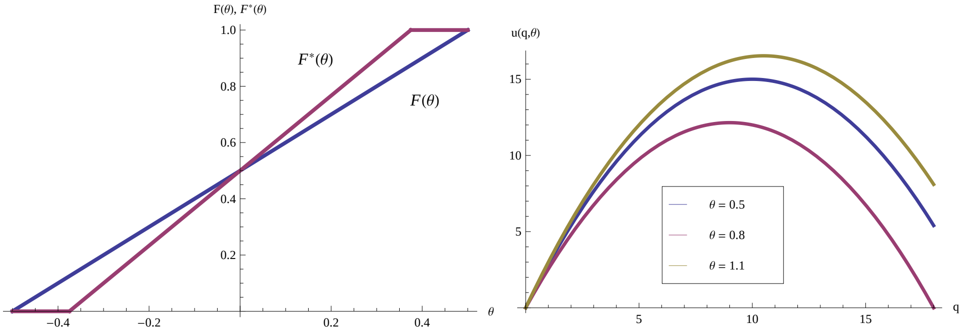
As the discussion above suggests, the formulation of this extension already makes determining the equilibrium of the market share game complicated. To understand the structure of the problem better we present, and attempt to solve, a simple illustrative example.

Example 1. *Let F be the uniform distribution on $[-\frac{1}{2}, \frac{1}{2}]$, and let F^* be the uniform distribution on $[\frac{3}{8}, \frac{3}{8}]$. The type space $\Theta := [-\frac{1}{2}, \frac{1}{2}]$. These distributions ensure that the monotonicity constraint is never binding (Grubb, 2009, p. 1783.). Let the utility function be $u(q,\theta) \equiv \frac{3}{2}(1+\theta)q - \frac{3}{20}q^2$. The costs of the firms are symmetric: $c_{\text{on}} = \frac{1}{5}$ and $c_{\text{off}} = \frac{1}{4}$ for both firms.*[19]

Since the derivative $u_{\theta qq}$ is zero, and $u_{qq} < 0$, this utility function also satisfies Assumption 3.3. This way, assuming that the satiation constraint is not binding, we can explicitly calculate the optimal quantity $\hat{q}(\theta, \bar{V}^i, \bar{V}^{1-i})$ from Equation 3.21. This is given by the following formula:

$$\hat{q}(\theta, \bar{V}^i, \bar{V}^{1-i}) = \frac{120\theta\tau + 81\tau + \bar{V}^i - \bar{V}^{1-i}}{12\tau} - \begin{cases} \frac{20\theta}{3} + \frac{1}{2}, & \text{if } \theta \in [-\frac{3}{8}, \frac{3}{8}] \\ 5 & \text{else} \end{cases}. \tag{3.22}$$

[19] Calculations were carried out in *Mathematica*.



(a) Distributions satisfying Assumption 3.1

(b) Utility functions

Figure 3.2: Functions in Example 1

Naturally, the optimal quantity also depends on the exogenous τ (through the market share). The case distinction in the curly bracket has to be made because $F(\theta) - F^*(\theta)$ is not precisely given by the two functional values. The reason for this is that the support of F^* is essentially $[\frac{3}{8}, \frac{3}{8}]$. In a more exact formulation, therefore, $F^*(\theta) = 0$, if $\theta \in [-\frac{1}{2}, -\frac{3}{8}]$, $F^*(\theta) = 1$, if $\theta \in [\frac{3}{8}, \frac{1}{2}]$, and given by the uniform formula else. The result of Proposition 3.5 is also nicely illustrated by the positive derivative of $\hat{q}$ with respect to $\bar{V}^i$.

The optimal price schedule is already very complicated. Nevertheless, its derivative with respect to $\bar{V}^i$ is contained in the appendix and is given by equation A.11. From this we can see that it still depends on the particular $\bar{V}^i$ where we evaluate it. Unfortunately, we cannot tell an explicit sign. The second derivative, however, is clearly negative indicating that an increase in $\bar{V}^i$ diminishes the marginal increase (or decrease) in $\hat{P}$.

After finding the optimal contract for each $\bar{V}^i$ (and $\bar{V}^{1-i}$)[20] we can move one step further finding the payoff function for the game described in Proposition 3.3. Since the profits are evaluated in expectations, the case distinctions will not appear in the payoff function that is given by the following equation:

$$\frac{1}{\tau^3}\Big\{0.5(\tau + \bar{V}^i - \bar{V}^{1-i})\Big[4.81719\tau^2 - 0.003125(\bar{V}^i)^2 \\ + \bar{V}^i(0.00625\bar{V}^{1-i} - \tau^2 - 0.06875\tau) - 0.003125(\bar{V}^{1-i})^2 + 0.06875\tau\bar{V}^{1-i}\Big]\Big\} \tag{3.23}$$

Observe, that this equation is cubic in $\bar{V}^i$, therefore we have to pay attention when using the FOC maximisation approach. There is one additional problem that is illustrated by the following figure.

[20] Remember: we are still assuming that the satiation and nonnegativity constraints do not bind for each θ.

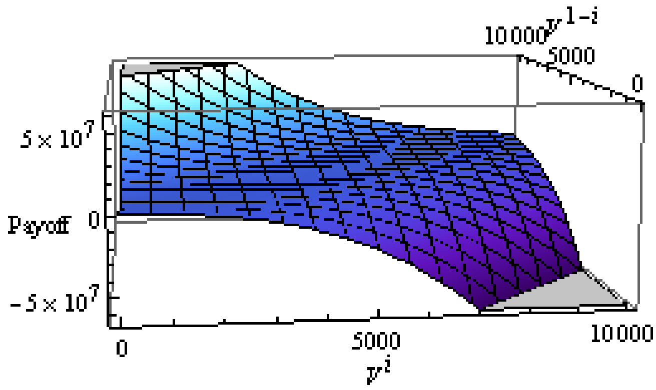


Figure 3.3: Payoff function for each pair $(\bar{V}^i, \bar{V}^{1-i})$ for $\tau = 2.7$

Figure 3.3 shows the payoff for each pair $(\bar{V}^i, \bar{V}^{1-i})$ for $\tau = 2.17$, which can be considered relatively high.[21] From this plot we can see, that for $\bar{V}^{1-i}$ high enough, the maximal $\bar{V}^i$ need not be easy to find through a simple FOC approach, corner solutions might appear.

Note, however, that in the derivation so far we did not take into account the following two factors: 1. The market share function is bounded for large differences in $\bar{V}$, and 2. the satiation constraint. The former does gives less incentives to offer greatly different utilities[22] and the latter gives a bound on the utility that can be offered to consumers.

Seeing the problem from this perspective, for relatively small $\bar{V}^{1-i}$ we can hope for a unique best response (possibly found through an FOC approach). This is illustrated in Figure 3.4.

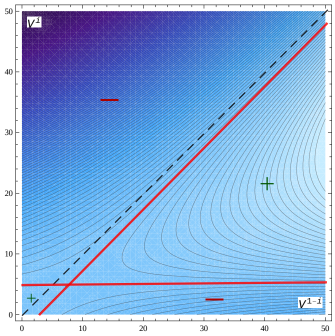
In this figure we can see the isopayoff levels for each pair $(\bar{V}^{1-i}, \bar{V}^i)$ (now the opponent's $\bar{V}^{1-i}$ displayed on the x axis). In Subfigures (a)-(b): The red lines indicate the zero isopayoff level; The black dashed line is the line with unity slope. Interestingly, the ridge does not lie on the line that designates the symmetric equilibria (black dashed line). Indeed, for sufficiently big $\bar{V}^{1-i}$ it is a dominant strategy for firm i to offer less utility down to a certain extent.[23] Since the game is symmetric, the same is true for the opponent. Therefore, the plot suggests that for high $\bar{V}^i$ and $\bar{V}^{1-i}$ no Nash-equilibrium might exist.

Note, however, that in the lower-left corner the payoffs are again positive, and increasing. Subfigure (c), which is the combination of (a) and (b) can be viewed as

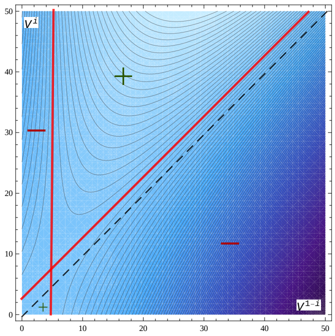
[21]The "flat", grey parts show regions where the machine precision cannot calculate further.

[22]In fact, if the difference is so high, that one of the firms covers the whole market, then this firm has the incentive to reduce the utility offered to its customers, as (given $\bar{V}^{i-1}$) the market shares won't change and the profits will be higher.

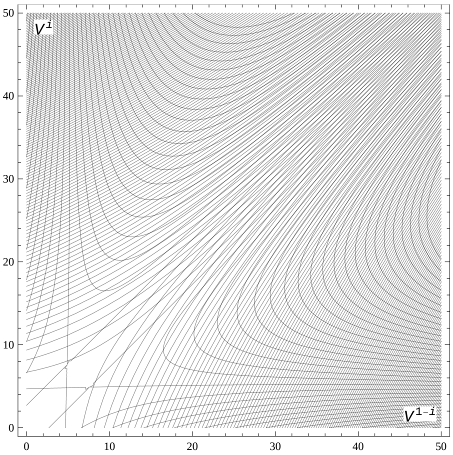
[23]This crucial nature of this plot does not change significantly with varying τ. If we expand the set of offerable utilities, then—as Figure 3.3 suggests—for a high choice of the opponent's, firm i is more likely to choose the zero corner solution.



(a) Firm i's payoff levels



(b) Firm $1-i$'s payoff levels



(c) Payoff matrix

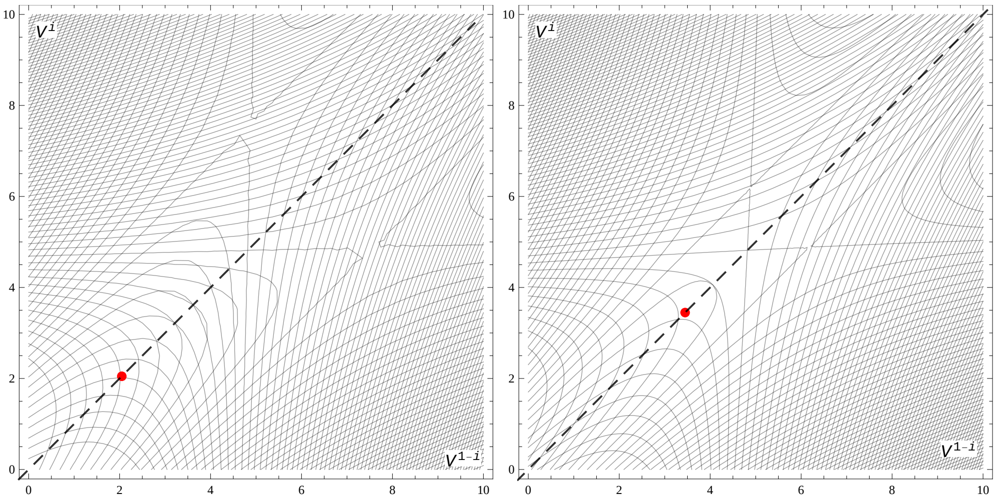
Figure 3.4: Contour plots for small $\bar{V}$ values

the payoff matrix in a standard normal form game: for each pair $(\bar{V}^{1-i}, \bar{V}^i)$ it shows the payoff levels for both firms. Observing this subfigure we can see, that for sufficiently small offered utility levels an equilibrium may exist. The symmetry of the equilibrium and the small absolute level of the utilities offered make this equilibrium plausible, even if we have not characterised the full solution yet.

Figure 3.5 illustrates these equilibira. For respective choices of $\tau^H = 2.7$ and $\tau^L = 1.3$, the following fixed points of firm i's best reply function can be calculated: $\hat{\bar{V}}^i = 2.0484$ and $\hat{\bar{V}}^i = 3.4484$ respectively. Since the game is symmetric, the payoff functions of the two firms are analogous, and also the reaction functions are the same. Hence, if $BR_i(x,\cdot) = x$, then also $BR_{1-i}(x,\cdot) = x$, and the pair (x,x) constitutes an NE. This way, we have found symmetric equilibrium candidates for the utility space game.

Now we make the following claim:

Claim 3.6. *In Example 1, if the satiation constraints are not binding for all θ, then $\hat{\bar{V}} = 2.0484$ and $\hat{\bar{V}} = 3.4484$ constitute a symmetric NE for $\tau = 2.7$ and $\tau = 1.3$ respectively.*



(a) $\hat{\bar{V}} = 2.0484$ ($\tau = 2.7$)   (b) $\hat{\bar{V}} = 3.4484$ ($\tau = 1.3$)

Figure 3.5: Fixed points in the utility space

Proof. The possible deviations from the symmetric equilibrium are analogous to those in the proof of Proposition 3.3.

For deviation $D \in (-\tau, \tau)$ the market share function is not kinked, thus the payoff is given by Equation 3.23. By looking at Figure 3.5 it is evident that every such deviation is strictly dominated by the symmetric NE strategy.

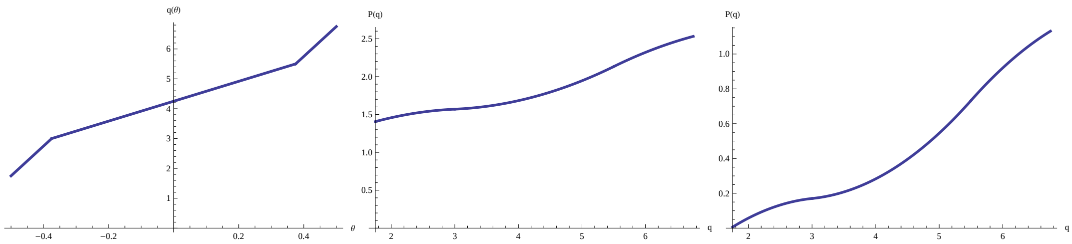
For deviation $D \leq -\tau$, the payoffs are zero, which is strictly worse than the equilibrium payoff.

For deviation $D \geq \tau$, the market share of firm i is going to be 1. Hence, if we take a look at Equation 3.21, we can see that the optimal quantity schedule (hence the variable price schedule) will be the same for all offered utility levels as long as the market share is. Hence it is a strictly dominant strategy to offer $\bar{V}^i$ such that $\bar{V}^i - \bar{V}^{1-i} = \tau$. In this case, however, the opponent firm has the incentive to raise its utility. Hence it cannot be an equilibrium.

Exchanging the role of the firms and repeating the argument concludes the proof. $\square$

From the above figures it is also intuitively clear that this NE that we have found for each τ is unique.

From this example interesting conclusions can be drawn. First of all, observe that the difference between the equilibrium utilities offered is precisely the difference in the trasport cost parameter, τ. Moreover, the equilibrium quantities offered are the same for both transport cost parameters. As can be seen in Figure 3.6, the optimal price schedules also differ only by a fixed fee. Hence, our general conclusions apply in this example. Note, however, that the three-part tariff structure does not carry on to this model extension, even if we let c_{on} be zero.



(a) Quantities offered (b) Price schedule ($\tau = 2.7$) (c) Price schedule ($\tau = 1.3$)

Figure 3.6: Optimal quantity and price schedules

Since, this example has simple, well-defined structure, we are not able to draw general conclusions about the proposed extension itself. Nevertheless, having understood the possible complications arising in this model setting, we can leave the problems for future research.

Chapter 4

Conclusion

The present thesis aims at examining the strategic mechanisms in the imperfectly competitive environment of the telecommunications industries. To this end, we systematically reviewed competition in linear, two-part and three-part tariffs. As Section 1.1 illustrates, the telecommunications industry plays an important role both at the macro- and the microeconomic level. Further, we have given real-world examples for the existence and relevance of all price schedules that we investigate.

In the case of linear and two-part tariffs, we could build on the well-developed literature of two-way interconnection (Armstrong and Wright, 2009; Laffont, Rey, and Tirole, 1998a,b). In the case of three-part tariffs, however, a monopolistic non-linear pricing model (Grubb (2009) – M) had to be extended to a duopolistic model (H). Furthermore, an extension of H had been proposed (H^+) in order to match the model more closely with the telecommunications sector.

Our results can be interpreted at two levels. At the general level, we are able to conclude that a general pattern in all models under scrutiny exists. First, if the setting of a fixed fee is allowed, then due to an increase in competition, the fixed fee is going to drop. Unit prices will be charged at marginal costs. Second, if a fixed fee is not allowed to be set, then competition decreases the unit price both on- and off-net. An exception to this latter rule is a model setting when utility can be gained from being called by an other party (Berger, 2004a).

On the specific level, we were able to successfully generalise the consumer overconfidence model of Grubb (2009) (M) to a duopolistic setting (H). To this end, we assumed Hotelling product differentiation and applied the competition in utility space approach proposed by Armstrong and Vickers (2001). Our main results are Proposition 3.3 and Corollary 3.4 which show that the main results of M will be flanking outcomes of H. In particular, a unique symmetric equilibrium in the utility game exists, and varying the transport cost parameter τ to its two extremes (0 and ∞) gives back the perfect competitive and the monopolistic pricing results of M respectively. Furthermore, we were able to conclude that H yields the same marginal pricing structure as M—which is a three part tariff. This way we were able to verify that the general pattern found in simpler models also applies to the more complex three-part tariff model.

Finally, an extension of H to H^+ has been proposed, which was motivated by our understanding of the two-way interconnection literature: variable costs differ based on the network on which the call terminates. Thus, we have assumed that 1. off-net costs are higher than on net costs and 2. the probability of a call terminating on-net equals the relative market share of the home provider. This formulation resulted in a more complex optimal contract problem than that of H. While under certain strong assumptions we were able to give an implicit characterisation of the optimal contract, we were not able to formulate the utility space game explicitly. However, through the means of a simple example and computer simulations we illustrated the interesting aspects of the problem, and drew two conclusions: First, the three-part tariff structure of H is not likely to carry on to H^+; Second, while a symmetric equilibrium of the utility game might not exist in general, we are able to find a symmetric NE in this simple setting. Examining comparative statics in the symmetric equilibrium of this example shows that our general level conclusions apply to this particularly simple setting.

The full characterisation of the solution to H^+ is an interesting area of further investigation that we leave for future research.

Bibliography

ARMSTRONG, M. (1998): "Network Interconnection in Telecommunications," *The Economic Journal*, 108(448), 545–564.

ARMSTRONG, M., AND J. VICKERS (2001): "Competitive Price Discrimination," *RAND Journal of Economics*, 32(4), 579–605.

——— (2010): "Competitive Non-linear Pricing and Bundling," *Review of Economic Studies*, 77(1), 30–60.

ARMSTRONG, M., AND J. WRIGHT (2009): "Mobile Call Termination," *Economic Journal*, 119(538), F270–F307.

BERGER, U. (2004a): "Access Charge in the Presence of Call Externalities," *Contributions to Economic Analysis and Policy*, 3(1), Article 21.

——— (2004b): "The Economics of Two–Way Interconnection," Doctoral dissertation, Vienna University of Business and Economics.

——— (2005): "Bill-and-Keep vs. Cost-Based Access Pricing Revisited," *Economics Letters*, 86(1), 107–112.

BÖRGERS, T. (2010): "An Introduction to the Theory of Mechanism Design," Manuscript.

CARROLL, G. (2012): "When are Local Incentive Constraints Sufficient?," *Econometrica*, 80(2), 661–686.

COURTY, P., AND L. HAO (2000): "Sequential Screening," *Review of Economic Studies*, 67(4), 697–717.

DESSEIN, W. (2003): "Network Competition in Nonlinear Pricing," *RAND Journal of Economics*, 34(4), 593–611.

FUDENBERG, D., AND J. TIROLE (1991): *Game Theory*. MIT Press, Cambridge, MA.

GANS, J. S., AND S. P. KING (2001): "Using 'bill and keep' interconnect arrangements to soften network competition," *Economics Letters*, 71(3), 413–420.

GENAKOS, C., AND T. VALLETTI (2011a): "Seesaw in the Air: Interconnection Regulation and the Structure of Mobile Tariffs," *Information Economics and Policy*, 23(2), 159–170.

——— (2011b): "Testing the "Waterbed" Effect in Mobile Telephony.," *Journal of the European Economic Association*, 9(6), 1114–1142.

GRUBB, M. D. (2009): "Selling to Overconfident Consumers," *American Economic Review*, 99(5), 1770–1807.

——— (2012): "Dynamic nonlinear pricing: Biased expectations, inattention, and bill shock," *International Journal of Industrial Organization*, 30(3), 287–290.

HAHN, J.-H. (2004): "Network Competition and Interconnection with Heterogeneous Subscribers," *International Journal of Industrial Organization*, 22(1), 611–631.

HERWEG, F., AND K. MIERENDORFF (2012): "Uncertain Demand, Consumer Loss Aversion, and Flat-rate Tariffs," *Journal of the European Economic Association*, forthcoming.

JENSEN, S. (2006): "Implementation of competitive nonlinear pricing: tariffs with inclusive consumption," *Review of Economic Design*, 10(1), 9–29.

JULLIEN, B., P. REY, AND W. SAND-ZANTMAN (2010): "Mobile Call Termination Revisited," Working Paper 551, Institut d'Économie Industrielle (IDEI), Toulouse.

KRÄHMER, D., AND R. STRAUSZ (2011): "The Benefits of Sequential Screening," CEPR Discussion Paper 8629, Centre for Economic Policy Research.

LAFFONT, J.-J., P. REY, AND J. TIROLE (1998a): "Network Competition: I. Overview and Nondiscriminatory Pricing," *RAND Journal of Economics*, 29(1), 1–37.

——— (1998b): "Network Competition: II. Price Discrimination," *RAND Journal of Economics*, 29(1), 38–56.

LAFFONT, J.-J., AND J. TIROLE (2000): *Competition in Telecommunications*. MIT Press, Cambridge, MA.

LAPPÖHN, S., A. POHL, AND B. ZUCKER (2011): "Preisentwicklung von Telekommunikationsdienstleistungen in Österreich – Empirische Darstellung," Research report, Insitute for Advanced Studies, Vienna.

MILGROM, P., AND I. SEGAL (2002): "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70(2), 583–601.

OECD (2011): "Communications Outlook 2011," Biennial report, Organization for Economic Co-operation and Development, ISBN: 9789264098381.

ROCHET, J.-C., AND L. A. STOLE (2002): "Nonlinear Pricing with Random Participation," *Review of Economic Studies*, 69(1), 277–311.

SHY, O. (1995): *Industrial Organization – Theory and Applications*. MIT Press, Cambridge, MA.

STOLE, L. A. (1995): "Nonlinear Pricing and Oligopoly," *Journal of Economics & Management Strategy*, 4(4), 529–562.

SUNDARARAJAN, A. (2004): "Nonlinear Pricing of Information goods," *Management Science*, 50(12), 1660–1673.

TIROLE, J. (1988): *The Theory of Industrial Organization*. MIT Press, Cambridge, MA.

WILSON, R. (1993): *Nonlinear Pricing*. Oxford University Press, Oxford.

A Calculations

Calculations for the canonical model of interconnection

We are solving the game described in section 2.1 and taken from Armstrong and Wright (2009) backwards. We have already established an implicit equation for market shares as one equilibrium of the customers' subgame.

Write the subscriber utility as

$$w_i = s_i v(p_i) + (1 - s_i) v(\hat{p}_i) - F_i \;, \tag{A.1}$$

further write the equilibrium market shares as

$$s_i = \frac{1}{2} + \frac{w_i - w_j}{2\tau} \;. \tag{A.2}$$

First, firm i given j's action determines the optimal prices. The profit of firm i that it seeks to maximise can be written as

$$\pi_i = s_i \cdot \left(\underbrace{s_i q(p_i)(p_i - c_{\text{on}}) + (1 - s_i) q(\hat{p}_i)(\hat{p}_i - c_{\text{off}}) + F_i - f}_{\text{revenues from own customers}} \right) + \underbrace{(1 - s_i) \cdot s_i (a - c_T) q(\hat{p}_j)}_{\text{termination revenues}} \;. \tag{A.3}$$

Suppose that $s_i = s_j = 1/2$. Suppose further that $(p_i, \hat{p}_i, F_i) = (p_j, \hat{p}_j, F_j)$. A firm offers the utility w_i to its customers. In this symmetric equilibrium, the triple $(p_i, \hat{p}_i, F_i)$ must therefore be such that they satisfy A.1. Given this symmetric equilibrium (hence dropping firm subscripts), firm profit can be written as

$$\pi = \frac{1}{4} \left[q(p)(p - c_{\text{on}}) + q(\hat{p})(p - c_{\text{off}}) \right] + \frac{1}{2} \left(\underbrace{\frac{1}{2} v(p) + \frac{1}{2} v(\hat{p}) - w}_{F} \right) + K \;, \tag{A.4}$$

where K is a constant. The first order conditions for p and $\hat{p}$ are

$$0 = \frac{\partial \pi}{\partial p} = \frac{1}{4}\left(pq'(p) + q(p) - c_{\text{on}}q'(p) + v'(p)\right)$$
$$0 = \frac{\partial \pi}{\partial \hat{p}} = \frac{1}{4}\left(\hat{p}q'(\hat{p}) + q(\hat{p}) - c_{\text{off}}q'(\hat{p}) + v'(\hat{p})\right) ,$$

and noting that $v'(p) = -q(p)$, it immediately follows that either $q'(p)$ and $q'(\hat{p})$ must be zero, or $p = c_{\text{on}}$ and $\hat{p} = c_{\text{off}}$ in optimum.

Next, firms act strategically, simultaneously to determine the fixed fees. Given the cost-based prices, the decision over F_i is simplified to the maximisation of the following function:

$$\pi_i = s_i \cdot [F_i - f + (1 - s_i)(a - c_T)q(c_{\text{off}})] . \qquad \text{(A.5)}$$

From A.2 and A.1 given the prices found above for the symmetric equilibrium, we can express the equilibrium market shares as

$$\begin{aligned}
s_i &= \frac{1}{2} - \frac{s_i(v(c_{\text{on}}) - v(c_{\text{off}})) + v(c_{\text{off}}) - F_i - s_j(v(c_{\text{on}}) - v(c_{\text{off}})) - v(c_{\text{off}}) + F_j}{2\tau} \\
&= \frac{1}{2} - \frac{(2s_i - 1)(v(c_{\text{on}}) - v(c_{\text{off}})) + F_j - F_i}{2\tau} \\
&= \frac{\tau - v(c_{\text{on}}) - v(c_{\text{off}}) + F_j - F_i}{2(\tau - v(c_{\text{on}}) + v(c_{\text{off}}))} \\
&\Downarrow \\
s_i &= \frac{1}{2} - \frac{F_i - F_j}{2(\tau - v(c_{\text{on}}) + v(c_{\text{off}}))}
\end{aligned}$$

where the second line follows from $s_j = 1 - s_i$, and the third is the explicit expression of s_i.

Denote $\Phi \equiv 2(\tau - v(c_{\text{on}}) + v(c_{\text{off}}))$, and $\Sigma \equiv (a - c_T)q(c_{\text{off}})$. These are both independent of F_i or F_j, thus are constants in the optimisation. Inserting the above explicit s_i into A.5 gives

$$\frac{F_i}{2} - \frac{F_i^2}{\Phi} + \frac{F_i F_j}{\Phi} - \frac{f}{2} + \frac{F_i f}{\Phi} - \frac{F_j f}{\Phi} + \frac{\Sigma}{2} - \frac{F_i \Sigma}{\Phi} + \frac{F_j \Sigma}{\Phi} - \left[\frac{1}{2} - \frac{F_i - F_j}{\Phi}\right]^2 \Sigma . \qquad \text{(A.6)}$$

Maximising this expression w.r.t. F_i (given F_j) results in

$$\frac{1}{2} - \frac{2F_i}{\Phi} + \frac{F_j}{\Phi} + \frac{f}{\Phi} - \frac{\Sigma}{\Phi} + 2\left[\frac{1}{2} - \frac{F_i - F_j}{\Phi}\right]\frac{\Sigma}{\Phi} . \qquad \text{(A.7)}$$

Substituting for $F_i = F_j$, i.e. investigating the symmetric equilibrium, the following result holds

$$F_i = f + \frac{\Phi}{2} = f + \tau - v(c_{\text{on}}) + v(c_{\text{off}})\,. \tag{A.8}$$

In the first stage, firms choose a. This they do by maximising the industry profits, i.e. $\pi_1 + \pi_2$. Noticing that Σ depends on a, and that $c_{\text{off}} = c_O + a$ substituting A.8 into A.5 gives

$$\begin{aligned}\Pi = 2\pi &= f + \tau - v(c_{\text{off}}) + v(c_O + a) - f + \frac{1}{2}\Sigma(a) \\ &= \frac{1}{2}\Sigma(a) + \tau - v(c_{\text{on}}) + v(c_O + a)\,. \end{aligned} \tag{A.9}$$

Noting that $v'(p) = -q(p)$, The derivative of this profit function w.r.t. a is

$$\begin{aligned}\Pi'(a) &= \frac{1}{2}[(a - c_T)q'(c_O + a) + q(c_O + a)] - q(c_O + a) \\ &= \frac{1}{2}[(a - c_T)q'(c_O + a) - q(c_O + a)]\,, \end{aligned} \tag{A.10}$$

which is negative in the ε-neighborhood of c_T given that assumption that $a \geq 0$ (thus $c_O + a \geq 0$). Naturally, the actual value depends on the demand function and its derivatives.

Supplement to Example 1

Recall, that in Example 1 $\theta \in [-\frac{1}{2}, \frac{1}{2}]$ by assumption.

The derivative of the optimal price schedule $\hat{P}$ with respect to $\bar{V}^i$ is given by:

$$\begin{aligned}\frac{60}{480\tau}\Bigg[\Bigg(\begin{cases} \frac{4\theta}{3} + \frac{1}{2} & -\frac{3}{8} \leq \theta \leq \frac{3}{8} \\ 1 & 8\theta > 3 \end{cases}\Bigg) \\ - 8\tau\Bigg(\begin{cases} \frac{0.125\theta + 0.0625}{\tau} & \theta > -0.5 \\ 0 & \text{else} \end{cases}\Bigg)\Bigg] \\ + \frac{1}{480\tau^2}[-60\theta\tau + 9\tau - \bar{V}^i + \bar{V}^{1-i}]\,. \end{aligned} \tag{A.11}$$

The derivative of this with respect to $\bar{V}^i$ is $-\frac{1}{480\tau^2} < 0$.

***