Universitat Autònoma de Barcelona

Ph.D. Thesis
Economics

Essays on the Industrial Organization of Telecommunications

Michele Pistollato

April, 2012

Supervisor: Prof. David Pérez Castrillo

DEPARTAMENT D'ECONOMIA I D'HISTÒRIA ECONÒMICA

Facultat de Economia i Empresa
Edifici B - Campus de Bellaterra
08193 Bellaterra (Barcelona)

A Paula y su familia.
Que me han apoyado mucho.

E alla mia famiglia.
Che mi è sempre stata vicino.

Acknowledgments

It is a pleasure to thank those who made this thesis possible.

First and foremost, I would like to express my sincerest gratitude to my advisor Professor David Pérez Castrillo for his never ending patience, motivation, enthusiasm, and knowledge. His guidance and continuous support has been invaluable and I could not have imagined a better mentor.

In addition, I am indebted to Professors Inés Macho Stadler and Xavier Martínez Giralt, who provided many significant contributions to the essays within my thesis. My sincerest thanks also goes to Professors Joan Calzada, Matthew Ellman, Steffen Hoernig, Sjaak Hurkens, Doh-Shin Jeon, Francesc Trillas Jané, and Tommaso Valletti.

I am grateful to all the professors and the administrative staff of the Department of Economics at Universitat Autònoma de Barcelona. I would particularly like thank to Ángels López Garcia and Mercè Vicente, who have the special ability of making all the procedures that are otherwise discouraging and time consuming seem easy and possible.

During one of my early days in the department, a professor told me that "I'll learn as much from the class hours as from my fellow students". No other words can better express why I am so grateful to all my colleagues. Among others, I would particularly thank my co-authors Tommaso and António.

Finally, I would like to acknowledge some very special people, whose unconditional support has been essential. My deep gratitude goes to my parents, Flavio and Caterina, for their emotional nearness and practical support. I am very grateful to my brothers, Carlo and Fabio, for inspiring and encouraging. And I would like to thank my grandparents, for teaching me the important things in life.

My sincerest thanks go to Paula and her family. Her parents, José Maria y Lourdes, have welcomed me as a son and at every moment made me feel like I was home. Paula, who with her joy makes me smile every day, has motivated me to finish my dissertation and move on to bigger and better things. Thanks for being by my side when I needed it most and for giving me the strength to carry on.

Contents

1 Introduction **1**

2 One-way access pricing **7**

2.1 Introduction . . . 7
2.2 The framework . . . 11
2.3 The model . . . 12
2.4 Welfare Analysis . . . 16
2.5 Optimal access prices . . . 18
2.5.1 A benchmark: non-discrimination between entrants . . . 18
2.5.2 Constraints . . . 20
2.5.3 The maximization problem . . . 22
2.5.4 Results . . . 25
2.6 Asymmetric Information . . . 28
2.6.1 Non-discriminating operators . . . 28
2.6.2 Discriminating operators . . . 30
2.7 Final remarks . . . 31
References . . . 34
Appendix 2.A Equilibrium and Welfare . . . 35
Appendix 2.B Optimal access prices . . . 37
Appendix 2.C Asymmetric Information . . . 42

3 CPP vs. RPP **45**

3.1 Introduction . . . 45
3.2 The Model . . . 50

3.3 The equilibrium . . . 52
3.3.1 Receiver sovereignty: $\beta p < r$ . . . 54
3.3.2 Caller sovereignty: $\beta p > r$ . . . 56
3.3.3 Comments . . . 58
3.4 Comparison . . . 59
3.4.1 Welfare analysis . . . 64
3.5 Conclusions . . . 65
References . . . 70
Appendix 3.A Proofs . . . 71
Appendix 3.B Results . . . 75

4 Network Neutrality **79**
4.1 The Model . . . 83
4.1.1 Basic Model . . . 83
4.1.2 Preliminaries: Congestion in the network system . . . 88
4.2 Monopolistic ISP . . . 91
4.2.1 Short run Analysis . . . 91
4.2.2 Long Run Analysis . . . 92
4.3 Benchmark: ISP competition with Network Neutrality . . . 93
4.4 ISP competition with Network Discrimination . . . 96
4.4.1 Consumers' choice . . . 96
4.4.2 Market shares . . . 104
4.4.3 Priority pricing . . . 107
4.4.4 End-users pricing . . . 109
4.5 Investment incentives . . . 112
4.5.1 Equilibrium profits . . . 113
4.5.2 Long run analysis . . . 115
4.6 Policy implications and conclusions . . . 117
References . . . 121
Appendix 4.A Network Neutrality . . . 122
Appendix 4.B Network Discrimination . . . 123

List of Figures

2.1 The triangular Hotelling . . . 13
2.2 Incumbent's iso-profits . . . 21
2.3 Incumbent's profits . . . 25
2.4 Welfare comparison . . . 27
2.5 Incomplete information about fixed cost. . . . 29
2.6 Incomplete information about transportation cost. . . . 30

3.1 Access price . . . 58
3.2 Usage prices p and r. . . . 59
3.3 Length of a call q. . . . 61
3.4 Fixed part F. . . . 62
3.5 Market penetration N. . . . 64
3.6 Total profits π_i. . . . 65
3.7 Total welfare W. . . . 66

4.1 Market representation . . . 88
4.2 Net neutrality: market sharing. . . . 95
4.3 Indifferent users by type . . . 101
4.4 Indifferent users when priorities are sold to different CPs . . . 103
4.5 Representation of market shares . . . 106
4.6 Users gained by a CP with priority under ISP_A . . . 108
4.7 Users gained by a CP with priority under ISP_B . . . 125

Chapter 1

Introduction

In the European countries, telecommunications has been a public owned industry for a long time. Many reasons have prompted governments to regulate the sector through public monopoly. First, the high fixed costs of infrastructures make telecommunications an industry where natural monopolies are likely to arise. Therefore, public ownership was a solution to correct this market failure. Second, the industry had a strategic importance for the economy of a country: a modern and technological telecommunications sector does not only contribute to the overall growth of the economy but it is a necessary condition to guarantee the development of other sectors as well. Hence, a public monopoly was also a solution to pursue a form of industrial policy and to control one key sector of the economy. Third, access to telecommunication services is considered a consumers' right, and as such, it needs to be protected. Providing telecommunications as a public service was the European solution to ensure a universal telecommunication service.

In the USA, the regulation of a private monopoly was preferred to a public ownership of the national telecom operator. However, the regulator's objectives were the same as those of the European counterpart. Prices were regulated either to mimic competitive outcomes or to achieve some industrial policy objective. The monopolist "voluntarily" provided a universal service as compensation for service exclusivity.

Economic theory provided low support to regulators' activity. On one hand (demand side theories), utilities regulation was regarded as a political rather than an economic issue. Universal coverage, service quality, and strategic industrial policies were the main objectives that regulators aimed to achieve. Monopoly was considered the best market outcome to a static industry where innovations only improved services quality but did not increase the range of the actual services that were offered. Competitive solutions were improbable due to the high fixed costs, which made inefficient and unattainable the replication of the network. On the other hand offer side theories), industrial theories assumed that telecommunications were a natural monopoly and did not consider competitive frameworks as possible solutions to the monopolistic market failure. The main focus was on providing sophisticated tools to address manifold economic and political issues. Regulation of natural monopolies contemplated issues like optimal pricing, investment incentives and optimal size of the utility.

The 1980s marked the major regulatory changes in the telecommunications industry. In the previous decade, the Governments' expenditure had grown at a great rate, rising public debts at worrying levels. It was necessary to cut public expenditure, and the utilities sector was one of the best candidates for cost savings. In particular, the public management of utilities was very inefficient, causing wastes that needed to be covered through transfers from central budget. In Europe, most countries decided to privatize utilities, mainly to avoid bad public management costs but also with the hope of boosting the economy (and hence tax incomes) through the growth of new private utility sectors. Telecommunications, along with other utilities (i.e. distribution of electricity and gas, water and sewerage, railways) were characterized by expensive networks, whose replication was unaffordable by entrants that could undermine competition. The solution has been to create a heavily regulated private monopoly for telecommunications long-distance services and to promote the growth of a competitive industry for local services.

The European solution was very similar to the US telecommunications market regulation, where the long-distance services were offered by a regulated monopoly (AT&T), which faced minor competition. However, three fundamental differences have marked/characterised European long-distance telecommunications services and offered the economists the opportunity to develop new regulatory tools to support policy makers. First, European regulators were very concerned about incumbent inefficiencies. While the US regulator controlled the incumbent prices to prevent the incumbent company from earning excessive profits due to its privileged position, the European counterpart aimed to regulate prices to achieve efficiency gains. Second, the European incumbents kept the right to compete with entrants in the provision of local services. To avoid regulating a vertical integrated company, the US telecommunications regulator (the Federal Communications Commission, FCC) separated local from long-distance services. Therefore, AT&T could not compete with regional companies in the supply of regional calls. Third, European regulators allowed competition for regional services provision. This decision was an implicit consequence of allowing the incumbent to compete with entrants in the regional markets. In the US, regional calls were provided by local monopolies while interregional calls were supplied by AT&T. In 1996, the FCC adopted a different approach by allowing local and long distance competition between AT&T and regional companies. In summary, the US vertical separation approach aimed to prevent the regulation of dominant market positions by separating markets and controlling prices to avoid monopolistic profits, while the European vertical regulated integration approach aimed to control prices to promote competitive markets where otherwise the rise of monopolistic companies would be the market outcome. Interestingly, the vertical separation approach has largely been adopted in Europe to regulate other utilities except telecommunication companies.

After privatization, the telecommunications industry has grown beyond expectation. However, despite to economists and regulators' efforts to develop and apply regulatory tools to foster the industry, the main driver of the sector growth has been technological progress. In the 1990s, innovations in telecommunications have resulted in the provision of new services (information services, i.e. internet, being the most important) rather than contributing to improve existent services. This probably represents the main difference between telecommunications and other industries, where innovation improved supply quality but did not contribute to the same extent to the creation of new services that are highly valuable to consumers.

While the privatization and liberalization of the telecommunications market are unanimously recognized as factors that boosted technological innovation, economists still do not agree on whether regulatory policies (such as price regulation, providers separation or integration, subsidies etc.) have promoted technological development or have just followed it. Evidence can come to hand to clear this issue. In many cases, new technologies have been helping economists to provide new theories to approach market regulation. In other cases, economists' indications and regulators' consequent policies have stimulated the adoption of new technologies. However, in many other cases, the apple of discord is represented by the discrepancies between theoretical models and regulatory practices. Economists developed analytical tools responding to regulators' needs but that were necessarily approximations of reality. More realistic models were requiring too many information about the market structure or providing fine tuning regulations. Policy makers have often found very complicated to apply economists' indications and have rather adopted a "rule of thumb" approach.

The purpose of this thesis is to address three important regulatory issues that have not yet been satisfactorily solved by current theoretical models. The aim is to improve the economic tools to the disposal of regulators and provide new indications to policy makers so they can more adequately undertake telecommunications market regulation complexities.

In Chapter 2, it is considered the typical problem of an incumbent that competes with entrants in the retail market. The incumbent is also the owner of the backbone infrastructure that entrants need to access final users. Therefore, the incumbent could opportunistically set the landline-leasing price (access price) to take advantage of its vertical integrated position in an anticompetitive way. This possible issue, called one-way access pricing, has widely been discussed by economists and many effective solutions have been proposed. However, technological progress and Internet advent have raised new challenges. Usually, Internet connection suppliers charge a fixed fee to provide the access to end-users. One-way access models mainly focus on linear tariff pricing regimes, where consumption is charged on minutes (or seconds) of connection bases and few applications have been studied to the case of one-way access flat pricing. The main reason is that when an incumbent competes with an entrant on flat tariff, the overall welfare is unaffected by access price choice. Therefore, it is suggested to set a low access price in the early stages of entry, to promote competition. Then, regulators apply a non-discrimination principle, on the basis of which all entrants pay the same access price. However, as shown in the chapter, a non-discrimination principle is not welfare maximizing if there are two asymmetric entrants. In this case, the best policy is to set a lower access price to the most efficient entrant.

In Chapter 3, it is considered the case where two network operators own the backbone infrastructure. In this case, callers of the one network need to be connected to both on and off network receivers. Therefore, networks have to provide mutual (two-way) access. To connect the rival's users to its own, each network charges a regulated access fee to the competitor. Only recently, police makers have considered the fact that also the receiver derives utility from receiving a call and have included it in the optimal access fee. Nowadays, the main concern of regulators is on how the regulated access fee affects the tariff regime choice and its impact on market penetration. In particular, Europe market is characterized by a tariff regime where the cost of the call is entirely paid by the caller (Calling Party Pays, CPP). Conversely, in the US, caller and receiver share the call cost (Receiving Party Pays, RPP). In the chapter, it is shown how a tariff regime and the resulting market penetration are an outcome of access fee regulation. In particular, the key factor that drives networks' pricing choices and consequent market size is the value of the receiver externality. In the studied model, a high access fee is desirable when the receiver externality is low, ending up with a CPP regime. Otherwise, if the receiver externality is high, the optimal policy is to set a close-to-zero access fee, determining an RPP regime.

In Chapter 4, it is analyzed one of the most debated topic regarding the Internet: the network neutrality. In this context, networks are regarded as platforms that serve a two-sided market. On one side, consumers connected to a network use network's services and browse Internet contents. On the other side of the market, Internet content providers develop digital contents, using the network platform to distribute them to consumers, and make profits by including banner ads on their web pages. Users' decision to browse one content depends on the content characteristics and the browsing speed. Under a network neutrality regime, networks cannot discriminate between contents by prioritizing one of them. However, networks claim the right to provide priority services, charging a fee to the prioritized content, to cover infrastructure costs to build and expand the network capacity. The European Commission is called to take a position to this regard. On one hand, an adequate level of investments in the next generation of networks is considered an important driver for economic growth. On the other hand, it is not yet clear whether network neutrality is a regime compatible with an appropriate investments level, or if allowing network discrimination could lead to an improving of the welfare. The actual policy aims to provide an optimal outcome by fostering network competition. However, the lack of economic models tackling the issue makes it hard to policy makers to deal with it. In this chapter, the problem of investment incentives and its linkage with the network neutrality regime is considered in detail. As a result, it is shown that a competitive market of networks is not a sufficient condition to achieve a high level of investments.

Chapter 2

One-way access pricing when oligopolies compete on flat tariffs

2.1 Introduction

Flat tariffs are daily becoming a more common pricing scheme in the telecommunication sector.[1] One of the most important services typically offered at a flat fee is internet access. Given the large number of hours that users spend browsing internet pages every day, consumers tend to choose to pay a flat tariff for access.

While in the early days of the Internet access was supplied by a single provider (typically the now privatized national telecom company), in just a few years, this sector has seen the emergence of several competitors. This growth has been driven by increasing demand and the reduction of costs because of the rapid development of new, cheaper technologies. Nevertheless, regulators have played a fundamental role in the growth of the industry.[2]

[1]See the statistics in the national regulators' annual reports, for instance CMT (2009) or Ofcom (2010).

[2]Subsequent the liberalization of the telecommunication industry, national authorities have faced several regulatory issues and shaped, through their decisions, the evolution of the market. See Armstrong & Sappington (2006) or Waverman & Sirel (1997).

One of the first policy problems they considered was the regulation of so-called "one-way access".[3] Due to the significant investments required to replicate the network (especially the "last mile" network), entrants were unable to compete directly with the incumbent (the former public telephone company). To foster competition, the regulator forced the incumbent to provide access to the entrants' end-users by renting out the "local loop": instead of duplicating the lines from the long-distance backbone to the end-users (local line), the entrants lease the incumbent's local wireline to provide internet access. The regulation of the rental price paid by the entrants to the incumbent is one of the most important policy instruments at the disposal of telecommunications authorities.

The strategic and economic importance of the telecommunications industry has attracted the attention of many economists, who have rapidly contributed to the creation of a flourishing literature regarding telecom regulation.[4] In particular, they have widely treated the regulation of rental prices by proposing several access price rules.[5] However, all of these rules are tailored to a pricing scheme where users are charged according to usage (linear price), despite to the fact that almost all internet services are offered at a flat tariff.[6] The lack of analysis on one-way pricing in a flat tariff setting is because access pricing policy, in this context, is not likely to have any impact on social welfare. In a market with one incumbent and one entrant, de Bijl & Peitz (2002) show that *a higher access price is transmitted one-to-one into higher flat fees for both operators*, and therefore, the marginal consumer is not affected by the lease price so that market shares do not respond to access price regulation. This implies that the entrant's profits do not respond to the lease price, hence lease price regulation cannot stimulate entry if firms compete on flat tariffs. So they conclude that, when there are only two operators in the retail

[3]This topic is widely treated by Armstrong (1998), Vogelsang (2003) and in a book by Laffont & Tirole (2000).

[4]Excellent reviews of the early literature on telecommunications regulation are provided by Armstrong et al. (1996), Kridel et al. (1996) and Laffont & Tirole (2000).

[5]See Laffont et al. (1998) and Vickers (1997) for some examples and for references.

[6]The early literature ignores flat rates because entrants initially offered telephone services, which are usually priced per minute.

market, *there are no economic benefits of such a regulatory policy.*

In this paper I consider three firms competing for final users (one incumbent and two asymmetric entrants). My main result is that, under these assumptions, the marginal consumer (and therefore the entrants' profits) depends on the regulation of the access price. Hence the telecommunications authority can affect the total surplus, even if the operators compete on flat tariffs. In particular, in the social optimum the more efficient entrant pays a lower access price than the less efficient one, and the incumbent can gain positive profits if it is sufficiently efficient. I also check for possible information problems: if the industry costs are unknown to the regulator, the operators can take advantage by misrepresenting them without any possibility of the authority preventing this behavior. Conversely, if the degree of competition in the retail market is unknown, the regulator can set a contract that prevents opportunistic behaviors.

After a brief literature review, the remainder of the paper is organized as follows. In Section 2 I describe the market and the industry. Section 3 defines the model (horizontal competition among three firms) and derives the equilibrium depending on the access price. In Section 4 I characterize social welfare, while in Section 5 I compute the optimal values of the access prices and compare them to a benchmark. Section 6 extends the previous analysis to a case where there is an information asymmetry between the regulator and the operators. Section 7 concludes.

Related literature. Unlike de Bijl & Peitz (2002), this paper considers other issues (unit demand with more than one entrant, asymmetric competition, and asymmetric information) that have already been studied in other contexts. However, while the one-way access problem has been widely discussed in the literature, little attention has been paid to the possibility of asymmetric entrants.

Armstrong (2002) considers two cases of the problem of foreclosure in a market with a unit demand model. In the first case he assumes that there is only one entrant, and he finds that the incumbent *has no incentive to distort competition downstream*, and therefore, regulation is not needed. In the second case he assumes that entrants constitute a competitive fringe. Here the findings are that it is optimal for the incumbent *to cause there to be a degree of productive inefficiency within the fringe in order to drive up the retail price.*[7] I extend his models by considering oligopolistic competition in the retail market.

Valletti (1998) develops a model to analyze the effects of imperfect competition in a Cournot setup where the incumbent is not operating in the downstream market. In his model, there are two downstream operators with different technologies, and the incumbent provides them the final access and charges a two-part tariff. When the regulator can set discriminatory access prices, two different welfare maximizing equilibria can arise: one where all the firms break even and one where all the firms make positive profits. He finds that *price discrimination induces the more productive firm to produce more. Discriminatory access tariffs, possibly below usage cost, are used to reduce inefficiencies without leaving downstream firms with extra profits.* Unlike him, I allow the incumbent to compete with the entrants in the final market.

In a later paper, de Bijl & Peitz (2004) extend their previous work by considering the dynamic evolution of the market in a repeated setting. The authors conclude that *the exact level of the lease price should not be too low, taking into account that the incumbent must be able to finance investments in the maintenance of the local access network* and that *consumer surplus is maximized by imposing a cost-based lease price, but the welfare result is unchanged.*

Finally, Peitz (2005) considers an incumbent-entrant setup where asymmetries are due to market power and not cost structure. He shows that *asymmetric access price regulation can stimulate entry and increase consumer surplus*, but, again, *total surplus increases with asymmetric access prices and remains constant under flat fees.*

[7] Some of Armstrong's results are related to those of Laffont et al. (1998). They found that the entrant should pay the full cost of connection.

2.2 The framework

I consider a case where a single operator owns the network in a region (usually, the formerly public telephone company). This network provides broadband access to the consumers living in the region. This includes various telecommunications services such as fixed telephony (for simplicity I ignore the connection to mobile networks) and internet access (it could also include television). The connection is offered to the end-users at a flat tariff that allows them to consume the service as much as they like. Hence consumers are interested in purchasing at most one unit of telecommunication service (broadband access).[8] To provide and maintain the physical connections to consumers, the owner of the network bears a fixed cost per user (the cost to install a wired connection from the main backbone to the home of the user and its maintenance). The operator also faces a cost to provide the service to the consumers (this cost depends on the time the user spends connected to the network).

The flat tariff represents a sunk cost to consumers, and once they have paid it, their consumption decisions do not depend on the price they paid but only on the utility they derive from using the telecommunication service. Hence, if all consumers have the same preferences, they will choose to consume the same amount of bandwidth. This amount of bandwidth will correspond to a cost for the operator that will decline as the operator becomes more efficient at providing the service. Notice that this is a constant per user marginal cost.

Once this operator has been established in the region, providing connections to all of the consumers, it may be the case that one or more entrants would like to compete with the incumbent in the retail market for broadband. To do so, they need to either build their own networks or purchase access to the end-users from the incumbent. If replication of the network is not feasible, the incumbent can act as a monopolist in the wholesale market for the broadband by arbitrarily setting the price for access to the final users (the so-called access price). A regulator could

[8]Although a single user could buy multiple connections, here I consider the most common case where each consumer chooses to have only one connection.

consider the behavior of the incumbent anticompetitive and welfare damaging and therefore, the regulator may want to regulate the access price in a welfare maximizing way.

2.3 The model

The model considers a telecommunications market with three firms: an incumbent I, which is the sole owner of the network, and two possibly asymmetric entrants, firm A and firm B, which compete with the incumbent in the end-users market. The incumbent pays a fixed cost f to supply network access to the end-users, and the other producers, A and B, are not able to replicate the network structure and need to purchase access from the monopolist by paying a regulated price a.

The final market is characterized by differentiated users, each of whom has a preferred variety of the final service. I assume that there is one user of each possible variety; hence the number of users is the same as the number of varieties. A type of consumer purchases one unit of the service from one firm i (where $i \in \{I, A, B\}$) and pays price p_i.

Varieties of the final service may be represented in an equilateral triangle with a perimeter normalized to 1. Each point x on the border of the triangle denotes a different variety and an associated type of user who prefers it to all others. Points closer to x are similar but different varieties. This allows for imperfect competition that I model through a triangular Hotelling model, where operators act as Nash competitors using price strategies. The users' preference space is represented by the triangle. When a generic user x purchases the service from operator i, offering service type x_i, she receives utility

$$u_i = v_0 - p_i - t|x_i - x|, \tag{2.1}$$

where $t > 0$ is a transportation cost, and the term $t|x_i - x|$ reflects how much consumer i dislikes service x compared to her most preferred service x_i. The parameter v_0 denotes how much a user appreciates the unit of telecom service she buys (v_0 is large enough to guarantee that all consumers choose to be connected to the network).[9]

I consider that firms differentiate themselves by selling differentiated services to the end-users, and, in particular, I avoid the location problem by assuming that there are only three, equidistant locations available: $x = 0$ (which is the same as $x = 1$), $x = 1/3$, and $x = 2/3$. Moreover, each firm chooses one of these positions, and, without any loss of generality, I set I to position 0, A to position 2/3, and B to position 1/3. Notice that the utility perceived by a consumer located at ϵ (where $0 < \epsilon < 1/2$) when she purchases the service provided by firm I is the same utility perceived by a consumer located at $1 - \epsilon$.

A user type of x decides to purchase the services of operator i only if her perceived utility is greater than the utility perceived from purchasing from j, where $j \neq i$, that is, when $u_i > u_j$. I assume that all of the operators are active in the market, and that the market is fully covered; therefore, there must always be an indifferent consumer between I and A, between I and B, and between A and B. A consumer who is indifferent between the services provided by i and j is denoted by c_{ij}. The spatial competition is depicted in Figure 2.1.

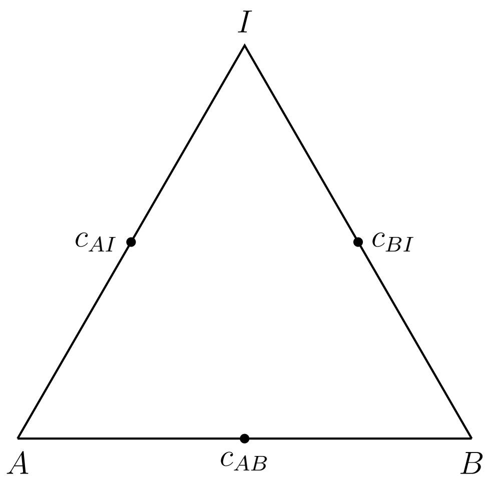


Figure 2.1: The triangular Hotelling

[9]The adoption of this particular utility function is made for technical convenience, as it considerably simplifies the study of the equilibrium.

Therefore, I can define the final market demand of each operator i as the number d_i of users who prefer the variety offered by i to the varieties offered by the two competitors.

Finally, to provide the final service, an operator incurs a cost K_i that is increasing in the number of users (which also denotes the number of units of final service that are produced). In particular, in a case where the industry presents constant returns to scale, the production costs for operator i when serving d_i users are $K_i \equiv k_i d_i$. Parameter $k_i > 0$ represents the productive efficiency of operator i, i.e., operator i is more efficient than operator j if $k_i < k_j$. Without any loss of generality, I assume that $k_A \leqslant k_B$.

The entrants must purchase access to the network from the incumbent, paying price a per user. Usually this price is assumed to be the same for all the entrants, but I do not impose such a condition here. Hence I have a_A and a_B, which are the respective access prices that operator A and operator B pay to the incumbent I. Therefore the profits of the operators are

$$\pi_I = p_I d_I - k_I d_I + a_A d_A + a_B d_B - f, \tag{2.2a}$$

$$\pi_A = (p_A - a_A) d_A - k_A d_A, \tag{2.2b}$$

$$\pi_B = (p_B - a_B) d_B - k_B d_B. \tag{2.2c}$$

The market demands. When computing the indifferent users, I assume for the moment that both adjoining networks on each segment have customers. In what follows, I will provide analytical conditions for this assumption to hold. Hence, indifferent users are

$$c_{AI}(a_A, a_B) = \frac{5}{6} + \frac{p_I - p_A}{2t}, \tag{2.3a}$$

$$c_{BI}(a_A, a_B) = \frac{1}{6} + \frac{p_B - p_I}{2t}, \tag{2.3b}$$

$$c_{AB}(a_A, a_B) = \frac{1}{2} + \frac{p_A - p_B}{2t}. \tag{2.3c}$$

The market demand of an operator, d_i, is given by the types of users who prefer the service it provides

$$d_I = \frac{1}{3} + \frac{p_A + p_B}{2t} - \frac{p_I}{t}, \tag{2.4a}$$

$$d_A = \frac{1}{3} + \frac{p_B + p_I}{2t} - \frac{p_A}{t}, \tag{2.4b}$$

$$d_B = \frac{1}{3} + \frac{p_A + p_I}{2t} - \frac{p_B}{t}. \tag{2.4c}$$

The industry equilibrium. Here I provide the industry equilibrium for any given access price a.[10] The equilibrium demands are

$$d_I^* = \frac{1}{3} + \frac{k_A + k_B - 2k_I}{5t}, \tag{2.5a}$$

$$d_A^* = \frac{1}{3} + \frac{-2k_A + k_B + k_I}{5t} - \frac{3a_A - 3a_B}{10t}, \tag{2.5b}$$

$$d_B^* = \frac{1}{3} + \frac{k_A - 2k_B + k_I}{5t} - \frac{3a_B - 3a_A}{10t}. \tag{2.5c}$$

Notice that the demands are linear functions of access prices. The equilibrium prices are

$$p_I^* = \frac{2k_A + 2k_B + 6k_I + 5a_A + 5a_B}{10} + \frac{t}{3}, \tag{2.6a}$$

$$p_A^* = \frac{6k_A + 2k_B + 2k_I + 7a_A + 3a_B}{10} + \frac{t}{3}, \tag{2.6b}$$

$$p_B^* = \frac{2k_A + 6k_B + 2k_I + 3a_A + 7a_B}{10} + \frac{t}{3}. \tag{2.6c}$$

[10] Intermediate steps can be found in Appendix 2.A.

By equation (2.6), it is easy to see that in equilibrium, the retail prices are also linear in the access prices. Hence, the profits in equilibrium are

$$\begin{aligned}\pi_I^* =& \frac{45a_A(5t-3k_A+3k_B)+45a_B(5t+3k_A-3k_B)+270a_Aa_B}{450t}+ \\ &+\frac{2(5t+3k_A+3k_B-6k_I)^2-135(a_A^2+a_B^2)}{450t}-f,\end{aligned} \tag{2.7a}$$

$$\pi_A^* = \frac{(10t-12k_A+6k_B+6k_I-9a_A+9a_B)^2}{900t}, \tag{2.7b}$$

$$\pi_B^* = \frac{(10t+6k_A-12k_B+6k_I+9a_A-9a_B)^2}{900t}. \tag{2.7c}$$

Existence of the equilibrium. While the second order conditions for a maximum are readily satisfied (see Appendix 2.A for further details), it remains to be seen whether the equilibrium is indeed an interior solution, i.e., if, given the equilibrium prices, there exists an indifferent consumer in each segment of the market. As entrant A is more efficient than B, the existence of c_{BI} (the consumer indifferent between B and I) also implies the existence of c_{AI}.

Lemma 2.1. *For any set of access prices (a_A, a_B), the industry has an interior equilibrium with three operators if the transportation cost is sufficiently large*

$$t > \tau(a_A,a_B) \equiv \max\left\{\frac{3}{5}|a_A-a_B+2k_A-2k_I|, \frac{3}{5}|a_A-a_B-2k_B+2k_I|, \frac{6}{5}|a_A-a_B+k_A-k_B|\right\}$$

Proof. See Appendix 2.A. □

2.4 Welfare Analysis

Total welfare in this economy is given by the weighted sum of consumer surplus plus industry profits. Hence I can define

$$W = \lambda CS + \Pi \tag{2.8}$$

where CS is the consumer surplus and $\Pi = \pi_I + \pi_A + \pi_B$. Parameter λ is a nonnegative number that denotes the importance that the regulator assigns to consumers.

Consumer surplus. The consumer surplus is the value that users perceive when they purchase a unit of telecommunications service and pay a price p_i. The total gross utility of purchasing the service, V, is:[11]

$$V = v_0 - p_I d_I - p_A d_A - p_B d_B. \tag{2.9}$$

Because the type of service may differ from the user's most preferred one, the utility of obtaining the service is reduced by the "cost" in terms of disutility of obtaining a service other than the most preferred one: $t|x - x_i|$. The total amount of this disutility is

$$\begin{aligned} T(a_A, a_B) =& \frac{t}{2}\left((1 - c^*_{AI}(a_A, a_B))^2 + (c^*_{BI}(a_A, a_B))^2\right) + \\ &+ \frac{t}{2}\left(\left(\frac{1}{3} - c^*_{BI}(a_A, a_B)\right)^2 + \left(c^*_{AB}(a_A, a_B) - \frac{1}{3}\right)^2\right) + \\ &+ \frac{t}{2}\left(\left(\frac{2}{3} - c^*_{AB}(a_A, a_B)\right)^2 + \left(c^*_{AI}(a_A, a_B) - \frac{2}{3}\right)^2\right) \end{aligned} \tag{2.10}$$

where $c^*_{ij}(a_A, a_B)$ are the indifferent users, defined by equations (2.3) in the equilibrium. By considering equations (2.9) and (2.10), the consumers' surplus is

$$CS = v_0 - p^*_I d^*_I - p^*_A d^*_A - p^*_B d^*_B - T(a_A, a_B). \tag{2.11}$$

Industry profit. By considering that the sum of the entrants' access costs is equal to the incumbent's revenue from the entrants, and by rearranging terms, the

[11] The total gross utility is the sum of the utilities the users get by purchasing telecommunications service from each operator

$$V = v_I \cdot d_I + v_A \cdot d_A + v_B \cdot d_B$$

where d_i are the market demands as defined by equations (2.5). Recalling that $v_i = v_0 - p_i$, I can rewrite the previous expression as

$$V = v_0(d_I + d_A + d_B) - p_I d_I - p_A d_A - p_B d_B.$$

Noting that $d_I + d_A + d_B = 1$, I get the above expression.

total profits are

$$\Pi = p_I^* d_I^* - k_I d_I^* - f + p_A^* d_A^* - k_A d_A^* + p_B^* d_B^* - k_B d_B^* \tag{2.12}$$

Welfare. Total welfare is given by the sum of equations (2.11) and (2.12):

$$W = \lambda(v_0 - T(a_A, a_B)) + (1-\lambda)p_I^* d_I^* + (1-\lambda)p_A^* d_A^* + (1-\lambda)p_B^* d_B^* - k_I d_I^* - k_A d_A^* - k_B d_B^* - f.$$

It depends on the access prices through their effects on the industry equilibrium prices and demands. While the convexity of $T(a_A, a_B)$ is indefinite, the incumbent's revenue, $p_I^* d_I^*$, and the industry costs ($k_I d_I^*$, $k_A d_A^*$ and $k_B d_B^*$) are linear functions in the access prices. The convexity of the entrants' revenue, $p_A^* d_A^* + p_B^* d_B^*$, is also indefinite.[12] Therefore, in an unconstrained welfare maximization problem, the classification of the stationary points has to be done by examining the high order terms (or by performing local exploration).

2.5 Optimal access prices

In this section I find the regulator's optimal choice of access prices. This choice is made by solving the welfare maximization problem while taking into account the possible constraints. However, before proceeding with this, it is useful to consider a more traditional problem.

2.5.1 A benchmark: non-discrimination between entrants

In this section I consider what the regulated access price would be if the regulator did discriminate between the entrants, that is, if it imposes the same access price, a, on both entrants regardless of their respective efficiencies. This approach differs from that in the existing literature because it assumes that all entrants are equal. However, here I want to show what the results could be if the regulator were to impose unique access prices for all entrants without discriminating among them with respect to their relative efficiency levels.

[12]See Appendix 2.A for further details.

In this case, an examination of the market equilibrium makes it clear that market shares do not depend on the access price:

$$d_I^{ND} = \frac{1}{3} + \frac{k_A + k_B - 2k_I}{5t},$$
$$d_A^{ND} = \frac{1}{3} + \frac{-2k_A + k_B + k_I}{5t},$$
$$d_B^{ND} = \frac{1}{3} + \frac{k_A - 2k_B + k_I}{5t},$$

but only on the relative efficiency of the operator, as in de Bijl & Peitz (2002).[13] Hence the access price is directly charged to final users leaving competition among the operators unaffected. The profits of the operators given by equation (2.7) become

$$\pi_I^{ND} = a - f + \frac{2(5t + 3k_A + 3k_B - 6k_I)^2}{450t},$$
$$\pi_A^{ND} = \frac{2(5t - 6k_A + 3k_B + 3k_I)^2}{450t},$$
$$\pi_B^{ND} = \frac{2(5t + 3k_A - 6k_B + 3k_I)^2}{450t}.$$

If the market is unregulated, the incumbent sets the highest possible access price to obtain higher profits. The entrants' profits remain unchanged. A regulator that assigns the same weight to the consumer surplus and industry profit would be indifferent with respect to the choice of a and would accept that chosen by the incumbent.[14] If the regulator cares more about consumers, the optimal access

[13] For further details see Appendix 2.B.

[14] The industry profit is

$$\Pi^{ND} = a - f + \frac{6(k_A^2 - k_A k_B + k_B^2 - (k_A + k_B)k_I + k_I^2)}{25t} + \frac{t}{3}$$

and the consumer surplus is

$$CS^{ND} = v_0 - a + \frac{24(k_A^2 - k_A k_B + k_B^2 - (k_A + k_B)k_I + k_I^2) - 100(k_A + k_B + k_I)t - 125t^2}{300t}.$$

price is that one which makes the incumbent break even, as it is a positive value

$$a^{ND} = \max\left\{0, f - \frac{2(5t + 3k_A + 3k_B - 6k_I)^2}{450t}\right\}.$$

Note that in this case the regulator is unable to minimize either the industry costs or the consumers' transportation costs.

2.5.2 Constraints

When the regulator chooses the values of a_A and a_B, it has to consider that operators cannot incur negative profits. In particular, given the assumption of positive demands for all firms in each segment of the market, the only operator that could incur negative profits is the incumbent due to the fixed cost f to provide access to each user. Given the market equilibrium, the profits of the incumbent are positive when:

$$\frac{45a_A(5t - 3k_A + 3k_B) + 45a_B(5t + 3k_A - 3k_B) + 270a_Aa_B}{450t} + \\ + \frac{2(5t + 3k_A + 3k_B - 6k_I)^2 - 135(a_A^2 + a_B^2)}{450t} - f \geqslant 0.$$

Note that the incumbent's profits are only an increasing function of access prices when they assume values that are relatively close to each other. Otherwise, when access prices are different enough, the incumbent faces negative profits. The reason for this is in the sensitivity of the total revenue that incumbent collects from the sale of access to the entrants $(a_Ad_A^* + a_Bd_B^*)$ to the differences between the access prices $(a_B - a_B)$. While it is clear that the retail revenue of the incumbent is always increasing in access prices, it can be shown that the incumbent's wholesale revenue is only an increasing function of access prices when the difference between the two lies within a bounded interval, i.e.,

$$\frac{2k_A - k_B - k_I}{3} - \frac{5}{9}t \leqslant a_B - a_A \leqslant \frac{k_A - 2k_B + k_I}{3} + \frac{5}{9}t.$$

Moreover, if the competition in the market is fierce enough $(t \geqslant 3(k_A+k_B-2k_I)/10)$ such an interval does not exist, and wholesale revenue decreases with increasing access prices. These effects can be observed in Figure 2.2, where the incumbent's profits are represented in the plane a_A, a_B. The shaded area represents the pairs of access prices that allow the incumbent to obtain positive profits given the market parameters. As access prices increase, retail revenue also increases and this can compensate for greater divergence in access prices (this is the reason that the positive profits are spreading).

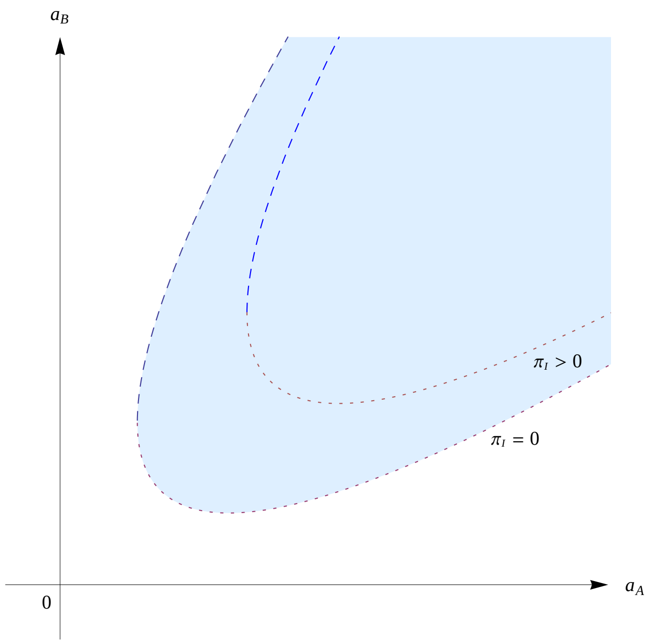


Figure 2.2: Incumbent's iso-profits for $f = 1$, $k_A = 0.15$, $k_B = 0.25$, $k_I = 0.2$, and $t = 1$.

Hence, I can define the region of feasible values for a_A and a_B (that is, the values that allow the incumbent to obtain nonnegative profits) through the pairs of access prices that satisfy the following restrictions:

$$a_B \geqslant a_A + \frac{75t + 45k_A - 45k_B - \sqrt{15\phi(a_A)}}{90}, \qquad (\check{\eta})$$

$$a_B \leqslant a_A + \frac{75t + 45k_A - 45k_B + \sqrt{15\phi(a_A)}}{90}. \qquad (\hat{\eta})$$

where[15]

$$\phi(a_A) = 1800ta_A + 207k_A^2 - 126k_Ak_B + 207k_B^2 - 288k_Ak_I - 288k_Bk_I + \\ + 288k_I^2 + 690k_At - 210k_Bt - 480k_It - 1800ft + 575t^2.$$

In particular, when the constraint ($\check{\eta}$) holds as an equality, the zero isoprofit curve is the dotted curve in the figure. When the constraint ($\hat{\eta}$) holds as an equality it is represented by the dashed curve. It is worth noting that the zero profits curve moves away from the origin as fixed cost increases: more inefficient incumbents require higher levels of access prices to cover network maintenance.

Finally, remember the assumption that an indifferent user always exists in each segment of the triangular Hotelling market. It can be shown that under the access prices that make the constraint ($\hat{\eta}$ binding, such an indifferent consumer would not exist in the market segment between entrant A and entrant B: operator A would always be able to exclude operator B from the market (indeed, operator A always finds this profitable), and therefore, an equilibrium where ($\hat{\eta}$) is binding cannot exist.[16]

2.5.3 The maximization problem

When the regulator has to set the access prices a_A and a_B, its objective is to maximize social welfare defined by equation (2.8). The restriction to its choice is given by constraint ($\check{\eta}$). Therefore, the problem that the regulator has to solve is

$$\max_{a_A, a_B} W(a_A, a_B) \quad \text{subject to} \, (\check{\eta}). \tag{2.13}$$

[15]Notice that incumbent's constraints ($\check{\eta}$) and ($\hat{\eta}$) depend both on term $\phi(a_A)$ and, in particular, that they could only both be binding simultaneously when $\phi(a_A) = 0$. This occurs when

$$a_A = f + \frac{9(-23k_A^2 + 14k_Ak_B - 23k_B^2 + 32(k_A + k_B)k_I - 32k_I^2)}{1800t} + \\ + \frac{30(-23k_A + 7k_B + 16k_I)t - 575t^2}{1800t}.$$

[16]For further details, see Appendix 2.B.

The cases where $0 < \lambda < 1$ are trivial. It can be shown that no real values optimize the welfare function. The optimal access prices tend toward infinity: the reason for this is that it is assumed that the market is fully covered, and the utility of purchasing the service, v_0, is high enough to make it optimal for users to always consume the service, independent of the prices. Therefore, by choosing arbitrarily large access prices, the regulator could increase the profits of the operators while keeping their market shares unchanged. By maintaining the proportionality between the access prices while increasing them to infinity, the result is that profits tend toward infinity as well.

The interesting case is when $\lambda \geq 1$, that is, when the regulator is assigning a higher weight to the consumer side of the market. Here, there is only a solution when $\check{\eta} > 0$, that is, when the break-even constraint of the incumbent is binding. The solution is given by values of the access prices a_A and a_B, which are functions of the marginal and fixed costs, transportation costs and consumer surplus weight. I denote them by $\check{a}_A$ and $\check{a}_B$, respectively.[17] Hence I can state the following proposition.

Proposition 2.1. *In a telecommunications market where entrants have different marginal costs, the optimal access price for the less efficient entrant is always*

[17]Once computed, the access prices turn out to be

$$\check{a}_A = f + \frac{-9(k_A^2 + k_B^2)(72 + \lambda(353\lambda - 642)) - 8(3-4\lambda)^2(6k_I - 5t)^2}{1800(3-4\lambda)^2 t} + \\ + \frac{6k_B(4\lambda - 3)(48k_I(4\lambda - 3) + 5(13\lambda - 66)t) + 6k_A(3k_B((97\lambda - 258)\lambda - 72\lambda))}{1800(3-4\lambda)^2 t} + \\ + \frac{6k_A(4\lambda - 3)(48k_I(4\lambda - 3) + 5(114 - 77\lambda)t)}{1800(4\lambda - 3)^2 t};$$

$$\check{a}_B = f + \frac{-9(k_A^2 + k_B^2)(72 + \lambda(353\lambda - 642)) - 8(3-4\lambda)^2(6k_I - 5t)^2}{1800(3-4\lambda)^2 t} + \\ + \frac{6k_B(4\lambda - 3)(48k_I(4\lambda - 3) + 5(114 - 77\lambda)t) + 6k_A(3k_B((97\lambda - 258)\lambda - 72\lambda))}{1800(3-4\lambda)^2 t} + \\ + \frac{6k_A(4\lambda - 3)(48k_I(4\lambda - 3) + 5(13\lambda - 66)t)}{1800(4\lambda - 3)^2 t}.$$

higher than the optimal access price for the more efficient one. In particular, the optimal access prices satisfy the following condition:

$$\breve{a}_B = \breve{a}_A + \frac{3(k_B - k_A)(2-\lambda)}{2(4\lambda - 3)}.$$

Proof. See Appendix 2.B. □

When I examine whether $\breve{a}_A$ and $\breve{a}_B$ satisfy the assumption about the existence of indifferent consumers (Lemma 2.1), it turns out that they only constitute an equilibrium for this market if the parameters satisfy some restrictions. In particular, $(\breve{a}_A, \breve{a}_B)$ is an equilibrium if

$$t > \begin{cases} -\frac{3}{5}\left(2k_I - 2k_B - \frac{3(k_B-k_A)(2-\lambda)}{8\lambda-6}\right), & \text{if } k_I < \dot{k}_I, \\ \frac{3(k_B-k_A)\lambda}{4\lambda-3}, & \text{if } \dot{k}_I \leq k_I \leq \ddot{k}_I \\ -\frac{3}{5}\left(2k_A - 2k_I - \frac{3(k_B-k_A)(2-\lambda)}{8\lambda-6}\right), & \text{if } k_I > \ddot{k}_I, \end{cases} \tag{2.14}$$

where $\dot{k}_I \equiv \frac{(13k_A+3k_B)\lambda-6(k_A+k_B)}{4(4\lambda-3)}$ and $\ddot{k}_I \equiv \frac{(3k_A+13k_B)\lambda-6(k_A+k_B)}{4(4\lambda-3)}$.[18]

Finally, the constraint being binding means that the optimal a_A and a_B should always be chosen in such a way that the incumbent could not gain positive profits, independent of its efficiency. However, this would imply that, for very low levels of f, the optimal access prices could be negative, meaning a subsidy from the incumbent to the entrants. An example of this is given in Figure 2.3. Here, the set of optimal access price combinations is given by the dashed straight line, while the incumbent's constraint is given by the curve. In the graphic, the pair of optimal access prices is the intersection of the incumbent's constraint and the set of optima. Given the parameters, the optimum a_A (given by point E in the picture) ends up being negative.

This would not make any sense in the real world, and I therefore exclude this outcome by imposing a restriction that the access prices have to be non-negative. Hence, a value of the per-user fixed cost $\bar{f}$ can be defined that is considered to be the minimum level of efficiency required on the part of the incumbent to make

[18] For the proof of this statement, see Appendix 2.B.

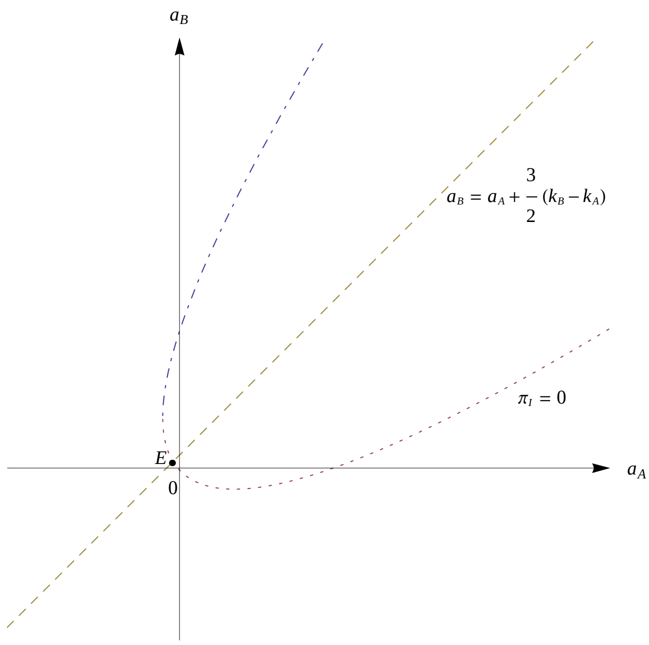


Figure 2.3: Incumbent's profits for $f = 0.1$, $k_A = 0.15$, $k_B = 0.25$, $k_I = 0.2$, $t = 1$, and $\lambda = 1.5$.

it to obtain positive profits.[19] Hence for values of f greater than $\bar{f}$, the solutions are $\check{a}_A$ and $\check{a}_B$, and the incumbent is breaking even. For values of f lower than $\bar{f}$, the incumbent gains positive profits, while the access prices are $a_A = 0$ and $a_B = 3(k_B - k_A)/2$.

2.5.4 Results

In the previous section I have fully characterized the equilibrium optimal access prices. Here I am going to derive some additional results. First, I return to the solution to the regulator's maximization problem in a case where the constraint is binding.

[19]Such value is

$$\bar{f} = \frac{-1953(k_A^2 + k_B^2) + 6k_A(699k_B - 48k_I - 185t) + 8(6k_I - 5t)^2 + 6k_B(-48k_I^2 + 265t)}{1800t}.$$

Proposition 2.2. *When a regulator wants to maximize social welfare in a telecommunications market, the optimal access prices are the following:*

$$(a_A^*, a_B^*) = \begin{cases} (\breve{a}_A, \breve{a}_B), & \text{if } f > \bar{f}, \\ (0, \frac{3}{2}(k_B - k_A)), & \text{otherwise} \end{cases} \tag{2.15}$$

and they constitute an equilibrium when one of the following conditions holds:

$$t > -\frac{3}{5}\left(2k_I - 2k_B - \frac{3(k_B - k_A)(2-\lambda)}{8\lambda - 6}\right), \qquad \text{if } k_I < \dot{k}_I, \tag{2.16a}$$

$$t > \frac{3(k_B - k_A)\lambda}{4\lambda - 3}, \qquad \text{if } \dot{k}_I \leq k_I \leq \ddot{k}_I \tag{2.16b}$$

$$t > -\frac{3}{5}\left(2k_A - 2k_I - \frac{3(k_B - k_A)(2-\lambda)}{8\lambda - 6}\right), \qquad \text{if } k_I > \ddot{k}_I. \tag{2.16c}$$

These access prices always imply a higher welfare than a case when a common access price for two different entrants is considered, and therefore, the regulator cannot ignore possible asymmetries of the entrants. In Figure 2.4, I offer two examples that show how social welfare is higher under discrimination than when it the regulator does not discriminate.

Conditions (2.16a) to (2.16c) assure that it is not in the interest of the regulator to have a telecommunications market with only two operators, even if there is a possibility of causing the less efficient one to exit: allowing a reduction in the industry costs by concentrating the production in only the most efficient operators does not compensate for the lost of variety for the end-users or the increased final prices due to decreased competition.
In particular, condition (2.16a) shows how a very efficient incumbent does not force entrant B out of the market if there is limited substitutability among the offered services. A higher degree of competition is admitted if the incumbent's efficiency is in line with that of one of the other producers: condition (2.16b) guarantees that entrant A does not force entrant B. Finally, if the incumbent is very inefficient, it is not forced out by entrant A if the transportation cost is sufficiently high (condition (2.16c)).

Finally, I want to consider how the minimum level of efficiency required for the incumbent to obtain positive profits, $\bar{f}$, is affected by the parameters of the

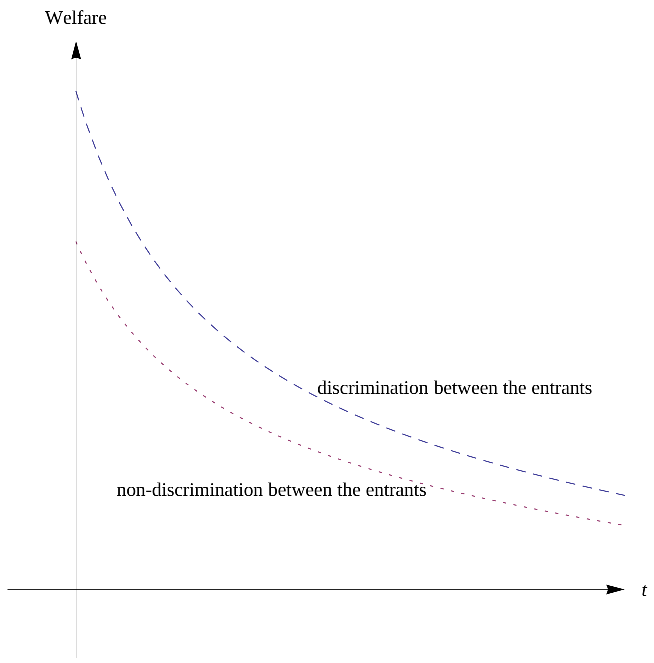


Figure 2.4: Welfare comparison with discriminated (dashed line) vs. non-discriminated access prices (dotted line).

model. In particular, I define $\hat{k}_{AB} \equiv k_B - k_A > 0$ and $\hat{k}_{AI} \equiv k_A - k_I$ as the difference between the marginal costs of the entrants and the difference between the marginal costs of the more efficient entrant and the incumbent, respectively. I want to explore how $\bar{f}$ is affected by the level of asymmetry between the entrants. I have

$$\frac{\partial \bar{f}}{\partial \hat{k}_{AB}} \leqslant 0 \quad \text{if} \quad \hat{k}_{AB} \leqslant \frac{48\hat{k}_{AI} + 10\sqrt{648\hat{k}_{AI}^2 - 434t^2}}{651};$$
$$\frac{\partial \bar{f}}{\partial \hat{k}_{AB}} > 0 \quad \text{otherwise.}$$

Hence an incumbent that is inefficient will still obtain positive profits in a regulated market if the degree of asymmetry between the entrants is sufficiently high. Paradoxically, if the entrants have the same technologies and the incumbent has a more efficient technology when serving the end-user, the regulator requires a higher efficiency in the maintenance of the network. This is because higher access prices are reflected in higher retail prices, which harm consumers: the regulator wants to keep the final prices as low as possible to maximize social welfare.

2.6 Asymmetric Information

In this section, I consider how asymmetric information affects the the regulator's choice and the market equilibrium. In a telecommunication market, the operators may hold private information. In particular, to set the optimal access prices, the regulator needs to know: the marginal costs, the fixed cost to connect each user to the network, and the degree of competition in the retail market. It is possible to imagine different scenarios where one or more of these parameters is unknown. In these cases, the regulator needs to design specific contracts to avoid operators' opportunistic behaviors that could harm social welfare. However, in some cases the regulator is unable to discriminate between different types of operators (mainly, when the regulator does not know the fixed or the marginal costs). In other cases (namely, when the degree of competition in the market is unknown), the regulator can separate different types of operators.

2.6.1 Non-discriminating operators

If the fixed cost is unknown to the regulator, the incumbent could take advantage by declaring a higher f than the true one.[20] However it is first worth noting that an information problem only arises when the incumbent has a sufficiently high fixed cost ($f > \bar{f}$).[21] Hence, suppose that the incumbent could be one of two types: one associated with a low fixed cost, f^L, and another less efficient type, f^H, where $\bar{f} < f^L < f^H$. In a symmetric case regulation, would be constituted by

[20]Here I only provide an intuition about the reason that it is impossible to discriminate among incumbents with different fixed costs. The reasoning for other possible information problems in the costs is analogous.

[21]For any $f \leqslant \bar{f}$, the access prices of the incumbent will be always the same and, therefore, it is not meaningful to declare a fixed cost other than the true one.

two different pairs of access prices where $\check{a}_A(f^L) < \check{a}_A(f^H)$ and $\check{a}_B(f^L) < \check{a}_B(f^H)$. Both types would obtain zero profits. I have an example of this in Figure 2.5,

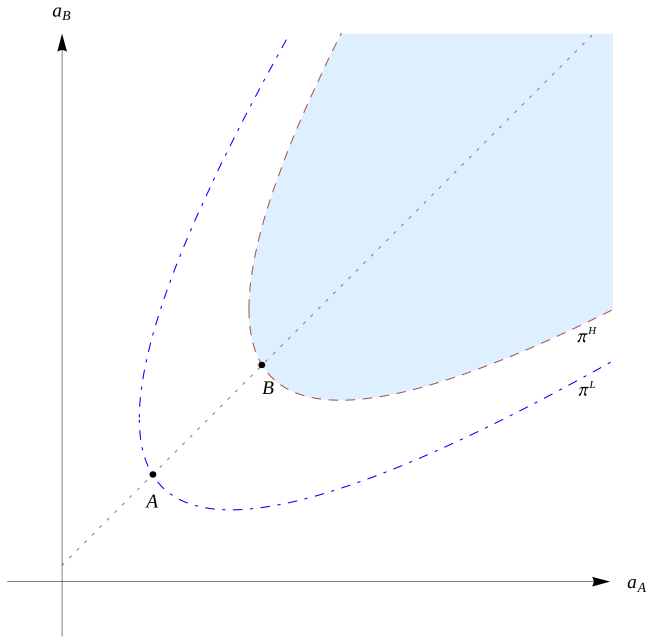


Figure 2.5: Incomplete information about fixed cost.

where the zero profit condition for type L is given by the dot-dashed line and that for type H by the dashed line. The optimal contract for L lies at point A and the optimal contract for H at point B.

In a case where the fixed cost is the private information of the incumbent, I denote by a_i^j (where $j = L, H$) the access prices paid to incumbent type j by entrant i. It can be shown that, if the access prices for the efficient type are lower than the access prices for the inefficient type, then $\pi^L(a_A^L, a_B^L) < \pi^L(a_A^H, a_B^H)$: the more efficient incumbent always has an incentive to misrepresent its type.[22] More particularly, every contract (a_A^H, a_B^H) that makes type H at least break even (contracts in the shaded area) is always preferred by type L to any other contract that does not. Because each feasible contract for H makes L better off than any other contract that is feasible for L but not for H, then it is not possible to offer separate contracts

[22]In other words, the iso profit functions of the two types never intersect (see Appendix 2.C).

to the two types of incumbents, and therefore, the regulator cannot discriminate between them.

2.6.2 Discriminating operators

If the transportation cost is unknown to the regulator but is known to the operators, the only firm that could profitably declare a false value is the incumbent.[23] In particular, denoting by t^L and t^H, with $t^L < t^H$, two different transportation costs, an incumbent of type H could gain higher profits by misrepresenting its type and declaring that it is type L. In Figure 2.6, the complete information

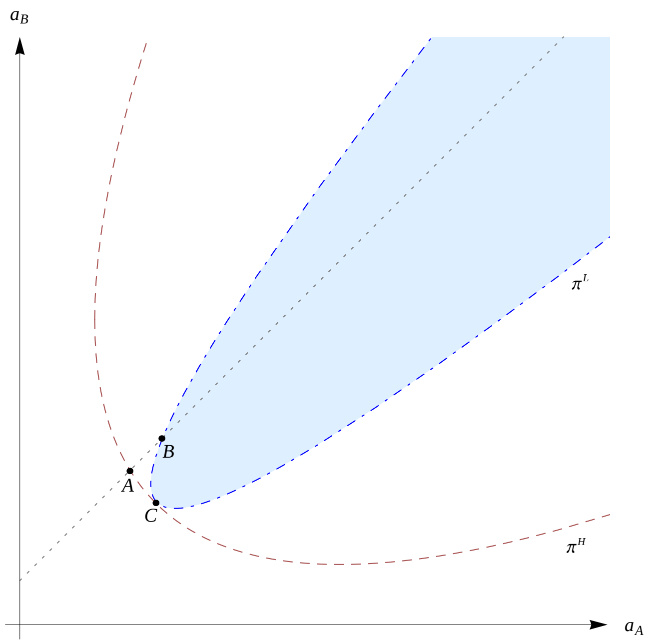


Figure 2.6: Incomplete information about transportation cost.

contract for H is A and for L is B. However, by choosing contract B, incumbent type H would obtain positive profits. The shaded region represents all of the

[23]The market shares and the profits of the entrants are always the same under different transportation costs since they depend only on the difference between the two access prices and such difference is invariant with respect to the transportation cost (for further details see Appendix 2.C).

contracts that satisfy a participation constraint of type L and, therefore, where a contract for L should lie. Note that almost all of these contracts make type H strictly better off. The only contracts that simultaneously satisfy the incentive compatibility constraints of both types are $(\check{a}_A^L, \check{a}_B^L)$ and $a_A^H = a_B^H = a^*$, where a^* is the optimal access price in a case where the regulator does not discriminate between the entrants, as is shown above (point C in the figure). Under these contracts, in a case where the transportation cost is H, the incumbent still earns zero profits, as the case under complete information, while the market share (and the profits) of the less efficient entrant increase to the detriment of the more efficient one.

2.7 Final remarks

I have modeled a telecommunications market with unit demand and asymmetric oligopoly and studied how the regulator should act in this type of market. The unit demand can be considered an important feature of particular markets (broadband connections, fixed line calls) derived from the usual practice of charging flat tariffs to the end-users. I found that the main implication of this assumption is that the market share of the incumbent cannot be affected by access price regulation. Moreover, the access price is directly transmitted to the retail price, leaving the profits of the entrants unchanged.

I also introduced asymmetries between the entrants: I considered an oligopolistic market where operators may differ in their marginal costs. In this context, a unique access price is not the best instrument because it would not correct for the eventual inefficiencies of an entrant. Hence I found the optimal access price for each type of entrant and gave the conditions under which these prices constitute an equilibrium in the market. The implication for a regulator that cares about consumer surplus is to set the access prices at exactly the level that makes the incumbent break even.

Finally, I observed that many of the information problems in this context cannot be resolved by the regulator: if operators have private, unobservable information about the structure of their costs, the regulator cannot prevent their using this information to obtain more favorable access prices, and it cannot design any contract that makes them reveal their true costs. In particular, if the incumbent is more efficient than the regulator believes it is, this efficiency would only be reflected in higher profits for the incumbent, while consumers would end up paying the same prices they paid when the incumbent was inefficient. The only case of asymmetric information that the regulator could resolve is a case where the degree of competition in the market is unknown to it: an appropriate choice of contracts by the regulator can make it possible to discriminate between different incumbents and achieve a second-best solution.

References

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

Armstrong, M. (2002). *The Theory of Access Pricing and Interconnection*, volume 1, chapter 8, (pp. 295–386). North-Holland.

Armstrong, M., Doyle, C., & Vickers, J. (1996). The Access Pricing Problem: A Synthesis. *The Journal of Industrial Economics*, *44*(2), 131–150.

Armstrong, M. & Sappington, D. E. (2006). Regulation, Competition, and Liberalization. *Journal of Economic Literature*, *XLIV*, 325–366.

CMT (2009). Informe Anual de la Comisión del Mercado de las Telecomunicaciones [in Spanish].

de Bijl, P. W. & Peitz, M. (2002). *Regulation and Entry into Telecommunications Markets*. Cambridge University Press.

de Bijl, P. W. & Peitz, M. (2004). Dynamic regulation and entry in telecommunications markets: a policy framework. *Information Economics and Policy*, *16*, 411–437.

Kridel, D. J., Sappington, D. E., & Weisman, D. L. (1996). The effects of incentive regulation in the telecommunications industry: A survey. *Journal of Regulatory Economics*, *9*(3), 269–306.

Laffont, J.-J., Rey, P., & Tirole, J. (1998). Network Competition: I. Overview and Nondiscriminatory Pricing. *The RAND Journal of Economics*, *29*(1), 1–37.

Laffont, J.-J. & Tirole, J. (2000). *Competition in telecommunications*. MIT Press, Cambridge.

Ofcom (2010). Annual Report and Accounts. Technical report.

Peitz, M. (2005). Asymmetric access price regulation in telecommunications markets. *European Economic Review*, *49*, 341–358.

Valletti, T. (1998). Two-part access pricing and imperfect competition. *Information Economics and Policy*, *10*, 305–323.

Vickers, J. (1997). Regulation, Competition, and the Structure of Prices. *Oxford Review of Economic Policy*, *13*(1), 15–26.

Vogelsang, I. (2003). Price Regulation of Access to Telecommunications Networks. *Journal of Economic Literature*, *XLI*, 830–862.

Waverman, L. & Sirel, E. (1997). European Telecommunications Markets on the Verge of Full Liberalization. *Journal of Economic Perspectives*, *11*(4), 113–126.

Appendix 2.A Equilibrium and Welfare

Equilibrium

The first order conditions for each operator are:

$$\frac{\partial \pi_I}{\partial p_I} = 0 \iff \frac{1}{3} + \frac{p_A + p_B}{2t} - \frac{2p_I}{t} + \frac{a_A + a_B + 2k_I}{2t} = 0 \tag{2.17a}$$
$$\frac{\partial \pi_A}{\partial p_A} = 0 \iff \frac{1}{3} + \frac{p_B + p_I}{2t} - \frac{2p_A}{t} + \frac{a_A + k_A}{t} = 0 \tag{2.17b}$$
$$\frac{\partial \pi_B}{\partial p_B} = 0 \iff \frac{1}{3} + \frac{p_A + p_I}{2t} - \frac{2p_B}{t} + \frac{a_B + k_B}{t} = 0 \tag{2.17c}$$

The second order conditions are trivial: $\frac{\partial^2 \pi_I}{\partial p_I^2} = \frac{\partial^2 \pi_A}{\partial p_A^2} = \frac{\partial^2 \pi_B}{\partial p_B^2} = -\frac{2}{t} < 0$ and the Hessian is negative definite ($|H| = -\frac{8}{t^3} < 0$). So the reaction functions are:

$$p_I^R = \frac{p_A + p_B}{4} + \frac{k_I}{2} + \frac{a_A + a_B}{4} + \frac{t}{6}, \tag{2.18a}$$
$$p_A^R = \frac{p_B + p_I}{4} + \frac{a_A + k_A}{2} + \frac{t}{6}, \tag{2.18b}$$
$$p_B^R = \frac{p_A + p_I}{4} + \frac{a_B + k_B}{2} + \frac{t}{6}. \tag{2.18c}$$

Indifferent users

In equilibrium the indifferent users are:

$$c_{AI}^* = \frac{5}{6} + \frac{-a_A + a_B - 2k_A + 2k_I}{10t} \tag{2.19a}$$
$$c_{BI}^* = \frac{1}{6} + \frac{-a_A + a_B + 2k_B - 2k_I}{10t} \tag{2.19b}$$
$$c_{AB}^* = \frac{1}{2} + \frac{a_A - a_B + k_A - k_B}{5t} \tag{2.19c}$$

Notice that the positions of the indifferent users are linear functions of the access prices. The indifferent users in each segment of the market exist if

$$\frac{2}{3} < c^*_{AI} < 1 \Longleftrightarrow t > \left|\frac{3(a_A - a_B + 2k_A - 2k_I)}{5}\right| \tag{2.20a}$$

$$0 < c^*_{BI} < \frac{1}{3} \Longleftrightarrow t > \left|\frac{3(a_A - a_B - 2k_B + 2k_I)}{5}\right| \tag{2.20b}$$

$$\frac{1}{3} < c^*_{AB} < \frac{2}{3} \Longleftrightarrow t > \left|\frac{6(a_A - a_B + k_A - k_B)}{5}\right| \tag{2.20c}$$

Welfare analysis

The equilibrium demands, d^*_I, d^*_A, and d^*_B, are linear functions of the access prices (as already stated in the text). By observing equations (2.5) and (2.6), it is immediate to see that also $p^*_I d^*_I$ is a linear function in the access prices.

When looking at the entrants' revenue, after simplifications, it holds:

$$p^*_A d^*_A + p^*_B d^*_B = \frac{(9(a_B - a_A) - 12k_A + 6(k_B + k_I) + 10t)(21a_A + 9a_B + 6(3k_A + k_B + k_I) + 10t)}{900t} + \\ + \frac{(9(a_A - a_B) + 6(k_A - 2k_B + k_I) + 10t)(9a_A + 21a_B + 6(k_A + 3k_B + k_I) + 10t)}{900t}$$

The Hessian is null and therefore the second order derivatives test is inconclusive: higher order terms may have to be examined, or local exploration can be performed.

Similarly, it can be shown that also the transportation cost is an indefinite function in the access prices.

Appendix 2.B Optimal access prices

Non-discrimination between entrants

In this case $a_A = a_B \equiv a$ so, given the profit functions in equations (2.2), the reaction functions in terms of prices are:

$$\frac{\partial \pi_I}{\partial p_I} = 0 \Longleftrightarrow p_I = \frac{t}{6} + \frac{2a + 2k_I + p_A + p_B}{4}$$
$$\frac{\partial \pi_A}{\partial p_A} = 0 \Longleftrightarrow p_A = \frac{t}{6} + \frac{2a + 2k_A + p_B + p_I}{4}$$
$$\frac{\partial \pi_B}{\partial p_B} = 0 \Longleftrightarrow p_B = \frac{t}{6} + \frac{2a + 2k_B + p_A + p_I}{4}$$

So the equilibrium prices are:

$$p_I^{ND} = \frac{k_A + k_B + 3k_I}{5} + \frac{t}{3} + a,$$
$$p_A^{ND} = \frac{3k_A + k_B + k_I}{5} + \frac{t}{3} + a,$$
$$p_B^{ND} = \frac{k_A + 3k_B + k_I}{5} + \frac{t}{3} + a.$$

The indifferent users are:

$$c_{AI}^{ND} = \frac{5}{6} + \frac{k_I - k_A}{5t}, \quad c_{BI}^{ND} = \frac{1}{6} + \frac{k_B - k_I}{5t}, \quad c_{AB}^{ND} = \frac{1}{2} + \frac{k_A - k_B}{5t}.$$

and they always exist when:

$$t > \frac{6}{5} \max\{|k_I - k_A|, |k_I - k_B|, |k_A - k_B|\},$$

which is the same condition required in Lemma 2.1.

Constraints

The indifferent user between entrant A and entrant B does not exist if $c_{AB} \leq \frac{1}{3}$ (since A is assumed to be more efficient than B, if the indifferent consumer does not exist, it is because A is covering the whole market of B). This occurs when:

$$\frac{1}{2} + \frac{a_A - a_B + k_A - k_B}{5t} \leq \frac{1}{3} \Longleftrightarrow a_B - a_A \geq k_A - k_B + \frac{5}{6}t$$

If constraint $(\hat{\eta})$ is binding, then

$$a_B - a_A = \frac{75t + 45k_A - 45k_B + \sqrt{15\phi(a_A)}}{90}$$

and

$$\frac{75t + 45k_A - 45k_B + \sqrt{15\phi(a_A)}}{90} > k_A - k_B + \frac{5}{6}t \Longrightarrow \frac{\sqrt{15\phi(a_A)}}{45} > k_A - k_B$$

which is always true since $k_A \leq k_B$: i.e., the indifferent consumer would never exist.

Welfare maximization

Setting μ as the Lagrangian multiplier of the constraint ($L(a_A, a_B) = W(a_A, a_B) + \mu \cdot \check{\eta}(a_A, a_B)$), the first-order conditions for a maximum require:

$$\frac{\partial L(a_A, a_B)}{\partial a_A} = 0 \Longleftrightarrow \mu \cdot \varphi(a_A, a_B) + \frac{1 - \lambda - 2\mu}{2} + \\ - \frac{3\left(2(\lambda - 2)(a_B - a_A) + (2\lambda + 1)\left(k_B - k_A\right)\right)}{50t} = 0 \qquad (2.21)$$

$$\frac{\partial L(a_A, a_B)}{\partial a_B} = 0 \Longleftrightarrow \frac{6(\lambda - 2)(a_B - a_A) + 3(2\lambda + 1)(k_B - k_A) - 25(\lambda - 2\mu - 1)t}{50t} = 0 \qquad (2.22)$$

where $\varphi(a_A, a_B) \equiv \sqrt{\frac{1500t^2}{3\left(600ta_A - 32k_I(3k_A + 3k_B + 5t) - 42k_A k_B + 230tk_A + 69k_A^2 - 70tk_B + 69k_B^2 + 96k_I^2\right) + 25t(23t - 72f)}}$.

Notice that, if the constraint is not binding (i.e., $\mu = 0$), the conditions become:

$$a_B - a_A = \frac{25(1 - \lambda)t - 3(2\lambda + 1)(k_B - k_A)}{6(\lambda - 2)} \equiv \Delta_1 \qquad (2.23)$$

$$a_B - a_A = \frac{25(\lambda - 1)t - 3(2\lambda + 1)(k_B - k_A)}{6(\lambda - 2)} \equiv \Delta_2 \qquad (2.24)$$

Since in this case $\lambda \geq 1$ (and since the concavity of the welfare function requires $\lambda < 2$) it is easy to see that $\Delta_1 \geq \Delta_2$ (and $\Delta_1 = \Delta_2$ if $\lambda = 1$). This implies that the constraint is broken if:

$$\Delta_2 < \frac{75t + 45k_A - 45k_B - \sqrt{15\phi(a_A)}}{90}$$

that is, if:

$$\frac{3(3+\lambda)(k_B - k_A) - 5(4\lambda - 3)t}{6(2-\lambda)} - \frac{\sqrt{15\phi(a_A)}}{90} \equiv \delta(\lambda) < 0$$

Notice that:

$$\frac{\partial \delta(\lambda)}{\partial \lambda} = \frac{15(k_B - k_A) - 25t}{6(\lambda - 2)^2} < 0 \Longleftrightarrow t > \frac{3}{5}(k_B - k_A).$$

Under the requirements of Lemma 2.1 (i.e., when the transportation cost is large enough so the equilibrium I am looking for could exist) the above inequality is always satisfied and therefore $\delta(\lambda)$ is decreasing in λ.
Moreover, considering again the requirements of Lemma 2.1:

$$\delta(1) = \frac{4(k_B - k_A) - 5t}{6} - \frac{\sqrt{15\phi(a_A)}}{90}.$$

It can be shown that $\delta(1) < 0$ if $\mu > 0$ and $\delta(1) = 0$ if $\mu = 0$. This means that for $\lambda > 1$ the constraint is always binding but a special case arises when $\lambda = 1$: the constraint is not binding. In the latter case both first-order conditions imply:

$$a_B - a_A = \frac{3}{2}(k_B - k_A)$$

that is, a potential solution of multiple equilibria satisfying the above equation.

The general analytical solution of the system given by the binding constraint (that is, $(\check{\eta})$ satisfied as an equality) and conditions (2.21) and (2.22) provided the pair of optimal access prices given in the text when $\lambda > 1$, i.e., when the regulator assigns a higher weight to the consumers.

The sufficient condition for $(\check{a}_A, \check{a}_B)$ to be a maximum is that the determinant of the bordered Hessian be positive, that is:

$$\frac{4(4\lambda-3)^3(3(3+\lambda)(k_B-k_A)+5(4\lambda-3)t)t}{(3(3+\lambda)(k_A-k_B)+5(4\lambda-3)t)^2((3+\lambda)(k_A-k_B)+5(4\lambda-3)t)} > 0.$$

Simple numerical checks allow me to state that the above condition is always verified within the feasibility region for the equilibrium. Although the solution of the system given by the first-order conditions and the calculation of the determinant of the bordered Hessian are performed as usually, the necessary steps are quite complicated and large so they are not included here but they are available upon request.

Two further remarks can be done on the equilibrium access prices. Firstly, notice that:

$$\phi(\check{a}_A) = \frac{15(5(3-4\lambda)t-3(3+\lambda)(k_B-k_A))^2}{(3-4\lambda)^2}$$

This result will be useful in a while. Secondly, after some calculations it results that:

$$\check{a}_B - \check{a}_A = \frac{3(k_B-k_A)(2-\lambda)}{2(4\lambda-3)}. \tag{2.25}$$

Given that $k_B > k_A$ and that $1 \leq \lambda < 2$, it turns out that, in a welfare maximising equilibrium, $\check{a}_B > \check{a}_A$.

Existence of the equilibrium

The equilibrium access prices need to satisfy the assumption of Lemma 2.1. Given the difference between the optimal access prices in equation (2.25), the thresholds for the transportation cost in Lemma 2.1 become:

$$\tau_1 = \frac{3}{5}\left|2k_A - 2k_I - \frac{3(k_B - k_A)(2-\lambda)}{8\lambda - 6}\right|$$
$$\tau_2 = \frac{3}{5}\left|2k_I - 2k_B - \frac{3(k_B - k_A)(2-\lambda)}{8\lambda - 6}\right|$$
$$\tau_3 = \left|\frac{3(k_B - k_A)\lambda}{4\lambda - 3}\right|$$

Given that $1 \leq \lambda < 2$ and $k_A \leq k_B$, notice that:

$$\frac{3}{5}\left(2k_A - 2k_I - \frac{3(k_B - k_A)(2-\lambda)}{8\lambda - 6}\right) + \frac{3}{5}\left(2k_I - 2k_B - \frac{3(k_B - k_A)(2-\lambda)}{8\lambda - 6}\right) = \frac{3(k_A - k_B)\lambda}{4\lambda - 3} < 0$$

so the relevant thresholds for τ_1 and τ_2 are:

$$\check{\tau}_1 = -\frac{3}{5}\left(2k_A - 2k_I - \frac{3(k_B - k_A)(2-\lambda)}{8\lambda - 6}\right);$$
$$\check{\tau}_2 = -\frac{3}{5}\left(2k_I - 2k_B - \frac{3(k_B - k_A)(2-\lambda)}{8\lambda - 6}\right).$$

Moreover:

$$\check{\tau}_3 = \frac{3(k_B - k_A)\lambda}{4\lambda - 3}.$$

It is easy to see that:

$$\frac{\partial \check{\tau}_1}{\partial k_I} = \frac{6}{5} \quad \text{and} \quad \frac{\partial \check{\tau}_2}{\partial k_I} = -\frac{6}{5} \quad \text{and} \quad \frac{\partial \check{\tau}_1}{\partial k_I} = 0.$$

It can be shown that:

$$\check{\tau}_1 = \check{\tau}_3 \Longleftrightarrow k_I = \ddot{k}_I \equiv \frac{(3k_A + 13k_B)\lambda - 6(k_A + k_B)}{4(4\lambda - 3)};$$
$$\check{\tau}_2 = \check{\tau}_3 \Longleftrightarrow k_I = \dot{k}_I \equiv \frac{(13k_A + 3k_B)\lambda - 6(k_A + k_B)}{4(4\lambda - 3)}.$$

It is immediate to see that $\ddot{k}_I \geq \dot{k}_I$ (and that $\ddot{k}_I = \dot{k}_I$ if $k_A = k_B$). So it is proven that the minimum transportation cost for the equilibrium to exist is:

$$t > \begin{cases} -\frac{3}{5}\left(2k_I - 2k_B - \frac{3(k_B - k_A)(2-\lambda)}{8\lambda - 6}\right), & \text{if } k_I < \dot{k}_I, \\ \frac{3(k_B - k_A)\lambda}{4\lambda - 3}, & \text{if } \dot{k}_I \leq k_I \leq \ddot{k}_I \\ -\frac{3}{5}\left(2k_A - 2k_I - \frac{3(k_B - k_A)(2-\lambda)}{8\lambda - 6}\right), & \text{if } k_I > \ddot{k}_I. \end{cases}$$

Appendix 2.C Asymmetric Information

Non-discriminating operators

If the incumbent misrepresents its fixed cost by declaring f^D (with $f^D \neq f$), then the access prices chosen by the regulator are $\check{a}_A^D(f^D)$ and $\check{a}_B^D(f^D)$. While the prices are affected by this misrepresentation, the equilibrium market shares do not (calculations are available upon request). The incumbent's profits turn out to be $\check{\pi}^D = f^D - f \neq \check{\pi} = 0$.

Discriminating operators

Given the real transportation cost t and the possibility to misrepresent it by signalling a different transportation cost t^D, the access prices chosen by the regulator (if it believes the firms declaration) are $\check{a}_A(t^D)$ and $\check{a}_B(t^D)$. The resulting optimised retail prices are $p_i(t, t^D)$ and the market shares are:

$$
\begin{aligned}
d_I^D =& \frac{1}{3} + \frac{k_A + k_B - 2k_I}{5t} \\
d_A^D =& \frac{(18 - 69\lambda)k_A + 3(6 + 7\lambda)k_B + 4(3k_I + 5t)}{60(4\lambda)t} \\
d_B^D =& \frac{(18 - 69\lambda)k_B + 3(6 + 7\lambda)k_A + 4(3k_I + 5t)}{60(4\lambda)t}
\end{aligned}
$$

that is, they do not depend on the misrepresenting value of the transportation cost.

The corresponding equilibrium profits are:

$$\pi_I^D = \frac{t - t^D}{1800(4\lambda - 3)^2 t t^D} \cdot \Omega(t, t^D) \neq \check{\pi}_I = 0$$
$$\pi_A^D = \frac{((18 - 69\lambda)k_A + 3(7\lambda + 6)k_B + 4(4\lambda - 3)(3k_I + 5t))^2}{3600(4\lambda - 3)^2 t} = \check{\pi}_A$$
$$\pi_B^D = \frac{((18 - 69\lambda)k_B + 3(7\lambda + 6)k_A + 4(4\lambda - 3)(3k_I + 5t))^2}{3600(4\lambda - 3)^2 t} = \check{\pi}_B$$

where $\Omega(t, t^D) = -9(k_A^2 + k_B^2)(72 + \lambda(353\lambda - 642)) + 18k_A(16k_I(4\lambda - 3)^2 + k_B(\lambda(97\lambda - 258) - 72)) + (4\lambda - 3)^2(288k_B k_I - 8(36k_I^2 - 25tt^D))$.

Hence, the only operator that could improve its profits by misrepresenting the transportation cost is the incumbent I.

Chapter 3

Calling vs. Receiving Party Pays: Market Penetration and Importance of the Call Externalities[1]

3.1 Introduction

To provide connection between all users, telecom networks need access to the consumers of their rival. Access is provided after the payment of a termination charge (or access price). This charge is a part of the cost of off-net calls and consequently affects the price of calls.

Over the last few years, there has been a growing discussion among regulatory authorities on the regulation of termination charges or access prices in Europe. The European Commission (2008, 2009) recommended that national regulatory authorities lower termination charges to lower the average price per minute. In March 2011, Ofcom proposed a cap on mobile termination rates (MTRs) based on

[1] Joint with Tommaso Majer

the long run incremental cost of terminating a call (i.e., pure LRIC) that would lower MTRs from 4.18 to 0.69 pence per minute over four years.

Frequently, countries that adopt low access charges as an interconnection arrangement happen to be RPP countries. This correlation may be explained by the following: a connection is provided after an access price is paid that covers the cost borne by the receiving network to terminate the incoming call. When a country adopts a very low access price (or a zero access price, the so-called Bill & Keep), the receiving network cannot recover the cost of terminating the call from the originating network and the receiving network charges the called party as a result.

In the past, Ofcom has expressed several concerns about the introduction of a RPP tariff regime.[2] The main objections of the UK telecommunications regulatory authority are that it would be disruptive to customers, it would meet with consumer resistance and it might also lead to customers turning off their mobile phones. There are many studies regarding this last concern (e.g. Bomsel et al. (2003), Cadman (2007) and Samarajiva & Melody (2000)) but they do not provide a theoretical background for their analyses. In particular, there is no model that explains how market penetration and welfare would change due to a shift from one regime to the other. Our intention is to model the two regimes and provide a theoretical framework to compare them.

Countries with different regimes present very different characteristics, as described by Littlechild (2006).

In RPP countries, minutes of usage are more than in CPP countries. The reason is because in RPP countries, where Bill & Keep (BaK) is common, network operators pay a price equal (or close to) zero to terminate a call. Therefore the marginal cost of a call is reduced, and in turn, the usage price is reduced. This leads to higher usage.

The data on market penetration are diverse. In 2005, market penetration in the US and Canada was far below that of the EU, but in other BaK countries, such as Hong Kong, penetration is above the EU average. Data for 2008 in ERG

[2]See Oftel (2002) and Ofcom (2005).

(2009)[3] show that the 87% penetration in the US was lower than the EU average of 123%. However, penetration in Hong Kong and Singapore was above EU average. From these data, it is difficult to determine whether market penetration is lower or higher in BaK countries, but market penetration in CPP countries may be overstated because of the traditionally greater number of prepaid schemes and multiple SIM cards.[4]

In this paper, we provide a theoretical analysis to show how the choice of access price determines the retail pricing regime. Our model confirms that the telecommunications industry chooses CPP in response to high access prices. Conversely, for low access charges, providers prefer to charge both the caller and receiver. Moreover, our simulation in Section 3.4 allows us to compare our equilibrium predictions for the two regimes and we provide some recommendations for a regulator that must choose the level of the MTRs. We find that with a large call externality (the utility that consumers obtain from receiving a call), a BaK policy (which is associated with RPP regimes) implies higher usage and higher market penetration.[5] This is because higher usage increases the utility of joining a mobile network, and consequently, more people would like to join a network. Moreover, BaK maximizes social welfare when compared to any other policy.

Our model stresses the relevance of the call externality. Indeed, as the European Commission (2008) notes, the welfare maximizing policy regarding the access

---

[3] See ERG (2009), Next Generation Networks Future Charging Mechanisms / Long Term Termination Issues.

[4] A report by Analysis Mason (2008, page 8) requested by Ofcom says:

> While looking at the comparative statics, it is important to note that the standard penetration data [...] measures the number of subscriptions in circulation, and not the number of users who hold mobile subscriptions, which in the case of Hong Kong, Singapore and the UK is much lower [...].

[5] Notice that the latter result is driven by (but does not rely only on) the assumption that all the consumers are homogenous. If we introduce heterogeneity across consumers (with low users and high users) we would expect lower market penetration in an RPP regime compared to a CPP regime. In fact, low users may not subscribe if they have to pay to receive calls.

price may be very different for different values of the call externality. For low values our model suggests that the optimal policy is to set the access price close to the termination cost and, consequently, to induce the industry to adopt a CPP price regime. Conversely, for high values of the call externality, the optimal policy should be BaK (in this case, the industry would adopt an RPP regime). The reason is that RPP regimes internalize the call externality by making the receiver paying to receive the call.[6] Hence, when this externality is relevant, RPP regimes are more efficient.

Related literature. The main contribution to the literature on telecommunications is provided by the seminal papers by Armstrong (1998) and Laffont et al. (1998a,b). These papers model telecommunications competition between two network operators that compete for consumers that only obtain utility from making calls. Laffont et al. (1998a) analyze network interconnection in an unregulated environment where price discrimination is excluded. They show that for non-linear retail prices, high interconnection tariffs raise final retail prices and reduce social welfare. Gans & King (2001) improve on the above analysis and find that, under price discrimination and non-linear pricing, providers prefer an access price below cost.[7]

These papers have inspired many works. Jeon et al. (2004) extend these models and allow consumers to obtain utility from receiving calls. In the usual setup of two horizontally differentiated networks with full coverage of the market, they introduce the possibility that operators may also charge customers for receiving a call. Hence receivers may affect the volume of the calls by hanging up first. The authors derive equilibrium usage prices under different off-net pricing tariffs. On the one hand, without network-based discrimination, networks set prices equal to the perceived marginal cost. On the other hand, in the presence of network-based discrimination, networks set high off-net prices (for high values of the externality, interconnection breaks down) and on-net prices lower than the marginal cost. To

[6]For a discussion of this topic, see BEREC (2010b).

[7]For a good survey of the literature, see Armstrong (2002).

avoid a multiplicity of equilibria, they introduce a noise term in the utility function of the receiver.[8]

Cambini & Valletti (2008) use a model where the demand for phone calls between each pair of customers is jointly determined. They show that under certain conditions, the connectivity breakdown is eliminated. Moreover, they explain the relationship between the access charge and the structure of the retail prices selected by the network operators (operators only choose to charge the receiver if the access charge is sufficiently low). López (2011) extends Jeon et al. (2004) in another direction. He introduces an additional random variable in the utility function of the caller. In this framework, operators set prices equal to the perceived marginal cost. Moreover, he shows that a firm's profits do not depend on the access charge.

Dessein (2003) builds a model with heterogenous customers. He considers a case with elastic participation and finds that networks prices equal marginal cost. When there is elastic participation, the industry exhibits positive network externalities, firms prefer access prices below marginal cost, and consequently, in equilibrium prices are below cost. This causes customers to make more calls and consequently increases the value of subscription.

Finally, Hermalin & Katz (2009) allow consumers to obtain utility from receiving calls, but in contrast to the previous papers, they assume that networks compete on quantity. The study demonstrates that a regulator cannot induce efficient off-net prices via the access charge.

In this paper, we modify the framework described in Jeon et al. (2004) and incorporate market expansion in the benchmark model to compare equilibrium prices (including the fixed part), market penetration and profits under the two different tariff regimes. Assuming no network-based discrimination, we consider a case where providers charge a strictly positive charge to the receivers (Receiver Party

[8]Notice that the hypothesis of full coverage prevents the possibility of analyzing the effects of different access price policies on market size and their consequences for welfare. Indeed, in one paragraph, the authors study incomplete coverage, but they limit their analysis to a definition of the equilibrium usage prices.

Pays regime) and a case where networks do not charge consumers for receiving a call (Caller Party Pays regime).

In Section 3.2, we present the model. In Section 3.3, we characterize the equilibrium prices and quantities in the receiver sovereignty case (when the receiver determines the length of the call) and in the caller sovereignty case (when the caller determines the length of the call). In Section 3.4, we simulate the equilibrium results and compare the solutions of the two cases. Section 3.5 concludes.

3.2 The Model

Networks. We consider two mobile networks $i = 1, 2$ located at two points x_1 and x_2 on an infinite Hotelling line. We normalize the distance between the two networks to one. Mobile networks have on-net call cost of $c = 2c_0 + c_1$, where c_0 is the marginal cost of originating or terminating a call and c_1 is the marginal cost of transmitting a call. Let a denote the access or termination charge. The marginal cost of an off-net call is therefore $c + (a - c_0)$ for the caller's network and $c_0 - a$ for the receiver's network.

Tariffs. Mobile network i offers a multi-part tariff (p_i, r_i, F_i) where p_i is the caller's usage price (notice that we only consider the case of non-network-based discrimination), r_i is the receiver's usage price and F_i is the fixed part.

Consumers. Consumers are differentiated along the Hotelling line. This line represents the preferences of the consumers over one characteristic of the networks. For instance, consumers may prefer a well know phone operator to a new one.

A consumer located at x and selecting network i incurs a transport cost equal to $t|x - x_i|$, with $t > 0$ representing the importance users assign to not being connected to their favorite network or, equivalently, the inverse of the degree of competition in the market.

The utility of placing a call is $u(q)$, where q denotes the length of the call. In a receiving party pays regime, the duration of the call may depend either on the calling price or on the receiving price. If the receiving price is very high compared to the calling price, the receiver will decide to hang up first, and determine the length of the call (the receiver sovereignty case). Otherwise, when the calling price is higher than the receiving price, the caller determines the duration of the call (the caller sovereignty case). Therefore, the quantity q is a piecewise function of the receiving and calling prices:[9]

$$q_i = \begin{cases} q(r_i) & \text{when receiver sovereignty} \\ q(p_i) & \text{when caller sovereignty} \end{cases} \tag{3.1}$$

The utility for a consumer joining network i when he calls a consumer in network j is

$$u_i = \begin{cases} u(q(r_i)) & \text{when receiver sovereignty} \\ u(q(p_i)) & \text{when caller sovereignty} \end{cases} \tag{3.2}$$

A consumer's utility for receiving a call from a consumer who joined network j is

$$\tilde{u}_i = \begin{cases} \tilde{u}(q(r_i)) = \beta u(r_i) & \text{when receiver sovereignty} \\ \tilde{u}(q(p_i)) = \beta u(p_i) & \text{when caller sovereignty} \end{cases} \tag{3.3}$$

Thus, the net surplus of a consumer who joined network i is

$$w_i = v_0 + n_i u_i + n_j u_i + n_i \tilde{u}_i + n_j \tilde{u}_j - p_i(n_i q_i + n_j q_i) - r_i(n_i q_i + n_j q_j) - F_i \tag{3.4}$$

where v_0 is a subscriber's utility from other mobile services, $n_i u_i$ is the utility from making calls to the n_i subscribers of network i, $n_i \tilde{u}_i$ is the utility from receiving calls from the n_i subscribers of network i, $p_i(n_i q_i + n_j q_i)$ is the cost of making calls and $r_i(n_i q_i + n_j q_j)$ is the cost of receiving calls. The profits of network i are given by

$$\pi_i = n_i[n_i(p_i - c)q_i + n_j[p_i - c - (a - c_0)]q_i + n_j(a - c_0)q_j + r_i(n_i q_i + n_j q_j) + F_i] \tag{3.5}$$

[9] Remember that because we consider a case with non-network based discrimination a customer of network i pays the same amount when he calls a customer of network i or a customer of network j (i.e. $p_{ii} = p_{ij} = p_i$ and $r_{ii} = r_{ij} = r_i$)

Notice that the expression for profits takes different forms depending on whether the caller or the receiver determines the length of the call. On the one hand, when $\beta p_i < r_j$ the receiver will hang up first and the length of the call depends only on the receiving price r. On the other hand, when $\beta p_i > r_j$ the caller will hang up first and the length of the call depends only on the price calling p.

3.3 The equilibrium

To analyze market penetration we assume elastic subscriber participation. Specifically, we model consumer demand using a Hotelling model with hinterlands.[10] If the two networks offer utilities w_1 and w_2, then network i attracts

$$n_i = \frac{1}{2} + \frac{w_i - w_j}{2t} + \lambda w_i \tag{3.6}$$

where $\lambda > 0$ represents the magnitude of market expansion possibilities. This is one of the novelties we introduce in our model compared to Jeon et al. (2004). This allows us to analyze how different values of the access price affect the equilibrium market penetration and the effects of the latter on welfare.

Notice that the market share in equation (3.6) is an implicit function: both w_i and w_j depend on n_i and n_j. To find the equilibria we need to explicitly express the market share of each operator i as a function n of the prices. This is performed by applying the fixed point theorem and by solving the system of the type

$$\begin{cases} n_i = \frac{1}{2} + \frac{w_i(n_i,n_j) - w_j(n_i,n_j)}{2t} + \lambda w_i(n_i, n_j) \\ n_j = \frac{1}{2} + \frac{w_j(n_j,n_i) - w_i(n_j,n_i)}{2t} + \lambda w_j(n_j, n_i) \end{cases}$$

The unique solution to the system gives the market share of operator i in the form $n_i = n(p_i, p_j, r_i, r_j, F_i, F_j)$.

Hence, we look for the symmetric equilibria in the cases where the length of the call is determined by the receiver (receiver sovereignty) or by the caller (caller sovereignty), given the assumption of a balanced calling pattern.[11]

[10] For more details see Armstrong & Wright (2009).

[11] This assumption means that the percentage of calls originated and terminated on a given network reflects the market share of this network.

Equilibrium selection. We find a set of infinite equilibria given by a vector of prices (p_i, r_i, F_i) that maximizes operator i's profits as defined by equation (3.5). Our objective is to characterize a unique combination of the equilibrium prices given the access charge. In particular, we want the calling price to increase with the access price, and the receiving price to decrease. These properties reflect the empirical evidence. Then we define the further criteria satisfied by the selected equilibrium as follows .

In a market where the number of users is constant, Jeon et al. (2004) find that the optimal prices are $p^* = c - (a - c_0)$ and $r^* = c_0 - a$. Where $a < a^* \equiv c_0 - c\frac{\beta}{1+\beta}$, they get that the receiver determines the length of the call. Conversely, where $a > a^*$, there is caller sovereignty. In our setup with elastic demand, the selection procedure used by Jeon et al. (2004) is only feasible when $r = \beta p$. The resulting prices are $r^* = c\beta/(1+\beta)$ and $p^* = c/(1+\beta)$, and only constitute an equilibrium where $a = a^* \equiv c_0 - c\frac{\beta}{1+\beta}$. We aim to select their same equilibrium for $a = a^*$, and their same optimal price for the party who determines the length of the call. The price of the other party must be consistent with the type of the sovereignty case.

Condition on the market expansion parameter λ. Note that in a symmetric equilibrium, the market share of network i becomes

$$n = \frac{1 - 2\lambda(F - v_0)}{2 - 4\lambda[(1+\beta)u - (p+r)q]}$$

Let us find the derivative of the market share with respect to the fixed part of the retail tariff. We obtain

$$\frac{\partial n}{\partial F} = \frac{\lambda}{2\lambda[(1+\beta)u - (p+r)q] - 1} \tag{3.7}$$

We need this derivative to be negative; therefore we impose a negative denominator. From this, we obtain the following condition on λ:

$$\lambda < \bar{\lambda} \equiv \frac{1}{2[(1+\beta)u - (p+r)q]} \tag{3.8}$$

We impose condition (3.8) to ensure non-explosive market shares. If this condition did not hold, a network could increase the fixed part of the tariff, and as a consequence, more customers would join the network. This condition reflects a similar condition found in Armstrong & Wright (2009); their results only hold when λ is sufficiently small.

Maximization procedure. When firms set a two-part tariff, the maximization method used in most telecommunications papers is a two-step procedure. First, a firm maximizes profits with respect to the usage prices keeping constant the market shares. Second, the firm selects the fixed part that maintains its market share constant. Note that, since the total size of the market is constant, therefore also the market share of the rival remains constant.

With elastic participation this method is no longer usable. The intuition is the following. First, a change in the usage prices p_i and r_i must be compensated by a change in F_i that maintains the market share of network i constant. Finally, to keep the market share of network j constant, a change in p_i and in r_i must not change the surplus of the people who join network j. This is only possible for certain values of the access price. A more formal suggestion of this reasoning can be found in Jeon et al. (2004, page 101).

In the next sections we consider two cases: first the case of receiver sovereignty, when the receiver determines the length of the call and second the case of caller sovereignty, when the caller determines the length of the call.

3.3.1 Receiver sovereignty: $\beta p < r$

The receiver determines the length of the call q when she hangs up first, i.e. when usage prices are such that $\beta p < r$. In this case the profit function is given by equation (3.5) and the length of the call (i.e. the quantity of minutes demanded) depends on the receiving price r. We maximize profits with respect to the calling price p_i, to the receiving price r_i and to the fixed part F_i. We find the following symmetric prices:

Proposition 3.1 (Equilibrium retail prices). *When the receiver determines the length of a call, a symmetric equilibrium is characterized by the following retail* *prices:*

$$p^{rs} = a + c - c_0 - \frac{2\lambda t[\beta c - (1+\beta)(c_0 - a)]}{2\lambda(1+\lambda t)(1+\beta)[\beta u - (c_0 - a)q] - \beta}$$

$$r^{rs} = c_0 - a$$

and $\lambda < \bar{\lambda}$ *and* $t < \bar{t}^{rs} \equiv \frac{\beta}{2\lambda^2(1+\beta)[\beta u - (c_0 - a)q]} - \frac{1}{\lambda}$. *In this case the sum of the usage prices are below the marginal cost* c *and the receiving price* r *is decreasing in* a.

Proof. See appendix. □

If the conditions on λ and t do not hold, then $\beta p^{rs} > r^{rs}$ and, therefore, the pair p^{rs}, r^{rs} would not constitute an equilibrium in receiver sovereignty because the caller is hanging up first.

Sum of usage prices. Notice that, when there is no possibility for market expansion ($\lambda \to 0$), prices simplify and are equal to the perceived marginal cost (as in Jeon et al. (2004)). In this case, the sum of the prices is c. With market expansion, the sum of the usage prices is

$$p^{rs} + r^{rs} = c - \frac{2\lambda t[\beta c - (1+\beta)(c_0 - a)]}{2\lambda(1+\lambda t)(1+\beta)[\beta u - (c_0 - a)q] - \beta} \tag{3.9}$$

When $\lambda < \bar{\lambda}$ and $t < \bar{t}^{rs}$ (i.e., when p^{rs} and r^{rs} constitute an equilibrium), the fraction in the above expression is positive (and therefore the sum of the prices is below the marginal cost c).[12]

Case $a = a^*$. Finally, notice that when $a = a^* \equiv c_0 - c\frac{\beta}{1+\beta}$, the prices in equation (3.18) are equal to the equilibrium prices in a case with non-elastic market size.

$$r^{rs} = c\frac{\beta}{1+\beta} \tag{3.10}$$

$$p^{rs} = c - c\frac{\beta}{1+\beta} \tag{3.11}$$

[12]Notice that to ensure that $\bar{t}$ is positive, we must assume that $\beta u(q) > (c_0 - a)q$, which means that the utility a consumer obtains from receiving a call is greater than the cost she pays for that call.

Moreover, these prices satisfy the condition $\beta p = r$. This means that the caller and receiver obtain the same utility from making and receiving the call, and therefore, they will hang up at the same moment.

3.3.2 Caller sovereignty: $\beta p > r$

The caller determines the length of the call q when she hangs up first, i.e. when usage prices are such that $\beta p > r$. The profit function is given by equation (3.5). To find the equilibrium multi-party tariff, we maximize the profits function with respect to the three parts of the tariff.

Proposition 3.2 (Equilibrium retail prices). *When the caller determines the length of a call, a symmetric equilibrium is characterized by the following retail prices:*

$$p^{cs} = a + c - c_0$$
$$r^{cs} = c_0 - a - \frac{2\lambda t[\beta c - (1+\beta)(c_0 - a)]}{1 - 2\lambda(1+\lambda t)(1+\beta)[u - (a + c - c_0)q]}$$

and $\lambda < \bar{\lambda}$ *and* $t < \bar{t}^{cs} \equiv \frac{1}{2\lambda^2(1+\beta)[u-(a+c-c_0)q]} - \frac{1}{\lambda}$. *In this case, usage prices are below the marginal cost, and the calling price p increases in the access charge a.*

Proof. See appendix. □

The conditions on λ and t must hold to have an equilibrium in this case as well.

Sum of the usage prices. The sum of the two prices is

$$p^{cs} + r^{cs} = c - \frac{2\lambda t[\beta c - (1+\beta)(c_0 - a)]}{1 - 2\lambda(1+\lambda t)(1+\beta)[u - (a + c - c_0)q]} \tag{3.12}$$

In caller sovereignty, the total level of the prices depends on the sign of the following fraction:

$$\frac{2\lambda t[\beta c - (1+\beta)(c_0 - a)]}{1 - 2\lambda(1+\lambda t)(1+\beta)[u - (a + c - c_0)q]} \tag{3.13}$$

In Appendix 3.A, we show that p^{cs} and r^{cs} are equilibrium prices if $\lambda < \bar{\lambda}$ and $t < \bar{t}^{cs}$. When these conditions hold, expression (3.13) is positive, and, therefore, the sum of the usage prices is smaller than the marginal cost c.[13] With elastic participation the industry exhibits positive network externalities, and prices below costs cause customers to make more calls, which consequently increases the value of subscription. This is optimal for a network that obtain more profits by extracting the surplus through the fixed monthly fee.

Case $a = a^*$. When $a = a^*$ the prices are equal to

$$r^{cs} = c\frac{\beta}{1+\beta} \tag{3.14}$$

$$p^{cs} = c - c\frac{\beta}{1+\beta}. \tag{3.15}$$

Notice that prices in (3.14) and (3.15) satisfy

$$\beta p = r.$$

With these prices, the caller and receiver obtain the same utility from making and receiving a call; therefore, the consumers want to hang up at the same moment. This means that when $a = a^*$, prices are such that callers and receivers jointly determine the length of the call.

It is important to note that for access prices lower than a^*, we are in a receiver sovereignty regime and with access prices greater than a^*, we have a caller sovereignty regime. In fact p^{cs} increases in a and r^{rs} decreases in a.[14]

Case $\beta = 0$. When there are not externalities from receiving a call ($\beta = 0$), we obtain the same result found by Dessein (2003). Indeed, consider equation (3.35), which represents all of the multiple equilibria under caller sovereignty. If we substitute $\beta = 0$, equation (3.35) represents the set of multiple equilibria, when

[13] To ensure that the upper bounds for t and λ are positive, we have to impose that $u > (a + c - c_0)q$, which means that the utility that consumers obtain from making a call is greater than the cost they pay.

[14] Notice that we impose $p^{cs} > r^{cs}$ and $r^{rs} > p^{rs}$; therefore, we simply compare p^{cs} and r^{rs}.

there are not externalities from receiving a call. If we choose a pair (p, r) from this set such that the receiving price r is zero, we obtain the following calling price:

$$p = c + \frac{a - c_0}{2}. \tag{3.16}$$

3.3.3 Comments

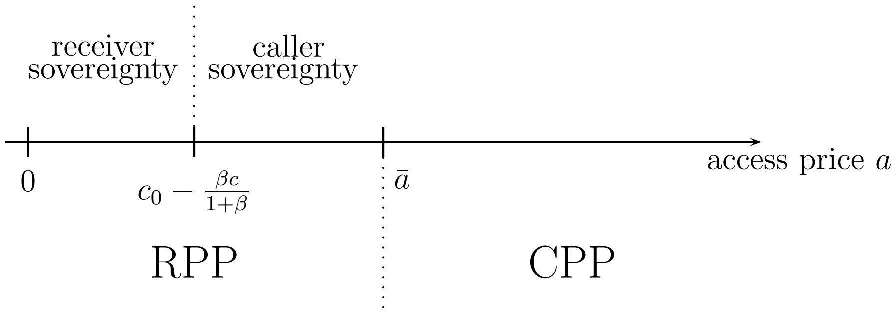


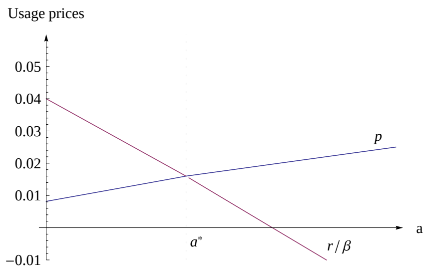
Figure 3.1: Access price

In our model, the telecom industry chooses a retail pricing regime according to the level of the access price. In Fig. 3.1 we summarize some results. When the access price is low, providers want to charge customers for receiving a call. The sum of the equilibrium usage prices is smaller than the marginal cost c. Moreover, callers and receivers only decide to hang up at the same time when the access price is $c_0 - \frac{\beta}{1+\beta}c$. For lower values, the price to receive a call is higher than the price of placing a call; therefore, the receiver will hang up first (receiver sovereignty). For higher values, the caller will hang up first (caller sovereignty). Moreover, for high values of the access price ($a > \bar{a}$), providers only charge the caller: the termination cost is fully paid by the calling network and the receiving provider does not need to recover it from the receiver.[15]

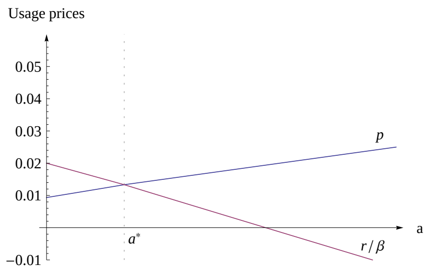
[15]The price for receiving a call, r, decreases in a and becomes 0 when $a = \bar{a}$. When $a > \bar{a}$, the receiver is subsidized for receiving a call. Therefore, only the caller pays. Notice also that $\bar{a}$ is necessarily larger than a^* until $\beta \leq 1$.

3.4 Comparison

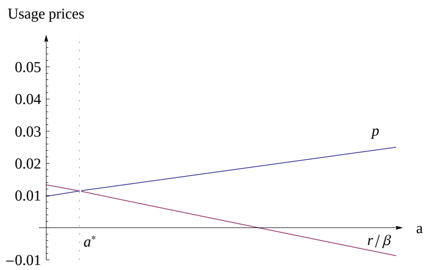
In this section we compare the two regimes and state the implications of different access price policies.



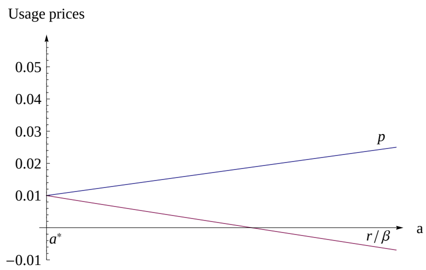
(a) $\beta = 0.25$



(b) $\beta = 0.5$



(c) $\beta = 0.75$



(d) $\beta = 1$

Figure 3.2: Usage prices p and r. Parameter values: $c_0 = 0.01$, $c = 0.02$, $\eta = 3$.

Usage price. Fig. 3.2 illustrates the calling and receiving prices. We represent the optimal prices p and r/β according to the possible levels of the access price a. We select the cost parameters following Hoernig & Harbord (2010). Moreover, from now on, we use a constant elasticity demand function $q(p) = p^{-\eta}$ (as in Hoernig (2007)) where $\eta > 1$ and $u(q) = \frac{\eta}{\eta-1} q^{\frac{\eta-1}{\eta}}$. The highest perceived price determines the sovereignty regime; the dotted line a^* marks the threshold between the RS regime (to the left) and the CS regime (to the right).

The four graphs show how increasing values of the call externality β affect the implication of different policies for the access price a. When the externality is low, BaK leads to a receiver sovereignty regime. When the impact of the externality is high, a Bill & Keep policy (BaK) implies a caller sovereignty regime. Moreover, as noted previously, the sum of the usage prices is below the marginal cost c. As a is set closer to a^*, the sum of the prices approaches the marginal cost. Therefore, a BaK policy when the externality is low would imply lower values for the usage prices than would be the case with a high externality.

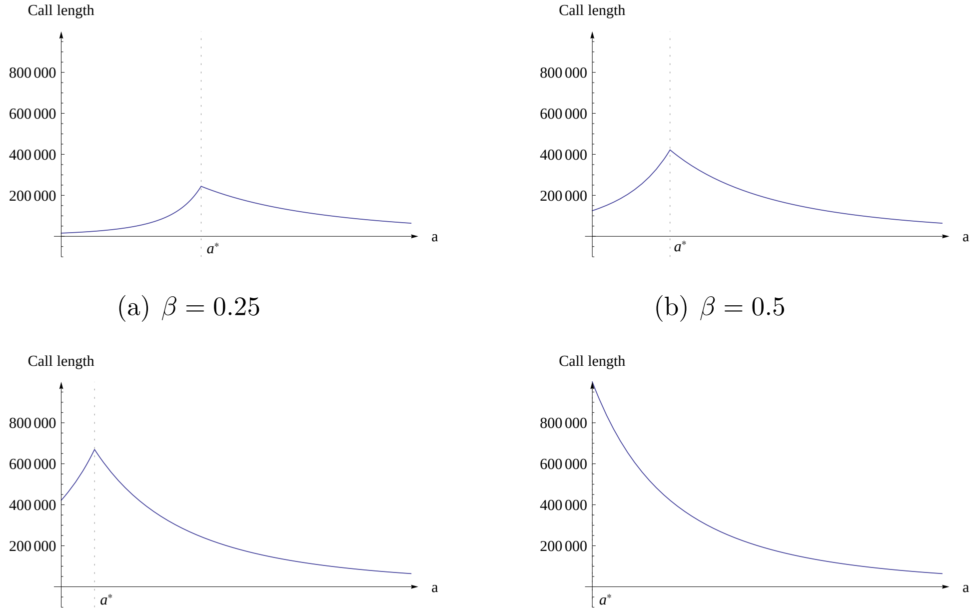
Finally, notice that the calling price is increasing in the access charge while the receiving price is decreasing. We allow the receiving price to assume negative values; in many countries the operators subsidize the receiver to take the calls, a negative value of r reflects this commercial policy. Obviously, when r is negative only the caller pays for taking part in a call, that is, the pricing scheme is calling party pays. When the receiver also pays for a call ($r > 0$) we have instead a receiving party pays regime. Therefore it is clear that RPP regimes are associated with low values of the access price, while CPP regimes are associated with high values of a.

Length of a call. As usual, the demand function is expressed in terms of the length of a call, $q(\cdot)$, and is a decreasing function of the retail prices ($q' < 0$). We represent the length of a call for different values of β in Fig. 3.3. The dotted line still represents the threshold between receiver sovereignty and caller sovereignty.

The longest length is attained for a^* when both caller and receiver want to hang up at the same time. The reason is because the price of a call is shared between the caller and receiver, taking into account the positive externality on the receiver, and therefore, the calls tend to be longer. However this is only true when both caller and receiver are eager to hang up at approximately the same time. Otherwise, whoever faces the higher price prefers to end the call earlier.

For low values of the call externality β, BaK produces shorter calls rather than higher values of the access charge.

For high values of β (Fig. 3.3(c) and 3.3(d)), access prices close (or equal) to zero imply longer calls under RPP than CPP. Because the receiver is eager to pay to receiver a call under higher values of the externality, the value of a that makes both caller and receiver to hang up at the same time shifts towards zero where the



(a) $\beta = 0.25$ (b) $\beta = 0.5$

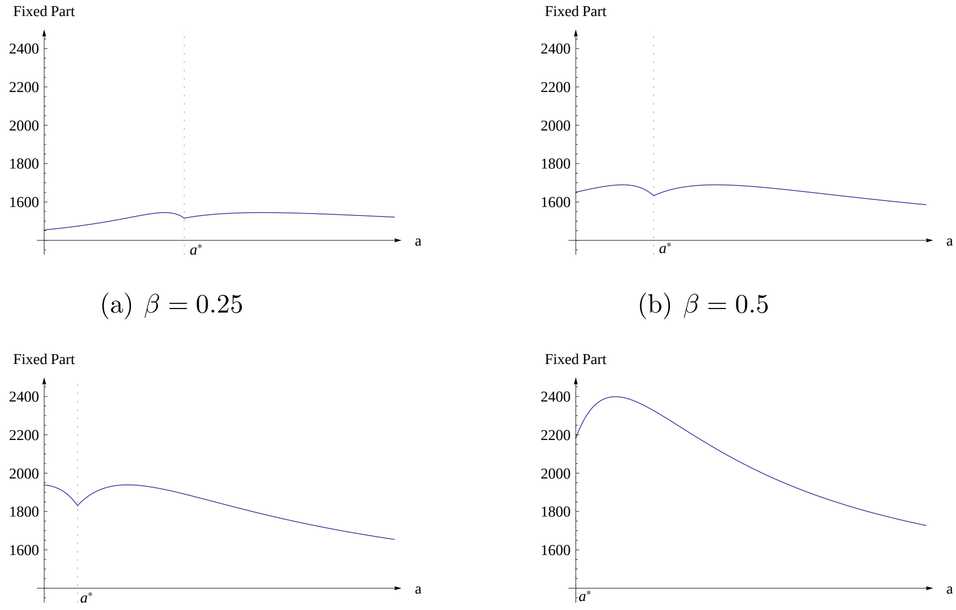
(c) $\beta = 0.75$ (d) $\beta = 1$

Figure 3.3: Length of a call q. Parameter values: $c_0 = 0.01$, $c = 0.02$, $t = 1500$ and $\eta = 3$.

associated retail prices are higher for the receiver. This allows the regulator to set a zero access price and makes calls longer than under CPP regimes. This confirms the expectations of Ofcom (2009, page 37):

> [...] international comparisons provide evidence that this relationship between termination rates, and take-up and usage, exists. A simple analysis of cross-country data [...] suggests that countries that have, broadly speaking, systems that adopt reciprocity or "bill and keep"-like arrangements – US, Hong Kong and Singapore (and to a lesser degree Canada) have higher usage than countries with "Calling Party's Network Pays" regimes.

Fixed part. Fig. 3.4 illustrates a comparison of the fixed parts of the two price regimes. First, notice that the value of λ is chosen according to equation (3.8). It is worthwhile to observe that an $a = a^*$ does not induce the highest fixed part. The



(a) $\beta = 0.25$ (b) $\beta = 0.5$

(c) $\beta = 0.75$ (d) $\beta = 1$

Figure 3.4: Fixed part F. Parameter values: $c_0 = 0.01$, $c = 0.02$, $t = 1500$, $\lambda = 0.00002$, $\eta = 3$ and $v_0 = 750$.

reason is straightforward: although a^* maximizes the consumer surplus of joining a call, the networks can extract a higher surplus through the fixed part. When $a \neq a^*$, the operators charge usage prices below the marginal cost, and, therefore, they need to increase the fixed part to cover them. However, as a departs from a^* the length of the call and the consumer surplus are lower. This reduces the extractable surplus and, consequently, the value of F.

For low values of the call externality, the relationship between the equilibrium values of F in RPP (implied by a low a) and in CPP (when a is large) is not univocally determined.

For high values of β, the fixed part in RPP is higher than the fixed part in CPP. In particular, BaK leads to a higher fixed fee than any other value of the access charge. The reason is that, when the access charge is below cost, calls last longer. Therefore, the consumer surplus that providers can extract is higher. This result coincides with many empirical observations. For instance, Ofcom (2009, page 37) expects the following:

> High termination rates tend to lead to a retail price structure with relatively high off-net call charges (since operators 'cover' their wholesale cost of each minute of a call with a corresponding retail charge) and lower subscription charges (since subscribers generate incoming calls that provide call termination revenue). [...] Equally, if termination rates are low, consumers will tend to face higher subscription fees but lower or no charges to make (or receive) calls.

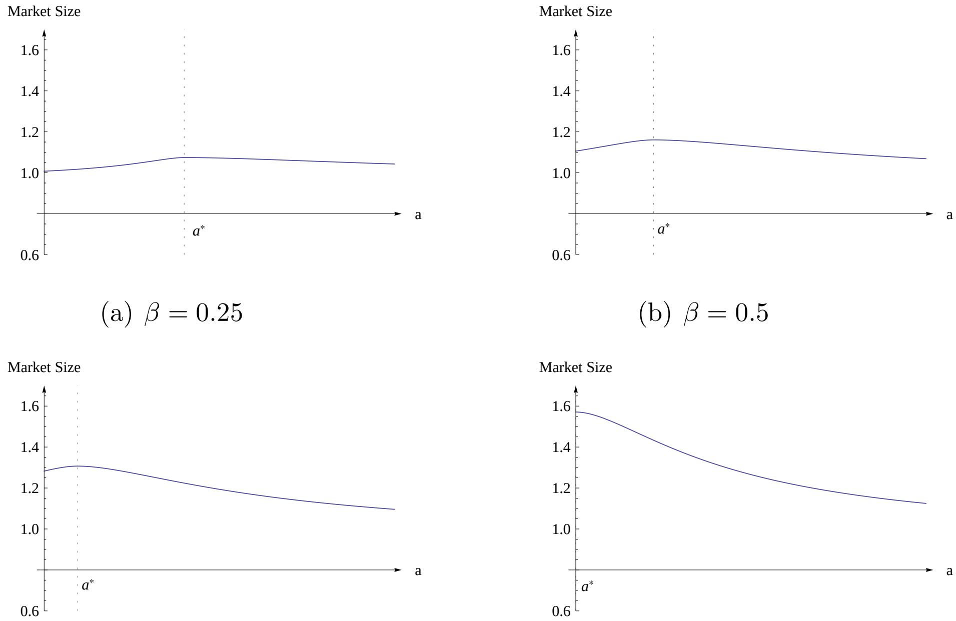
Market penetration. Fig. 3.5 illustrates that there is not a clear relationship between the degrees of market penetration in the two regimes. For low values of the receiver externality, the values of access charge may be such that penetration is higher in CPP. Conversely, for a high receiver externality, RPP regimes exhibit a high number of subscribers. This indeterminacy is also present in the empirical evidence. On the one hand, Littlechild (2006) shows how CPP are distinguished by higher market penetration. On the other hand, Analysis Mason (2008) states that the actual data misrepresent the true values of penetration by overestimating penetration in CPP countries. Moreover, high penetration is explained by through the higher surplus that consumers receive. Once again, we have the highest level of penetration under RPP at $a = a^*$.

The graphs in Fig. 3.5 again show that RPP regimes are more sensitive to variations in the perceived externality; market size is increasing in β.

Profits. Industry profits are maximized when $a = a^* \leq c_0$: the consumer surplus is maximized, and therefore, the rent the operators can extract is higher. Notice that BaK is more attractive to the industry when the call externality is higher.

With high values of call externalities, networks prefer an access charge below marginal cost. When call externalities are negligible and close to zero, networks prefer an access price close to the marginal cost c_0.

Notice also that profits are increasing in a when $a < a^*$ (i.e., under receiver sovereignty), while they decrease in the access price when $a > a^*$ (caller sovereignty).



(a) $\beta = 0.25$ (b) $\beta = 0.5$

(c) $\beta = 0.75$ (d) $\beta = 1$

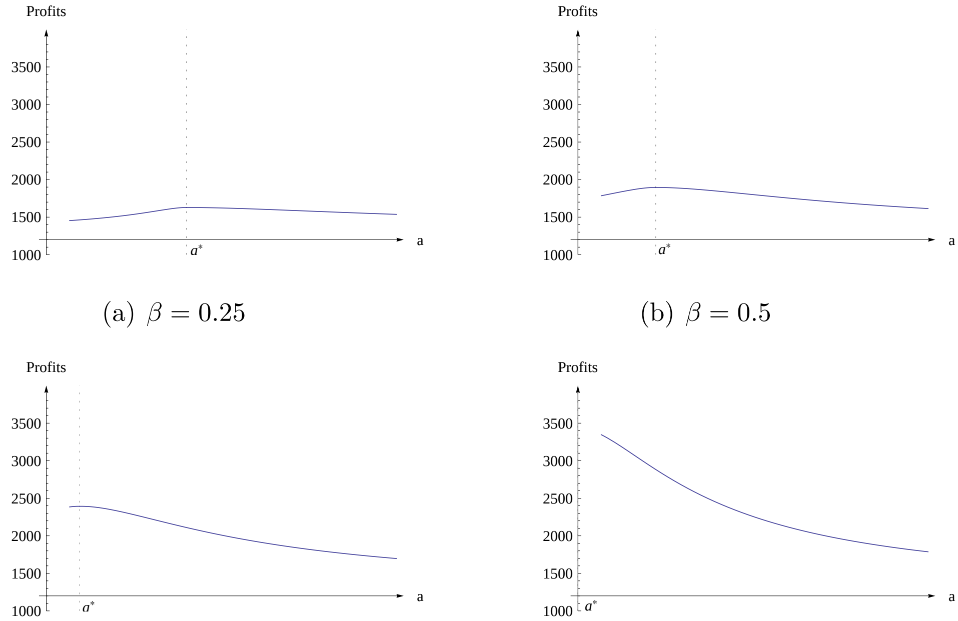
Figure 3.5: Market penetration N. Parameter values: $c_0 = 0.01$, $c = 0.02$, $t = 1500$, $\lambda = 0.00002$, $\eta = 3$ and $v_0 = 750$.

3.4.1 Welfare analysis

We compare welfare between the two regimes in Fig. 3.7. Total welfare is given by a weighted sum of consumer surplus and industry profits.It is clear, that the greatest welfare is attained at $a = a^*$. At this value of the access price, consumer surplus of a call is maximized and the network can obtain the highest profits by extracting it.

Notice that as the receiver externality increases, welfare increases in RPP regimes. This fact again highlights the importance for a regulator having a very precise knowledge of the values of β when choosing the access price: very low values of a (accompanied by an RPP regime) are socially optimal only if the receiving externality is high. The European Commission (2008) arrives at the same conclusion:

> RPP might not be efficient if the calling party values the call highly



(a) $\beta = 0.25$ (b) $\beta = 0.5$

(c) $\beta = 0.75$ (d) $\beta = 1$

Figure 3.6: Profits Π. Parameter values: $c_0 = 0.01$, $c = 0.02$, $t = 1500$, $\lambda = 0.00002$, $\eta = 3$ and $v_0 = 750$.

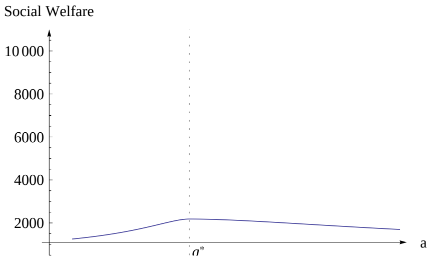
> but the called party does not and, as a result, an efficient call might not be completed. The reverse issue may arise in the CPP system, where an efficient call may not be initiated even if the called party values it highly but the calling party does not.

Indeed, if a regulator considers that the externality is very low in its country, BaK is not the welfare-maximizing policy.

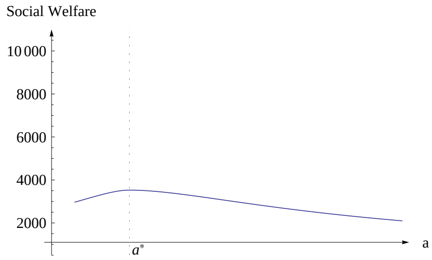
Finally, assigning different weights to consumer surplus and industry profits does not qualitatively change the results.

3.5 Conclusions

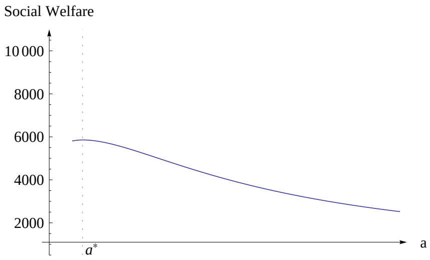
Regulatory authorities are concerned about reducing mobile termination rates but there is a lack of theoretical analysis that could give them information about the



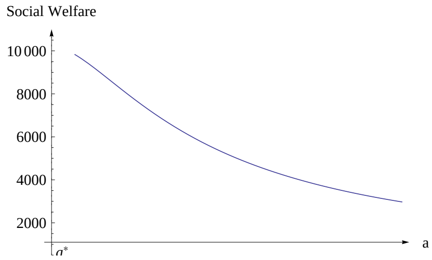
(a) $\beta = 0.25$



(b) $\beta = 0.5$



(c) $\beta = 0.75$



(d) $\beta = 1$

Figure 3.7: Total welfare W. Parameter values: $c_0 = 0.01$, $c = 0.02$, $t = 1500$, $\lambda = 0.00002$, $\eta = 3$ and $v_0 = 750$.

consequences of such a policy.

The European Commission (2008, 2009) proposed a drastic reduction of the mobile termination rates over the coming years. This policy, according to empirical evidence and companies' predictions, would imply charging consumers for receiving calls to cover the termination cost of a call. The European Commission (2008, page 26) noticed that "RPP may evolve after a reduction of the regulated termination charge or as a response to a Bill and Keep system". Ofcom (2005) warned that RPP regimes could be meet with opposition by consumers who do not want to be charged for incoming calls.

In our paper, we provide a theoretical framework that allows us to compare the two tariff regimes. We confirm that a relationship between interconnection arrangements and retail price structure exists. We showed that one tariff regime is not necessarily superior to the other in terms of retail prices, usage, market penetration and overall welfare for all values of the access price.

Using realistic values of the industry parameters, we find that the level of the call externality is crucial. When this level takes high values, market penetration and total welfare are higher in an RPP regime with access charges close to zero. This suggests that a BaK policy (that results in the adoption of an RPP regime) should be implemented only once the presence of a high call externality is proven. Otherwise access pricing at the termination cost would be a better policy.
To our knowledge, there are no estimates of the call externalities. On the one hand, the Body of European Regulators for Electronic Communications (BEREC (2010b)) noted that it seems reasonable to assume that the utility of the receiver is lower than that of the caller, but that the difference is not very significant. On the other hand, in BEREC (2010a), several phone companies claimed that the call externalities are very low or even equal to zero.

References

Analysis Mason (2008). Case studies of mobile termination regimes in Canada, Hong Kong, Singapore and the USA: report for Ofcom. Technical report, Analysis Mason.

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

Armstrong, M. (2002). *The Theory of Access Pricing and Interconnection*, volume 1, chapter 8, (pp. 295–386). North-Holland.

Armstrong, M. & Wright, J. (2009). Mobile Call Termination. *The Economic Journal*, *199*, 270–307.

BEREC (2010a). Consultation Report to the Common Statement “Next Generation Networks Future Charging Mechanisms / Long Term Termination Issues”. Technical report, BEREC. available at `http://www.erg.eu.int/doc/berec/bor_10_24b.pdf`.

BEREC (2010b). Next Generation Networks Future Charging Mechanisms / Long Term Termination Issues. Technical report, BEREC. available at `http://www.erg.eu.int/doc/berec/bor_10_24_ngn.pdf`.

Bomsel, O., Cave, M., Blanc, G. L., & Neumann, K.-H. (2003). How mobile termination charges shape the dynamics of the telecom sector. Technical report, University of Warwik and Wik Consult.

Cadman, R. (2007). NGN Interconnection: Charging Principles and Economic Efficiency. Technical report, NGN UK, London.

Cambini, C. & Valletti, T. (2008). Information Exchange and Competition in Communicatios Networks. *The Journal of Industrial Economics*, *22*, 1821.

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

ERG (2009). Next Generation Networks Future Charging Mechanisms / Long Term Termination Issues. Technical report, European Regulatory Group.

European Commission (2008). Explanatory note accompanying document to the Commission Reccomandations on the Regulatory Treatment of Fixed and Mobile Termination Rates in the EU. Technical report, European Commission, Brussels.

European Commission (2009). Commission Recommentation of 7 May 2009 on the Regulatory Treatment of Fixed and Mobile Termination Rates in the EU (2009/396/EC). Technical report, European Commission, Brussels.

Gans, J. S. & King, S. P. (2001). Using 'bill and keep' interconnect arrangements to soften network competition. *Economics Letters*, *71*(3), 413 – 420.

Hermalin, B. E. & Katz, M. L. (2009). Customer or Complementor? Intercarrier Compensation with Two-Sided Benefits. Working Paper.

Hoernig, S. (2007). On-net and off-net pricing on asymmetric telecommunications networks. *Information Economics and Policy*, *19*(2), 171 – 188.

Hoernig, S. & Harbord, D. (2010). Welfare Analysis of Regulating Mobile Termination Rates in the UK (with an Application to the Orange/T-Mobile Merger). Working Paper.

Jeon, D.-S., Laffont, J.-J., & Tirole, J. (2004). On the "Receiver-Pays" Principle. *The RAND Journal of Economics*, *35*(1), 85–110.

Laffont, J.-J., Rey, P., & Tirole, J. (1998a). Network Competition: I. Overview and Nondiscriminatory Pricing. *The RAND Journal of Economics*, *29*(1), 1–37.

Laffont, J.-J., Rey, P., & Tirole, J. (1998b). Network Competition: II. Price Discrimination. *The RAND Journal of Economics*, *29*(1), 38–56.

Littlechild, S. (2006). Mobile termination charges: Calling Party Pays versus Receiving Party Pays. *Telecommunications Policy*, *30*(5-6), 242 – 277.

López, Á. L. (2011). Mobile termination rates and the receiver-pays regime. *Information Economics and Policy*, *23*(2), 171–181.

Ofcom (2005). Wholesale mobile voice call termination. Preliminary consultation on future regulation. Technical report, Ofcom.

Ofcom (2009). Wholesale mobile voice call termination. Preliminary consultation on future regulation. Technical report, Ofcom.

Oftel (2002). Receiving Party Pays compared to Calling Party Pays. Technical report, Oftel.

Samarajiva, R. & Melody, W. H. (2000). Briefing paper. In *Fixed-Mobile Interconnection Workshop*. available at `http://www.itu.int/osg/spu/ni/fmi/workshop/`.

Appendix 3.A Proofs

Proof of Proposition 3.1. The first order conditions for operator i are:

$$\frac{\partial \pi_i^{rs}}{\partial p_i^{rs}} = 0 \tag{3.17a}$$

$$\frac{\partial \pi_i^{rs}}{\partial r_i^{rs}} = 0 \tag{3.17b}$$

$$\frac{\partial \pi_i^{rs}}{\partial F_i^{rs}} = 0 \tag{3.17c}$$

Since we restrict our analysis to symmetric equilibria we impose: $p_i = p_j = p^{rs}$, $r_i = r_j = r^{rs}$, $F_i = F_j = F^{rs}$.

First, we solve (3.17c) with respect to F^{rs} and we obtain the optimal fixed part as a function F of the usage prices, $F^{rs} = F(p^{rs}, r^{rs})$.[16]

When $F^{rs} = F(p^{rs}, r^{rs})$, equation (3.17a) is always satisfied, so we focus our attention on the equation (3.17b). Intuitively, when the receiver determines the length of the call and the fixed part is optimally chosen, the changes in the calling price do not affect the profits.[17]

Given condition $F^{rs} = F(p^{rs}, r^{rs})$, equation (3.17b) is now a function of p^{rs} and r^{rs}, as equation (3.32) in Appendix 3.B shows. Solving it with respect to p^{rs} we obtain two solutions, one of which makes the market shares equal to zero.

So the solutions to the maximization problem are given by all the pairs (p^{rs}, r^{rs}) that satisfy condition (3.33) in Appendix 3.B. Among the multiple equilibria, we select a pair of prices continuous in a that allows us to see how the prices change when the access price changes. Consider the following pair of prices:

$$r^{rs} = c_0 - a \tag{3.18a}$$

$$p^{rs} = a + c - c_0 - \frac{2\lambda t[\beta c - (1+\beta)(c_0 - a)]}{2\lambda(1+\lambda t)(1+\beta)[\beta u - (c_0 - a)q] - \beta} \tag{3.18b}$$

[16] The complete condition is given by equation (3.31) in Appendix 3.B.

[17] In receiving sovereignty only the receiving price determines the duration of the call. Therefore the calling price has an effect only on market participation. The optimal fixed part compensates a change in the calling price so that market size is kept constant. Hence, in receiving sovereignty profits do not depend on the calling price.

Consistency conditions on λ. We need to check that the prices we found are consistent with the case of receiving sovereignty (i.e. the receiver hangs up first). The above prices are consistent with the case of receiver sovereignty when $\beta p^{rs} < r^{rs}$. This condition is satisfied when

$$\lambda < \frac{\beta}{2(1+\beta)[\beta u(q)-(c_0-a)q]} \quad \text{and}$$
$$t < \bar{t} \equiv \frac{\beta}{2\lambda^2(1+\beta)[\beta u(q)-(c_0-a)q]} - \frac{1}{\lambda} \tag{3.19}$$

or[18]

$$\frac{\beta}{2(1+\beta)[\beta u-(c_0-a)q]} < \lambda < \frac{\beta}{(1+\beta)[\beta u-(c_0-a)q]} \tag{3.20}$$
$$\text{and} \qquad t < \frac{\beta - 2(1+\beta)[\beta u(q)-(c_0-a)q]\lambda}{2\lambda((1+\beta)[\beta u-(c_0-a)q]\lambda-\beta)} \tag{3.21}$$

or

$$\lambda > \frac{\beta}{(1+\beta)[\beta u(q)-(c_0-a)q]} \tag{3.22}$$

If parameters do not satisfy conditions (3.19), (3.20), (3.21) and (3.22), then in the equilibrium the calling price would be higher than the receiving price and this would implies that the caller hang up first. But in this case the profits function that the operator maximises takes a different form (the equilibrium with caller sovereignty is analysed in the following subsection).

Finally, we need to check that the these conditions on λ are compatible with condition (3.8) that ensures that market expansion is not explosive. It is easy to show that condition (3.8) implies $\lambda < \bar{\lambda}$. When $\lambda < \bar{\lambda}$, by consistency we need also to impose $t < \bar{t}$. The two conditions on λ and on t imply that the sum of the usage prices is below the marginal cost and, in an equilibrium, the receiver is hanging up first.[19]

[18] Notice that the upper bound for t is always positive.

[19] To see it, notice that

$$\frac{1}{2((1+\beta)u(q)-(p+r)q)} < \frac{\beta}{2(1+\beta)(\beta u(q)-(c_0-a)q)} \Longleftrightarrow (1+\beta)(c_0-a)q > \beta(p+r)q$$

Proof of Proposition 3.2. As before, we maximize the profits function deriving the profits with respect to the usage prices and to the fixed part. The first order condition for operator i are:

$$\frac{\partial \pi_i^{cs}}{\partial p_i^{cs}} = 0 \tag{3.23a}$$

$$\frac{\partial \pi_i^{cs}}{\partial r_i^{cs}} = 0 \tag{3.23b}$$

$$\frac{\partial \pi_i^{cs}}{\partial F_i^{cs}} = 0 \tag{3.23c}$$

We restrict our analysis to symmetric equilibria and we impose: $p_i = p_j = p^{cs}$, $r_i = r_j = r^{cs}$, $F_i = F_j = F^{cs}$.

First we solve (3.23c) with respect to F^{cs} and we obtain the optimal fixed part as a function F of the usage prices, $F^{cs} = F(p^{cs}, r^{cs})$, which expression is given in Appendix 3.B by equation (3.34).

When $F^{cs} = F(p^{cs}, r^{cs})$, the equation (3.23a) is always satisfied. Again, intuitively this happens because when the caller determines the length of the call and the fixed part is optimally chosen, the changes in the calling price do not affect the profits. Therefore, we focus our attention on the equation (3.23b). Given condition $F^{cs} = F(p^{cs}, r^{cs})$, equation (3.23b) is now a function of p^{cs} and r^{cs}, as shown by equation (3.35).

Solving the latter with respect to r^{cs} we obtain two solutions: one that makes the market shares equal to zero, and another one that establishes a relationship between p^{cs} and r^{cs}. All the pair (p^{cs}, r^{cs}) that maximize the profits function have to satisfy equation (3.36).

Among the multiple equilibria satisfying this condition we choose a pair of prices continuous in a that allows us to see how the prices change when the access charge changes. Furthermore, we will show that the calling price is increasing in

---

In the selected equilibrium, $c_0 - a = r$, therefore the inequality becomes

$$(1+\beta)rq > \beta(p+r)q \Longleftrightarrow r > \beta p$$

The inequality is the condition we require to be consistent with the receiver sovereignty case.

a and the receiving price is decreasing in a. Consider the following pair of prices:

$$r^{cs} = c_0 - a - \frac{2\lambda t[\beta c - (1+\beta)(c_0 - a)]}{1 + 2\lambda(1+\lambda t)[(1+\beta)(a+c-c_0)q - (1+\beta)u]} \tag{3.24}$$

$$p^{cs} = a + c - c_0 \tag{3.25}$$

These two prices satisfy (3.36).

Consistency condition on λ. As we are in caller sovereignty case, we have to impose conditions such that the caller is willing to hang up first and determines the duration of the call. Hence, we impose that

$$\beta p^{cs} > r^{cs}.$$

This condition is satisfied when:

$$\lambda < \frac{1}{2(1+\beta)[u(q) - (a+c-c_0)q]} \qquad \text{and} \tag{3.26}$$

$$t < \bar{t}^{cs} \equiv \frac{1}{2\lambda^2(1+\beta)[u(q) - (a+c-c_0)q]} - \frac{1}{\lambda} \tag{3.27}$$

or

$$\frac{1}{2(1+\beta)[u - (a+c-c_0)q]} < \lambda < \frac{1}{(1+\beta)[u - (a+c-c_0)q]} \tag{3.28}$$

$$\text{and} \qquad t < \frac{1 - 2(1+\beta)\lambda[u - (a+c-c_0)q]}{2\lambda\{1 + (1+\beta)\lambda[u - (a+c-c_0)q]\}} \tag{3.29}$$

or

$$\lambda > \frac{1}{(1+\beta)[u(q) - (a+c-c_0)q]}. \tag{3.30}$$

When the above conditions are not satisfied then the equilibrium prices do not satisfy the condition of characterizing a caller sovereignty outcome. Notice also that, similarly to the previous proof, the condition (3.8) implies that the parameters have to satisfy $\lambda < \bar{\lambda}$ and $t < \bar{t}^{cs}$.[20]

[20]That is, condition (3.8) implies that we have to consider the first range of parameters:

$$\frac{1}{2((1+\beta)u - (p+r)q)} < \frac{1}{2(1+\beta)(u - (a+c-c_0)q)} \Longleftrightarrow (1+\beta)(a+c-c_0)q > (p+r)q$$

Appendix 3.B Results

Receiver sovereignty. In receiver sovereignty, given symmetry, the first order condition with respect to the fixed part implies

$$\frac{\partial \pi^{rs}}{\partial F^{rs}} = 0 \Longleftrightarrow F^{rs} = \frac{K_1 K_2}{K_3 + K_4} \tag{3.31}$$

where:

$$\begin{aligned}
K_1 \equiv & 1 + 2\lambda v_0 \\
K_2 \equiv & 2\lambda(1+t\lambda)q^2(p+r)^2 - (1+t\lambda)(2\lambda(1+\beta)u-1)q(p+r) + t(2u(1+\beta)\lambda - 1) \\
& + c(\lambda(2u(1+\beta) - 2(1+t\lambda)q(p+r) + (2u(1+\beta)\lambda - 3)) - 1)q \\
K_3 \equiv & (1-2u(1+\beta)\lambda)(2\lambda((1+\beta)u + t\lambda(1+\beta)u - 2t - 1) \\
K_4 \equiv & 2\lambda(2\lambda((1+\beta)u + t\lambda(1+\beta)u - t) - 1)q(p+r) \\
& + 2c\lambda(\lambda(2(1+\beta)u - 2q(p+r)(1+t\lambda) + t(2\lambda(1+\beta)u - 3)) - 1))q
\end{aligned}$$

Given equation (3.31), the first order condition with respect the receiving price becomes

$$\left.\frac{\partial \pi^{rs}}{\partial r^{rs}}\right|_{F^{rs}=\frac{K_1K_2}{K_3+K_4}} = 0 \Longleftrightarrow \frac{q_r' K_1 K_5 K_6}{K_7} = 0 \tag{3.32}$$

where

$$\begin{aligned}
K_5 \equiv & 1 + 2q(p+r)\lambda(1+t\lambda) - 2\lambda((1+\beta)u + t(\lambda(1+\beta)u - 1)) \\
K_6 \equiv & 2t\lambda^2(c(1+\beta)(-qr+u\beta) + q(p+r)(r+\beta(a+2r-c_0)) \\
& - u\beta(1+\beta)(a+p+2r-c_0)) + \beta(a-c+p+2r-c_0) \\
& + 2\lambda(r(q(p+r)+t) + ((a+2r)(q(p+r)+t) - (a+p+2r)u)\beta \\
& - (a+p+2r)u\beta^2 - c(qr(1+\beta) - \beta(-t+u+u\beta)) + \beta(-q(p+r) - t + u + u\beta)c_0) \\
K_7 \equiv & 4\beta((-1+2u(1+\beta)\lambda)(-1+2\lambda(-2t+u+u\beta+tu(1+\beta)\lambda)) \\
& - 2q(p+r)\lambda(-1+2\lambda(-t+u+u\beta+tu(1+\beta)\lambda)) \\
& + 2cq\lambda(1+\lambda(-2u(1+\beta) + 2q(p+r)(1+t\lambda) + t(3-2u(1+\beta)\lambda))))^2
\end{aligned}$$

In the selected equilibrium, $a + c - c_0 = p$, therefore the inequality becomes

$$(1+\beta)pq > (p+r)q \Longleftrightarrow \beta p > r$$

The inequality is the condition we require to be consistent with the caller sovereignty case.

Condition (3.32) is satisfied when

$$p = \frac{K_8 + K_9 - K_{10}}{K_{11}} \tag{3.33}$$

where

$$\begin{aligned} K_8 \equiv & 2\left(r(-cq+qr+t)+(a-c+2r)(qr+t-u)\beta+(-a+c-2r)u\beta^2\right)\lambda \\ & +(a-c+2r)\beta \\ K_9 \equiv & \beta(-1+2\lambda(-t+u+u\beta+tu(1+\beta)\lambda-q(r+rt\lambda)))c_0 \\ K_{10} \equiv & 2t((a+2r)u\beta(1+\beta)-qr(r+a\beta+2r\beta)+c(1+\beta)(qr-u\beta))\lambda^2 \\ K_{11} \equiv & -2qr\lambda(1+t\lambda)+2u\beta^2\lambda(1+t\lambda)+ \\ & +\beta(-1-2(aq+2qr-u)\lambda(1+t\lambda))+2q\beta\lambda(1+t\lambda)c_0 \end{aligned}$$

Caller sovereignty. In caller sovereignty, given symmetry, the first order condition with respect to the fixed part implies

$$\frac{\partial \pi^{cs}}{\partial F^{cs}} = 0 \Longleftrightarrow F^{cs} = \frac{K_1 K_2}{K_3 + K_4} \tag{3.34}$$

Notice that this condition is the same than in receiver sovereignty. Given equation (3.34), the first order condition with respect the calling price becomes

$$\left.\frac{\partial \pi^{cs}}{\partial p^{cs}}\right|_{F^{cs}=\frac{K_1K_2}{K_3+K_4}} = 0 \Longleftrightarrow \frac{q_p' K_1 K_5 H_6}{H_7} = 0 \tag{3.35}$$

where

$$\begin{aligned} H_6 \equiv & -2c+2p+r+a(-1+2\lambda(u+u\beta-q(p+r)(1+t\lambda)+t(-1+u(1+\beta)\lambda)))+c_0 \\ & +2\lambda t(-c(qr+pq(2+\beta)+2(t-u(1+\beta)))-ct(qr-2u(1+\beta)+pq(2+\beta))\lambda \\ & -r(1+\beta)u(1+t\lambda)+p^2q(2+\beta)(1+t\lambda)+p(-2u(1+\beta)+qr(2+\beta)(1+t\lambda) \\ & +t(2+\beta-2u(1+\beta)\lambda))+(q(p+r)+t-u(1+\beta)+t(q(p+r)-u(1+\beta))\lambda)c_0) \\ & -2q(p+r)\lambda(-1+2\lambda(-t+u+u\beta+tu(1+\beta)\lambda)) \\ & +2cq\lambda(1+\lambda(-2u(1+\beta)+2q(p+r)(1+t\lambda)+t(3-2u(1+\beta)\lambda))))^2 \\ H_7 \equiv & 4((-1+2u(1+\beta)\lambda)(-1+2\lambda(-2t+u+u\beta+tu(1+\beta)\lambda)) \\ & -2q(p+r)\lambda(-1+2\lambda(-t+u+u\beta+tu(1+\beta)\lambda)) \\ & +2cq\lambda(1+\lambda(-2u(1+\beta)+2q(p+r)(1+t\lambda)+t(3-2u(1+\beta)\lambda))))^2 \end{aligned}$$

Condition (3.35) is satisfied when

$$r = \frac{H_8 + H_9 + H_{10}}{H_{11}} \tag{3.36}$$

where

$$H_8 \equiv -a - 2c + 2p + 2(-(a + 2c - 2p)(pq + t - u) + (p(pq + t - 2u) + au + c(-pq + 2u))\beta)\lambda$$
$$H_9 \equiv +2t(a(-pq + u + u\beta) - (c - p)(-2u(1 + \beta) + pq(2 + \beta)))\lambda^2$$
$$H_{10} \equiv (1 + 2\lambda(t - u(1 + \beta) - tu(1 + \beta)\lambda + p(q + qt\lambda)))c_0$$
$$H_{11} \equiv -1 + 2(aq + cq - 2pq + u - pq\beta + u\beta)\lambda(1 + t\lambda) - 2q\lambda(1 + t\lambda)c_0$$

Chapter 4

Asymmetric ISP competition and Network Neutrality[1]

This chapter performs an analysis on the investment incentives of Internet Service Providers (ISPs) under two different network regimes, network neutrality and network discrimination, taking into account ISP competition and congestion in online traffic.

Internet is one of the main network of communications nowadays, offering different types of services to users, such as email, browsing, or access to information services. The internet market can be seen as a two-sided market, where two sides of a market use a platform to interact and exchange information.[2] An ISP is a platform that brings consumers of contents and content providers together through an internet connection. ISPs connect consumers, also called end-users, that become able to access online content on the internet. The online contents that are requested by end-users are produced by Content Providers (CP). Examples of ISPs are AT&T or Verizon in the U.S. market, Skycom or O_2 in the European market, and examples of CPs are Yahoo, Google, Youtube, Facebook, Skype, or Messen-

[1] Joint with António Freitas

[2] For a survey of two-sided markets, see for example, Rochet & Tirole (2006) and Armstrong (2006).

ger,. . . among others.

Due to the privileged role that ISPs play in this market as the managers in data traffic, a large debate both in the U.S. and in Europe has arisen in recent years over what type of control can ISPs have with respect to data traffic management on the internet. One of the principles that have ruled in internet until recent years is nondiscrimination requirements, as in the telecommunications market. *Network neutrality* is the principle that all internet traffic should be treated equally. There are two possible interpretations of what net neutrality is. According to one interpretation, it means that ISPs cannot distinguish between data packets and therefore cannot determine their origin, hence they cannot charge a fee to content providers for delivering it to end-users. According to another interpretation, network neutrality means that ISPs cannot engage in traffic management by, for example, prioritizing traffic and favoring certain data packets over others (Schuett (2010)).

In this paper we adopt the second interpretation of network neutrality, where ISPs cannot prioritize certain traffic and only sell one type of internet service to content providers. One of the main points in the net neutrality debate is the innovation incentives of players in the market. Some net neutrality supporters, including consumer advocates, online companies and some technology companies, argue that discrimination on the internet will put newer CPs at a disadvantage and slow innovation in online services. On the other side of the debate, opponents of net neutrality, such as telecommunications providers, argue that prioritization of bandwidth is necessary to add revenue and to invest in increasing the capacity of the network to provide a wider access to more consumers. They also argue that net neutrality can bring negative consequences for innovation and competition by making it more difficult for ISPs to recoup their investments in broadband networks.

We study the network capacity investment incentives of ISPs when competing under a neutral regime versus a non-neutral regime, i.e., being able to sell different qualities of internet services to content providers, perform second-degree price discrimination. Also, we introduce a measure of discrimination to capture the quality degradation of traffic and how its intensity can affect the results.

Several authors have already focused on this topic under different frameworks. Among others, Hermalin & Katz (2007) follow a contractual approach by considering a monopolist ISP who charges different fees to CPs for different qualities of online connection, where CPs differ in the attractiveness of their content but the ISP has no information on the type of CPs, offering a menu of contracts to screen CPs. Choi & Kim (2010), Cheng et al. (2011), and Krämer & Wiewiorra (2009) take into account network congestion following queuing theory when evaluating the investment incentives both in the short and long run. However, these authors have only considered a monopoly setting.[3] The main contribution of our paper is to introduce ISP competition with internet congestion. This way we extend the approach of Cheng et al. and of Choi & Kim. We adopt a simpler congestion approach than queuing theory that delivers the same qualitative results as Choi & Kim's model under a monopoly setting.

The model of internet and content service developed in our paper is a two-sided market where consumers have heterogeneous preferences with respect to asymmetric ISPs and CPs. ISPs and CPs compete for end-users in a duopoly. We assume multi-homing of CPs, that is, content providers can provide content to users through more than one ISP. There are other network structures over which contents are delivered to consumers, however we chose multi-homing since it is possible to capture the direct interaction between each content provider and each internet service provider.

We consider internet congestion a necessary element to the analysis since otherwise all the effects of a discriminatory regime on the shift of users between contents

[3]Other authors consider ISPs competition but in different frameworks. For instance, Baake & Mitusch (2007) study ISPs Cournot and Bertrand competition when consumers face congestion externalities. However, in their paper, the authors do not consider the network discrimination problem: ISPs are not allowed to prioritize one of the contents. In another paper, Musacchio et al. (2009) model competition between several ISPs. In this case, the focus in not on consumers externality due to congestion but on the value consumers attribute to CPs investment in content quality.

could not correctly be accounted for. Also, introducing competition among ISPs is an important contribution to the literature for the two following reasons. Firstly, although an ISP monopoly may be realistic in some geographic areas of lower population density and where the costs to penetrate are higher, the offer of ISP connection is large and ISPs compete over end-user subscription in large urban areas. Secondly, the European legal framework is set such that the net neutrality debate is not regarded as being a problem. In Europe it is argued that the market for internet service is set to be sufficiently competitive to solve unreasonable network management, that is, consumers that are not satisfied with discriminatory practices of their ISP can easily switch to another ISP which does not discriminate or discriminates less. The regulation only intervenes under situations of unacceptable degradation of services by implementing measures to protect consumers (Sluijs (2010)). Hence, we approach this scenario as well.

Our main result is that competition between networks provides lower investments in capacity when network discrimination is allowed. The result holds if ISPs can charge a high fee to CPs. Contrary to what happens in a monopolistic network, ISPs market size is not fixed but can be increased to the detriment of the competitor. Network discrimination harms part of the consumers, hence end-users migrate to the network who penalize them less. If networks have asymmetric capacities, this translates into a transfer of consumer from the larger to the smaller network. As a result, the smaller network has lower incentives to invest because network discrimination partially reduces the gap between the two networks without requiring to increase capacity. If the larger network invests in capacity, it can mitigate users outflow but this would reduce the revenue from the priority fee. However, the loss of consumer due to network discrimination can be compensated if the fee charged to CPs is high enough. The overall effect is that ISPs prefer to discriminate between content and have lower incentive to expand the capacity of their networks.

Outline. The rest of the paper is organized as follows. We introduce the model in Section 2 and briefly illustrate how congestion works in the network system.

Section 3 shows how our work is related to the seminal paper of Choi & Kim (2010). In Section 4, we present the equilibrium when there is duopoly competition among ISPs under a neutral regime. This constitutes a benchmark for future comparisons. In section 5, we determine the equilibrium outcomes where network discrimination is allowed. In Section 6, we compare investments incentives between the two regimes described in the previous sections. Section 6 provides conclusions and some policy indications. All proofs are in the Appendix.

4.1 The Model

4.1.1 Basic Model

Internet users have access to online contents through an existing broadband network. CPs are the producers and deliverers of online content to end-users. To deliver the content, CPs must use a broadband network which is provided by ISPs. The network structure we assume is such that each CP is directly associated to more than one ISP, called multi-homing. We assume this network structure throughout the paper. Also, we assume that all consumers are single-homing, i.e., they choose only one ISP with which to connect to.[4] We assume that both the market for online service and the market for content are duopolies. Hence, we denote the ISPs in the market as ISP_A and ISP_B, and the CPs as CP_1 and CP_2.

Concept of Net Neutrality. ISPs are the connection between CPs and end-users and they may have the possibility to manage how contents are delivered to internet users. Under net neutrality, ISPs are not allowed to discriminate between content providers over the speed at which contents are delivered, i.e., the time users have to wait to access all contents is the same, nor any price is charged to CPs for the service. Under network discrimination, each ISP is able to sell a priority service to one of the CPs and charge a fee f. We assume this fee is attained by

[4] Although in practice the possibility of multi-homing exists, it is not common in end-user behavior. The great majority of users purchases internet services from one provider only.

Nash bargaining solution, where θ is the bargaining power of the ISP selling the priority service. The fee calculated is always an amount between the maximum willingness to pay of the more efficient CP and the maximum willingness to pay of the less efficient CP. This priority gives the content's end-user the right to receive it ahead of any other user requesting a non-priority content. Therefore, the time end-users have to wait for the content requested varies and this affects the utility they derive from accessing the content.

Consumers. There is a mass (1×1) of consumers with heterogeneous preferences with respect to: (i) the two contents being delivered; and (ii) the two internet services that provide the online connection. Each user is described by the pair (x, y), where $x \in [0, 1]$ denotes the preferred type of content and $y \in [0, 1]$ the preferred type of service. The intrinsic value of the content and the service that each user derives is $v > 0$, and it is assumed that v is high enough so the market is always covered. If a consumer is not able to access to her preferred type of content, the loss of utility she faces is represented by a linear transport cost $t > 0$ which multiplies the distance between the preferred and the received content. Similarly, if a consumer cannot connect to her preferred service, the associated disutility is given by a linear transport cost $s > 0$ which multiplies the distance between the preferred and the received service.

We assume each consumer has demands for one of the two CPs and for one of the two ISPs (single-homing demand). When a consumer decides which type of content to browse, her utility is also influenced by the time she has to wait to browse the content. Such waiting disutility, w, is assumed to be independent on the content type and is determined by the degree of congestion of the network. However, when she decides which ISP joining to, the user is not able to determine the exact degree of congestion because, as a matter of fact, it is impossible to know it before joining a network and browsing a content. So she uses the network capacity of the ISP hosting the content, her perception of connected users, and the eventual prioritization/discrimination of the content as congestion proxy parameters to establish her waiting disutility.

In practice, a user can observe the network capacity and the priority choices of an ISP, but she does not know how many users in her neighborhood, block, building, or house, are simultaneously connected to the same content through the same ISP. And it is the number of these users which ultimately determines congestion. Only once she has connected to an ISP, she can observe the actual waiting disutility of browsing a given CP and eventually choose to browse the other if it provides a higher utility. However, she cannot switch to the other ISP due to high switching costs (usually due to a permanency obligation). So the ISP choice is done before knowing the exact waiting disutility.

Users request for contents is captured by demand intensity parameter λ, representing a demand intensity which is the same for all the consumers. It can represent the time a user spends browsing a content or the number of clicks in the page of the content. Finally, a consumer pays a price p to the ISP that give her the access.

Therefore, the utility of a consumer (x, y) when she browses content type $\bar{x}$ through an Internet service type $\bar{y}$ has the form

$$u_{x,y} \equiv v - t \cdot |x - \bar{x}| - s \cdot |y - \bar{y}| - w - p.$$

Internet Service Providers. ISP_A and ISP_B compete to sell the internet connection service to end-users by setting prices p_j, with $j = A, B$. They compete in the Hotelling manner, by offering two different types of services. For instance, one could provide an Internet access service which includes free movies and the other an Internet access service including free sport programs. Ultimately one of them could provide Internet and television access, and the other Internet and phone calls access.

We suppose the two ISPs are exogenously located at the extreme points of the possible consumers service preference line. That is, ISP_A offers an Internet service type $y = 0$ while ISP_B offers a service type $y = 1$. So consumer y faces a transport cost sy from choosing ISP_A and $s(1 - y)$ from choosing ISP_B.

Each ISP has a network capacity of μ_j. In the short run, μ is fixed. In the long run, it is endogenous to the model. Larger network capacity is associated with a shorter period of time that the end-user waits for the content to be delivered.

If it is allowed to discriminate, each ISP decides which CP to sell the priority to through the payment of a fee f_j. Managing the traffic (i.e., assigning a priority) does not imply any cost. Moreover, we assume that all the costs to provide internet services are already sunk. So ISP_j profits are given by

$$\pi_j = \sigma_j p_j + f_j$$

where σ_j denotes the total market share of an ISP. When network discrimination is not allowed, $f_j = 0$.

Content Providers. CP_1 and CP_2 compete to deliver content to end-users. Consumers are interested in browsing only one of the two contents because lack of time (for instance, a British user could decide if watching news in the BBC or in the ITV website), because their friends use mainly one chat (i.e. Messenger vs. Skype), because they like only American (vs. British) movies (e.g. in the UK, this means to choose Netflix or Amazon/LoveFilm). CP_1 is assumed to be exogenously located at point 0 while CP_1 at point 1 of contents preference space. An end-user located at x incurs a transport cost tx to consume CP_1's services, while she incurs $t(1-x)$ to consume contents from CP_2.

Each content provider i adopts a business model that delivers contents without receiving any payment from end-users. The revenues are generated exclusively through advertisement, obtaining r_i from advertisers for each content requested.[5] Each content provider faces a request serving cost of c_i so it has a mark-up of $m_i = r_i - c_i$. Hence the profit of each CP_i is given by $m_i \lambda \sigma_i$, where λ is the demand intensity of requests and σ_i is the market share of content provider i.

Timing. The sequence of players' choices is the following.

1. Each ISP decides the capacity of its network (μ_j). This stage can only be played in the long run. In the short run, capacity is given.

[5]Each request is measured by one user's clicks on the internet page associated to the content, the λ parameter of our model.

2. ISPs decide their pricing strategies, p_j.

3. Each ISP negotiates the priority price, f_j with a CP. This stage is not feasible when network discrimination is not allowed.

4. Consumers observe prices and capacities (and, eventually, priorities) and pick an ISP.

5. Given their ISP choice, consumers pick one CP.

The model is solved by backward induction and the equilibrium concept is the sub-game perfect Nash equilibrium.

Market representation. The market we consider can be represented as a variation of a two-dimensional Hotelling model where consumers first choose one characteristic (the ISP) of the final product (the service and content provision) and, then, they choose the other characteristic (the CP). In Figure 4.1, we represent one example of market sharing.

ISP_A provides a service type $y = 0$, so all users type $(x, 0)$, for all $x \in [0, 1]$, do not face any transport cost if they connect to ISP_A. Similarly, ISP_B provides the preferred service of users type $(x, 1)$. CP_1 distributes a content type $x = 0$, which exactly meets the content preference of consumers type $(0, y)$, for all $y \in [0, 1]$. CP_2 distributes a content type $x = 1$ that is the most preferred of users type $(1, y)$.

Users type (x_A, y) are indifferent between the two CPs if they connect to ISP_A. Similarly, users type (x_B, y) are indifferent between the CPs when they connect to ISP_A. Users type (x, y_1) are indifferent between the two ISPs if they browse CP_1 while users type (x, y_2) are indifferent if they browse CP_2.

Region $(1, A)$ represents the market share of users who connect to ISP_A and browse CP_1. Similarly, area $(2, A)$ represents the market share of users who connect to ISP_A and browse CP_2. The remaining areas refer to ISP_B. So, for instance, the sum of $(1, A)$ and $(2, A)$ gives the total market share of ISP_A, while the sum of $(1, A)$ and $(1, B)$ gives the total market share of CP_1.

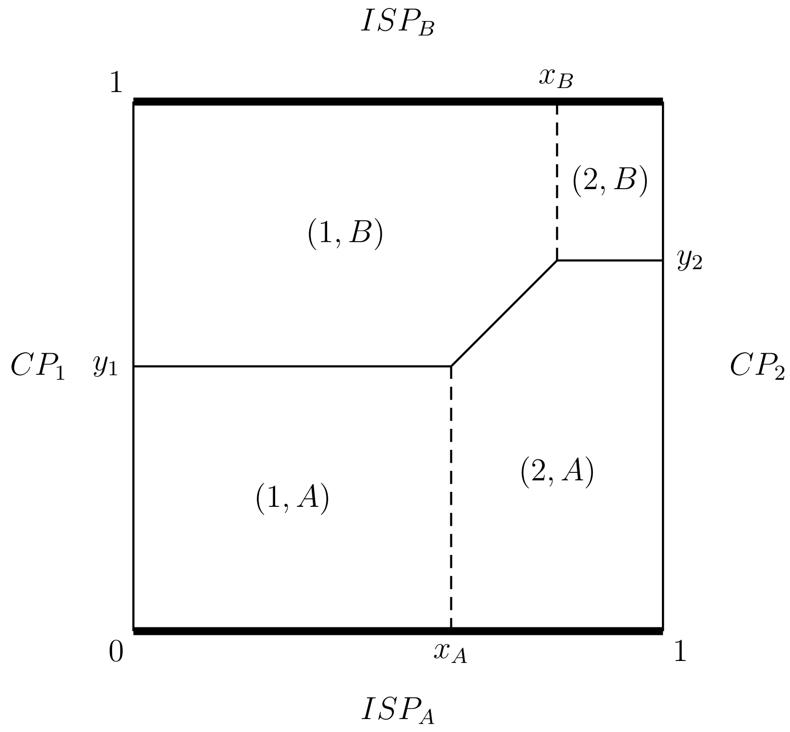


Figure 4.1: Market representation

4.1.2 Preliminaries: Congestion in the network system

In order to model internet network congestion and use it in a framework of ISP competition, we set up an approach that is able to capture some of the qualitative features of the queuing theory that are of interest to our study. This modeling choice allows us to simplify the analysis and to obtain tractable solutions when introducing ISP competition.

Under network neutrality requests of contents are treated equally. Each end-user that subscribes an online service network from one of the two ISPs, has the expected waiting disutility of w, given by:

$$w = y - (\mu_j - \lambda)$$

where y is the consumer perception of ISP_j's market share the end-user subscribes to, λ is the demand intensity of content and μ_j is the ISP_j's network capacity, where $\mu_j > \lambda$. The waiting disutility increases with the perceived market share of the ISP. A higher market share perception of one ISP implies more people are accessing the network. Also, waiting disutility increases with the demand intensity and decreases with the broadband network capacity of the ISP.

By consumer perception of an ISP's market share, we mean that consumers cannot forecast the exact market share of an ISP (which ultimately determines the actual level of congestion) before connecting to it. However, consumers could have some "perception" of the possible market share of an ISP when comparing ISPs with one another. Given a type x-users, here we set the share y of that type x-users as a proxy of the market share of ISP_A. The rationale is that a user could know the share y of users of her same type x which connect to ISP_A. But she cannot know the share of other types x who connect to the same ISP.[6]

In the discriminatory network, a user who requests content from a priority class has an expected waiting disutility of w^p, given by:

$$\begin{aligned} w^p &= w - \alpha(1-x)\left(\frac{1}{\mu_j} - \lambda\right) = \\ &= y - (\mu_j - \lambda) - \alpha(1-x)\left(\frac{1}{\mu_j} - \lambda\right). \end{aligned} \tag{4.1}$$

Waiting disutility w^p is the difference between the waiting time w that users face under network neutrality and an amount of time that depends on the share of users that request content from the priority class x, $\alpha > 0$ which is the degree of priority that the ISP may impose to the content. In contrast, the user that requests content without priority faces an expected waiting disutility of w^d, given

[6] Literature in industrial organization provides many cases where the rationality hypothesis is dropped. See, for instance, the review of Ellison (2006). In our context, for instance, consumers are assumed to be able to solve a massive game theory problem with all other consumers. In the telecommunication framework, an interesting example of how the bounded rationality assumption better fits with empirical evidence, is provided in Möbius (2001).

by:

$$w^d = w + \alpha x \left(\frac{1}{\mu_j} - \lambda \right) =$$
$$= y - (\mu_j - \lambda) + \alpha x \left(\frac{1}{\mu_j} - \lambda \right) \tag{4.2}$$

where here α reflects the extra disutility caused by requesting a discriminated content. If α is set at 0, this means that the priority effect is null and both priority and non-priority expected waiting time are equal to w and we would be in the neutral regime.

Assumption 4.1. *We assume that $\alpha > 0$ and $\lambda\mu_j < 1$. This implies that network capacity is always enough to serve online users' requests.*

As in Choi & Kim's model for internet congestion, there are three properties that our approach to congestion satisfies. First, the end-user faces a higher waiting cost by requesting a second-priority content instead of a prioritized one, that is,

Property 4.1. $w^d > w > w^p$

This property is established by computing the difference between waiting times w^p and w^d. Also, the difference between waiting times is constant regardless of the distribution of total traffic across different priority classes.

Second, we find that the difference in waiting time becomes smaller as the network capacity increases, that is,

Property 4.2. $\frac{\partial(w^d - w^p)}{\partial\mu} < 0$

The marginal reduction in waiting time for the priority service from an expansion in ISP capacity expansion decreases as the network capacity level becomes high. The intuition is that as an ISP increases capacity, other things being equal, the priority service becomes relatively less attractive with respect to the non-priority service.

Third, network discrimination does not change the share-weighted average waiting disutility:

Property 4.3. $x \cdot w^p + (1 - x) \cdot w^d = w$

Assigning priorities does not improve aggregated waiting disutility but it only is a way of managing contents speed. Some users improve their utility, others get worse due to higher waiting disutility. On average (indeed, on an average weighted by the proportions of users who get better and get worse), the waiting disutility of a network does not change.

4.2 Monopolistic ISP

Given that our model is similar to Choi & Kim (2010), this section analyzes the case where there is an ISP monopoly and users have heterogeneous preferences with respect to contents à la Hotelling. When using our approach to network congestion, we attain the same qualitative results as Choi & Kim with respect to the long run incentives of ISPs to invest in network capacity, both under a neutral network and a discriminatory network regime.

4.2.1 Short run Analysis

In a neutral network regime, end-users pay a subscription price p to the ISP (with monopoly, $y = 1$) and have preferences over contents à la Hotelling to choose content of only one CP. The indifferent user x^* between the two content providers is defined as

$$v - w - tx^* - p = v - w - t(1 - x^*) - p$$

where users $x \leq x^*$ choose CP_1's content and those $x > x^*$ choose CP_2's. Since CPs are symmetrically located, the solution is $x^* = \frac{1}{2}$, i.e., the two CPs share the market equally. The ISP sets the price p that maximizes its profit, π_M, conditional on full market coverage, implying a positive utility for the indifferent user $x = \frac{1}{2}$. The equilibrium profit of the ISP is $\pi_M^* = p^* = v - 1 + \mu - \lambda - \frac{t}{2}$.

Under a discriminatory regime, the ISP is allowed to charge a fee f to a CP for providing a priority service, so the delivery time of a content depends on the content a user chooses. The indifferent user $\tilde{x}$ between the priority service and the non-priority service is

$$v - w^p(\tilde{x}) - t\tilde{x} - p = v - w^d(\tilde{x}) - t(1 - \tilde{x}) - p.$$

where the tilde is used to denote the variables under the regime with discrimination. The solution to this problem is $\tilde{x} = \frac{1}{2} + \alpha\frac{1-\mu\lambda}{2t\mu}$ and we easily observe that the priority CP has larger market share than the non-priority CP, since $\tilde{x} > \frac{1}{2}$. We focus on interior solution where both CPs stay in the market assuming a sufficiently high transport cost $t > \alpha\frac{1-\mu\lambda}{\mu}$. In this case, the market share of the content provider with the first priority is stable and decreases as the ISP's capacity increases, that is, $\frac{\partial\tilde{x}}{\partial\mu} < 0$. This result follows Choi & Kim (2010). The ISP fixes the subscription price to maximize its profit, given by $\tilde{p} + f$, conditional on covering the market. The priority fee f is calculated through Nash bargaining, and is $f = [m_2 + \theta(m_1 - m_2)](2\tilde{x} - 1)\lambda$.[7] Therefore, the ISP's profit in the discriminatory network is $\tilde{\pi}_M = \tilde{p} + f = (v - 1 + \mu - \lambda\tilde{x} - t\tilde{x}) + [m_2 + \theta(m_1 - m_2)](2\tilde{x} - 1)\lambda$.

4.2.2 Long Run Analysis

In the long run, the incentives to invest in network capacity are reflected in the partial derivatives of profits on the capacity. Under network neutrality, $\frac{\partial\pi_M^*}{\partial\mu} = 1$, so there is always incentive to invest. Under network discrimination, we obtain:

$$\frac{\partial\tilde{\pi}_m}{\partial\mu} = 1 - \frac{\alpha(1-\tilde{x})}{\mu^2} - \left(\frac{\alpha}{\mu} - \alpha\lambda + t\right)\frac{\partial\tilde{x}}{\partial\mu} + 2\lambda(m_1 + \theta(m_1 - m_2))\frac{\partial\tilde{x}}{\partial\mu}$$

To see whether the incentives to invest are higher under the discriminatory regime than under the neutral regime, we need to study the sign of the difference ($\frac{\partial\tilde{\pi}_M}{\partial\mu} - \frac{\partial\pi_M^*}{\partial\mu}$). As in Choi & Kim, the sign is undetermined. The effect of capacity expansion on the sale price of the priority is negative ($2\lambda(m_1 + \theta(m_1 - m_2))\frac{\partial\tilde{x}}{\partial\mu} < 0$), while the effect of capacity expansion on end-user subscription price due to discrimination is undetermined ($\frac{\alpha(1-\tilde{x})}{\mu^2} - (\frac{\alpha}{\mu} - \alpha\lambda + t)\frac{\partial\tilde{x}}{\partial\mu}$).

Therefore, for interior solutions where both CPs remain in the market under network discrimination, the incentives to invest under the discriminatory regime

[7] The fee is an amount set between CP_2's and CP_1's willingness to pay for the priority service.

than under the neutral regime, and the overall effect remains undefined, such as in Choi & Kim. On one hand, under network neutrality the ISP always has incentives to invest in network capacity. On the other hand, under a discriminatory regime, the ISP continues to face two effects that may go in the opposite direction. The effect of capacity expansion on the sale price of priority is negative, since there are less users that choose the priority content, but the effect on the subscription fee of the end-user in undetermined. So, under our linear approach there may exist cases where the incentives to invest under network discrimination are higher than under network neutrality.

4.3 Benchmark: ISP competition with Network Neutrality

In this section and in the sections that follow, we analyze the market when ISP compete for the subscription of users. In order to reduce the number of cases to analyze, we assume that ISP_A has a greater network capacity, that is $\mu_A > \mu_B$. In particular, in the current section we analyze the equilibrium outcome when ISPs are not allowed to discriminate between contents. This represents a benchmark to assess if allowing network discrimination leads to lower or higher incentives to invest in capacity expansion.

In a neutral network, the waiting disutility of each user connected to ISP_j, $j = A, B$, is the same independently of the content requested. So, denoting by $w_{i,j}$ the waiting disutility of a user who browses CP_i under ISP_j, we have

$$w_{1,j} = w_{2,j} = y - (\mu_j - \lambda).$$

The delivery speed may not be the same in the whole online network, but it is so in each network. The indifferent user between contents subscribing an internet connection to ISP_A is defined as:

$$v - y + (\mu_A - \lambda) - tx - sy - p_A = v - y + (\mu_A - \lambda) - t(1 - x) - sy - p_A$$

which provides exactly the same solution as in the monopolist case, $x^* = \frac{1}{2}$. The indifferent user subscribing ISP_B is the same.

When choosing which ISP to subscribe to, the indifferent consumer is characterized by:

$$v - y + (\mu_A - \lambda) - sy - p_A = v - (1 - y) + (\mu_B - \lambda) - s(1 - y) - p_B. \quad (4.3)$$

The solution is:

$$y^* = \frac{1}{2} + \frac{p_B - p_A}{2(1+s)} + \frac{\mu_A - \mu_B}{2(1+s)}. \quad (4.4)$$

The market share of users who are connected to ISP_A and browse CP_1 is $\sigma^*_{1,A} \equiv x^* y^* = \frac{y^*}{2}$. Similarly, the market share of users who connect to ISP_B and browse CP_2 is $\sigma^*_{2,A} \equiv (1 - x^*)y^* = \frac{y^*}{2}$. The total market share of ISP_A is given by $\sigma^*_A \equiv \sigma^*_{1,A} + \sigma^*_{2,A} = y^*$ and is dependent on the network capacity differential, as well as the price differential. It depends positively on its own network capacity but negatively on price, and the opposite with respect to the competitor decisions on capacity and price. In Figure 4.2, we represent the market sharing in case of net neutrality.

The equilibrium prices set by the ISPs are the result of the best responses of each ISP. In the case of ISP_A, profit is now the subscription price it sets times the market share it captures. Prices set by each ISP are the solution of:

$$\max_{p_A} \pi_A = y^*(p_A, p_B)\, p_A \qquad \text{s.t.} \quad v - y^* + (\mu_A - \lambda) - sy^* - p_A \geq 0$$

And the same applies to ISP_B:

$$\max_{p} \pi_B = (1 - y^*(p_A, p_B))\, p_B \qquad \text{s.t.} \quad v - (1 - y^*) + (\mu_B - \lambda) - s(1 - y^*) - p_B \geq 0$$

The (Nash) equilibrium prices are the following:

$$p^*_A = 1 + s + \frac{\mu_A - \mu_B}{3}$$
$$p^*_B = 1 + s - \frac{\mu_A - \mu_B}{3}$$

In order to guarantee that both ISPs stay in the market we assume their products are sufficiently differentiated:

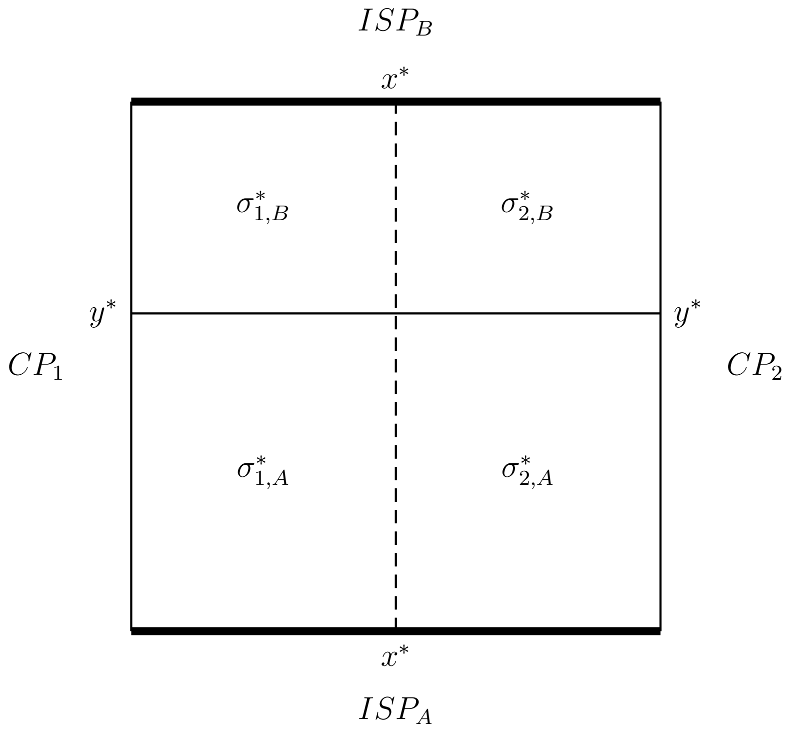


Figure 4.2: Net neutrality: market sharing.

Assumption 4.2. $s > \frac{(\mu_A - \mu_B)}{3} - 1$

Assumption 2 means that the internet service providers must be sufficiently differentiated depending on the network capacity differential. Hence the equilibrium market share is

$$\sigma^*_A = \frac{1}{2} + \frac{\mu_A - \mu_B}{6(1+s)} \tag{4.5}$$

The profit of ISP_A is given by:

$$\pi^*_A = \frac{(1+s)}{2} + \frac{(\mu_A - \mu_B)}{3} + \frac{(\mu_A - \mu_B)^2}{18(1+s)}$$

Hence, in the long run the incentives to invest in network capacity are dependent on the profit derivative with respect to μ_A

$$\frac{\partial \pi^*_A}{\partial \mu_A} = \frac{1}{3} + \frac{\mu_A - \mu_B}{9(1+s)}$$

which is always positive for an interior equilibrium solution. Hence, under the neutral network, ISP_A always has incentive to invest in network capacity. This is not surprising since more capacity always provides a higher market share of end-users. Lets now look at ISP_B's profit and at the long run incentives to invest in network capacity:

$$\pi_B^* = \frac{(1+s)}{2} - \frac{(\mu_A - \mu_B)}{3} + \frac{(\mu_A - \mu_B)^2}{18(1+s)}$$

$$\frac{\partial \pi_B^*}{\partial \mu_B} = \frac{1}{3} - \frac{\mu_A - \mu_B}{9(1+s)}$$

which is always positive under Assumption 4.2. So, in the neutral network under competition, both ISPs have an incentive to invest in network capacity.

4.4 ISP competition with Network Discrimination

In this section, we analyze ISP competition when networks are allowed to discriminate contents. The equilibrium outcomes will be compared with the ones of network neutrality in the following Section 4.5. Here we look for interior equilibria and solve by backward induction the ISPs maximization problem considering the timing presented in Section 4.1. We first consider the last stages of the game, where consumers choose which ISP to connect and which ISP to browse. Then, in previous stages, we analyze how ISPs assign the priority service to one of the CP on a fee payment and set prices to provide connection to end-users.

4.4.1 Consumers' choice

Users observe the announced priority choice of each ISP, the connection prices, and the capacity of the two networks. Then, they (1) subscribe their preferred ISP and (2) decide which CP to browse. User type (x, y) receives a utility $u_{i,j}$

when she browses CP_i under network j:

$$u_{i,j} = \begin{cases} v - w_{1,A}(\cdot) - xt - ys - p_A, & \text{if her choice is } (1,A) \\ v - w_{2,A}(\cdot) - (1-x)t - ys - p_A, & \text{if her choice is } (2,A) \\ v - w_{1,B}(\cdot) - xt - (1-y)s - p_B, & \text{if her choice is } (1,B) \\ v - w_{2,B}(\cdot) - (1-x)t - (1-y)s - p_B, & \text{if her choice is } (2,B) \end{cases}$$

where waiting disutilities depend on the prioritized content of a network.

2nd stage: CP choice

Depending on the ISP subscribed in the previous stage, users decide which CP to browse. Among all the users (x, y) with a given preference y for an ISP, a user who subscribes ISP_j is indifferent between CP_1 and CP_2 when $u_{1,j} = u_{2,j}$. For instance, if ISP_A prioritizes CP_1, a user who has connected to network A in the previous stage has the following utilities:

$$u_{i,A} = \begin{cases} v - \left(y - (\mu_A - \lambda) - \alpha(1-x)\left(\frac{1}{\mu_A} - \lambda\right)\right) - xt - ys - p_A, & \text{if chooses } CP_1 \\ v - \left(y - (\mu_A - \lambda) + \alpha x\left(\frac{1}{\mu_A} - \lambda\right)\right) - (1-x)t - ys - p_A, & \text{if chooses } CP_2 \end{cases}$$

where waiting disutilities are defined according to equations (4.1) and (4.2). All users type $(\tilde{x}_A, y)$ are indifferent between CP_1 and CP_2 when connected to ISP_A; where

$$\tilde{x}_A = \frac{1}{2} + \alpha\frac{1 - \lambda\mu_A}{2t\mu_A}.$$

The second addend denotes the priority effect: for a given y-type users, a fraction $\alpha\frac{1-\lambda\mu_A}{2t\mu_A}$ switches to CP_1 because it implies a lower waiting disutility. Note that, from the users perspective, CPs are symmetric. Hence, if ISP_A prioritizes CP_2, the indifferent user would be $(1 - \tilde{x}_A, y)$, where

$$1 - \tilde{x}_A = \frac{1}{2} - \alpha\frac{1 - \lambda\mu_A}{2t\mu_A}.$$

Therefore, to avoid an excessive notation, we denote by $(1 - \tilde{x}_A, y)$ and $(1 - \tilde{x}_B, y)$ the indifferent users in networks A and B, respectively, when the prioritized content is CP_2.[8]

Proposition 4.1. *When ISP_B assigns the priority to a CP, that CP attracts a larger proportion of users of a given* y*-type than in the case where the priority is provided by ISP_A. That is,*

$$1 - \tilde{x}_B < 1 - \tilde{x}_A < x^* < \tilde{x}_A < \tilde{x}_B.$$

Proof. See Appendix 4.B. □

This result is due to the congestion externality: network A has higher capacity so its users assign a lower disutility to congestion. Hence, they are less likely to switch away from their preferred CP (we described this effect in Property 4.2).

In this case, where ISP_A sells the priority service to CP_1, the waiting disutilities of its users are[9]

$$\begin{aligned} w_{1,A}(\tilde{x}_A) =& y - \mu_A + \lambda - \alpha(1-\tilde{x}_A)\left(\frac{1}{\mu_A} - \lambda\right) \\ w_{2,A}(\tilde{x}_A) =& y - \mu_A + \lambda + \alpha\tilde{x}_A\left(\frac{1}{\mu_A} - \lambda\right). \end{aligned}$$

Similar disutilities are associated with ISP_B (when a user connects to ISP_B, waiting disutilities depend on its network capacity, μ_B, and on the perceived congestion, $1 - y$).

[8]We are looking for an interior equilibrium, that is an equilibrium where the users connected to each ISP are browsing both CP_1 and CP_2. One necessary condition is provided by $\tilde{x}_j \in \left(\frac{1}{2}, 1\right)$, $\forall j$, that is $t > \alpha\frac{1-\lambda\mu_B}{\mu_B}$. This implies the contents providers are sufficiently differentiated so there is always an indifferent x-type user between them.

[9]Otherwise, if it sells the priority service to CP_2, the waiting disutilities of its users are

$$\begin{aligned} w_{1,A}(1-\tilde{x}_A) =& y - \mu_A + \lambda + \alpha(1-\tilde{x}_A)\left(\frac{1}{\mu_A} - \lambda\right) \\ w_{2,A}(1-\tilde{x}_A) =& y - \mu_A + \lambda - \alpha\tilde{x}_A\left(\frac{1}{\mu_A} - \lambda\right). \end{aligned}$$

Proposition 4.2. *When an Internet Service Provider assigns a priority service to one content, the waiting disutility associated with the prioritized content improves while the disutility of the discriminated content worsens. In particular,*

a) if two ISPs prioritize the same CP, the disutility gap between priority and discriminated content increases less in the larger network (ISP_A) than in the smaller (ISP_B). In particular, the disutility gap of a user connected to the smaller network is greater than in the larger network by a proportion

$$\omega(\mu_A, \mu_B) \equiv \frac{\alpha^2(\mu_A - \mu_B)(\mu_A + \mu_B - 2\lambda\mu_A\mu_B)}{2t\mu_A^2\mu_B^2}. \tag{4.6}$$

b) If the two ISPs prioritize two different CPs, the disutility gap in the smaller network is greater by the same value $\omega(\mu_A, \mu_B)$.

Proof. See Appendix 4.B. □

This result is implied by Property 4.1 of the waiting function and follows the same intuition as Choi & Kim (2010). Prioritizing a content creates a disutility gap between the priority and the discriminated contents. However, the gap is larger in smaller networks because, as stated by Choi & Kim, higher network capacity makes congestion less important. Hence, less users switch to the prioritized service and the average gap is lower (Property 4.2).

This result also applies in the case ISPs prioritize different CPs: within a network, the effect of discrimination on the indifferent user is the same, independently on the prioritized content (given that contents are symmetric to the users). When a network discriminates between two CPs, total waiting disutility faced by its users does not change, given that a fraction of consumers switch from the discriminated CP (with higher waiting disutility) to the prioritized CP (with lower waiting disutility), as stated in Property 4.3.

1st stage: ISP choice

Depending on their CP choice in stage 2, users anticipate the waiting disutilities according to their y-type. Proceeding with the above example where both ISPs sell the priority to CP_1, we have:

- all users type $(x \leq \tilde{x}_A, y)$ always browse CP_1. They decide which ISP connect to according to their preference y. Among all these users, there is one type $(x \leq \tilde{x}_A, y = \tilde{y}_1)$ who is indifferent between which of the ISPs joining to (that is, $u_{1,A} = u_{1,B}$);

- all users type $(x > \tilde{x}_B, y)$ always browse CP_2. They decide according to their preference y which ISP to join to. Among all these users, there is one type $(x > \tilde{x}_B, y = \tilde{y}_2)$ who is indifferent between which of the ISPs joining to (that is, $u_{2,A} = u_{2,B}$);

- all users type $(x \in (\tilde{x}_A, \tilde{x}_B], y)$ browse CP_2 under ISP_A and CP_1 under ISP_B. Among them, there is one type $(x \in (\tilde{x}_A, \tilde{x}_B], y = \tilde{y}_m(x))$ who is indifferent between browsing CP_2 under ISP_A or CP_1 under ISP_B.

In Figure 4.3, we represent the indifferent users in this case. For instance, user $(\tilde{x}_A, \tilde{y}_1)$ receives the same utility if connecting to ISP_B and browsing CP_1 or if connecting to ISP_A and browsing either CP_1 or CP_2.

Both ISP sell the priority to the same CP

In particular, in case both ISPs sell the priority to CP_1, a user type $(x \leq \tilde{x}_A, y)$ always chooses CP_1, independently on the ISP she joins, and is indifferent between the two ISPs when $u_{1,A} = u_{1,B}$.[10] The indifferent users are $(x \leq \tilde{x}_A, y = \tilde{y}_1)$, where[11]

$$\tilde{y}_1 \equiv y^*(p_A, p_B, \mu_A, \mu_B) + \frac{\omega(\mu_A, \mu_B)}{2(1+s)} - \frac{\alpha(\mu_A - \mu_B)}{4(1+s)\mu_A \mu_B}.$$

[10]The case where both ISPs sell the priority to CP_2 is symmetric to this case.

[11]Indifferent users type $(x \leq \tilde{x}_A, y = \tilde{y}_1)$ solve the following equation:

$$v - w_{1,A}(\tilde{x}_A, y) - xt - ys - p_A = v - w_{1,B}(\tilde{x}_B, 1-y) - xt - (1-y)s - p_B.$$

Note we added the argument y in the waiting disutility: now y-type's perception of the congestion in the two networks affects her ISP choice.

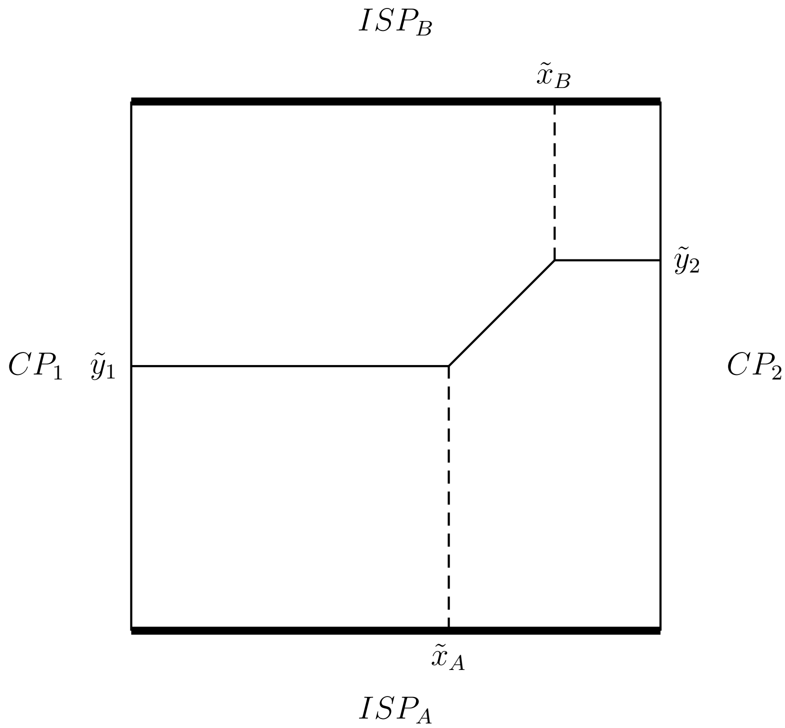


Figure 4.3: Indifferent users by type

A user type $(x > \tilde{x}_B, y)$ always chooses CP_2, independently on the ISP she joins, and is indifferent between the two ISPs when $u_{2,A} = u_{2,B}$. In this case, the indifferent users are $(x \leq \tilde{x}_A, y = \tilde{y}_2)$, where[12]

$$\tilde{y}_2 = y^*(p_A, p_B; \mu_A, \mu_B) + \frac{\omega(\mu_A, \mu_B)}{2(1+s)} + \frac{\alpha(\mu_A - \mu_B)}{4(1+s)\mu_A\mu_B}.$$

Comparing the indifferent user types $\tilde{y}_1$ and $\tilde{y}_2$, we observe that:

- the first term $y^*(p_A, p_B; \mu_A, \mu_B) = \frac{1}{2} + \frac{p_B - p_A + \mu_A - \mu_B}{2(1+s)}$ is the same as in net neutrality. It denotes how indifferent users are affected by price choices and network capacities;

[12] Indifferent users type $(x \leq \tilde{x}_A, y = \tilde{y}_2)$ solve the following equation:

$$v - w_{1,A}(1 - \tilde{x}_A, y) - (1-x)t - ys - p_A = v - w_{1,B}(1 - \tilde{x}_B, 1-y) - (1-x)t - (1-y)s - p_B.$$

- the term $\frac{\omega(\mu_A,\mu_B)}{2(1+s)} = \frac{\alpha^2(\mu_A-\mu_B)(\mu_A+\mu_B-2\lambda\mu_A\mu_B)}{4(1+s)t\mu_A^2\mu_B^2} > 0$ reflects the difference in the disutility gaps between the networks;

- the term $\frac{\alpha(\mu_A-\mu_B)}{4(1+s)\mu_A\mu_B} > 0$ denotes how much the difference in gaps in the two networks is due comparing discriminated (utility-worsened) users with prioritized (utility-improved) users across networks.

The case where ISPs sell priorities to different CPs

Indifferent users change if ISP_A sells the priority to one CP and ISP_B to the other. We consider the case where A prioritizes CP_2 while B prioritizes CP_1.[13] In this case, users type $(x \leq 1-\tilde{x}_A, y)$ always browse CP_1. Users type $(x > \tilde{x}_B, y)$ always browse CP_2, and users type $(x \in (1-\tilde{x}_A, \tilde{x}_B], y)$ browse CP_2 under A and CP_1 under B. Intuitively, compared to the case where both ISPs sold the priority to CP_1, now ISP_A gains more users who prefer CP_2: it prioritizes a content which is discriminated under ISP_B. Hence, the difference in waiting disutilities between A and B is more emphasized. Similarly, ISP_B gains more users who prefer CP_1.

Now users type $(x \leq \tilde{x}_A, y)$ always choose CP_1, independently on the ISP they join to. The indifferent users are[14]

$$\hat{y}_1 = y^*(p_A, p_B; \mu_A, \mu_B) + \frac{\omega(\mu_A, \mu_B)}{2(1+s)} - \frac{\alpha(\mu_A + \mu_B - 2\lambda\mu_A\mu_B)}{4(1+s)\mu_A\mu_B}$$

where we denote by $\hat{y}$ the fact that indifferent users are determined by priorities assigned to different CPs.

Users type $(x > \tilde{x}_B, y)$ always choose CP_2, independently on the ISP they join

[13] The other case, where CP_1 is prioritized by A and CP_2 by B, is symmetric.

[14] Indifferent users satisfy the following condition:

$$v - w_{1,A}(1-\tilde{x}_A, y) - xt - ys - p_A = v - w_{1,B}(\tilde{x}_B, 1-y) - xt - (1-y)s - p_B.$$

to. The indifferent users are[15]

$$\hat{y}_2 = y^*(p_A, p_B; \mu_A, \mu_B) + \frac{\omega(\mu_A, \mu_B)}{2(1+s)} + \frac{\alpha(\mu_A + \mu_B - 2\lambda\mu_A\mu_B)}{4(1+s)\mu_A\mu_B}$$

In Figure 4.4, we represent the indifferent users in this case.

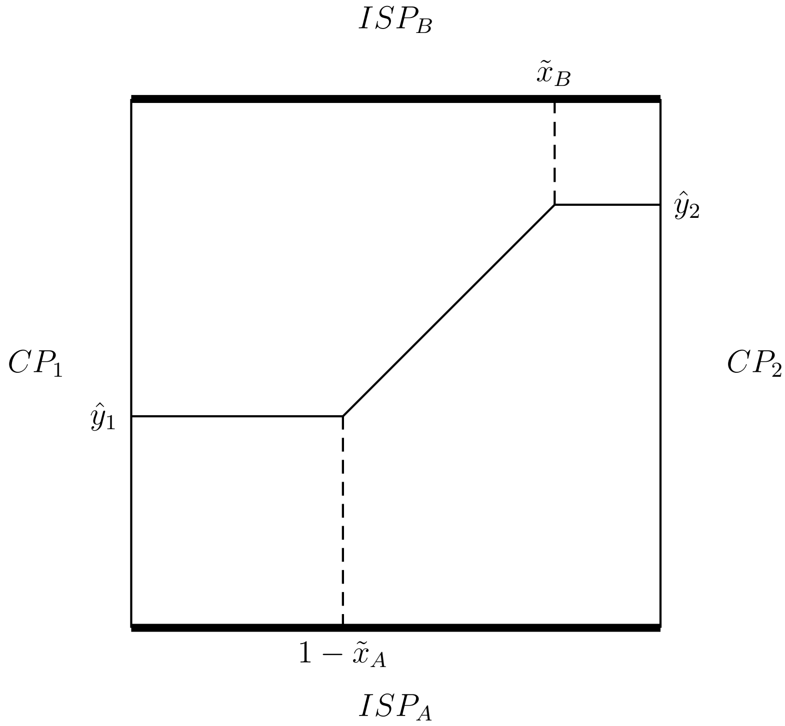


Figure 4.4: Indifferent users when priorities are sold to different CPs

Proposition 4.3. *A discriminating network attracts more users of the content it prioritizes when the other network prioritizes a different content. That is,*

$$\hat{y}_1 < \tilde{y}_1 < \tilde{y}_2 < \hat{y}_2$$

[15]Indifferent users satisfy the following condition:

$$v - w_{1,A}(\tilde{x}_A, y) - (1-x)t - ys - p_A = v - w_{1,B}(1-\tilde{x}_B, 1-y) - (1-x)t - (1-y)s - p_B.$$

Proof. See Appendix 4.B. □

ISP_A has the higher network capacity so, in its network, users face lower congestion externality. When it discriminates between contents, less users compared to ISP_B switch to the premium content. This result implies that users in network A are less penalized by browsing a content which does not correspond to their preferred type. Hence, ceteris paribus, ISP_B is less attractive than in case of network neutrality. The lower attractiveness of B is the same, either in case ISPs prioritize the same content or in case they prioritize different CPs. However, both ISPs attract more users of the content they prioritize if they sell the priority to different CPs. This effect is countervailed by a higher lost of users browsing the discriminated content.

4.4.2 Market shares

The market share of each ISP depends on which CP it has prioritized in the previous stage. Four different priority assignments can arise. However, given that CPs are symmetric from users' perspective, each ISP' market share is the same if both prioritize either CP_1 or CP_2. Similarly, an ISP's market share does not change either if it prioritizes CP_1 and the competitor prioritizes CP_2 or in the opposite case. Therefore, we study the case where both ISPs sell the priority to CP_1 and the case where ISP_A prioritizes CP_2 while ISP_B prioritizes CP_1. The remaining cases give the same market shares yielded in these two cases.

Both ISPs have given the priority to CP_1

The market share of ISP_A is given by the sum of users who browse CP_1 (denoted by $\tilde{\sigma}_{1,A}$) and CP_2 (denoted by $\tilde{\sigma}_{2,A}$) within its network.[16] Total users connected

[16] In particular, the total amount of users who always browse CP_1 is $\tilde{\sigma}_{1,A} = \tilde{x}_A \cdot \tilde{y}_1$ and the size of users who browse CP_2 is

$$\tilde{\sigma}_{2,A} = (1 - \tilde{x}_B) \cdot \tilde{y}_2 + \frac{(\tilde{y}_1 + \tilde{y}_2)(\tilde{x}_B - \tilde{x}_A)}{2}.$$

to A are $\tilde{\sigma}_A \equiv \tilde{\sigma}_{1,A} + \tilde{\sigma}_{2,A}$, that is

$$\tilde{\sigma}_A = \tilde{x}_A \cdot \tilde{y}_1 + (1 - \tilde{x}_B) \cdot \tilde{y}_2 + \frac{(\tilde{y}_1 + \tilde{y}_2)(\tilde{x}_B - \tilde{x}_A)}{2}.$$

Similarly, the total market share of ISP_B is

$$\tilde{\sigma}_B = \tilde{x}_A \cdot (1 - \tilde{y}_1) + (1 - \tilde{x}_B) \cdot (1 - \tilde{y}_2) + \frac{(2 - (\tilde{y}_1 + \tilde{y}_2))(\tilde{x}_B - \tilde{x}_A)}{2}.$$

Although not indicated, it is worthwhile to remember that users' choice of content is a function of network capacities, that is $\tilde{x}_i(\mu_A, \mu_B)$. Moreover, choice of service is a function of prices and capacities: $\tilde{y}_j(p_A, p_B; \mu_A, \mu_B)$. We suppressed these arguments to avoid excessive notation.

ISP_A has given the priority to CP_2 and ISP_B to CP_1

In this case, total market share of ISP_A is the sum of users who browse CP_1 (denoted by $\hat{\sigma}_{1,A}$) and CP_2 (denoted by $\hat{\sigma}_{2,A}$) when A prioritizes 2 while B prioritizes 1: $\hat{\sigma}_A \equiv \hat{\sigma}_{1,A} + \hat{\sigma}_{2,A}$.

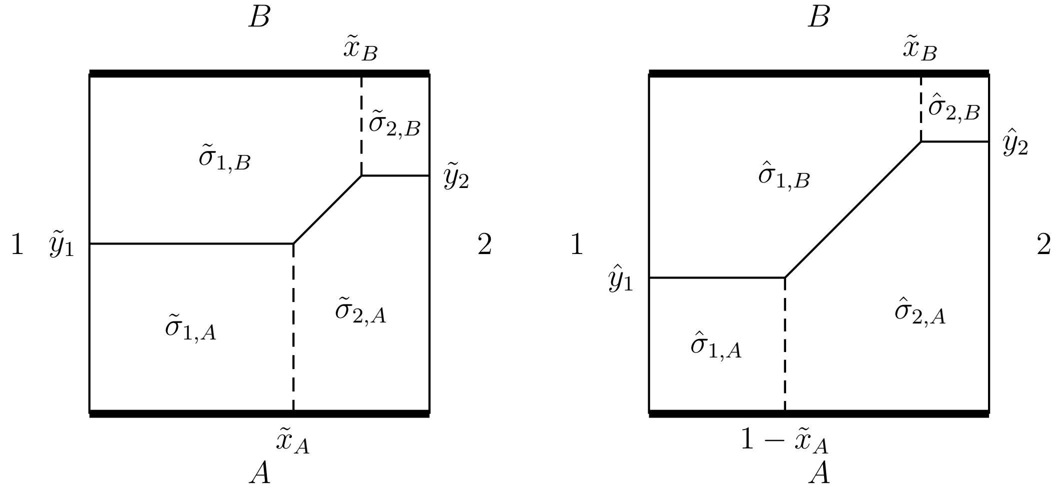
$$\hat{\sigma}_A = (1 - \tilde{x}_A) \cdot \hat{y}_1 + (1 - \tilde{x}_B) \cdot \hat{y}_2 + \frac{(\hat{y}_1 + \hat{y}_2)(\tilde{x}_B - (1 - \tilde{x}_A))}{2}.$$

Market size of ISP_B is

$$\hat{\sigma}_B = (1 - \tilde{x}_A) \cdot (1 - \hat{y}_1) + (1 - \tilde{x}_B) \cdot (1 - \hat{y}_2) + \frac{(2 - (\hat{y}_1 + \hat{y}_2))(\tilde{x}_B - (1 - \tilde{x}_A))}{2}.$$

In Figure 4.5, we provide an illustration of market shares. In case both ISPs prioritize the same content, ISP_A serves a larger proportion (compared to ISP_B) of users browsing the discriminated content, as Figure 4.5(a) shows. In case ISPs prioritize different contents, both serve a larger proportion of users browsing their respective premium contents, as shown in Figure 4.5(b).

The last fraction is the share of users type ($x \in (\tilde{x}_A, \tilde{x}_B], y$) who browse CP_2 when they join to ISP_A and CP_1 when they join to ISP_B.



(a) Both ISPs prioritize the same content (b) ISPs prioritize different contents

Figure 4.5: Representation of market shares

Market shares comparison

Before assigning priorities, ISPs consider what will be the impact on their market shares, i.e., if prioritizing the same content of the competitor attracts more or less users than prioritizing a different content. Interestingly but unsurprisingly, priority assignments do not matter for market sharing.

Proposition 4.4. *Under network discrimination, the total market share of each ISP is always the same, independently which CP the priority is sold to. In particular, with respect to the case of network neutrality, total market share of* ISP_A *increases by the waiting disutilities gaps difference:* $\tilde{\sigma}_A = y^* + \frac{\omega}{4(1+s)}$. *And, symmetrically, total market share of* ISP_B *decreases by the waiting disutilities gaps difference:* $\tilde{\sigma}_B = (1 - y^*) - \frac{\omega}{4(1+s)}$.

Proof. See Appendix 4.B. □

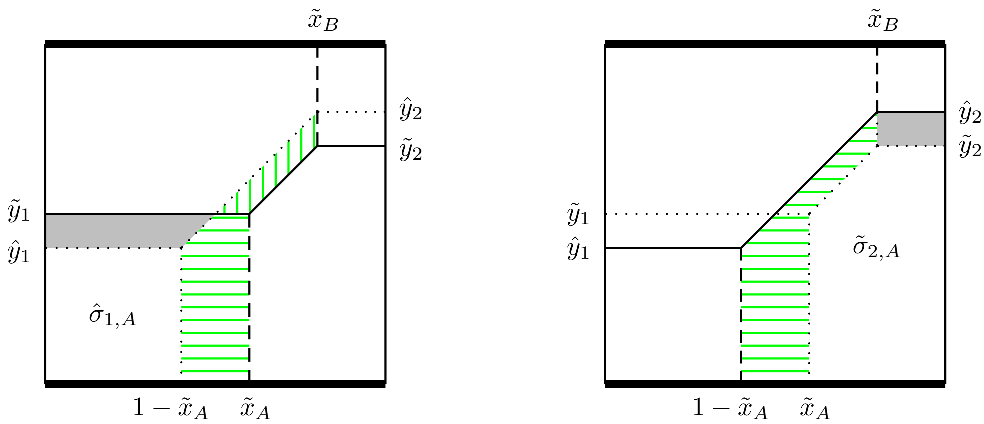
This result extends one of the outcomes of Choi & Kim (2010). In their model, when a monopolistic ISP discriminates between contents, the utility of users browsing the discriminated content deteriorates. Therefore, the monopolist has to decrease the end-users price to guarantee full coverage in a market with inelastic participation. In our model, market participation is still inelastic but providers' share are elastic. Hence, if discriminated users' utility deteriorates more in an ISP than in the other, users can switch to the other ISP. In particular, users in the smaller network would switch to the larger network. The smaller network can either reduce its price to maintain users who would migrate or keep a higher price and loose those consumers.

The fact that total market shares do not depend on priority assignments stems from Property 4.3. Aggregated disutility does not change when an ISP prioritizes a content: gaps difference is inversely proportional to the number of users who face worse quality. This implies that the number of users who would migrate to the larger network is constant.

4.4.3 Priority pricing

ISPs' priority strategies do not affect market shares outcomes. Therefore, each ISP prioritizes the content that provides the highest revenue. A CP purchases the priority service if the extra-profits are greater or equal to the premium fee paid for the priority. That is, the maximum willingness to pay of CP_i is given by the increment of its market share, denoted by $\Delta\tilde{\sigma}_i$, times per-user profit: $\lambda m_i \Delta\tilde{\sigma}_i$. Hence, we first determine CPs' market shares variation before analyzing how priority fees are fixed by Nash bargaining. Recalling the symmetries that arise between different priority assignments, we study the two representative cases.

If ISP_B assigns the priority service to CP_1, ISP_A can either assign the priority to CP_1 or sell it to CP_2. Both situations are represented in Figure 4.6. In Figure 4.6(a), ISP_A sells the priority service to CP_1; dotted lines represent market shares in case the priority be sold to CP_2. The market share gained by CP_1 (versus the case where ISP_A sells the priority to CP_2) is given by the hatched area. The horizontal lines represent the users of ISP_A who shift from CP_2 to CP_1 given the higher speed of the content provider. Vertical lines denote the users of ISP_A who were browsing CP_2 but, given the worsened speed, have shifted to ISP_B choosing CP_1. The shaded area represents the users of ISP_B browsing CP_1 and



(a) CP_1 has the priority (b) CP_2 has the priority

Figure 4.6: Users gained (area in green) by a CP with priority under ISP_A

who have now shifted to ISP_A because it is their preferred ISP; these users do not constitute an effective change in CP_1's market share. In Figure 4.6(b), ISP_A assigns the priority service to CP_2. The users who shift their choice are similar to the previous case. If ISP_B sells the priority to CP_2, the result is symmetric.

When ISP_B decides which CP to sell the priority to, the problem is similar to that one of ISP_A. We describe the problem in detail in Appendix 4.B.

Proposition 4.5. *The share of users gained by a CP that purchases the priority service from an ISP is always the same, independently on which type (*1 *or* 2*) it is and on the priority assignment decision of the other ISP. In particular, the shares gained by the content providers within the two network are*

$$\Delta\tilde{\sigma}_A = \frac{\alpha(1-\lambda\mu_A)}{t\mu_A}\left(y^* + \frac{\omega}{2(1+s)}\right) \qquad \text{under } ISP_A$$

$$\Delta\tilde{\sigma}_B = \frac{\alpha(1-\lambda\mu_B)}{t\mu_B}\left(1 - y^* - \frac{\omega}{2(1+s)}\right) \qquad \text{under } ISP_B$$

Proof. See Appendix 4.B. □

Note that the share gained by CP_1 is proportional to the respective market size of each ISP (the terms in brackets). Hence, ISP_A can provide a higher share, due to the larger capacity of its network. However, more users switch from CP_2 to CP_1 in network B because of the prioritized content is more valuated when capacity is low (the fraction before brackets). The overall effect depends on the end-users price decisions of the ISPs.

Given that CP_1's margin is higher than CP_2's, ISP_A always prefers to sell the priority right to CP_1. Priority price results from the Nash bargaining between the ISP and the CP: $(\theta m_1 + (1-\theta)m_2)\lambda\Delta\tilde{\sigma}_A$. Therefore, ISP_A charges a fee

$$f_A = \frac{\alpha(1-\lambda\mu_A)}{t\mu_A}\left(y^* + \frac{\omega}{2(1+s)}\right)\lambda(\theta m_1 + (1-\theta)m_2).$$

Similarly, ISP_B also sells the priority right to CP_1. The price resulting from the Nash bargaining is $(\theta m_1 + (1-\theta)m_2)\Delta\tilde{\sigma}_B$. Hence, it charges a fee

$$f_B = \frac{\alpha(1-\lambda\mu_B)}{t\mu_B}\left(1 - y^* - \frac{\omega}{2(1+s)}\right)\lambda(\theta m_1 + (1-\theta)m_2).$$

4.4.4 End-users pricing

In case of network discrimination, the profits of the two service providers are given by the price paid by the respective share end-users plus the fee charged to the prioritized content:

$$\begin{aligned}\pi_A =& p_A\tilde{\sigma}_A + f_A\\ \pi_B =& p_B\tilde{\sigma}_B + f_B\end{aligned}$$

Both ISPs maximize their profits by choosing their respective access prices, p_j, where the only restriction is represented by the full coverage of the market.[17] We assume the transport cost s is sufficiently large to avoid corner solutions.[18]

[17]Therefore, the indifferent consumer is always receiving a nonnegative surplus when joining to one ISP. However we assume v, the parameter denoting the benefit from joining a network and browsing a content, is sufficiently large to prevent the participation constraint to bind.

[18]There must be an indifferent user between ISP_A and ISP_B who browses CP_1 and an indifferent user who browses CP_2. This requires $\tilde{y}_1, \tilde{y}_2, \hat{y}_1, \hat{y}_2 \in (0,1)$, which depends on the equilibrium

Proposition 4.6. *The asymmetric ISPs market has a unique interior equilibrium characterized by the following prices:*

$$\tilde{p}_A = p_A^* + \frac{\omega}{6} - \frac{\alpha(\mu_A + 2\mu_B - 3\lambda\mu_A\mu_B)}{3t\mu_A\mu_B}\lambda(\theta m_1 + (1-\theta)m_2)$$
$$\tilde{p}_B = p_B^* - \frac{\omega}{6} - \frac{\alpha(2\mu_A + \mu_B - 3\lambda\mu_A\mu_B)}{3t\mu_A\mu_B}\lambda(\theta m_1 + (1-\theta)m_2).$$

When ISPs are allowed to discriminate, the gap between equilibrium prices increases with respect to the case of network neutrality:

$$\tilde{p}_A - \tilde{p}_B = p_A^* - p_B^* + \frac{1}{3}\left(\omega + \frac{\alpha(\mu_A - \mu_B)}{t\mu_A\mu_B}\lambda(\theta m_1 + (1-\theta)m_2)\right).$$

Proof. See Appendix 4.B. □

When networks are allowed to discriminate between contents, a proportion of the fees charged to content providers is deducted to end-users prices. This discount stems from the higher competition for users: a larger market share represents a higher revenue from the fee charged to the prioritized content provider. Moreover, users who browse the discriminated content are more penalized in the low capacity network (because of the higher waiting disutility). Hence, the smaller ISP needs to allow a higher discount (here denoted by ω) than the high capacity ISP.

For simplicity, we denote by

$$\varphi \equiv \frac{\alpha(\mu_A - \mu_B)}{t\mu_A\mu_B}\lambda(\theta m_1 + (1-\theta)m_2) > 0 \tag{4.7}$$

the extra-discount that ISP_B allows as a proportion to the fee charged to the content provider; that is, a higher share of CP's margin is transferred to the users in the smaller network. The other term in the difference between equilibrium prices, ω, represents the extra-discount that ISP_B needs to allow whatever the

prices. Given that $\mu_A > \mu_B$ the condition to be satisfied is $\tilde{y}_2 < 1$. Therefore, the transport cost s must satisfy

$$s > \frac{\mu_A - \mu_B}{3} + \frac{2}{3}\omega - \frac{\alpha(\mu_A - \mu_B)\lambda(\theta m_1 + (1-\theta)m_2)}{3t\mu_A\mu_B} + \frac{\alpha(\mu_A + \mu_B - 2\lambda\mu_A\mu_B)}{2\mu_A\mu_B} - 1.$$

margin of the CP is. In other words, both ISPs compete allowing higher discounts to end-users, in order to gain a larger market share to offer to the prioritized CP. In equilibrium, the smaller network provides a higher discount by φ. If the content provider has a high margin, ISPs competition for users is very fierce and the discount to compensate extra-disutility in the small network (ω) is less relevant to determine equilibrium prices. Otherwise, if the margin of the prioritized CP is insignificant, ISPs do not compete so fiercely through discounts. Therefore, the smaller network needs only to compensate the extra-utility loss ω.

Assumption 4.3. *We assume that $\varphi > \omega$. That is, the discount that ISP_B allows to its users to increase its market share is high enough to compensate discriminated users for the lower utility. This requires*

$$\lambda(\theta m_1 + (1-\theta)m_2) > \frac{\alpha(1-\lambda\mu_A)}{2\mu_A} + \frac{\alpha(1-\lambda\mu_B)}{2\mu_B}.$$

This assumption relates to the profitability of prioritization. If some users are penalized (due to increased waiting disutility), then ISPs find profitable to compensate them by a price discount if the fee paid by the prioritized content is sufficiently large.

Proposition 4.7. *When network discrimination is allowed, the equilibrium market share of the network with a smaller capacity increases with respect to the case of network neutrality. In particular, the equilibrium market shares of the two ISPs are:*

$$\tilde{\sigma}_A = \sigma_A^* - \frac{1}{6(1+s)}\left(\varphi - \frac{\omega}{2}\right)$$

$$\tilde{\sigma}_B = \sigma_B^* + \frac{1}{6(1+s)}\left(\varphi - \frac{\omega}{2}\right)$$

Proof. See Appendix 4.B. □

The reason why ISP_B is able to increase its market share is that, in its network, priority is more valuable. More users switch to prioritized content, allowing the ISP to charge a higher fee. In network A, the service provider has also the incentive to give a discount to attract more users. However, this incentive is lower because users perceive a lower disutility when browsing the discriminated content and, therefore, they are less eager to switch content. Note that if ISPs are sufficiently differentiated (i.e., the transport cost s satisfies the condition for an interior equilibrium), the total market share of ISP_A is still larger than that one of ISP_B.

Finally, equilibrium fees charged to the prioritized content provider are:

$$\tilde{f}_A = \frac{\alpha(1-\lambda\mu_A)}{t\mu_A}\left(\sigma_A^* - \frac{1}{3(1+s)}\left(\frac{\varphi}{2}-\omega\right)\right)\lambda(\theta m_1 + (1-\theta)m_2)$$

$$\tilde{f}_B = \frac{\alpha(1-\lambda\mu_B)}{t\mu_B}\left(\sigma_B^* + \frac{1}{3(1+s)}\left(\frac{\varphi}{2}-\omega\right)\right)\lambda(\theta m_1 + (1-\theta)m_2)$$

Again, the fees are positive if the transport cost is high enough.

4.5 Investment incentives

In the long run, ISPs can invest to increase network capacity μ_j. As in case of a monopolistic service provider, expanding capacity has two opposite effects on an ISP's revenues. On the one hand, higher capacity reduces the congestion in the network. Users achieve higher utility when connected (due to the reduced waiting disutility) and, therefore an ISP can charge a higher price to provide network connection. On the other hand, capacity expansion affects revenues from priority: lower congestion reduces the value for priority since users are less likely to switch to a prioritized content. Hence, a CP will pay a lower fee to purchase the premium service.

There are also other two countervailing effects that only arise in the case of competing ISPs. First, if a network expands its capacity more than its competitor, it becomes more attractive to marginal users, given that they face lower congestion (and disutility) if they migrate to the network that invests more in capacity. Hence, the increased market share allows an ISP to gain higher revenues from end-users. Second, the lower fee charged to the CP reduces the discount an ISP can allow to its users. Therefore, the competitor can contrast users migration by offering a greater discount.

4.5.1 Equilibrium profits

To analyze the overall effect of network discrimination on investment incentives, equilibrium profits of the two networks are compared to the case where network discrimination is not allowed (network neutrality). When an ISP plans investments in capacity expansion, it decides the optimal level of μ_j given the impact of capacity on equilibrium prices, fees, and market shares analyzed in the previous section. In particular, for any capacity level μ_A, equilibrium profits of ISP_A are

$$\tilde{\pi}_A = \pi_A^* + \Gamma_A(\mu_A, \mu_B) \tag{4.8}$$

where $\Gamma_A(\mu_A, \mu_B)$ is a function of the network capacities defined as

$$\begin{aligned}\Gamma_A(\mu_A, \mu_B) \equiv \frac{1}{18(1+s)}\left(\varphi - \frac{\omega}{2}\right)^2 - \left(\frac{1}{3} + \frac{\mu_A - \mu_B}{9(1+s)}\right)\left(\varphi - \frac{\omega}{2}\right) + \\ + \frac{(1-\lambda\mu_A)\mu_B}{2(1+s)(\mu_A - \mu_B)}\left(\varphi \cdot \frac{\omega}{2}\right).\end{aligned} \tag{4.9}$$

Similarly, for any capacity level μ_B, equilibrium profits of ISP_B are

$$\tilde{\pi}_B = \pi_B^* + \Gamma_B(\mu_A, \mu_B) \tag{4.10}$$

where $\Gamma_B(\mu_A, \mu_B)$ is

$$\begin{aligned}\Gamma_B(\mu_A, \mu_B) \equiv \frac{1}{18(1+s)}\left(\varphi - \frac{\omega}{2}\right)^2 + \left(\frac{1}{3} - \frac{\mu_A - \mu_B}{9(1+s)}\right)\left(\varphi - \frac{\omega}{2}\right) + \\ - \frac{(1-\lambda\mu_B)\mu_A}{2(1+s)(\mu_A - \mu_B)}\left(\varphi \cdot \frac{\omega}{2}\right).\end{aligned} \tag{4.11}$$

Equations (4.8) and (4.10) allow for a ready comparison of equilibrium profits between a regime network neutrality and a regime of network discrimination. In particular, ISP_A and ISP_B increment their profits by Γ_A and Γ_B, respectively. Both increments depend on the expressions φ and $\frac{\omega}{2}$. As we have seen in equation (4.7), φ is the extra-discount allowed by network B to its consumers by transferring a larger proportion of the fee paid by the prioritized content. In equation (4.6), we defined ω as the extra-disutility faced by users in the small network due to the higher congestion. Therefore, $\varphi - \frac{\omega}{2}$ represents the overall attractiveness gained by ISP_B because of the relative change of its end-user price with respect to network A. More in details, the profits increment (with respect to the case of network neutrality) in equations (4.9) and (4.11) can be decomposed in three terms:

- the first term, $\frac{1}{18(1+s)}\left(\varphi - \frac{\omega}{2}\right)^2$, is the same for both ISPs and denotes the average increment in total profits;

- the second term indicates to what extent the overall profits are due to consumers.

  - For ISP_A, the term $-\left(\frac{1}{3} + \frac{\mu_A - \mu_B}{9(1+s)}\right)\left(\varphi - \frac{\omega}{2}\right)$ is negative. Network A loses attractiveness with respect to B and, therefore, profits derived from consumers are less relevant.
  - For ISP_B, $\left(\frac{1}{3} - \frac{\mu_A - \mu_B}{9(1+s)}\right)\left(\varphi - \frac{\omega}{2}\right)$ is positive. Network B increases its market share and consumers provide a larger share of increased profits;

- the last term represents how profits change due to the fee charged to the prioritized content.

  - In network A, $\frac{(1-\lambda\mu_A)\mu_B}{2(1+s)(\mu_A - \mu_B)}\left(\varphi \cdot \frac{\omega}{2}\right) > 0$. ISP_A transfers to consumers (in form of a discount) a lower fraction of the fee charged to the prioritized CP.
  - In network B, $-\frac{(1-\lambda\mu_B)\mu_A}{2(1+s)(\mu_A - \mu_B)}\left(\varphi \cdot \frac{\omega}{2}\right) < 0$. ISP_B applies a higher discount to attract users by transferring a larger part of the fee. Therefore, the fee charged to the prioritized CP has a lower direct contribution to the extra-profits of network B.

Further details on profits derivation are provided in Appendix 4.B.

4.5.2 Long run analysis

We investigate the case whether in the long run (i.e., when network capacity is endogenous) Internet Service Providers have higher incentive to invest in capacity expansion in a discriminating regime rather than in a neutrality regime. When discrimination is allowed, equilibrium profits are given by equations (4.8) and (4.10). In particular, when both networks prioritize one content, ISP_A's profits increase by Γ_A and ISP_B's ones by Γ_B. Therefore, if Γ_j is increasing in μ_j (that is, if profits in network discrimination increases on one ISP capacity), then ISP_j has higher incentives to invest if it is allowed to discriminate.

Note that profits increment Γ_j depends on the change in relative ISPs' attractiveness $\varphi - \frac{\omega}{2}$. Remember that $\omega(\mu_A, \mu_B)$ is the function of capacities that denotes the difference in congestion disutilities between A and B, and that $\varphi(\mu_A, \mu_B)$ is the function of capacities that denotes the difference in the allowed user-price discount.

Lemma 4.1. *Functions $\omega(\mu_A, \mu_B)$ and $\varphi(\mu_A, \mu_B)$ are increasing in the capacity of the larger network (μ_A) and decreasing in the capacity of the smaller network (μ_B).*

Proof. First-order derivatives are

$$\frac{\partial \omega}{\partial \mu_A} = \frac{\alpha^2(1-\lambda\mu_A)}{t\mu_A^3} > 0 \qquad \frac{\partial \omega}{\partial \mu_B} = -\frac{\alpha^2(1-\lambda\mu_B)}{t\mu_B^3} < 0$$

$$\frac{\partial \varphi}{\partial \mu_A} = \frac{\alpha}{t\mu_A^2}\lambda(\theta m_1 + (1-\theta)m_2) > 0 \quad \frac{\partial \varphi}{\partial \mu_B} = -\frac{\alpha}{t\mu_B^2}\lambda(\theta m_1 + (1-\theta)m_2) < 0.$$

□

These results reflect the fact that, when ISP_A expands its capacity, the gap between disutilities of its users and ISP_B's users increases. Moreover, ISP_B needs to transfer a higher share to its users of the content fee to compensate for the utility loss. On the contrary, if ISP_B expands its capacity, the disutility gap diminishes and, therefore, the needing to compensate it. This condition is described in the following lemma.

Lemma 4.2. *$\varphi - \omega$ is an increasing function of the capacity of the larger network (μ_A) and a decreasing function of the capacity of the smaller network (μ_B).*

Proof. In particular, the derivatives of the difference $\varphi - \frac{\omega}{2}$ are

$$\frac{\partial\left(\varphi - \frac{\omega}{2}\right)}{\partial \mu_A} = \frac{\alpha}{t\mu_A^3}\left(\lambda(\theta m_1 + (1-\theta)m_2) - \frac{\alpha(1-\lambda\mu_A)}{2\mu_A}\right) > 0$$

$$\frac{\partial\left(\varphi - \frac{\omega}{2}\right)}{\partial \mu_B} = -\frac{\alpha}{t\mu_B^3}\left(\lambda(\theta m_1 + (1-\theta)m_2) - \frac{\alpha(1-\lambda\mu_B)}{2\mu_B}\right) < 0.$$

□

This results simply states that, when ISP_A (ISP_B) expands its capacity, it increases its attractiveness relatively to ISP_B (ISP_A).

Therefore, the overall effect of capacities expansion depends on the relationship between the extra-disutility created when discriminating and the margin that ISPs can extract from CP by charging a priority fee.

Proposition 4.8. *When ISPs are allowed to discriminate between content providers, and CPs gain a sufficiently high margin, both networks have less incentives to expand their capacities than in a neutrality regime.*

Proof. See Appendix 4.B. □

The intuition behind this result is the following. On the one hand, an ISP that invests in capacity increases its attractiveness due to lower congestion, increasing its total market share. On the other hand, higher capacity reduces congestion disutility to users, so less consumers switch to the prioritized CP. If an ISP expands its capacity more than its competitor, it increases its total market share. However, this happens by attracting discriminated users of the other ISP, since they are more penalized in the a network that invests less. Therefore, the ISP that invests more in capacity, increases its share of users browsing the discriminated content. This implies a lower fee charged to the prioritized content.

If CP's margin is high enough compared to users' extra-disutility (which is reflected in lower total market share or a lower end-users price), a network prefers to lose consumers and to charge a higher fee.

4.6 Policy implications and conclusions

Proposition 4.8 has important policy implications. Allowing network discrimination reduces ISPs' incentives to invest in larger network capacities. Hence, a regulator that aims to improve networks capacity should impose a neutrality regime. However, some further considerations need to be done about: (1) the implications of the assumptions we made, and (2) about the policy objectives of a regulator.

Two assumptions are particularly relevant: profitability of content providers and initial asymmetry of capacities. Assumption 4.3 was made to ensure that both ISPs find profitable to discriminate. If it is relaxed, that is, if CPs margins are low, the fee charged is low too; so, ISPs revenues from end-consumers are higher compared to revenues from CPs. Therefore, ISPs do not choose to discriminate, since, in case of discrimination, the loss of revenues from users who migrate to the other network (due to extra-disutility) is not compensated by the revenues earned from CPs. Hence, the outcome is the one of network neutrality.

If networks have symmetric capacities, the result would change. ISPs discount an amount of the charged fee to end-users. The discount is allowed to all the users while the fee is only paid for the users who switch from the discriminated content to the prioritized content.[19] The overall effect is profits neutral, in the sense that profits are the same as in network neutrality. Therefore, the incentives to invest would be the same as in the neutrality regime. However, this result would only hold under the hypothesis of no binding consumers participation.

We have seen that, when asymmetric ISPs discriminate between contents, incentives to invest in network expansion are lower. This conclusion suggests that regulators need to impose network neutrality to achieve larger network capacities. However, the European position is not completely clear to this regard. The European Commission and the Body of European Regulators for Electronic Communications (BEREC) recognize the importance of developing faster, next generation networks. They claim that providing conditions for a competitive ISPs market allows achieving proper incentives to invest and, as a consequence, network neu-

[19] More in details, ISPs discount $\frac{2}{3}$ of the fee to all their users.

trality.

In our setup, a competitive market for ISPs is intended as a market where migrating from a ISP to another is costless to consumers (i.e., the transport cost when connecting to an ISP is always zero). However, in practice, horizontal differentiation between ISPs is impossible to wipe out. For instance, Internet Service Providers continuously tie up extra-services to the internet connection service (modems, TV programs, phone calls,. . .). Hence, consumers face a cost if they cannot connect to the preferred Internet Service offer. Since each ISP has a captive market, the only way a Content Provider can cover the whole market is to contract access through all the ISPs. This gives some degree of bargaining power to ISPs, which can extract part of the CP surplus by charging a fee on the provision of priority service.

A policy that limits ISPs' differentiation capacity and makes Internet service provision a competitive market could probably achieve network neutrality as market outcome. If consumers can freely migrate from an ISP to another, they would be rather unwilling to accept congestion disutility associated to discrimination. Hence, more discriminated consumers would migrate to the competitor when an ISP prioritizes a content. This would make discrimination very costly to ISP, meaning a great loss in terms of consumers. To this extent, the European perspective would be correct: warranting ISP competition could allow to achieve network neutrality without imposing it by law. However, if an ISP can recover the lower revenue due to consumers loss through a high prioritization fee, network discrimination would still be profitable (and desirable from the ISP's point of view). In this case, as we have seen, incentives to invest in capacity expansion are lower.

Therefore, the main concern of the European Commission should not only be to guarantee a competitive ISP market for end-users but also to provide a competitive ISP market for Content Providers. If ISPs had lower bargaining power when they negotiate the prioritization fee, their capacity to extract CPs surplus would be lower as well. Therefore, the revenue from the prioritization fee would not compensate for the consumers loss due to discrimination. In this case, an ISP would prefer to not discriminate and competing for consumers through larger investments in capacity.

A final remark should be done as potential incipit for future research. While the attention is all focused on ISPs actions, very little has been said about the CP side of the market. We have seen that a low capacity of CP margin extraction (associated to higher ISPs competition) implies that network neutrality is more profitable for ISPs and leads to higher investments in capacity expansion. However, a similar outcome would yield if the CP margin from advertisement would be very small. Even though ISPs had a great bargaining power, if the margin they can extract is small, network neutrality would be more profitable. Therefore, a regulator should also investigate CPs' advertisement negotiation, assessing if high margins (when present) are justified through the nature of the market or derive from some market power practice.

References

Armstrong, M. (2006). Competition in Two-Sided Markets. *RAND Journal of Economics*, *37*(3), 668–691.

Baake, P. & Mitusch, K. (2007). Competition with Congestible Networks. *Journal of Economics*, *91*(2), 151–176.

Cheng, H. K., Bandyopadhyay, S., & Guo, H. (2011). The Debate on Net Neutrality: A Policy Perspective. *Information Systems Research*, *22*, 60–82.

Choi, J. P. & Kim, B.-C. (2010). Net Neutrality and Investment Incentives. *RAND Journal of Economics*, *41*(3), 446–471.

Ellison, G. (2006). Bounded Rationality in Industrial Organization. *Econometric Society Monographs*, *42*, 142.

Hermalin, B. E. & Katz, M. L. (2007). The Economics of Product-Line Restrictions with an Application to the Network Neutrality Debate. *Information Economics and Policy*, *19*(2), 215–248.

Krämer, J. & Wiewiorra, L. (2009). Network Neutrality and Congestion Sensitive Content Providers: Implications for Service Innovation, Broadband Investment and Regulation. MPRA Paper 16655, University Library of Munich, Germany.

Musacchio, J., Schwartz, G., & Walrand, J. (2009). A Two-Sided Market Analysis of Provider Investment Incentives with an Application to the Net-Neutrality Issue. *Review of Network Economics*, *8*(1), 3.

Möbius, M. (2001). Death through Success: The Rise and Fall of Local Service Competition at the Turn of the Century. Technical report, mimeo.

Rochet, J.-C. & Tirole, J. (2006). Two-Sided Markets: A Progress Report. *RAND Journal of Economics*, *37*(3), 645–667.

Schuett, F. (2010). Network Neutrality: A Survey of the Economic Literature. *Review of Network Economics*, *9*(2), 1–14.

Sluijs, J. P. (2010). Network Neutrality between False Positives and False Negatives: Introducing a European Approach to American Broadband Markets. *Federal Communications Law Journal*, *62*, 77–117.

Appendix 4.A Proofs in Section 4.3: Net Neutrality

i) Maximization problem of the ISP in the neutral network:

$$\max_{p} \ \pi_M = p \ s.t. \ v - w - tx^* - p \geq 0.$$

ii) Profit of the ISP in a discriminatory network:

$$\max_{\tilde{p}} \ \tilde{\pi}_M = \tilde{p} + f \qquad \text{s.t.} \quad v - w_1(\tilde{x}) - t\tilde{x} - \tilde{p} \geq 0.$$

iii) Calculation of the fee by Nash bargaining:

$$\max_{f} \ [f - m_2(2\tilde{x}-1)\lambda]^{\theta} \cdot [(m_1\tilde{x}\lambda - f) - m_1(1-\tilde{x})\lambda]^{1-\theta} \Rightarrow f = [m_2 + \theta(m_1 - m_2)](2\tilde{x}-1)\lambda$$

where $\theta \in [0,1]$ denotes the ISP's bargaining power.

iv) CP's profits under discrimination: when the ISP assigns the priority to CP_1, each content provider's profit will be respectively given by

$$\tilde{\pi}_1 = m_1\tilde{x}\lambda - [m_2 + \theta(m_1 - m_2)](2\tilde{x} - 1)\lambda; \qquad \tilde{\pi}_2 = m_2(1 - \tilde{x})\lambda.$$

v) First-order conditions of the problem of ISPs under the neutral network:

$$\frac{\partial \pi_A}{\partial p_A} = 0 \Longleftrightarrow \frac{1+s}{2} + (\frac{\mu_A - \mu_B + p_A - p_B}{2(1+s)} - \frac{p_A}{2(1+s)} = 0$$

$$\frac{\partial \pi_B}{\partial p_B} = 0 \Longleftrightarrow \frac{1}{2} - (\frac{\mu_A - \mu_B + p_B - p_A}{2(1+s)} - \frac{p_B}{2(1+s)} = 0$$

Appendix 4.B Proofs in Section 4.4: Net Discrimination

Proof of Proposition 4.1

Given the definitions of indifferent users, the difference between $\tilde{x}_B$ and $\tilde{x}_A$ is

$$\begin{aligned}\tilde{x}_B - \tilde{x}_A &= \frac{1}{2} + \alpha\frac{1-\lambda\mu_B}{2t\mu_B} - \left(\frac{1}{2} + \alpha\frac{1-\lambda\mu_A}{2t\mu_A}\right) \\ &= \alpha\frac{\mu_A(1-\lambda\mu_B) - \mu_B(1-\lambda\mu_A)}{2t\mu_A\mu_B} \\ &= \alpha\frac{\mu_A - \mu_B}{2t\mu_A\mu_B} > 0.\end{aligned}$$

Assumption 4.1 ($1-\lambda\mu_A > 0$) guarantees that $x^* < \tilde{x}_A$.

Proof of Proposition 4.2

Proposition 4.2.a

When both ISPs prioritize the same content, the difference between the disutility gaps in the two networks is $(w_{1,B}(\tilde{x}_B) - w_{2,B}(\tilde{x}_B)) - (w_{1,A}(\tilde{x}_A) + w_{2,A}(\tilde{x}_A))$, that is

$$\frac{\alpha^2(\mu_A - \mu_B)(\mu_B + \mu_A(1-2\lambda\mu_B))}{4(1+s)t\mu_A^2\mu_B^2} > 0.$$

Proposition 4.2.b

When the ISPs prioritize different content providers, the difference between the disutility gaps in the two networks is $(w_{1,B}(\tilde{x}_B) - w_{2,B}(\tilde{x}_B)) - (w_{2,A}(1-\tilde{x}_A) - w_{1,A}(1-\tilde{x}_A))$, that is

$$\frac{\alpha^2(\mu_A - \mu_B)(\mu_B + \mu_A(1-2\lambda\mu_B))}{4(1+s)t\mu_A^2\mu_B^2} > 0.$$

Proof of Proposition 4.3

The result of the comparison between indifferent users depends on the sign of the following difference:

$$\frac{\alpha(\mu_A + \mu_B - 2\lambda\mu_A\mu_B)}{4(1+s)\mu_A\mu_B} - \frac{\alpha(\mu_A - \mu_B)}{4(1+s)\mu_A\mu_B}$$

that is, on the sign of $\frac{\alpha(1-\lambda\mu_A)}{2(1+s)\mu_A}$, which is always positive (given our assumption $1 > \lambda\mu_A$).

Proof of Proposition 4.4

Let define $\tilde{\tau} \equiv \frac{\alpha(\mu_A - \mu_B)}{4(1+s)\mu_A\mu_B}$ and $\hat{\tau} \equiv \frac{\alpha(\mu_A+\mu_B-2\lambda\mu_A\mu_B)}{4(1+s)\mu_A\mu_B}$. In the case both ISPs prioritize CP_1, total market share of ISP_A is

$$\begin{aligned}\tilde{\sigma}_A &= \tilde{x}_A(y^* + \omega - \tilde{\tau}) + (1 - \tilde{x}_B)(y^* + \omega + \tilde{\tau}) + (y^* + \omega)(\tilde{x}_B - \tilde{x}_A) \\ &= y^* + \omega + \tilde{\tau}(1 - (\tilde{x}_A + \tilde{x}_B)) = y^* + \frac{\omega}{2}\end{aligned}$$

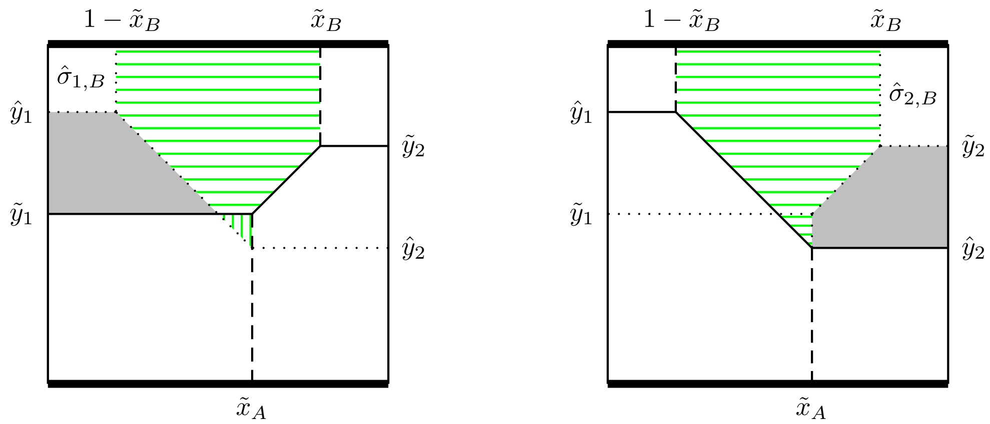
given that $\tilde{\tau}(1 - (\tilde{x}_A + \tilde{x}_B)) = -\frac{\omega}{2}$. In the case ISP_A prioritizes CP_2 and ISP_B prioritizes CP_1, total market share of A is

$$\begin{aligned}\hat{\sigma}_A &= (1 - \tilde{x}_A)(y^* + \omega - \hat{\tau}) + (1 - \tilde{x}_B)(y^* + \omega + \hat{\tau}) + (y^* + \omega)(\tilde{x}_B - (1 - \tilde{x}_A)) \\ &= y^* + \omega + \hat{\tau}(\tilde{x}_A - \tilde{x}_B) = y^* + \frac{\omega}{2}\end{aligned}$$

given that $\hat{\tau}(\tilde{x}_A - \tilde{x}_B) = -\frac{\omega}{2}$.

Users gained by prioritized CP under ISP_B

When ISP_B sells the priority service, the possible outcomes are represented in Figure 4.7. In particular, in Figure 4.7(a), the market share gained by CP_1 when it purchases the priority is represented by the (green) hatched area. Horizontal lines denote the users of ISP_B who switch from CP_2 to CP_1. Vertical lines represent users who were browsing CP_2 under B and now browse CP_1 under A. The (gray) shaded area define the users who switch from A to B but not constitute a market



(a) CP_1 has the priority (b) CP_2 has the priority

Figure 4.7: Users gained (area in green) by a CP with priority under ISP_B

gain for CP_1. In Figure 4.7(b), we represent the case where the priority is sold to CP_2. The meaning of the areas is the same as in the previous case.

Proof of Proposition 4.5

When both ISPs sell the priority to CP_1, total market share of the content is

$$\tilde{\sigma}_1 = \tilde{x}_A + (2 - \tilde{y}_1 - \tilde{y}_2)\frac{\tilde{x}_A - \tilde{x}_B}{2} = \tilde{x}_B - \frac{\tilde{x}_B - \tilde{x}_A}{2}\left(y^* + \frac{\omega}{2(1+s)}\right).$$

If ISP_A sells the priority to CP_2, total market share of CP_1 would be $\tilde{x}_B + (1 - \tilde{x}_A - \tilde{x}_B)\left(y^2 + \frac{\omega}{2}\right)$. If ISP_B sells the priority to CP_2, total market share of CP_1 would be $1 - \tilde{x}_B - (1 - \tilde{x}_A - \tilde{x}_B)\left(y^2 + \frac{\omega}{2}\right)$. Total market share of CP_1 diminishes by $(2\tilde{x}_A - 1)\left(y^* + \frac{\omega}{2(1+s)}\right)$ if A deviates prioritizing 2. If B deviates, market share of 1 diminishes by $(2\tilde{x}_B - 1)\left(1 - y^* - \frac{\omega}{2(1+s)}\right)$. Recalling that $(2\tilde{x}_A - 1) = \frac{\alpha(1-\lambda\mu_A)}{t\mu_A}$ and $(2\tilde{x}_B - 1) = \frac{\alpha(1-\lambda\mu_B)}{t\mu_B}$, we get the market share variations in the proposition.

Proof of Proposition 4.6

The first order conditions/reaction functions for a maximum are

$$\frac{\partial \pi_A}{\partial p_A} = 0 \Longrightarrow$$
$$\frac{1}{2} + \frac{p_B - 2p_A + \mu_A - \mu_B}{2(1+s)} + \frac{\omega(\mu_A, \mu_B)}{2} - \frac{\alpha(1-\lambda\mu_A)(m_2 - \theta(m_2 - m_1))}{2(1+s)t\mu_A} = 0$$

$$\frac{\partial \pi_B}{\partial p_B} = 0 \Longrightarrow$$
$$\frac{1}{2} - \frac{2p_B - p_A + \mu_A - \mu_B}{2(1+s)} - \frac{\omega(\mu_A, \mu_B)}{2} - \frac{\alpha(1-\lambda\mu_B)(m_2 - \theta(m_2 - m_1))}{2(1+s)t\mu_B} = 0$$

The second order conditions are readily satisfied, $\frac{\partial^2 \pi_A}{\partial p_A^2} = \frac{\partial^2 \pi_B}{\partial p_B^2} = -\frac{1}{1+s} < 0$.

Proof of Proposition 4.7

Recall that

$$y^*(\tilde{p}_A, \tilde{p}_B) = \sigma_A^* - \frac{1}{6(1+s)} \left(\omega + \frac{\alpha(\mu_A - \mu_B)}{t\mu_A\mu_B} \lambda(\theta m_1 + (1-\theta)m_2) \right).$$

Given our definition of φ and market shares defined in Proposition 4.4, we derive equilibrium market shares. The sign of the change in market shares with respect to the case of network neutrality is given by Assumption 4.3.

Equilibrium profits

Note that equilibrium prices can be written as

$$\begin{aligned}
\tilde{p}_A =& p_A^* - \frac{1}{3}\left(\varphi - \frac{\omega}{2}\right) - \frac{\alpha(1-\lambda\mu_A)}{t\mu_A}\lambda(\theta m_1 + (1-\theta)m_2) \\
\tilde{p}_B =& p_B^* + \frac{1}{3}\left(\varphi - \frac{\omega}{2}\right) - \frac{\alpha(1-\lambda\mu_B)}{t\mu_B}\lambda(\theta m_1 + (1-\theta)m_2).
\end{aligned}$$

Using the fact that $\sigma_A^* = \frac{p_A^*}{2(1+s)}$ and $\sigma_B^* = \frac{p_B^*}{2(1+s)}$, equilibrium market shares can be written as

$$\tilde{\sigma}_A = \sigma_A^* - \frac{\sigma_A^*}{3p_A^*}\left(\varphi - \frac{\omega}{2}\right)$$
$$\tilde{\sigma}_B = \sigma_B^* + \frac{\sigma_B^*}{3p_B^*}\left(\varphi - \frac{\omega}{2}\right).$$

Therefore, total revenues from users, $\tilde{p}_A\tilde{\sigma}_A$ and $\tilde{p}_B\tilde{\sigma}_B$, are

$$\begin{aligned}\tilde{p}_A\tilde{\sigma}_A =& p_A^*\sigma_A^* + \frac{1}{18(1+s)}\left(\varphi - \frac{\omega}{2}\right)^2 - \frac{2\sigma_A^*}{3}\left(\varphi - \frac{\omega}{2}\right) + \\ & - \frac{\alpha(1-\lambda\mu_A)}{t\mu_A}\left(\sigma_A^* - \frac{\sigma_A^*}{3p_A^*}\left(\varphi - \frac{\omega}{2}\right)\right)\lambda(\theta m_1 + (1-\theta)m_2)\end{aligned}$$

$$\begin{aligned}\tilde{p}_B\tilde{\sigma}_B =& p_B^*\sigma_B^* + \frac{1}{18(1+s)}\left(\varphi - \frac{\omega}{2}\right)^2 + \frac{2\sigma_B^*}{3}\left(\varphi - \frac{\omega}{2}\right) + \\ & - \frac{\alpha(1-\lambda\mu_B)}{t\mu_B}\left(\sigma_B^* + \frac{\sigma_B^*}{3p_B^*}\left(\varphi - \frac{\omega}{2}\right)\right)\lambda(\theta m_1 + (1-\theta)m_2)\end{aligned}$$

Proof of Proposition 4.8

The derivative of Γ_A with respect to μ_A is

$$\begin{aligned}\frac{\partial\Gamma_A}{\partial\mu_A} =& \frac{1}{9(1+s)}\left(\varphi - \frac{\omega}{2}\right)\frac{\partial\left(\varphi - \frac{\omega}{2}\right)}{\partial\mu_A} - \frac{\sigma_A^*}{3}\frac{\partial\left(\varphi - \frac{\omega}{2}\right)}{\partial\mu_A} + \\ & - \frac{1}{9(1+s)}\left(\varphi - \frac{\omega}{2}\right) + \frac{(1-\lambda\mu_A)\mu_B}{2(1+s)(\mu_A-\mu_B)}\frac{\partial\left(\varphi\cdot\frac{\omega}{2}\right)}{\partial\mu_A} + \\ & \frac{(1-\lambda\mu_B)\mu_B}{2(1+s)(\mu_A-\mu_B)^2}\left(\varphi\cdot\frac{\omega}{2}\right).\end{aligned}$$

Rearranging terms we get

$$\begin{aligned}\frac{\partial\Gamma_A}{\partial\mu_A} =& -\frac{1}{9(1+s)}\left(\varphi - \frac{\omega}{2}\right)\left(1 - \frac{\partial\left(\varphi - \frac{\omega}{2}\right)}{\partial\mu_A}\right) - \frac{\sigma_A^*}{3}\frac{\partial\left(\varphi - \frac{\omega}{2}\right)}{\partial\mu_A} + \\ & + \frac{\mu_B}{2(1+s)(\mu_A-\mu_B)}\left((1-\lambda\mu_A)\frac{\partial\left(\varphi\cdot\frac{\omega}{2}\right)}{\partial\mu_A} + \frac{(1-\lambda\mu_B)}{\mu_A-\mu_B}\left(\varphi\cdot\frac{\omega}{2}\right)\right).\end{aligned}$$

Given Assumption 4.3, we have

$$\frac{1}{9(1+s)}\left(\varphi-\frac{\omega}{2}\right)>\frac{\mu_B}{2(1+s)(\mu_A-\mu_B)},$$

while Lemmas 4.1 and 4.2 provide

$$\left(1-\frac{\partial\left(\varphi-\frac{\omega}{2}\right)}{\partial\mu_A}\right)>\left((1-\lambda\mu_A)\frac{\partial\left(\varphi\cdot\frac{\omega}{2}\right)}{\partial\mu_A}+\frac{(1-\lambda\mu_B)}{\mu_A-\mu_B}\left(\varphi\cdot\frac{\omega}{2}\right)\right)$$

therefore, $\frac{\partial\Gamma_A}{\partial\mu_A}<0$.

The same result is obtained when computing the sign of the derivative of Γ_B with respect to μ_B, given that

$$\begin{aligned}\frac{\partial\Gamma_B}{\partial\mu_B}=&\frac{1}{9(1+s)}\left(\varphi-\frac{\omega}{2}\right)\frac{\partial\left(\varphi-\frac{\omega}{2}\right)}{\partial\mu_B}-\frac{\sigma_B^*}{3}\frac{\partial\left(\varphi-\frac{\omega}{2}\right)}{\partial\mu_B}+\\&-\frac{1}{9(1+s)}\left(\varphi-\frac{\omega}{2}\right)+\frac{(1-\lambda\mu_B)\mu_A}{2(1+s)(\mu_A-\mu_B)}\frac{\partial\left(\varphi\cdot\frac{\omega}{2}\right)}{\partial\mu_B}+\\&\frac{(1-\lambda\mu_A)\mu_A}{2(1+s)(\mu_A-\mu_B)^2}\left(\varphi\cdot\frac{\omega}{2}\right).\end{aligned}$$

Bibliography

Analysis Mason (2008). Case studies of mobile termination regimes in Canada, Hong Kong, Singapore and the USA: report for Ofcom. Technical report, Analysis Mason.

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

Armstrong, M. (2002). *The Theory of Access Pricing and Interconnection*, volume 1, chapter 8, (pp. 295–386). North-Holland.

Armstrong, M. (2006). Competition in Two-Sided Markets. *RAND Journal of Economics*, *37*(3), 668–691.

Armstrong, M., Doyle, C., & Vickers, J. (1996). The Access Pricing Problem: A Synthesis. *The Journal of Industrial Economics*, *44*(2), 131–150.

Armstrong, M. & Sappington, D. E. (2006). Regulation, Competition, and Liberalization. *Journal of Economic Literature*, *XLIV*, 325–366.

Armstrong, M. & Wright, J. (2009). Mobile Call Termination. *The Economic Journal*, *199*, 270–307.

Baake, P. & Mitusch, K. (2007). Competition with Congestible Networks. *Journal of Economics*, *91*(2), 151–176.

BEREC (2010a). Consultation Report to the Common Statement "Next Generation Networks Future Charging Mechanisms / Long

Term Termination Issues". Technical report, BEREC. available at http://www.erg.eu.int/doc/berec/bor_10_24b.pdf.

BEREC (2010b). Next Generation Networks Future Charging Mechanisms / Long Term Termination Issues. Technical report, BEREC. available at http://www.erg.eu.int/doc/berec/bor_10_24_ngn.pdf.

Bomsel, O., Cave, M., Blanc, G. L., & Neumann, K.-H. (2003). How mobile termination charges shape the dynamics of the telecom sector. Technical report, University of Warwik and Wik Consult.

Cadman, R. (2007). NGN Interconnection: Charging Principles and Economic Efficiency. Technical report, NGN UK, London.

Cambini, C. & Valletti, T. (2008). Information Exchange and Competition in Communicatios Networks. *The Journal of Industrial Economics*, *22*, 1821.

Cheng, H. K., Bandyopadhyay, S., & Guo, H. (2011). The Debate on Net Neutrality: A Policy Perspective. *Information Systems Research*, *22*, 60–82.

Choi, J. P. & Kim, B.-C. (2010). Net Neutrality and Investment Incentives. *RAND Journal of Economics*, *41*(3), 446–471.

CMT (2009). Informe Anual de la Comisión del Mercado de las Telecomunicaciones [in Spanish].

de Bijl, P. W. & Peitz, M. (2002). *Regulation and Entry into Telecommunications Markets*. Cambridge University Press.

de Bijl, P. W. & Peitz, M. (2004). Dynamic regulation and entry in telecommunications markets: a policy framework. *Information Economics and Policy*, *16*, 411–437.

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

Ellison, G. (2006). Bounded Rationality in Industrial Organization. *Econometric Society Monographs*, *42*, 142.

ERG (2009). Next Generation Networks Future Charging Mechanisms / Long Term Termination Issues. Technical report, European Regulatory Group.

European Commission (2008). Explanatory note accompanying document to the Commission Reccomandations on the Regulatory Treatment of Fixed and Mobile Termination Rates in the EU. Technical report, European Commission, Brussels.

European Commission (2009). Commission Recommentation of 7 May 2009 on the Regulatory Treatment of Fixed and Mobile Termination Rates in the EU (2009/396/EC). Technical report, European Commission, Brussels.

Gans, J. S. & King, S. P. (2001). Using 'bill and keep' interconnect arrangements to soften network competition. *Economics Letters*, *71*(3), 413 – 420.

Hermalin, B. E. & Katz, M. L. (2007). The Economics of Product-Line Restrictions with an Application to the Network Neutrality Debate. *Information Economics and Policy*, *19*(2), 215–248.

Hermalin, B. E. & Katz, M. L. (2009). Customer or Complementor? Intercarrier Compensation with Two-Sided Benefits. Working Paper.

Hoernig, S. (2007). On-net and off-net pricing on asymmetric telecommunications networks. *Information Economics and Policy*, *19*(2), 171 – 188.

Hoernig, S. & Harbord, D. (2010). Welfare Analysis of Regulating Mobile Termination Rates in the UK (with an Application to the Orange/T-Mobile Merger). Working Paper.

Jeon, D.-S., Laffont, J.-J., & Tirole, J. (2004). On the "Receiver-Pays" Principle. *The RAND Journal of Economics*, *35*(1), 85–110.

Kridel, D. J., Sappington, D. E., & Weisman, D. L. (1996). The effects of incentive regulation in the telecommunications industry: A survey. *Journal of Regulatory Economics*, *9*(3), 269–306.

Krämer, J. & Wiewiorra, L. (2009). Network Neutrality and Congestion Sensitive Content Providers: Implications for Service Innovation, Broadband Investment and Regulation. MPRA Paper 16655, University Library of Munich, Germany.

Laffont, J.-J., Rey, P., & Tirole, J. (1998a). Network Competition: I. Overview and Nondiscriminatory Pricing. *The RAND Journal of Economics*, *29*(1), 1–37.

Laffont, J.-J., Rey, P., & Tirole, J. (1998b). Network Competition: II. Price Discrimination. *The RAND Journal of Economics*, *29*(1), 38–56.

Laffont, J.-J. & Tirole, J. (2000). *Competition in telecommunications*. MIT Press, Cambridge.

Littlechild, S. (2006). Mobile termination charges: Calling Party Pays versus Receiving Party Pays. *Telecommunications Policy*, *30*(5-6), 242 – 277.

López, Á. L. (2011). Mobile termination rates and the receiver-pays regime. *Information Economics and Policy*, *23*(2), 171–181.

Musacchio, J., Schwartz, G., & Walrand, J. (2009). A Two-Sided Market Analysis of Provider Investment Incentives with an Application to the Net-Neutrality Issue. *Review of Network Economics*, *8*(1), 3.

Möbius, M. (2001). Death through Success: The Rise and Fall of Local Service Competition at the Turn of the Century. Technical report, mimeo.

Ofcom (2005). Wholesale mobile voice call termination. Preliminary consultation on future regulation. Technical report, Ofcom.

Ofcom (2009). Wholesale mobile voice call termination. Preliminary consultation on future regulation. Technical report, Ofcom.

Ofcom (2010). Annual Report and Accounts. Technical report.

Oftel (2002). Receiving Party Pays compared to Calling Party Pays. Technical report, Oftel.

Peitz, M. (2005). Asymmetric access price regulation in telecommunications markets. *European Economic Review*, *49*, 341–358.

Rochet, J.-C. & Tirole, J. (2006). Two-Sided Markets: A Progress Report. *RAND Journal of Economics*, *37*(3), 645–667.

Samarajiva, R. & Melody, W. H. (2000). Briefing paper. In *Fixed-Mobile Interconnection Workshop.* available at `http://www.itu.int/osg/spu/ni/fmi/workshop/`.

Schuett, F. (2010). Network Neutrality: A Survey of the Economic Literature. *Review of Network Economics*, *9*(2), 1–14.

Sluijs, J. P. (2010). Network Neutrality between False Positives and False Negatives: Introducing a European Approach to American Broadband Markets. *Federal Communications Law Journal*, *62*, 77–117.

Valletti, T. (1998). Two-part access pricing and imperfect competition. *Information Economics and Policy*, *10*, 305–323.

Vickers, J. (1997). Regulation, Competition, and the Structure of Prices. *Oxford Review of Economic Policy*, *13*(1), 15–26.

Vogelsang, I. (2003). Price Regulation of Access to Telecommunications Networks. *Journal of Economic Literature*, *XLI*, 830–862.

Waverman, L. & Sirel, E. (1997). European Telecommunications Markets on the Verge of Full Liberalization. *Journal of Economic Perspectives*, *11*(4), 113–126.