

Mobile Termination, Network Externalities, and Consumer Expectations *

Sjaak Hurkens[†] Ángel L. López[‡]

March 9, 2010

Abstract

We re-examine the literature on mobile termination in the presence of network externalities. Externalities arise when firms discriminate between on- and off-net calls or when subscription demand is elastic. This literature predicts that profit decreases and consumer surplus increases in termination charge in a neighborhood of termination cost. This creates a puzzle since in reality we see regulators worldwide pushing termination rates down while being opposed by network operators. We show that this puzzle is resolved when consumers' expectations are assumed passive but required to be fulfilled in equilibrium (as defined by Katz and Shapiro, AER 1985), instead of being rationally responsive to non-equilibrium prices, as assumed until now.

Keywords: Networks, Rational Expectations, Access Pricing, Interconnection, Regulation, Telecommunications

JEL numbers: D4, K23, L51, L96

*We thank Luís Cabral, Giulio Federico, Doh-Shin Jeon and Bruno Jullien, and the seminar audiences at IAE (Barcelona) and University of Groningen for helpful comments. Sjaak Hurkens gratefully acknowledges financial support through grant ECO2009-12695 from the Spanish Ministry of Science and Innovation, through grant CONSOLIDER-INGENIO 2010 (CSD2006-00016), and thanks the support of the Barcelona GSE research network, and of the Generalitat de Catalunya. Ángel López gratefully acknowledges the financial support from the Spanish Ministry of Science and Technology under ECO2008-05155 and from Juan de la Cierva Program.

[†]Institute for Economic Analysis (CSIC) and Public-Private Sector Research Center (IESE Business School, University of Navarra, Spain). E-mail sjaak.hurkens@iae.csic.es.

[‡]Public-Private Sector Research Center, IESE Business School, University of Navarra, Spain. E-mail alopezr@iese.edu.

1 Introduction

The market for mobile telecommunication services is in many countries rather competitive, as consumers have often the choice between at least three or four network operators so that operators have to compete vigorously to attract (and retain) customers. An important technological aspect of telecommunications is that subscribers of different network operators can call each other, which means that different networks are compatible. The possibility (and even obligation) to interconnect two networks has been key to promote entry and competition. Surely, the value of owning a mobile would be much reduced if one could only call people that subscribe to the same network. Unfortunately, the existence of competition and the need to interconnect has created a new bottleneck. Since most people will only subscribe to one of the networks, each network holds a monopoly position over the so called market for termination of calls directed to its subscribers.¹ If each network is free to set the price it will charge other networks for terminating calls, this termination charge (sometimes referred to as access price) would be set at a (too) high level. This in turn would induce networks to price calls at inefficiently high levels, and thus make the retail market rather uncompetitive, at least in terms of efficiency and prices, despite the presence of multiple networks.²

Surprisingly, more competition in the retail market may actually exacerbate the problem. It is easily established that, under the usual assumption of a balanced calling pattern, the fraction of off-net calls equals $1 - HHI/10000$, where HHI denotes the Herfindahl-Hirschman Index. Hence, the less concentrated the market is, the larger the fraction of calls that will be off-net, which are the ones that have to pass through the bottleneck. High termination charges in relatively competitive retail markets could prove to be particularly harmful when operators charge different prices for on- and off-net calls, since then high termination charges may affect off-net prices more than on-net prices.

Regulators around the world, and especially in the European Union, are concerned about too high termination charges and intervene in the markets of termination. Initially regu-

¹European national regulatory authorities are required by the European Commission to periodically assess the markets for termination and typically find that all mobile network operators have significant market power in these markets.

²In countries where consumers are charged for receiving calls, such as the US, it is not so clear whether the terminating operator wants to set a very high termination charge to the originating operator. In these countries termination charges are often low (eg. *Bill and Keep*). However, it is not obvious whether low termination charges induce firms to charge consumers for receiving calls, or whether the fact that consumers are charged for receiving calls induces firms to set low termination charges. The current paper deals with the case where only the calling party pays, which happens to be the case in most countries, especially in the European Union.

lators allowed higher termination charges for late entrants in order for them to overcome start-up problems. More recently regulators tend to treat operators more equally and lowered termination charges for all of them considerably. At present, the European Commission recommends national regulators to push termination rates further down to the cost of terminating a call by the end of 2012. (EC 2009a).³ Mobile operators have repeatedly opposed the cuts in termination rates imposed by the national regulatory authorities (NRAs) during the last decade. This is of course a clear indication that mobile operators fear to see their profits reduced when termination charges are decreased.⁴ This seems inconsistent with the argument of some operators that excessive termination charges are irrelevant because these will be returned to consumers in the form of lower retail prices for some mobile services, such as hand-set subsidies.⁵ Some operators even warned regulators that reducing termination charges would distort competition and hurt consumers because increased subscription fees would reduce mobile penetration.⁶ Some NRAs did not believe that a reduction in termination charge would lead to an increase in retail price. Others, on the other hand, accepted the argument that above cost termination charges could be used to subsidize marginal consumers to join a network, increase mobile penetration and thereby internalize the network externality. The UK regulator Ofcom calculated the so called externality surcharge to be positive, but very mild and took it into account when determining the termination charge (Ofcom, 2007). The European Commission, however, recommends against applying a surcharge and aims for termination charges equal to cost (EC 2009b, par 5.2.4.).

The initial academic literature on two-way termination charges confirms the concerns of regulators that high termination charges lead to excessive profits and, more importantly, to an inefficient structure of retail prices.⁷ Armstrong (1998) and Laffont, Rey and Tirole (1998a) show that, in a setting where firms set one linear price, firms prefer to have above cost termination charges, whereas (total or consumer) welfare maximizing termination charges are below cost. Laffont, Rey, and Tirole (1998b) find that if firms set linear prices but can discriminate between on- and off-net calls, then above cost termination charges induce

³In the initial stages of wireless telecommunication the most important regulatory issue was the fixed-to-mobile (FTM) termination rate, i.e. the price to be paid by the incumbent land-line operator for calls terminating on a mobile network. High FTM charges acted as a subsidy for the mobile operators in the initial phase of development. At present the mobile market is rather competitive and well-developed in most European countries. There is less need to subsidize mobile operators and the main issue that regulators are concerned about is how mobile termination charges affect competition. In this paper we do not consider the fixed line operators and focus exclusively on the mobile market.

⁴T-Mobile made this concern explicit in response to the 2006 public consultation procedure in the UK. See Ofcom, 2006, par 7.12.

⁵See Ofcom, 2006, par. 7.7.

⁶Ofcom, 2007, par. 7.8.

⁷For an excellent literature review on two-way access pricing we refer the reader to Armstrong (2002).

firms to increase the price for off-net calls and reduce it for on-net calls. The effect on profits is ambiguous: If firms are sufficiently differentiated, joint profits are maximized by a termination charge above cost, while if firms are close substitutes, joint profits may be maximized by a termination charge below cost.⁸ Again, welfare maximizing termination charges are below cost.

In practice, many contracts for telecommunication services use non-linear prices since they include monthly subscription or minimum usage fees. Firms will then set the variable usage price efficiently so as to maximize consumer surplus, and extract rents and compete for market share by charging a fixed fee. Laffont, Rey and Tirole (1998a) consider two-part tariffs when firms cannot charge different prices for on- and off-net calls. The usage price will reflect the average marginal cost so that firms make profits from on-net calls but suffer a loss from outgoing off-net calls. This induces firms to compete more fiercely for market share by lowering their fixed fee. Laffont, Rey and Tirole (1998a) show that there is an exact profit neutrality result: Firms do not care about the level of termination charges since the extra profit earned on calls through higher call prices is exactly off-set by the lowered fixed fee. There is a 100 per cent waterbed effect.⁹ The profit neutrality result suggests that firms should not oppose the proposals of regulators to set termination charges at the efficient level (in this context, equal to the cost of termination). This seems quite at odds with the fact that operators usually do oppose cuts in termination rates.¹⁰ The profit neutrality result seems to be a knife-edge result and a minor change in the model could overturn it, in either direction.¹¹ Carter and Wright (2003) show that in an asymmetric duopoly the large firm always strictly prefers termination charge at cost while the small firm prefers termination charge at cost only if the asymmetry is large. Dessein (2003) shows that when subscription demand is elastic, a positive externality surcharge is typically welfare increasing, but that firms would obtain higher profits from *below* cost termination charges. The intuition for this result is that there exists a positive network externality since each consumer benefits from having an additional subscriber. Because of competition, each firm only partially internalizes this externality. However, by having low termination charges and usage prices, the externality

⁸Laffont, Rey, and Tirole (1998b, Proposition 2) only point out that the profit maximizing termination charge initially increases in the substitutability parameter but may decrease for high values of this parameter. They did not show that in fact the optimal termination charge may be below cost.

⁹The term *waterbed effect* was first coined by the late Prof. Paul Geroski during the investigations of the impact of fixed-to-mobile termination charges on retail competition. See also Genakos and Valletti (2009).

¹⁰The only case we are aware of a firm that favors low termination charges is the mobile phone company “3” (Hutchinson) in the UK. In May 2009 it started a campaign “Terminate the Rate” to lower termination rates.

¹¹It is worth noting though that neither increasing the number of firms (Tangerås, 2009) nor allowing for heterogeneity of customers (Dessein, 2003 and Hahn, 2004) suffices to overturn this result.

becomes larger and more of it will be internalized. These results are of course again at odds with the resistance of operators against reductions in termination charges although they do largely support the views of NRAs.

Another striking result applies when firms use non-linear prices and are allowed to charge different prices for on-net and off-net calls (as many operators in fact do). In this case, variable prices will be set equal to perceived marginal cost, so that equilibrium profits accrue from the collection of fixed fees and from the provision of termination services. Laffont, Rey and Tirole (1998b) show that the total profit of firms is strictly decreasing in termination charge. Building on this result, Gans and King (2001) show that firms strictly prefer below cost termination charges. The intuition behind this result is that when there is a price differential between on- and off-net calls, there are so-called tariff-mediated network externalities: consumers care about the size of each network. In particular, when termination charge is above cost, off-net calls will be more expensive than on-net calls so that consumers will then prefer to belong to the larger network. As a result, lowering the fixed fee will become a more effective competitive tool to increase market share and price competition is intensified.¹² Clearly, firms prefer instead to soften competition and this can be attained by having termination charge below cost, which comes at the expense of reduced social welfare and consumer surplus.

This result has been shown to be very robust. It holds for any number of networks (Calzada and Valletti, 2008), when call externalities are taken into account (Berger, 2005) and when networks are asymmetric (López and Rey, 2009). Hurkens and Jeon (2009) show that the result also holds when there are both direct network externalities (i.e., elastic subscription demand as in Dessein, 2003) and tariff-mediated network externalities (i.e. on- and off-net price differentiation as in Gans and King, 2001).

The theoretical results with non-linear pricing show that the collusion concern that firms can obtain excessive profits are not necessarily associated with high termination charges, as the models with linear pricing argue. From a policy perspective, however, excessive profits are only of a secondary concern. The primary concern should be the efficiency of the retail pricing structure. The above models predict that welfare maximizing termination charges are below cost (when prices are linear) and at cost (when prices are non-linear) with a positive or negative mark-up, depending on whether network externalities or call

¹²A similar intuition applies for the case of linear prices considered before. In that case, lowering the on-net price becomes a more effective competitive tool. As seen before, the effect of increasing termination charge on firms' profit is ambiguous in the case of linear discriminatory prices since it depends on the substitutability between firms.

externalities, respectively, are thought to be relevant. Reducing termination charges to cost seems therefore certainly welfare increasing and may be at least a good second-best policy. However, this conclusion is drawn from models that at the same time predict that operators would favor reductions in termination charges, which is certainly wrong. Hence, regulatory policy to date has been based on an incomplete understanding of strategic interaction in mobile markets.

Very recently, a few attempts have been made to reconcile theory with real world practice. Armstrong and Wright (2009)¹³, Jullien, Rey and Sand-Zantman (2009)¹⁴, and Hoernig, Inderst and Valletti (2009)¹⁵ have in common that they introduce additional realistic features of the telecommunication industry into the Laffont, Rey and Tirole (1998b) framework and then show that for some parameter range joint profits are maximized at termination charges above cost. Moreover, these papers conclude that the need to regulate termination charges is reduced since the socially optimal termination charge would also be above cost.

We present in this paper a rather different solution to the puzzle, and also come to a very different conclusion. Instead of adding one or more realistic features of telecommunication competition to the Laffont, Rey and Tirole (1998b) framework, we focus on an element that has always been present in this framework, but has practically gone unnoticed. Namely, we address the issue of how consumers form expectations about the size of networks. These expectations are crucial whenever externalities exist. We show that a seemingly innocuous twist of the modeling of such expectations predicts, quite generally, that (i) firms prefer termination charges above cost while (ii) socially optimal termination charges are below or at cost.

The literature thus far assumes that first firms compete in prices, then consumers form

¹³Armstrong and Wright (2009) argue that mobile-to-mobile charges cannot be seen independently from fixed-to-mobile charges. Mobile operators prefer high FTM charges as this provides a subsidy to mobile networks. If MTM and FTM charges must be chosen uniformly, as is in fact the case in most European countries, firms will trade off desirable high FTM and desirable low MTM charges and arrive at some intermediate level, which may well be above cost. However, for this argument to go through, one needs to assume that (i) there is some room for mobile market expansion, and that (ii) income from fixed lines is sufficiently important. If (i) does not hold, FTM subsidies would be completely competed away because of the waterbed effect and firms would then still prefer termination charges below cost. If (ii) does not hold, the Gans and King (2001) effect of softening competition between mobile operators would dominate and firms again would prefer termination charges below cost.

¹⁴Jullien, Rey and Sand-Zantman (2009) consider a model with heterogeneous consumers. “Small users” are assumed to have a more elastic demand than “large users” and not all small users necessarily participate. Assuming that firms do not differentiate between on- and off-net calls, termination charges above cost are both good for firms and for welfare, as it increases the competition for and the participation of small users.

¹⁵Hoernig, Inderst and Valletti (2009) assume the existence of calling clubs. That is, people are more likely to call friends (people similar to themselves). In this case the calling pattern is not uniform but skewed. They show that in this case usage prices will not be set equal to perceived marginal cost. In fact, on-net price will be above, and off-net price will be below perceived marginal cost. They show that if the calling pattern is very skewed, then firms prefer to have termination charge above cost.

expectations about network sizes (and these thus may depend on the prices chosen by firms) and finally consumers make optimal subscription decisions, given the prices and their expectations. A strong rationality condition is imposed on expectations. Namely, for *all* prices expectations are required to be self-fulfilling. We will refer to such expectations as rationally responsive. Consumers having rationally responsive expectations means that any change of a price, how tiny it may be, by one firm is assumed to lead to an instantaneous *rational* change in expectations of all consumers, such that, given these changed expectations, optimal subscription decisions will lead realized and expected network sizes to coincide. So a unilateral change in price does not lead only to a change in market shares, but it also leads consumers to accurately predict how market shares will change.

In this paper we propose to relax the assumption of rationally responsive expectations and to replace it by one of fulfilled equilibrium expectations. This concept was first proposed by Katz and Shapiro (1985). Katz and Shapiro (1985) assume that first consumers form expectations about network sizes, then firms compete (in their Cournot model by setting quantities), and finally consumers make optimal subscription or purchasing decisions, given the expectations. These decisions then lead to actual market shares and network sizes. Katz and Shapiro impose that, *in equilibrium*, realized and expected network sizes are the same. We will refer to such expectations throughout the paper as passive (self-fulfilled) expectations. They are passive as they do not respond to out of equilibrium deviations by firms.

We show that when expectations are assumed passive, as in Katz and Shapiro (1985), results about termination charges in mobile network industries are in fact in line with real world observations. Firms typically prefer above cost termination charges and regulators are justified in their efforts to push termination charges down. In particular, and most importantly, we overturn the Gans and King (2001) result. When firms compete in non-linear prices and can distinguish between on- and off-net calls (and subscription demand is inelastic), firms prefer termination charge above cost so that off-net calls are priced at monopoly prices. Fixed fees and on-net prices are not influenced by the termination charge and thus, in this model, there is no waterbed effect at all.¹⁶ Socially optimal termination charges are equal to cost. When subscription demand is elastic, a termination charge above cost reduces participation, consumer surplus and total welfare. From a social point of view it is thus optimal to set termination below cost, as it helps to internalize the network effect. Although *Bill*

¹⁶The complete absence of a waterbed effect depends on the assumption of duopoly. We show that in oligopolies with more than two firms a partial waterbed effect exists. In any case, firms prefer termination charges above cost.

and Keep (zero termination charges) is not necessarily optimal, it could perform better than cost-based termination charges. On the other hand, firms prefer termination charge above cost, unless the network effect is so strong that firms prefer to increase penetration rather than to increase fixed fees. This means that in most European countries — with penetration rates already close to 100 per cent — firms prefer above cost termination charges.

It turns out that characterizing equilibrium prices by means of first-order conditions is easier when expectations are assumed passive than when expectations are assumed rationally responsive. This allows us to consider many different extensions in a single paper. Our main result is shown to be robust to the inclusion of call externalities, as in Berger (2005). If the call externality is modest, firms prefer again above cost termination charges. If the call externality is very strong, however, firms prefer termination charge below cost in order to reduce connectivity breakdown. This is because in this case, even when termination is charged at cost, off-net call prices would be too high, above the monopoly level. We also re-examine Laffont, Rey and Tirole (1998b) when firms compete in linear prices. We find that on-net price is independent of termination charge, and that off-net price is increasing in termination charge. Consequently, profits are maximized by a termination charge above cost.¹⁷ Finally, we also consider the case of brand loyalty causing asymmetric networks and show that both networks will prefer above marginal cost termination charges. This happens despite the fact that the smaller firm will compete more aggressively for market share when consumers come with termination profits.

The rest of the paper is organized as follows. In the next section we briefly discuss the vices and virtues of the different assumptions about consumer expectations. Section 3 introduces the general model with passive expectations. Section 4 deals with the models in which all consumers subscribe to one of the networks. We start in subsection 4.1 with the case in which firms use non-linear prices and distinguish between on- and off-net calls, as in Gans and King (2001). In subsection 4.2 we extend the model to allow for call externalities. Next, in subsection 4.3 we re-examine Laffont, Rey and Tirole (1998b) where firms compete in linear prices and distinguish between on- and off-net calls. The last part of section 4 is dedicated to the case of two asymmetric networks as in Carter and Wright (1999, 2003), but where firms distinguish between on- and off-net prices (as in López and Rey, 2009). Section 5 deals with elastic subscription demand, so that the total number of subscribers is endogenous. Firms compete in non-linear prices. We examine both the case of termination-

¹⁷This result depends again on the duopoly assumption. With more than two firms on-net price is decreasing in termination charge but joint profits are still maximized by termination charge above cost.

based price discrimination (as in Hurkens and Jeon, 2009) and the case where firms must set the same price for on- and off-net calls (as in Dessein, 2003). Section 6 concludes. Proofs for sections 4 and 5 are collected in Appendix A and B, respectively.

2 Passive versus Responsive Expectations

Expectations are important in any model with network effects, not just in the case of telecommunication. Examples include two-sided markets such as newspapers or credit cards. Readers care about the number of ads and advertisers care about the number of readers. Merchants care about the number of users of a particular credit card and users care about the number of merchants accepting a particular credit card. Network effects can also occur in financial markets. The riskiness of a bank may very well depend on its size, that is, the number of depositors. Of course, depositors will care about the riskiness and thus about the number of other people who will deposit in a given bank. (See Matutes and Vives, 1996.) Expectations even play a role in markets without network effects. For example, consider a monopolistic upstream supplier of an input to several downstream firms that compete with each other in a final product market. The prices paid for the inputs determine the marginal costs for the downstream firms. If the prices of inputs are set secretly, each downstream firm needs to form expectations about the prices paid by its competitors in order to know how profitable competition will turn out to be and to determine the demand for inputs. In this context passive beliefs seem very reasonable and widely accepted.¹⁸

Many papers have been written on markets with network effects and some have modeled consumer expectations as passive and some have modeled them as rationally responsive.¹⁹ Very few papers justify or even discuss the assumption about expectations. Katz and Shapiro (1985) do mention the possibility of responsive beliefs in their Appendix, but in their quantity setting framework results are not altered in an important manner.²⁰ Lee and Mason (2001) take the issue of expectations serious and point out that the results change dramatically if rationally responsive beliefs are used in their pricing game. Matutes and Vives (1996) char-

¹⁸See Hart and Tirole (1990), O'Brien and Shaffer (1992), Fontenay and Gans (2005) and Rey and Tirole (2007). McAfee and Schwartz (1994) also consider so called "wary", non-passive, beliefs.

¹⁹Examples of the first include Katz and Shapiro (1985), Matutes and Vives (1996), Economides (1996b), Lee and Mason (2001) and Malueg and Schwartz (2006). Examples of the latter include Crémer, Rey and Tirole (2000), Caillaud and Jullien (2003) and Galeotti and Moraga-González (2009).

²⁰Hermalin and Katz (2009) also consider Cournot competition. They argue that "the Cournot model can be viewed as a means of approximating a dynamic process in which consumer expectations with respect to network sizes change slowly over time because consumers observe network sizes and predict that these sizes will remain stable."

acterize the equilibria under passive beliefs but do point out that with rationally responsive expectations any pair of deposit rates leading to non-negative profits can be sustained as an equilibrium. Finally, Griva and Vettas (2004) analyze price competition in a duopoly where products are both horizontally and vertically differentiated and exhibit positive, product-specific network effects. They do so both for the case where prices do not influence consumer expectations (passive) and for the case where firms can influence expectations through prices (rationally responsive). They point out that competition is more intense under the latter assumption.

In order to illustrate the difference between passive and responsive expectations, and to explain why rationally responsive expectations may intensify competition, let us consider a duopolistic industry with network effects. Each network is located at one end of the Hotelling interval $[0, 1]$ over which consumers are uniformly distributed. Suppose the value of subscribing to network i equals $v_i(\alpha_i)$, where α_i denotes the size of network i . We will assume here that network effects are positive (i.e., $v'_i > 0$) and that networks compete for consumers in flat fees, denoted by F_1 and F_2 . Given these fees, market shares are stable at $(\alpha_0, 1 - \alpha_0)$ if the consumer at location α_0 is exactly indifferent between the two networks, that is when

$$v_1(\alpha_0) - t\alpha_0 - F_1 = v_2(1 - \alpha_0) - t(1 - \alpha_0) - F_2,$$

where $t > 0$ denotes the Hotelling transportation cost. In other words, given fees F_1 and F_2 expectations $(\alpha_0, 1 - \alpha_0)$ are fulfilled.

Now let us investigate what happens when suddenly firm 1 lowers its price to $F_1 - \Delta$. How will consumers react?

If consumers take into account only the *direct* pecuniary effect of the lower price, the result will be that some consumers will switch to network 1. The consumers who will switch are those at locations $x \in (\alpha_0, \alpha_1)$, where α_1 is defined by

$$v_1(\alpha_0) - t\alpha_1 - (F_1 - \Delta) = v_2(1 - \alpha_0) - t(1 - \alpha_1) - F_2.$$

That is,

$$\alpha_1 = \alpha_0 + \Delta/(2t).$$

After these switches have occurred, network 1 has increased in size. Since network effects are assumed to be positive, the consumer located at α_1 will no longer be indifferent. In fact,

all consumers located at $x \in (\alpha_1, \alpha_2)$ will now switch to network 1, where

$$\alpha_2 = \alpha_1 + [v_1(\alpha_1) - v_1(\alpha_0) + v_2(1 - \alpha_0) - v_2(1 - \alpha_1)]/(2t).$$

Note that the right-hand side does not directly depend on Δ , but only *indirectly* through the changed network sizes. Of course, the story continues as network sizes have changed again. Defining recursively

$$\alpha_{k+1} = \alpha_k + [v_1(\alpha_k) - v_1(\alpha_{k-1}) + v_2(1 - \alpha_{k-1}) - v_2(1 - \alpha_k)]/(2t),$$

one observes that in consecutive steps consumers in (α_k, α_{k+1}) will switch to network 1. In the limit, $\lim_{k \rightarrow \infty} \alpha_k = \alpha_\infty$ where

$$v_1(\alpha_\infty) - t\alpha_\infty - (F_1 - \Delta) = v_2(1 - \alpha_\infty) - t(1 - \alpha_\infty) - F_2.$$

The limit market shares $(\alpha_\infty, 1 - \alpha_\infty)$ are equal to the rationally responsive expectations given fees $F_1 - \Delta$ and F_2 . In contrast, the market shares $(\alpha_1, 1 - \alpha_1)$ correspond to the (now no longer) fulfilled passive expectations $(\alpha_0, 1 - \alpha_0)$.

One thus concludes that lowering the fee is a more competitive tool for gaining market share when expectations are rationally responsive than when expectations are passive: a decrease of the fee by Δ increases market share of network 1 under rationally responsive expectations by $\alpha_\infty - \alpha_0$, while under passive beliefs, market share is only increased by $\alpha_1 - \alpha_0$. Beliefs are passive when consumers only take into account the direct pecuniary effect of a price decrease. Beliefs are rationally responsive when consumers also take into account all the indirect, higher order, effects that a decrease of the fee has on network size.

In the remainder of this section we briefly point out some of the virtues and vices of rationally responsive expectations. We do not pretend to provide a complete discussion nor do we intend to convince the reader that the assumption of passive beliefs is always more appropriate than the one of rationally responsive expectations. We hope to convince the reader that considering passive beliefs is not completely unreasonable and that exploring the consequences of employing this assumption in the context of telecom competition is very fruitful.

- ***Subgame Perfection***

The assumption of rationally responsive expectations is theoretically quite pleasing, be-

cause of its similarity with subgame perfection. Indeed, the assumption is equivalent to demanding that consumers play a Nash equilibrium of the subscription/purchasing/deposit game, once prices are set, for any prices. An equilibrium of the price setting game then corresponds to a subgame perfect equilibrium of the whole game. Instead, assuming passive self-fulfilling expectations will just correspond to a Nash equilibrium of this game. It should be noted though that subgame perfect equilibrium outcomes are not always more appealing than other Nash equilibrium outcomes. A clear example is the ultimatum game, where the first player proposes a division of a dollar and the second player can only accept or reject. The subgame perfect equilibrium predicts that the proposer will propose to take (almost) all and that the responder will accept. There are Nash equilibria in which a more even split is agreed upon, and such outcome can even be evolutionarily stable.²¹ Experimental evidence clearly shows very little support for the subgame perfect equilibrium outcome.

• *Multiplicity and Coordination*

A typical problem of models with network effects is the multiplicity of equilibria.²² For example, an equilibrium with positive sales often coexists with one where no trade takes place. The latter equilibrium is supported by having consumers believe that nobody else will participate in the market. Also, a shared market equilibrium may coexist with cornered market equilibria. For example, in the Hotelling model outlined above cornered market equilibria exist when $v_1(0) = v_2(0) = 0$: if all consumers expect one network to corner the market, this network will in fact corner the market. This type of multiplicity of equilibria can occur both with passive and with responsive expectations.

A potentially more severe problem — which only occurs with responsive beliefs — is that for given prices multiple rational expectations exist. This can occur because the rational expectations demand curve can have increasing parts, so that for the same price three or more rational demands exist.²³ Also, in the Hotelling model for some prices both shared market and cornered market expectations can be rational. In such a case the assumption that consumers are able to coordinate on one of the equilibria is very strong. Moreover, in such a case it is very easy to construct many different equilibria by threatening to continue with a very bad equilibrium in case some firm deviates from the candidate equilibrium prices. This occurs for example in the Matutes and Vives (1996) model of depositors and is also

²¹See Gale, Binmore, and Samuelson (1995).

²²See section 3.4 in Farrell and Klemperer (2007) for an extensive discussion of the coordination problem in network industries.

²³See for example Economides (1996a).

mentioned as a possibility in Galeotti and Moraga-González (2009).

It is also worthwhile to point out that in the case of multiplicity of rational expectations for given prices, some rational expectations are sometimes discarded as they are considered to be unstable. For example, the ones on the increasing part of the rational expectation demand curve are typically unstable in a tâtonnement process where consumers adapt their decisions based on previously realized network size. That is, some rationally responsive expectations are discarded (and of course, some other is selected) on the basis of an adaptive process where consumers have myopic, non-rational, expectations. The same occurs in Laffont, Rey and Tirole (1998b) when they discard the shared market equilibrium when also cornered market equilibria exist. This suggests that there seems to be some merit in assuming that consumers are myopic, naive or passive even for those that do employ the assumption of rationally responsive expectations.

• *Dynamics and Commitment*

It is well known that initial conditions and path dependence are important in markets with network effects. The most appropriate form to model such markets would inherently involve dynamics. Ideally, we would like to consider a truly dynamic model in which both consumers and firms can change their decisions over time. As long as relatively few consumers can switch in each period and consumers are impatient, the resulting prices and policy recommendations will probably not depend in very great detail on how expectations are exactly defined, because then expectations would be strongly correlated with previous period market shares in any case. Analyzing a truly dynamic model is beyond the scope of the present paper.²⁴

Calculating rationally responsive expectations involves solving a (possibly complex) fixed-point problem for any combination of prices. This requires not only a high level of analytical skill on the part of each consumer, it also requires that each consumer is confident that all other consumers will (be able to) make such calculations. An alternative, evolutive, method to arrive at the new market shares is to let consumers sequentially adapt their subscription responses to the price and the realized market shares. When no consumer wishes to change her decision any more, one has arrived at the new market shares, which would coincide with the rationally responsive expectations given the new prices.²⁵

²⁴Even within a dynamic model with network effects, consumer expectations can be modeled as myopic or as forward looking. Cabral (2009), Driskill (2007) and Laussel and Resende (2007) consider forward looking consumers. Doganoglu (2003) and Mitchell and Skrzypacz (2006) consider myopic consumers. Radner and Sundararajan (2006) allow for a mixture of forward looking and boundedly rational consumers.

²⁵See our discussion above of the Hotelling duopoly with network effects.

A potential problem with this is that it could take a very long time and, at least in the telecommunication sector, firms may want to change their prices before the consumers have fully adapted. It may be very hard for firms to commit to a price. For example, in the Spanish market, the recent entrant (Yoigo) offers a contract with free on-net calls. Clearly, it would be rational for all consumers in Spain to switch and make free calls for the rest of their lives. This does not and will not happen. Once the market share of the entrant passes a certain threshold, it will certainly withdraw the offer (or go bankrupt).

Notice that the difference between rationally responsive and self-fulfilling passive expectations does not occur for equilibrium prices (when all consumers have correct expectations in either case) but rather when prices are out of equilibrium. And exactly when prices are out of equilibrium, at least one firm has an incentive to deviate, which makes his price commitment not credible.

3 The Model

We consider competition between two full-coverage networks, 1 and 2, indexed by $i \neq j \in \{1, 2\}$. Each has the same cost structure. The marginal cost of a call equals $c = c_O + c_T$, where c_O and c_T denote the costs borne by the originating and terminating network, respectively. To terminate an off-net call, the originating network must pay a reciprocal and non-negative access charge a to the terminating network. The termination mark-up is equal to

$$m \equiv a - c_T.$$

Therefore, the perceived cost of calls is the true cost c for on-net calls, augmented by the termination mark-up for the off-net calls: $c + a - c_T = c + m$.

Networks (i.e., firms) offer differentiated but substitutable services. The two firms compete for a continuum of consumers of unit mass. Each firm i ($i = 1, 2$) charges a fixed fee F_i and may (or may not) discriminate between on-net and off-net calls. Firm i 's marginal on-net price is written p_i and off-net price is written \hat{p}_i . Consumer's utility from making calls of length q is given by a concave, increasing and bounded utility function $u(q)$. Demand $q(p)$ is defined by $u'(q(p)) = p$. The indirect utility derived from making calls at price p is $v(p) = u(q(p)) - pq(p)$. Note that $v'(p) = -q(p)$. For given prices p and \hat{p} , the profit earned on the on-net calls is

$$R(p) = (p - c)q(p),$$

whereas the profit earned on the off-net calls is

$$\hat{R}(\hat{p}) = (\hat{p} - c - m)q(\hat{p}).$$

We assume that $R(p)$ has a unique maximum at $p = p^M$, is increasing when $p < p^M$, and decreasing when $p > p^M$. That is, p^M denotes the monopoly price. We assume that $R(p^M) > f$, where f is the fixed cost per subscriber. This means that the market is viable. The Ramsey price p^R is defined as the lowest break-even price characterized by

$$R(p^R) = f.$$

We make the standard assumption of a balanced calling pattern, which means that the percentage of calls originating on a given network and completed on another given (including the same) network is equal to the fraction of consumers subscribing to the terminating network. Let α_i denote the market share of network i . The profit of network i is therefore equal to:

$$\pi_i \equiv \alpha_i \left(\alpha_i R(p_i) + \alpha_j \hat{R}(\hat{p}_i) + F_i - f \right) + \alpha_i \alpha_j m q(\hat{p}_j). \quad (1)$$

The first term represents the profit made on consumer services (on-net and off-net calls, fixed fee and cost), whereas the second term represents the profit generated by providing termination services.

We assume that the terms of interconnection are negotiated (or regulated) first. Then, for a given access charge a (or equivalently, a given m) the timing of the game is as follows:

1. Consumers form expectations about the number of subscribers of each network i (β_i) with $\beta_1 \geq 0$, $\beta_2 \geq 0$ and $\beta_1 + \beta_2 \leq 1$. We let $\beta_0 = 1 - \beta_1 - \beta_2$ denote the number of consumers that is expected to remain unsubscribed. In the case of full participation $\beta_0 = 0$ and $\beta_1 + \beta_2 = 1$.
2. Firms take these expectations as given and choose simultaneously retail tariffs $T_i = (F_i, p_i, \hat{p}_i)$ for $i = 1, 2$.
3. Consumers make rational subscription and consumption decisions, given their expectations and given the networks' tariffs.

Therefore, market share α_i is a function of prices and consumer expectations. Self-fulfilling expectations imply that at equilibrium $\beta_i = \alpha_i$.

4 Full Participation

In this section we assume that the networks are differentiated à la Hotelling. Consumers are uniformly located on the segment $[0, 1]$, whereas the two networks are located at the two ends of this segment ($x_1 = 0$ and $x_2 = 1$). For a given volume of calls q and a given income y , a consumer located at x and joining network i obtains a gross utility given by

$$y + u(q) - |x - x_i|/(2\sigma),$$

where $\sigma > 0$ measures the degree of substitutability between the two networks. To ensure full participation we will assume that the surplus derived from being connected to either network is sufficiently large. We will focus our attention on the properties of shared market equilibria, where both firms have strictly positive market shares.²⁶

4.1 Non-linear pricing and termination-based price discrimination

In this subsection we assume that firms can set a fixed fee, an on-net price and an off-net price, as in Gans and King (2001). We first characterize the prices in a shared market equilibrium. Afterwards we show that such an equilibrium indeed exists and is unique.

Given the balanced calling pattern assumption and consumer expectations β_1 and β_2 , the surplus from subscribing to network i (gross of transportation costs) equals:

$$w_i = \beta_i v(p_i) + \beta_j v(\hat{p}_i) - F_i.$$

Market share of network i is thus given by $\alpha_i = 1/2 + \sigma(w_i - w_j)$, whenever this is between 0 and 1.

Marginal cost pricing. As usual, at equilibrium with strictly positive market shares, network operators find it optimal to set cost-based usage prices. Adjusting F_i so as to maintain net surpluses w_1 and w_2 and thus market shares constant, leads network i to set p_i and \hat{p}_i so as to maximize

$$\alpha_i \left(\alpha_i R(p_i) + \alpha_j \hat{R}(\hat{p}_i) + \beta_i v(p_i) + \beta_j v(\hat{p}_i) - w_i - f \right) + \alpha_i \alpha_j m q (\hat{p}_j).$$

²⁶Cornered market equilibria, where one firm dominates the whole market, may exist, but are of little relevance in mature markets.

The first-order conditions are

$$(\alpha_i - \beta_i)q(p_i) + \alpha_i(p_i - c)q'(p_i) = 0 \quad (2)$$

and

$$(\alpha_j - \beta_j)q(\widehat{p}_i) + \alpha_j(\widehat{p}_i - c - m)q'(\widehat{p}_i) = 0. \quad (3)$$

At equilibrium, self-fulfilling expectations ($\beta_i = \alpha_i$) yield perceived marginal cost pricing as long as both firms have positive market share: $p_i = c$ and $\widehat{p}_i = c + m$. Note, however, that out of equilibrium firms do not necessarily want to set usage prices equal to marginal cost.

Market shares. If firms set usage prices equal to marginal cost, and if consumers expect market shares β_1 and β_2 , the actual market share, α_i , as a function of expectations and fixed fees F_1 and F_2 , is given by

$$\alpha_i(\beta_i, F_i, F_j) = \frac{1}{2} + \sigma(F_j - F_i) + 2\sigma\left(\beta_i - \frac{1}{2}\right)(v(c) - v(c + m)). \quad (4)$$

Equilibrium fixed fees. We now characterize the equilibrium fixed fees. Since in a shared market equilibrium network operators find it optimal to set cost-based usage prices, network i 's profit can be written as:

$$\pi_i = \alpha_i(\beta_i, F_i, F_j) [F_i - f + \alpha_j(\beta_j, F_j, F_i) R(c + m)], \quad (5)$$

where $R(c + m) = mq(c + m)$ is the profit, per incoming call, from providing termination services. In equilibrium, each firm i is optimizing given the fixed fee of the other network, F_j , and consumer expectations. Using $\partial\alpha_i/\partial F_i = -\sigma$, we have

$$\frac{d\pi_i}{dF_i} = -\sigma [F_i - f + \alpha_j(\beta_j, F_j, F_i) R(c + m)] + \alpha_i(\beta_i, F_i, F_j) [1 + \sigma R(c + m)].$$

Note that

$$\frac{d^2\pi}{dF_i^2} = -2\sigma(1 + \sigma R(c + m)).$$

This means that a necessary local second-order condition is that $1 + \sigma R(c + m) > 0$, which we will assume to hold.²⁷ Solving the first-order condition for F_i we obtain

$$F_i = \frac{f + \frac{1}{2\sigma} + (1 + 2\sigma R(c + m)) [(2\beta_i - 1)(v(c) - v(c + m)) + F_j]}{2(1 + \sigma R(c + m))}. \quad (6)$$

²⁷This condition holds for all $m \geq 0$ and also for $m < 0$ as long as $\sigma < -1/R(c + m)$.

Solving the pair of first-order conditions yields:

$$F_i = f + \frac{1}{2\sigma} + \left(\frac{1 + 2\sigma R(c + m)}{3 + 4\sigma R(c + m)} \right) (2\beta_i - 1) (v(c) - v(c + m)). \quad (7)$$

Substituting the expressions for F_1 and F_2 into Eq. (4) yields

$$\alpha_i = \frac{1}{2} + 2\sigma \left(\frac{1 + 2\sigma R(c + m)}{3 + 4\sigma R(c + m)} - \frac{1}{2} \right) (1 - 2\beta_i) (v(c) - v(c + m)). \quad (8)$$

Using the fulfilled expectations condition $\alpha_i = \beta_i$, Eq. (8) reduces to a linear equation in α_i with a unique solution: $\alpha_i = 1/2$. Note that the symmetry of the shared market equilibrium is due to the assumption of a symmetric duopoly.²⁸ There simply does not exist any asymmetric shared market equilibrium. It follows immediately that at the equilibrium

$$F^* = f + \frac{1}{2\sigma}.$$

Previous literature has suggested that lower access charges will result in higher retail prices for mobile subscribers, which is known as the waterbed effect. When consumer expectations are passive, we have that at the equilibrium the fixed fee is equal to the fixed cost f plus the Hotelling mark-up $1/2\sigma$. That is, the waterbed effect is not at work on the fixed component of the three-part tariff.

The analysis above has shown that there is a unique candidate for a shared market equilibrium. To establish the existence of such an equilibrium not only requires the local second-order condition mentioned, but also that the described strategies are in fact global maximizers. In particular, one needs to verify that no firm wants to try to corner the market, given the prices chosen by its competitor and given the expectations of consumers. Note that the firm that corners the market by lowering its fixed fee, will also want to adjust the on- and off-net prices. In particular, that firm will want to set the off-net price at zero.²⁹ This is not costly to the firm (as no off-net calls will be made in a cornered market) but will fool consumers with passive expectations who believe that half of their calls will be off-net. If the indirect utility of making calls at zero price were unbounded, the firm would be able to corner the market and make unbounded profits. Hence, existence of a shared market equilibrium requires an upper bound on the utility obtained from making calls at zero price, as assumed in section 3. This requirement is mild, as consumers can at most make calls

²⁸This assumption is relaxed in Section 4.4.

²⁹The firm will also increase the on-net price above cost.

24 hours a day, but is not met by the constant elasticity demand function. The following proposition establishes the conditions for the existence and uniqueness of the shared market equilibrium.³⁰

Proposition 1 *Any shared market equilibrium is symmetric and is characterized by $p_1 = p_2 = c$, $\hat{p}_1 = \hat{p}_2 = c + m$ and $F_1 = F_2 = f + \frac{1}{2\sigma}$. A necessary condition for existence is that $1 + \sigma R(c + m) > 0$. A sufficient condition is that σ is small enough.*

Comparative statics. The symmetric equilibrium profit is

$$\pi_1 = \pi_2 = \frac{1}{4\sigma} + \frac{1}{4}R(c + m).$$

That is, networks gain the full profit from providing termination services (without competing it away through lower fixed fees). The equilibrium profit is increasing in m when $c + m < p^M$ and decreasing when $c + m > p^M$. Equilibrium profits are maximized for the termination mark-up m^* that maximizes the termination profit:

$$\frac{d\pi_i}{dm} \equiv \frac{1}{4} \frac{dR}{dm}(c + m^*) = 0.$$

We thus have

Corollary 1 *Under non-linear pricing and termination-based price discrimination, shared-market equilibrium profits are maximized with the termination mark-up m^* that maximizes the termination profit, i.e. $m^* = p^M - c > 0$. Total welfare is maximized at $m^W = 0$.*

Laffont, Rey and Tirole (1998b) show that in the case of rationally responsive expectations, the first-order derivative with respect to the equilibrium profit is:

$$\begin{aligned} \frac{d\pi_i}{dm} &= \frac{1}{4} [-q(c + m) + mq'(c + m)] \\ &= \frac{1}{4} \frac{dR}{dm}(c + m) - \frac{1}{2}q(c + m). \end{aligned}$$

The additional term $-\frac{1}{2}q(c + m)$ is produced by the assumption that consumers change their expectations in response to any variation of prices such that they perfectly foresee realized market shares. As a result,

$$\left. \frac{d\pi_i}{dm} \right|_{m \geq 0} < 0,$$

³⁰For expositional purposes we prove the existence of equilibrium for this case only. Our subsequent results will focus on the characterization of equilibria only.

so that firms prefer below-cost access charges. Gans and King (2001) provides the intuition for this result:³¹ When m is positive, off-net calls are more expensive than on-net calls, so that users then wish to belong to the larger network (all else being equal). This makes it easier (or less costly) for firms to gain market share. This leads firms to compete more aggressively for market share (i.e., the reaction functions shift downward) which results in lower fixed fees in equilibrium. Firms prefer thus instead to have a negative termination mark-up.

Sometimes an alternative but somewhat misleading intuition has been used to explain the negative relation between the termination mark-up and the equilibrium profit. The reasoning goes as follows: When m increases, users become more profitable, in the sense that they bring with them higher termination profits, and this leads firms to compete more aggressively for market share. This is not an essentially correct explanation of the impact of m on equilibrium profits in the case of duopoly. Namely, termination profits are only made on calls originated on the rival network. The number of such calls depends on market shares. Firm i terminates $n_i = \alpha_i(1 - \alpha_i)$ of such calls. At the symmetric equilibrium $\alpha_i = 1/2$ and n_i is in fact maximized. This is independent of m so that an increase in m will not induce firms to fight more aggressively for consumers, at equilibrium, since a marginal change in the fixed fee will have no impact on n_i . What is true is that, *out of equilibrium*, the level of the termination mark-up influences the optimal fixed fee. For example, if firm i at current prices would have the smaller market share ($\alpha_i < 1/2$ and thus $F_i > F_j$), it would react by lowering its fixed fee when the termination mark-up increases. But at the same time, the rival firm with the higher market share ($\alpha_j > 1/2$) would react by increasing its fixed fee in order to reduce α_j and increase n_j . Hence, the reaction function of each firm rotates around the intersection point with the 45 degree line when m increases, and does not affect the equilibrium fixed fee.

Figure 1 illustrates the above findings. For usage prices fixed at perceived marginal cost, it shows the optimal fixed fee of firm i as a function of the fixed fee of firm j . From Eq. (6) we know this is a linear function with positive slope less than one. The symmetric equilibrium fixed fee is given by the intersection of the reaction function with the 45 degree line. An increase in the termination mark-up leads the smaller (larger) firm to compete more (less) aggressively, rotating the reaction function counterclockwise around the intersection point.

³¹Moreover, Gans and King (2001) building on this result show that the optimal termination mark-up from the operators viewpoint is not zero (as Proposition 5 of Laffont, Rey and Tirole, 1998b erroneously concludes) but below cost, so that maximum profits are above (and not bounded above by) the Hotelling level $\frac{1}{4\sigma}$, which is achieved with $m = 0$ (that is, with $a = c_T$).

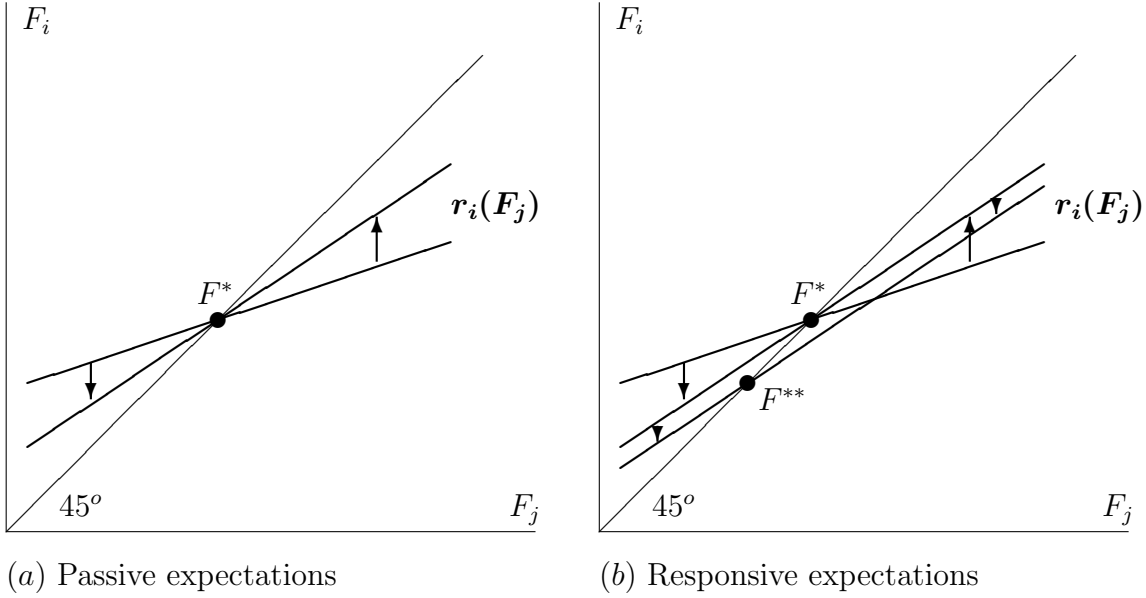


Figure 1: The effect of an increase of m on equilibrium fixed fees in duopoly.

(See Fig. 1a.) Moreover, in the case of rationally responsive expectations an increase in the termination mark-up shifts the reaction function downward. (See Fig. 1b.) This explains why only in this case an increase in the termination mark-up reduces the equilibrium fixed fee (from F^* till F^{**}) and consequently the equilibrium profit.

Oligopolistic competition. The discussion thus far has focussed on duopoly. When expectations are assumed passive, there is no waterbed effect as the equilibrium fixed fee is independent of the termination mark-up. When expectations are rationally responsive, on the other hand, there exists a very strong waterbed effect as the equilibrium fixed fee decreases so fast with m that profits in fact decrease. Nonetheless, if the number of competing networks is larger than two, then there will exist a partial waterbed effect on the fixed component of the three-part tariff even if consumer expectations are passive. To see this, assume that there exist $n \geq 3$ competing networks. The first-order condition is

$$0 = \frac{d\pi_i}{dF_i} = \left[\alpha_i + \frac{d\alpha_i}{dF_i}(F_i - f) \right] + \frac{d\alpha_i}{dF_i} (1 - 2\alpha_i) R(c + m).$$

In a symmetric equilibrium with two networks the second term of the right-hand side of the above equation disappears, thereby m has no impact on the equilibrium fixed fee. Instead, if $n \geq 3$, then at any symmetric equilibrium $\alpha_i = \frac{1}{n} < \frac{1}{2}$, implying that the equilibrium fixed fee will depend on m . Notice that at the symmetric equilibrium

$$F_i = f - \frac{1}{n} \left(\frac{d\alpha_i}{dF_i} \right)^{-1} - \left(1 - \frac{2}{n} \right) R(c + m),$$

thus F_i decreases with m as long as $m < m^*$. The reason is that the number of off-net calls

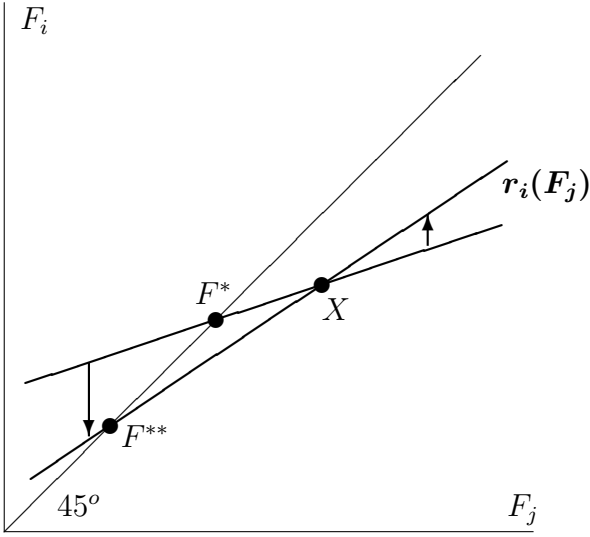


Figure 2: The effect of an increase of m on equilibrium fixed fees in oligopoly (passive expectations).

terminated on network i equals $n_i = \alpha_i(1 - \alpha_i)$ which is increasing in α_i at $\alpha_i = 1/n$. As m increases, the profit from terminating calls increases and each firm will compete more fiercely for market share. Yet, the waterbed effect is less than one hundred per cent as the equilibrium profit is still maximized with the termination mark-up that maximizes the termination profit per terminated call:

$$\pi_i = \frac{1}{n} \left[-\frac{1}{n} \left(\frac{d\alpha_i}{dF_i} \right)^{-1} + \frac{1}{n} R(c + m) \right].$$

That is, the equilibrium profit is still maximized with $m^* = p^M - c > 0$. Recall that under passive expectations $\frac{d\alpha_i}{dF_i}$ does not depend on the termination mark-up. Note that total welfare is maximized with termination charges at cost, independent of the number of firms while consumer surplus is maximized with a negative termination mark-up.

Figure 2 explains why there is a partial waterbed effect when expectations are passive and there are at least three firms. It shows the reaction function of firm i against the fixed fee F_j which is assumed to be the same for all firms $j \neq i$. Again, the intersection of this reaction function with the 45 degree line indicates the equilibrium fixed fee. An increase in termination mark-up above 0 leads the reaction function to rotate counterclockwise around the point X , defined as the point where firm i 's market share would be $1/2$. This is because the firm will fight more fiercely for market share when termination profit per call increases, as long as its market share is less than $1/2$. The equilibrium fixed fee thus decreases from F^* till F^{**} .

4.2 Call externalities

In this section we extend the model to consider call externalities, as in Berger (2005). A call externality exists if a consumer derives utility from receiving a call. It seems obvious that call externalities exist, since otherwise nobody would bother to answer a call. How strong such call externalities are is probably an empirical matter. Ofcom (2004) considers that “call externalities probably do not justify any adjustment to call prices. [...] these are likely to be effectively internalized by callers, as a high percentage of calls are from known parties and there are likely to be implicit or explicit agreements to split the origination of calls.” On the other hand, Harbord and Pagnozzi (2010) argue that call externalities are strong and that therefore *Bill and Keep* is the appropriate termination charge regime, both from a social and from a private perspective.

We assume that consumers derive utility $\bar{u}(q)$ from receiving calls of volume q , with $\bar{u} = \lambda u$, where $0 < \lambda < 1$ measures the strength of the call externality. If consumers expect market shares β_1 and β_2 , then they expect a net surplus

$$w_i = \beta_i [v(p_i) + \bar{u}(q(p_i))] + \beta_j [v(\hat{p}_i) + \bar{u}(q(\hat{p}_j))] - F_i$$

from subscribing to network i , for $i \neq j \in \{1, 2\}$. The actual market share, α_i , as a function of the consumer expectations and prices, is determined by the indifferent consumer:

$$\begin{aligned} \alpha_i &= \frac{1}{2} + \sigma \beta_i [v(p_i) - v(\hat{p}_j) + \bar{u}(q(p_i)) - \bar{u}(q(\hat{p}_i))] \\ &\quad - \sigma \beta_j [v(p_j) - v(\hat{p}_i) + \bar{u}(q(p_j)) - \bar{u}(q(\hat{p}_j))] + \sigma(F_j - F_i). \end{aligned} \quad (9)$$

Network i 's profit is given by Eq. (1). As in Berger (2005) we have

Lemma 1 *In a symmetric equilibrium with $\alpha_1 = \alpha_2 = 1/2$, networks set*

$$p_i = p^* \equiv \frac{c}{1 + \lambda} \quad \text{and} \quad \hat{p}_i = \hat{p}^* \equiv \frac{c + m}{1 - \lambda}.$$

Therefore, in a symmetric equilibrium $p_i < c$ and $\hat{p}_i > c + m$, i.e., usage prices do not reflect the perceived cost of calls. Networks find it optimal to internalize the call externality by setting the on-net price below the cost so as to extract the higher consumer surplus through the fixed fee. The off-net price, on the other hand, is set above the cost so as to reduce the utility of rival's customers from receiving calls. If λ tends to 1 (which amounts to say that callers and receivers obtain the same utility from a given call), then the off-net

price will tend to $+\infty$, resulting in connectivity breakdown (as shown in Jeon et al. 2004).

The first-order condition is

$$0 = \frac{d\pi_i}{dF_i} = -\sigma \left[\alpha_i R(p^*) + \alpha_j \widehat{R}(\widehat{p}^*) + F_i - f \right] + \alpha_i \left[-\sigma R(p^*) + \sigma \widehat{R}(\widehat{p}^*) + 1 \right] \\ + \sigma (\alpha_i - \alpha_j) m q(\widehat{p}^*),$$

which defines i 's reaction function. Hence, in a symmetric equilibrium ($\alpha_1 = \alpha_2 = 1/2$) the first-order condition is satisfied at $F_i = F^* \equiv f + \frac{1}{2\sigma} - R(p^*)$ or, equivalently,

$$F^* \equiv f + \frac{1}{2\sigma} + \lambda c q \left(\frac{c}{1 + \lambda} \right).$$

The equilibrium profit is thus

$$\pi_1 = \pi_2 = \frac{1}{4\sigma} + \frac{1}{4} [R(\widehat{p}^*) - R(p^*)].$$

The equilibrium fixed fee is independent of m while equilibrium profits depend on m through the off-net price.

Proposition 2 *Under non-linear pricing, termination-based price discrimination and call externalities, symmetric equilibrium profits are maximized with the termination mark-up m^* that maximizes the retail profit earned on the off-net calls (gross of termination payments): $m^* = \arg \max_{m \geq -c_T} R\left(\frac{c+m}{1-\lambda}\right) = \max\{(1-\lambda)p^M - c, -c_T\}$. Hence $m^* > 0$ if and only if $\lambda < \frac{p^M - c}{p^M}$.*

In contrast with Jeon et al. (2004) and Berger (2005), the termination mark-up does not affect the fixed fee. The reason is as before: if consumers do not change their expectations with price variations, then $\partial\alpha_i/\partial F_i = -\sigma$, i.e., the fixed fee affects the market share only through the direct pecuniary effect but not via a change in the expectations of consumers. Then, networks maximize shared-market equilibrium profits by setting the termination mark-up m^* that maximizes the retail profit from the off-net calls made by their subscribers. That is, m^* is so that $\widehat{p}^* = \frac{c+m^*}{1-\lambda}$ equals p^M . The equilibrium profits are therefore higher with an above cost access charge than with a below cost access charge when λ is relatively low. When λ is close to 1, there is risk of connectivity breakdown because then $\widehat{p} > p^M$, even if $m = 0$. In this case a below cost access charge brings \widehat{p}^* down towards p^M and increases profits.

The welfare maximizing termination mark-up m^W is such that \hat{p}^* satisfies the condition $\hat{p}^* = p = \frac{c}{1+\lambda}$.³² Therefore, we have

Proposition 3 *In the presence of call externalities, the socially optimal termination mark-up is negative ($a^W < c_T$) and given by $m^W = \max\{-\frac{2\lambda c}{1+\lambda}, -c_T\}$. Hence $m^W < 0 < m^*$ holds when λ is relatively low.*

This result is in contrast with Berger (2005), who shows that in the presence of rationally responsive expectations, the best termination charge from the operators' perspective is below cost and smaller than the socially optimal termination charge, that is, $m^* \leq m^W = \max\{-\frac{2\lambda c}{1+\lambda}, -c_T\} < 0$, where the inequality binds when *Bill and Keep* is socially and privately optimal, i.e. when $m^* = m^W = -c_T$. This occurs when externalities are relatively strong. Berger (2005) even argues that regulatory intervention is superfluous in this case since private and social incentives are then perfectly aligned. Our analysis shows that *Bill and Keep* may be optimal from a social point of view, but firms will most likely prefer termination charges above cost. Only if the call externality is extremely high, firms would also prefer *Bill and Keep*. This requires firms setting off-net prices above monopoly price in case termination charge is set at cost. Of course, if the externality is very strong, firms may be tempted to abandon the calling party pays regime and to start charging subscribers for receiving calls.

4.3 Linear pricing and termination-based price discrimination

In this section we analyze competition in linear prices – i.e., networks charge on- and off-net calls but not the fixed fee. Under linear pricing and termination-based price discrimination, and for some given expectations on market shares β_1 and β_2 , the variable net surplus offered to network i 's customers is

$$w(p_i, \hat{p}_i) \equiv \beta_i v(p_i) + (1 - \beta_i) v(\hat{p}_i). \quad (10)$$

³²As there is full participation and payments are only transfers from one agent to another, what matters is the utility that consumers derive from incoming and outgoing calls, and the cost of these calls. Given that $\bar{\mu} = \lambda\mu$, the socially optimal price maximizes the expression $u(q(p)) + \lambda u(q(p)) - cq(p)$. Hence, this price coincides with the equilibrium on-net price. (Recall that networks set the on-net price to maximize consumer surplus so as to extract it via the fixed fee.)

Market shares are determined by the indifferent customer:

$$\begin{aligned}\alpha_1 &= \frac{1}{2} + \sigma [w(p_1, \hat{p}_1) - w(p_2, \hat{p}_2)] \\ &= \frac{1}{2} + \sigma [\beta_1 (v(p_1) - v(\hat{p}_2)) - \beta_2 (v(p_2) - v(\hat{p}_1))].\end{aligned}\tag{11}$$

Differentiating Eq. (1) – where α_i is given by Eq. (11) and $F_i = 0$ – with respect to p_i and \hat{p}_i , we have that at a symmetric equilibrium ($p_1 = p_2 = p$, $\hat{p}_1 = \hat{p}_2 = \hat{p}$, $\alpha_i = \beta_i = 1/2$):

$$[R(p) - f] - \frac{R'(p)}{2\sigma q(p)} = 0,\tag{12}$$

$$[R(p) - f] - \frac{\hat{R}'(\hat{p})}{2\sigma q(\hat{p})} = 0,\tag{13}$$

Let p^D be the equilibrium price in a duopoly model where termination-based price discrimination is not allowed and $m = 0$. From Eq. (12) we have that the equilibrium on-net price p^* equals p^D and is therefore neutral with respect to the access charge. Using Eqs. (12) and (13) we obtain

$$\frac{q(p) + (p - c)q'(p)}{q(p)} = \frac{q(\hat{p}) + (\hat{p} - c - m)q'(\hat{p})}{q(\hat{p})}.\tag{14}$$

Assuming a constant elasticity demand function³³ ($\eta \equiv -q'(p)(p/q)$), we can rewrite Eq. (14) as follows

$$\frac{\hat{p}}{p} = \frac{c + m}{c},$$

which coincides with the proportionality rule derived in Laffont, Rey and Tirole (Lemma 1, 1998b). However, here the off-net price is increasing in the termination mark-up: $\frac{d\hat{p}}{dm} = \frac{p^*}{c} > 0$ (since $\frac{dp^*}{dm} = 0$). It is easily established that $p^R < p^* < p^M$.

In the symmetric equilibrium, i 's profit can be written as follows

$$\pi_i = \frac{1}{4} [R(p^D) + R(\hat{p}^*) - 2f],$$

where $\hat{p}^* = p^D \left(\frac{c+m}{c}\right)$. We summarize our results in the next Proposition.

Proposition 4 *Under linear pricing and termination-based price discrimination, in equilibrium the on-net price does not depend on the access charge. Moreover, for a constant elasticity call demand function*

(i) *the off-net price increases with the access charge;*

³³To guarantee existence of equilibrium we need $v(p)$ to be bounded so that we need to cap the demand function by setting $q(p) = \min\{\bar{q}, p^{-\eta}\}$ for some constant \bar{q} .

- (ii) the shared-market equilibrium profits are maximized with the termination mark-up m^* that maximizes the retail profit earned on the off-net calls $R(\hat{p}^*)$: $m^* = \left(\frac{p^M}{p^D} - 1\right) c > 0$;
- (iii) total welfare is maximized by a termination subsidy $m^W < 0$.

In the presence of rationally responsive expectations, and assuming a constant elasticity demand function, Laffont, Rey and Tirole (1998b) obtain that the on-net price decreases with m , and the off-net price increases with m if the degree of substitutability σ is small, though it may decrease with m otherwise. In addition, they obtain that if σ is small, then the access charge that maximizes profits is above the cost ($m > 0$) and higher than the access charge that is socially preferred, which is below the cost ($m < 0$). Nevertheless, they also find that if σ is not that small, then the access charge that maximizes profits may be lower. Indeed, it could be even below cost.³⁴

Under passive expectations we obtain different results. Proposition 4 states that the on-net price does not react to the level of the termination mark-up, whereas the off-net price always increases with the termination mark-up (independently of the degree of substitutability between the two networks). Consequently, networks find it profitable to increase the access charge above the cost as it exerts upward pressure on the off-net price (towards the monopoly level), which leads to higher profits.

Oligopolistic competition. The neutrality result on the on-net price is specific to the number of competing networks (as the neutrality result on the fixed fee in the case of three-part tariffs analyzed in section 4.1). If the number of competing networks is larger than two, then there will exist a partial waterbed effect on the on-net price even if consumer expectations are passive. As commented earlier, the number of terminated off-net calls increases with a firm's market share as long as it is below 1/2. Since an increase in the termination mark-up increases the profit per terminated call, firms will fight more intensively to gain market share. They do this by lowering the on-net price. Yet, simulations show that the waterbed effect is less than one hundred percent. In particular, assuming a Logit model with $n \geq 3$ competing networks we obtain that: (i) on-net prices decrease and off-net prices increase with m (in a neighborhood of $m = 0$); (ii) the profit maximizing termination mark-up is positive and increasing in the number of networks; (iii) consumer surplus is maximized with minimum mark-up $m^{CS} = -c_T$ (i.e., with $a = 0$); (iv) total surplus is maximized with a positive termination charge below cost (i.e. $-c_T < m^W < 0$). Table 1 reports optimal

³⁴Simulations show that for $\eta = 1.2$, $c_T = 0.5$, $c = 2$, $f = 0$ and $\sigma = 1$, the profit maximizing termination mark-up equals $m^* = -0.10$, whereas the total welfare maximizing termination mark-up is positive.

termination charges for oligopolies with 2, 3, 4 and 5 firms.³⁵ More details about prices, profits and welfare can be found in Fig. 3.

Termination markup	$n = 2$	$n = 3$	$n = 4$	$n = 5$
m^*	0.414	0.538	0.61	0.66
m^{CS}	-0.50	-0.50	-0.50	-0.50
m^W	-0.293	-0.264	-0.265	-0.274

Table 1: Optimal termination mark-up for firms (m^*), consumer surplus (m^{CS}) and total welfare (m^W). Simulation parameters: $\eta = 2$, $c_T = 0.5$, $c = 1$, $f = 0$, $\mu = 0.25$.

4.4 Asymmetric networks

In this section we analyze the competition between two asymmetric networks, an incumbent and a new entrant. We allow for brand loyalty as in Carter and Wright (1999, 2003) but allow firms to use termination-based price discrimination. The parameter $\gamma > 0$ measures the degree of asymmetry between the networks. The net surplus from subscribing to network 1 is

$$w_1 = \gamma t + \beta_1 v(p_1) + \beta_2 v(\hat{p}_1) - F_1,$$

whereas the net surplus offered to network 2's consumers is

$$w_2 = \beta_2 v(p_2) + \beta_1 v(\hat{p}_2) - F_2.$$

Network i 's profit is given by Eq. (1). Thus, as in Section 4.1, in an equilibrium where firms share the market, it is optimal to adopt cost-based usage prices: $p_i = c$ and $\hat{p}_i = c + m$. The market share of network 1 is thus given by

$$\alpha_1 = 1 - \alpha_2 = \frac{1 + \gamma}{2} + \sigma (F_2 - F_1) + 2\sigma \left(\beta_1 - \frac{1}{2} \right) (v(c) - v(c + m)). \quad (15)$$

The first-order condition yields

$$F_i = f + \frac{\alpha_i}{\sigma} + 2 \left(\alpha_i - \frac{1}{2} \right) R(c + m). \quad (16)$$

³⁵The parameter μ is the degree of product differentiation in a Logit model (see section 5).

Equilibrium profit of firm i is thus

$$\pi_i = \alpha_i^2 \left(\frac{1}{\sigma} + R(c + m) \right), \quad (17)$$

where α_i is given by Eqs. (15) and (16).

Proposition 5 *In the presence of two asymmetric networks and starting from cost-based termination charges ($m = 0$), in any shared-market equilibrium a small increase in the termination charge:*

- (i) raises the fixed fee of the large network and lowers the fixed fee of the small network,*
- (ii) reduces the difference in market shares between the two networks,*
- (iii) leads to higher equilibrium profits for both the large and small network,*
- (iv) reduces total welfare.*

López and Rey (2009) analyze competition between two asymmetric networks (an incumbent and a new entrant) in the presence of rationally responsive expectations. They find that in the shared-market equilibrium a below-cost access charge generates higher equilibrium profits (for the large and small network) than any above-cost access charge. Their finding therefore extends the insight of Gans and King to asymmetric networks. In contrast, Proposition 5 states that in the presence of passive expectations, increasing the access charge above the cost, raises the equilibrium profit for both networks (as in the case of symmetric networks studied in Section 4.1) while it reduces total welfare. In this asymmetric case an increase in termination profit makes the large firm compete less fiercely for market share as this in fact increases the number of calls to be terminated. The small firm will compete more fiercely for market share. This makes equilibrium market shares less asymmetric, and thus reduces the HHI index. Nevertheless, taking into account that network 1 provides higher value ($\gamma > 0$), this distortion actually reduces total and consumer welfare.

5 Voluntary Subscription

In this section we do not assume that all consumers will subscribe to one of the two networks. Consumers have the option to stay unsubscribed. Since consumers can only call to subscribers, consumers will care about the total number of people that will subscribe to *some* network. In the case of termination-based price discrimination consumers will care about the number of subscribers to *each* network. The addition of an extra subscriber has

a positive benefit for all subscribers. The nature of competition impedes firms to fully internalize this externality. It has been argued by some mobile operators and regulators that the termination charge should include a network externality surcharge so as to facilitate the internalization of the externality. Dessein (2003) and Hurkens and Jeon (2009) show indeed that when subscription demand is elastic, a surcharge may increase penetration and improve total welfare. However, these models assume again rationally responsive expectations. Moreover, these models also predict that firms prefer not to have a surcharge, since profits are higher with termination charges below cost. We will now review this issue under the assumption of passive expectations.

The Hotelling framework is not very well suited to address the issue of elastic subscription demand. Namely, if some consumers in the center of the interval do not subscribe, networks would operate like local monopolists, rather than as competitors.³⁶ We therefore will use a Logit model in which consumers have random utility.³⁷

We consider competition between two networks. Each firm i ($i = 1, 2$) charges a fixed fee F_i and may be allowed (or may not be allowed) to discriminate between on-net and off-net calls. For ease of exposition we will continue to use the notation p_i and \hat{p}_i for on- and off-net call prices of firm i . When termination-based price discrimination is not allowed we impose that $p_i = \hat{p}_i$. Notation and definition of call demand is as before. In particular, given some expectations β_1 and β_2 , utility from subscribing to network 1 equals

$$w_1 = \beta_1 v(p_1) + \beta_2 v(\hat{p}_1) - F_1,$$

while subscribing to network 2 yields

$$w_2 = \beta_2 v(p_2) + \beta_1 v(\hat{p}_2) - F_2.$$

Finally, not subscribing at all yields utility w_0 .

We now add a random noise term and define $U_1 = w_1 + \mu\varepsilon_1$, $U_2 = w_2 + \mu\varepsilon_2$, and $U_0 = w_0 + \mu\varepsilon_0$. The parameter $\mu > 0$ reflects the degree of product differentiation in a Logit model. A high value of μ implies that most of the value is determined by a random draw so that competition between the firms is rather weak. The noise terms ε_k are random variables of zero mean and unit variance, identically and independently double exponentially

³⁶Armstrong and Wright (2009) consider a Hotelling model with hinterlands to address the possibility of expansion.

³⁷See Anderson and de Palma (1992) and Anderson, de Palma and Thisse (1992) for more details about the Logit model.

distributed. They reflect consumers' preference for one good over another. A consumer will subscribe to network 1 if and only if $U_1 > U_2$ and $U_1 > U_0$; he will subscribe to network 2 if and only if $U_2 > U_1$ and $U_2 > U_0$; he will not subscribe to any network otherwise. The probability of subscribing to network i is denoted by α_i where α_0 represents the probability to remain unsubscribed. The probabilities are given by

$$\alpha_i = \frac{\exp[w_i/\mu]}{\sum_{k=0}^2 \exp[w_k/\mu]}. \quad (18)$$

Note that for $i = 1, 2$

$$\frac{\partial \alpha_i}{\partial F_i} = -\frac{\alpha_i(1 - \alpha_i)}{\mu}, \quad (19)$$

while for $j \in \{0, 1, 2\} \setminus \{i\}$

$$\frac{\partial \alpha_j}{\partial F_i} = \frac{\alpha_i \alpha_j}{\mu}. \quad (20)$$

Consumer surplus in the Logit model has been derived by Small and Rosen (1981) as (up to a constant)

$$CS = \mu \ln \left(\sum_{k=0}^2 \exp(w_k/\mu) \right) = w_0 - \mu \ln(\alpha_0), \quad (21)$$

where the right-hand side follows from (18). Clearly, consumer surplus is increasing in market penetration $1 - \alpha_0$.

5.1 Equilibrium

We will first establish that also in a setting with voluntary participation firms will set variable price equal to perceived marginal cost. The reason is simply that a firm can offer the same consumer surplus by setting variable price closer to perceived marginal cost while adjusting the fixed fee accordingly. This will keep the number of subscribers to each firm constant, but improve the profit of the firm. The reasoning is valid both for the case where firms are not allowed to charge different prices for on- and off-net calls and for the case where this is allowed. Of course, the notion of perceived marginal cost depends on the case under consideration. When firms can price discriminate, perceived marginal cost for on-net calls equals c and perceived marginal cost for off-net calls equals $c + m$. In this case profit is given by Eq. (1).

In the case where price discrimination is not allowed, we denote

$$\tilde{c}_i = \frac{\alpha_i c + \alpha_j (c + m)}{\alpha_i + \alpha_j}$$

as the weighted average marginal cost of calls. Now, i 's profit can be rewritten as

$$\pi_i = \alpha_i [(\alpha_i + \alpha_j)(p_i - \tilde{c}_i)q(p_i) + F_i - f] + \alpha_i \alpha_j m q(\hat{p}_j).$$

Using these expressions for profit, it is easy to establish the following perceived marginal cost pricing result.

Lemma 2 (i) *When firms can price discriminate between on- and off-net calls, in equilibrium firm i will set $p_i = c$ and $\hat{p}_i = c + m$.*

(ii) *When firms cannot price discriminate between on- and off-net calls, in equilibrium firm i will set $p_i = \hat{p}_i = \tilde{c}_i$. In a symmetric equilibrium $\tilde{c}_i = c + m/2$.*

Note that when $m = 0$, firms will not set distinct prices for on- and off-net calls even when they are allowed to do so.

Given the perceived marginal cost pricing result, profits can be rewritten (in both cases) as:

$$\pi_i = \alpha_i(F_i - f) + \alpha_i \alpha_j m q(\hat{p}_j).$$

Profits stem from the fixed fee and from termination services. The necessary first-order condition with respect to the fixed fee thus gives³⁸

$$0 = \frac{\partial \alpha_i}{\partial F_i}(F_i - f) + \alpha_i + [\alpha_i \frac{\partial \alpha_j}{\partial F_i} + \alpha_j \frac{\partial \alpha_i}{\partial F_i}] m q(\hat{p}_j).$$

Substituting (19) and (20) and re-arranging yields

$$F_i = f + \frac{\mu}{1 - \alpha_i} - q(\hat{p}_j) \frac{\alpha_j(1 - 2\alpha_i)}{1 - \alpha_i}.$$

Looking for a symmetric solution with $\alpha_i = \alpha_j = \alpha$, we find the following relation between equilibrium fixed fee and equilibrium number of subscribers per firm:

$$F = f + \frac{\mu}{1 - \alpha} - m q(\hat{p}) \frac{\alpha(1 - 2\alpha)}{1 - \alpha}, \quad (22)$$

where \hat{p} is the equilibrium off-net price.³⁹ We will denote the right-hand side of equation (22) by $F^{FOC}(\alpha, m)$ and we will refer to this curve as the equilibrium curve.

³⁸Firm i keeps \hat{p}_j constant when choosing fixed fee, although we know that in equilibrium, in the case of no termination-based price discrimination, $p_j = \tilde{c}_j$ depends on subscription rates.

³⁹Thus $\hat{p} = c + m$ when termination-based price discrimination is allowed for, and $\hat{p} = c + m/2$ when this is not.

From (18) we know that expectations being fulfilled in the case of a symmetric solution (F, p, \hat{p}) requires that the number of subscribers per firm (denoted by α), must satisfy

$$\alpha = \frac{\exp[(\alpha v(p) + \alpha v(\hat{p}) - F)/\mu]}{2 \exp[(\alpha v(p) + \alpha v(\hat{p}) - F)/\mu] + \exp[w_0/\mu]}.$$

This can be rewritten as

$$F = \alpha v(p) + \alpha v(\hat{p}) - w_0 - \mu \log \left(\frac{\alpha}{1 - 2\alpha} \right). \quad (23)$$

We will denote the right-hand side of equation (23) by $F^{RE}(\alpha, m)$ and we will refer to this curve as the rational expectations curve.

A symmetric equilibrium with fulfilled expectations is thus found by solving the system of equations (22) and (23). It is easily verified that this system of equations always admits a solution. Namely, for any given and fixed m , the (continuous) equilibrium curve is bounded on the interval $[0, 1/2]$ while the rational expectation curve approaches $+\infty$ as $\alpha \downarrow 0$ while it approaches $-\infty$ as $\alpha \uparrow 1/2$. The following lemma gives a sufficient condition for the uniqueness of such a solution. We will denote this solution (indexed by m) as $(F(m), \alpha(m))$. Since we will be particularly interested in how profits and welfare behave in a neighborhood of $m = 0$, we introduce $(F^*, \alpha^*) = (F(0), \alpha(0))$. Recall that for $m = 0$, it does not matter whether termination-based price discrimination is allowed or not, since firms do not find it optimal to discriminate. However, when doing comparative statics with respect to termination mark-up m around 0, one should distinguish between the case of termination-based price discrimination and uniform call prices.

Lemma 3 *For $|m|$ small enough and $\mu > v(c)/4$, the system of equations (22) and (23) has a unique solution.*

5.2 Comparative statics

We now investigate how the equilibrium behaves in a neighborhood of $m = 0$. We first establish that an increase in the termination mark-up above 0 reduces the number of subscribers and reduces equilibrium fixed fees.

Proposition 6 *A marginal increase in the termination mark-up above 0 lowers overall subscription and equilibrium fixed fees.*

From Lemma 2 we know that an increase in the termination fee raises usage price⁴⁰, since termination costs are passed onto consumers. This also means that consumers bring with them termination profits. Competition for customers becomes fiercer and this leads firms to charge lower fixed fees in equilibrium. This means that there is a waterbed effect at play. However, the effect is less than one hundred per cent. Consumers are not fully compensated by lower fixed fee for the higher call price and this is why less consumers will subscribe. It is not obvious how profits are affected by an increase in termination mark-up. Namely, an increase in the termination charge improves termination profits. We will now analyze the total effect on equilibrium profits (and on welfare). In equilibrium, profit equals

$$\pi(m, F, \alpha) = \alpha(F - f) + \alpha^2 m q(\hat{p}).$$

We first analyze how profits change along the rational expectations curve $F^{RE}(\alpha, 0)$ (when termination charge is fixed at $a = c_T$) as market penetration is varied. Note that profits, in this case, are just equal to $\Pi = \alpha(F^{RE}(\alpha, 0) - f)$ so that

$$\frac{\partial \pi}{\partial \alpha} = F^{RE}(\alpha, 0) - f + \alpha \frac{\partial F^{RE}}{\partial \alpha}.$$

Using that at $m = 0$, $F^{RE}(\alpha, 0) = F^{FOC}(\alpha, 0) = \mu/(1 - \alpha) + f$, one obtains

$$\frac{\partial \pi}{\partial \alpha} = \alpha \frac{2v(c)(1 - \alpha)(1 - 2\alpha) - \mu}{(1 - \alpha)(1 - 2\alpha)}.$$

The sign is ambiguous since it is negative for mature markets (when $\alpha \approx 1/2$) while it is positive for $\alpha \approx 0$ and $\mu < 2v(c)$. If the sign is negative, it means that colluding networks would prefer to increase fixed fees and lower market penetration. This is the case if competition is effective in terms of boosting penetration and lowering prices. We will refer to this case as one of effective competition. If the sign is positive, it means the opposite, that is, firms would prefer to increase penetration and reduce fixed fees. This would be the case in which externalities are important and are not well internalized under competition. We will refer to this case as one of strong network externalities.

Next, we consider how the profit changes as the termination charge is changed, keeping market penetration constant. An increase in m increases termination profits, but decreases the fixed fee. At $m = 0$ these effects exactly cancel out.

⁴⁰When termination-based price discrimination is allowed, only off-net usage price increases.

$$\left. \frac{\partial \pi}{\partial m} \right|_{m=0} = \alpha \frac{\partial F^{RE}}{\partial m} + \alpha^2 q(c) = 0,$$

where the second equality follows from Eq. (23) and $v'(c) = -q(c)$. Putting the two effects together shows that

$$\begin{aligned} \left. \frac{d\pi}{dm} \right|_{m=0} &= \frac{\partial \pi}{\partial m} + \frac{\partial \pi}{\partial F} \frac{dF}{dm} + \frac{\partial \pi}{\partial \alpha} \frac{\partial \alpha}{\partial m} \\ &= \alpha^2 q(c) + \alpha \left(\frac{\partial F}{\partial m} + \frac{\partial F}{\partial \alpha} \frac{\partial \alpha}{\partial m} \right) + (F - f) \frac{\partial \alpha}{\partial m} \\ &= \alpha^2 q(c) + \alpha \left(-\alpha q(c) + \frac{\partial F}{\partial \alpha} \frac{\partial \alpha}{\partial m} \right) + (F - f) \frac{\partial \alpha}{\partial m} \\ &= \frac{\partial \alpha}{\partial m} \left(F - f + \alpha \frac{\partial F}{\partial \alpha} \right). \end{aligned}$$

Note that the expression between brackets in the last line is just the derivative with respect to α of profits along the rational expectations curve. If profits increase along the rational expectations curve (low μ and low market penetration), then an increase of the termination charge lowers profits, since $\partial \alpha / \partial m < 0$ from Proposition 6. On the other hand, if profits decrease along the rational expectations curve, then an increase in the termination charge increases the profits. This case happens exactly when colluding firms would set a higher fixed fee than what they would choose under competition, and when market penetration would be lower under monopoly than under duopoly. The latter situation is the more likely scenario, especially for European countries where penetration rates are close to 100 per cent.

We now turn our attention to the effects of termination charges on consumer and total surplus. Note that total surplus is just the sum of consumer surplus and industry profit:

$$TS = CS + 2\pi.$$

From (21) we know that $dCS/dm = (\partial \alpha / \partial m)(2\mu / (1 - 2\alpha))$. Hence

$$\begin{aligned} \left. \frac{dTS}{dm} \right|_{m=0} &= \left(\frac{\partial \alpha}{\partial m} \right) \left(\frac{2\mu}{1 - 2\alpha} + 2 \left(F - f + \alpha \frac{\partial F}{\partial \alpha} \right) \right) \\ &= 2 \left(\frac{\partial \alpha}{\partial m} \right) \left(2\alpha v(c) + \frac{\mu}{1 - \alpha} \right). \end{aligned}$$

Since the second factor is positive, total surplus increases whenever market penetration (or consumer surplus) increases.

Proposition 7 *In order to maximize either consumer surplus or total surplus, the termination charge has to be set strictly below the cost of termination. Firms' profits are maximized by a termination charge above the cost of termination if and only if $\mu > 2v(c)(1-\alpha^*)(1-2\alpha^*)$ (that is, if and only if competition is effective).*

This result is in stark contrast with Dessein (2003) and Hurkens and Jeon (2009). Dessein (2003) does not allow for termination-based price discrimination while Hurkens and Jeon (2009) do. Both papers find that firms always prefer termination charge below the cost of termination. Moreover, they both find that in the (plausible) case of effective competition, consumer surplus and total welfare are maximized when termination charge is above cost. Instead, we find that an externality surcharge does not improve penetration. The argument that some operators successfully used to convince Ofcom and other NRAs to allow for a mark-up above cost has thus shown to be misleading. Our result thus underpins the Recommendation of the European Commission that externality surcharges in fact distort competition.

6 Conclusion

This article has studied how consumer expectations affect retail competition when network externalities exist. As in Katz and Shapiro (1985) and related literature on network externalities, we assume that first consumers form expectations about network sizes, then firms compete, and last consumers make rational subscription and consumption decisions based on their expectations and the chosen prices. Expectations must be fulfilled in equilibrium. Instead, in the literature on termination charges and tariff-mediated network externalities (starting from Laffont, Rey and Tirole, 1998b), rational expectations are imposed interim: any change of a price by one firm leads to a *rational* change in consumer expectations for which subscription decisions are such that realized and expected network sizes coincide. We have shown that the way consumers form expectations and how these react to price variations have important implications in terms of the impact of termination charges on retail competition. We have shown that if expectations are modeled as in Katz and Shapiro (1985), results regarding the impact of mobile termination rates on retail competition are in fact in line with real world observations. In particular, we overturn the Gans and King (2001)

result: in our model when firms compete in fixed fees and charge different prices for on- and off-net calls, they prefer a termination charge above cost so that off-net calls are priced at monopoly levels (socially optimal termination charges are equal to cost). Moreover, fixed fees and on-net prices are neutral with respect to the termination charge and thus, in the case of full participation and two networks, there is no waterbed effect at all. If the number of competing networks is larger than two, then there exists a partial waterbed effect on the subscription fee. Our theoretical results are thus in line with the empirical evidence of the existence of a waterbed effect that is not full, provided by Genakos and Valletti (2009).

Our results provide formal support to the relatively commonly-held view in the decision practice on mobile markets that firms benefit from high termination rates. Given the current debate on the optimal level of mobile termination rates, our results have direct policy implications. Mobile network operators (MNOs) have opposed cuts in termination charges over the past decade, and they keep doing so. This is of course a clear sign that mobile operators fear to see their profits reduced. The arguments these operators employ to defend their opposition against lowering termination charges sometimes make reference to the existence of a waterbed effect. They warn regulators that cutting termination rates may lead to higher prices and that may hurt consumers. Regulators have not been very empathetic to this argument and sometimes even denied the existence of a waterbed effect. For example, the Australian Competition and Consumer Commission (ACCC, 2007), wrote “The Commission considers that these trends of lower average retail prices [...] demonstrate that the converse of the ‘waterbed’ effect has been in operation.” The New Zealand Commerce Commission (NZCC, 2006) initially discarded the existence of a waterbed effect, but after the intervention by Prof. Hausman, acting on behalf of one of the MNOs, acknowledged the possible existence of a waterbed effect. However, the NZCC noted that to the extent that there is a waterbed effect, whereby retail mobile prices are adjusted in some way in response to regulation, it considered it likely that mobile prices will decline under regulation but at a slower rate than without. The UK regulator (Ofcom 2004) accepts the existence of a waterbed effect, but does not believe it is full because the retail market is not yet fully competitive. On the other hand, Ofcom (2004, 2007) and some other NRAs did accept the suggestion that a externality surcharge to promote subscription was appropriate. Our model shows that this conclusion is not warranted and that the Recommendation of the European Commission to not allow for such a mark-up is correct.

The present paper has considered a number of theoretical models (linear and non-linear pricing, duopoly and oligopoly, symmetric and asymmetric firms, elastic and inelastic sub-

scription demand, and call externalities) and shows that a waterbed effect often does exist but that it is always less than full, so that consumer welfare is improved when termination charges are reduced toward or even below cost. A further important lesson from our paper is that more competition in the telecommunication market may not be effective if it is not accompanied by continued adequate regulation of the monopolistic bottlenecks. In fact, regulation may be even more important in these cases since the number of off-net calls decreases with the HHI index.

We have assumed that the expectations of consumers do not change with price variations (off the equilibrium path) and that expectations are fulfilled in equilibrium. We believe this to be a plausible assumption in the context of telecommunication markets. Notwithstanding, we believe it will be important to consider a truly dynamic model where consumers face switching costs and where expectations are formed endogenously. A key question would be whether the results in such a dynamic model resemble the ones obtained in the static model with passive or with responsive expectations. We hope that this article will stimulate further research extending the analysis in this direction. Although the current paper has already considered a wide range of models, a number of issues that deserve further investigation has remained unaddressed. For example, how do passive expectations affect equilibrium outcomes when (i) different types of consumers are allowed for?; (ii) the called party also pays? (as is the case in Canada, Hong Kong, Singapore and the US); (iii) when both fixed and mobile operators compete with each other?

References

- ACCC. "MTAS Pricing Principles Determination: 1 July 2007 to 31 December 2008." November 2007.
- Anderson, S.P. and A. de Palma. "The Logit as a Model of product Differentiation," *Oxford Economic Papers*, Vol. 44 (1992), 51-67.
- Anderson, S.P, A. de Palma and J.-F. Thisse. *Discrete Choice Theory of Product Differentiation*. (1992). The MIT Press, Cambridge (MA).
- Armstrong, M. "Network Interconnection." *Economic Journal*, Vol. 108 (1998), 545-564.
- Armstrong, M. "The Theory of Access Pricing and Interconnection." in M. Cave, S. Majumdar and I. Vogelsang (eds.), *Handbook of Telecommunications Eco-*

- nomics*, chap. 8, (2002). North Holland.
- Armstrong, M. and J. Wright. “Mobile Call Termination.” *Economic Journal*, Vol. 119 (2009): F270-F307.
- Berger, U. “Bill-and-Keep vs. Cost-Based Access Pricing.” *Economics Letters*, Vol. 86 (2005), 107-112.
- Cabral, L. “Dynamic Pricing with Network Effects,” NYU DP (2009).
- Caillaud, B. and B. Jullien. “Chicken & Egg: Competition among Intermediation Service Providers,” *RAND Journal of Economics*, Vol. 34 (2003), 309-328.
- Calzada, J. and T. Valletti. “Network Competition and Entry Deterrence,” *Economic Journal*, Vol. 118 (2008), 1223-1244.
- Carter, M. and J. Wright. “Interconnection in Network Industries.” *Review of Industrial Organization*, Vol. 14 (1999), 1-25.
- Carter, M. and J. Wright. “Asymmetric Network Interconnection.” *Review of Industrial Organization*, Vol. 22 (2003), 27-46.
- Crémer, J., P. Rey, and J. Tirole. “Connectivity in the commercial Internet,” *Journal of Industrial Economics* Vol. 48 (2000), 433–472.
- Dessein, W. “Network Competition in Nonlinear Pricing.” *Rand Journal of Economics*, Vol. 34 (2003), 593-611.
- Doganoglu, T. “Dynamic Price Competition with Consumption Externalities,” *Netnomics* 5, 43-69 (2003).
- Driskill, R., “Monopoly and Oligopoly Supply of a Good with Dynamic Network Externalities,” Vanderbilt University (2007).
- EC 2009a. *Commission Recommendation on the Regulatory Treatment of Fixed and Mobile Termination rates in the EU.* (2009/396/EC). *Official Journal of the European Union* (May 20,2009), L124/67-L124/74.
- EC 2009b. *Explanatory Note, Commission Staff Working Document*, accompanying the *Commission Recommendation on the Regulatory Treatment of Fixed and Mobile Termination rates in the EU..* European Commission, Brussels. 2009.
- Economides, N. “The Economics of Networks,” *International Journal of Industrial Organization*, Vol. 14 (1996a), 673-699.

- Economides, N. “Network Externalities, Complementarities, and Invitations to Enter,” *European Journal of Political Economy*, Vol. 12 (1996b), 211-232.
- Farrell, J. and P. Klemperer. “Coordination and Lock-in: Competition with Switching Costs and Network Effects,” in *Handbook of Industrial Organization*, Vol. III, (eds. Mark Armstrong and Rob Porter) (2007). Elsevier.
- de Fontenay, C.C. and J.S. Gans “Vertical Integration in the Presence of Upstream Competition,” *RAND Journal of Economics*, Vol. 36, No. 3 (Autumn, 2005), 544-572.
- Gale, J., K.G. Binmore, and L. Samuelson. “Learning to be imperfect: The ultimatum game,” *Games and Economic Behavior*, Vol. 8 (1995), 56-90.
- Galeotti, A. and J.L. Moraga-González. “Platform intermediation in a market for differentiated products,” *European Economic Review*, Vol. 53 (2009), 417-428.
- Gans, J.S., and S.P. King. “Using ‘Bill and Keep’ Interconnect Arrangements to Soften Network Competition.” *Economics Letters*, Vol. 71 (2001), 413-420.
- Genakos, C. and T. Valletti. “Testing the “Waterbed” Effect in Mobile Telephony”. Mimeo. (2009).
- Griva, K. and N. Vettas. “Price competition in a differentiated products duopoly under network effects.” CEPR DP 4574. (2004).
- Hahn, J.-H. “Network competition and interconnection with heterogeneous subscribers,” *International Journal of Industrial Organization*, Vol. 22 (2004), 611-631.
- Harbord, D. and M. Pagnozzi. “Network-Based Price Discrimination and Bill and Keep vs. Cost-Based Regulation of Mobile Termination Rates,” *Review of Network Economics* Vol. 9 (2010), Iss. 1, Article 1.
- Hart, O. and J. Tirole. “Vertical Integration and Market Foreclosure,” *Brookings Papers on Economic Activity*, (1990), 205-285.
- Hermalin, B.E. and M.L. Katz. “Customer or Complementor? Intercarrier Compensation with Two-Sided Benefits.” Mimeo (2009).
- Hoernig, S., R. Inderst and T. Valletti. “Calling Clubs: Network Competition with Non-Uniform Calling Patterns.” Mimeo (2009).

- Hurkens, S. and D.S. Jeon. "Mobile Termination and Mobile Penetration." Mimeo. (2009).
- Jeon, D.S., J.J. Laffont and J. Tirole. "On the "Receiver-Pays" Principle." *Rand Journal of Economics*, Vol. 35 (2004), 85-110.
- Jullien, B., P. Rey and W. Sand-Zantman "Mobile Call Termination revisited", *IDEI Working Paper*, n° 551, April 2009.
- Katz, M.L. and C. S. Shapiro. "Network Externalities, Competition, and Compatibility," *American Economic Review*, Vol. 75 (1985), 424-440.
- Laffont, J.J., P. Rey and J. Tirole. "Network Competition I: Overview and Nondiscriminatory Pricing." *Rand Journal of Economics*, Vol. 29 (1998a), 1-37.
- Laffont, J.J., P. Rey and J. Tirole. "Network Competition II: Price Discrimination." *Rand Journal of Economics*, Vol. 29 (1998b), 38-56.
- Laussel, D. and J. Resende, "Competition in the Market vs Competition for the Market: The Case of Cell Phone Networks," GREQAM and CORE, November 2007.
- Lee, I.H. and R. Mason. "Market structure in congestible markets," *European Economic Review*, Vol. 45 (2001), 809-818.
- López, A.L. and P. Rey. "Foreclosing Competition through Access Charges and Price Discrimination." *IESE Business School Working Paper*, n° 801, July 2009.
- Malueg, D.A. and M. Schwartz. "Compatibility Incentives of a Large Network Facing Multiple Rivals," *Journal of Industrial Economics*, Vol. 54 (2006), 527-564.
- Matutes, C. and X. Vives. "Competition for Deposits, Fragility, and Insurance," *Journal of Financial Intermediation*, Vol. 5 (1996), 184-216.
- McAfee, R.P. and M. Schwartz. "Opportunism in Multilateral Vertical Contracting: Nondiscrimination, Exclusivity, and Uniformity," *American Economic Review*, Vol. 84 (1994), 210-230.
- Mitchell, M.F. and A. Skrzypacz. "Network Externalities and Long-run Market Shares," *Economic Theory*, Vol. 29 (2006), 621-648.

NZCC. Telecommunications Act 2001: Schedule 3 Investigation into regulation of mobile termination. Reconsideration Final Report. April 2006.

O'Brien, D.P. and G. Shaffer. "Vertical Control with Bilateral Contracts," *RAND Journal of Economics*, Vol. 23 (1992), 299-308.

Ofcom *Wholesale Mobile Voice Call Termination: Statement*. June 2004.

Ofcom *Mobile Call Termination. Proposals for consultation*. September 2006.

Ofcom *Mobile Call Termination: Statement*. March 2007.

Radner R. and A. Sundararajan. "Dynamic Pricing of Network Goods with Boundedly Rational Consumers," CeDER working paper 06-03 (2006).

Rey, P. and J. Tirole. "A Primer on Foreclosure," in Mark Armstrong and Rob Porter (eds.) *Handbook of Industrial Organization*, Vol. III, (2007). Elsevier.

Small, K.A. and H.S. Rosen. "Applied Welfare Economics with Discrete Choice Models," *Econometrica*, Vol. 49, (1981), 105-130.

Tangerås, T. "On the profit neutrality of access charges in network competition." Mimeo (2009).

Appendix A

Proof of Proposition 1:

We have already established that there is a unique candidate for a shared market equilibrium. We need to show that for σ small enough, this candidate solution is indeed an equilibrium. We fix the strategy for firm 2 as $p_2 = c$, $\hat{p}_2 = c + m$ and $F_2 = f + 1/(2\sigma)$. Moreover, we fix consumer expectations at $\beta_1 = \beta_2 = 1/2$. We need to calculate the optimal response of firm 1. Recall from our discussion of the perceived marginal cost pricing that firm 1, in order to maximize its profit, can adjust its fixed fee F_1 so as to keep market share constant at α_1 . The on- and off-net price have to satisfy the first-order conditions (2) and (3), respectively. Denote these prices by $p_1(\alpha_1)$ and $\hat{p}_1(\alpha_1)$. One derives immediately that for $0 < \alpha_1 < 1$

$$R'(p_1(\alpha_1)) = \frac{q(p_1(\alpha_1))}{2\alpha_1} > 0 \quad (24)$$

and

$$\widehat{R}'(\hat{p}_1(\alpha_1)) = \frac{q(\hat{p}_1(\alpha_1))}{2(1-\alpha_1)} > 0. \quad (25)$$

Hence, $p_1(\alpha_1) < p^M$ and $\hat{p}_1(\alpha_1) < p^M$.

Let $F_1(\alpha_1)$ denote the fixed fee that yields firm 1 indeed a market share of α_1 . That is,

$$F_1(\alpha_1) = f + \frac{1 - \alpha_1}{\sigma} + \frac{1}{2}[v(p_1(\alpha_1)) + v(\hat{p}_1(\alpha_1)) - v(c) - v(c + m)].$$

Finding the optimal response for firm 1 boils down to finding the optimal market share.

The profit of firm 1, as a function of chosen market share, is

$$\Pi_1(\alpha_1) = \alpha_1 \left(\alpha_1 R(p_1(\alpha_1)) + (1 - \alpha_1) \widehat{R}(\hat{p}_1(\alpha_1)) - f + F_1(\alpha_1) \right) + \alpha_1(1 - \alpha_1)mq(c + m).$$

Note that $\Pi_1(0) = 0$ and that, since $F_1(1) < f + v(0)$, for σ small enough

$$\Pi_1(1) < R(p^M) + v(0) < 1/(4\sigma) + mq(c + m)/4 = \Pi_1(1/2).$$

Because the profit function is continuous, there exist $\underline{\alpha}$ and $\bar{\alpha}$ with $0 < \underline{\alpha} < \bar{\alpha} < 1$ so that the profit function will be maximized on the interval $[\underline{\alpha}, \bar{\alpha}]$.

Because of the envelope theorem, the partial derivatives with respect to on-net and off-net price are equal to zero, so that

$$\frac{d\Pi_1}{d\alpha_1} = 2\alpha_1 R(p_1(\alpha_1)) + (1 - 2\alpha_1) \widehat{R}(\hat{p}_1(\alpha_1)) - f + F_1(\alpha_1) - \frac{\alpha_1}{\sigma} + (1 - 2\alpha_1)mq(c + m).$$

Note that at $\alpha_1 = 1/2$ the first order derivative indeed equals zero since $p_1(1/2) = c$, $\hat{p}_1(1/2) = c + m$, and $F(1/2) = f + 1/(2\sigma)$. Using expressions (24) and (25) we can write the second-order derivative as

$$\frac{d^2\Pi_1}{d\alpha_1^2} = 2R(p_1(\alpha_1)) + \frac{q(p_1(\alpha_1))p_1'(\alpha_1)}{2} - 2\widehat{R}(\hat{p}_1(\alpha_1)) - \frac{\alpha_1 q(\hat{p}_1(\alpha_1))\hat{p}_1'(\alpha_1)}{2(1 - \alpha_1)} - 2\left(\frac{1}{\sigma} + mq(c + m)\right).$$

Clearly, for small enough σ this is strictly negative as the first 4 terms of this expression are bounded on the interval $[\underline{\alpha}, \bar{\alpha}]$.

■

Proof of Lemma 1:

For given rival strategies, maximizing π_i with respect to p_i , while adapting F_i so as to keep market shares constant, yields

$$\alpha_i \left[\alpha_i (q(p_i) + (p_i - c) q'(p_i)) + \frac{dF_i}{dp_i} \right] = 0. \quad (26)$$

For a constant α_i , differentiating Eq. (9) with respect to p_i yields

$$\sigma \left[\beta_i (q(p_i) - \bar{u}'(q(p_i)))q'(p_i) + \frac{dF_i}{dp_i} \right] = 0. \quad (27)$$

In equilibrium, expectations are fulfilled ($\beta_i = \alpha_i$), then from Eqs. (26) and (27) we have that $c - p_i = \bar{u}'(q(p_i))$. Since $\bar{u}(q) = \lambda u(q)$ and $u'(q) = p$, it follows that

$$p_i = p^* \equiv \frac{c}{1 + \lambda}. \quad (28)$$

Similarly, for given rival strategies, the first-order derivative of i 's profit with respect to \hat{p}_i , while adapting F_i so as to maintain market shares constant, yields

$$\alpha_i \left[\alpha_j (q(\hat{p}_i) + (\hat{p}_i - c)q'(\hat{p}_i)) - \alpha_j m q'(\hat{p}_i) + \frac{dF_i}{d\hat{p}_i} \right] = 0. \quad (29)$$

By differentiating α_i with respect to \hat{p}_i we obtain

$$-\sigma \beta_i \bar{u}'(q(\hat{p}_i))q'(\hat{p}_i) - \sigma \beta_j q(\hat{p}_i) - \sigma \frac{dF_i}{d\hat{p}_i} = 0. \quad (30)$$

Comparing Eqs. (29) and (30), we have that $\beta_i \bar{u}'(q(\hat{p}_i))q'(\hat{p}_i) = \alpha_j (\hat{p}_i - c - m)q'(\hat{p}_i)$. Using $\bar{u}'(q) = \lambda \hat{p}_i$, we obtain $\beta_i \lambda \hat{p}_i = \alpha_j (\hat{p}_i - c - m)$, where $\beta_i = \alpha_i$. Hence

$$\hat{p}_i = \hat{p}^*(\alpha_i) \equiv \frac{(1 - \alpha_i)(c + m)}{1 - \alpha_i(1 + \lambda)}. \quad (31)$$

■

Proof of Proposition 5:

Profits in equilibrium are given by (17). Totally differentiating with respect to m gives

$$\frac{d\pi_i}{dm} = 2\alpha_i \frac{d\alpha_i}{dm} \left(\frac{1}{\sigma} + m q(c + m) \right) + \alpha_i^2 (q(c + m) + m q'(c + m)).$$

Evaluating this derivative at $m = 0$ yields

$$\left. \frac{d\pi_i}{dm} \right|_{m=0} = \frac{2\alpha_i}{\sigma} \left. \frac{d\alpha_i}{dm} \right|_{m=0} + \alpha_i^2 q(c). \quad (32)$$

Totally differentiating (15) and (16), using $\alpha_2 = 1 - \alpha_1$, gives

$$\frac{d\alpha_i}{dm} = \sigma \left(\frac{dF_j}{dm} - \frac{dF_i}{dm} \right) + 2\sigma \left(\beta_i - \frac{1}{2} \right) q(c + m)$$

and

$$\frac{dF_i}{dm} = \frac{1}{\sigma} \frac{d\alpha_i}{dm} + 2 \frac{d\alpha_i}{dm} m q(c+m) + 2 \left(\alpha_i - \frac{1}{2} \right) (q(c+m) + m q'(c+m)).$$

Evaluating this derivative at $m = 0$ yields

$$\left. \frac{dF_i}{dm} \right|_{m=0} = \frac{1}{\sigma} \left. \frac{d\alpha_i}{dm} \right|_{m=0} + 2 \left(\alpha_i - \frac{1}{2} \right) q(c).$$

Thus, we have that

$$\left. \frac{d\alpha_i}{dm} \right|_{m=0} = \frac{2\sigma}{3} \left[2 \left(\frac{1}{2} - \alpha_i \right) + \left(\beta_i - \frac{1}{2} \right) \right] q(c).$$

Self-fulfilling expectations imply that at equilibrium $\beta_i = \alpha_i$, thus

$$\left. \frac{d\alpha_i}{dm} \right|_{m=0} = -\frac{2\sigma}{3} \left(\alpha_i - \frac{1}{2} \right) q(c) \quad i = 1, 2, \quad (33)$$

and

$$\left. \frac{dF_i}{dm} \right|_{m=0} = \frac{4}{3} \left(\alpha_i - \frac{1}{2} \right) q(c).$$

That is, starting from cost-based access charges, a slight increase in m , raises F_1 (lowers α_1) and lowers F_2 (raises α_2), which in turn reduces the asymmetry between the networks. Substituting Eq. (33) into Eq. (32) we get that

$$\left. \frac{d\pi_i}{dm} \right|_{m=0} = \frac{2}{3} \alpha_i \left(1 - \frac{\alpha_i}{2} \right) q(c) > 0 \quad i = 1, 2.$$

Finally, total surplus equals

$$\begin{aligned} TS(m) &= \alpha_1 [\gamma / (2\sigma) + \alpha_1 v(c) + (1 - \alpha_1) (u(q(c+m)) - cq(c+m))] \\ &\quad + (1 - \alpha_1) [(1 - \alpha_1) v(c) + \alpha_1 (u(q(c+m)) - cq(c+m))] \\ &\quad - \frac{1}{2\sigma} \left[\frac{\alpha_1}{2} \alpha_1 + \frac{1 - \alpha_1}{2} (1 - \alpha_1) \right]. \end{aligned}$$

Using that at $m = 0$, $\alpha_1 = (\gamma + 3)/6$, one easily verifies that

$$\left. \frac{dTS}{dm} \right|_{m=0} = -q(c) \frac{\gamma^2}{27} < 0.$$

■

Appendix B

Proof of Lemma 2:

- (i) Suppose that $(p_i, \hat{p}_i) \neq (c, c + m)$. We claim that firm i can improve its profit by changing its tariff from (p_i, \hat{p}_i, F_i) to $(c, c + m, \tilde{F}_i)$ where \tilde{F}_i is defined by

$$\beta_i v(p_i) + \beta_j v(\hat{p}_i) - F_i = \beta_i v(c) + \beta_j v(c + m) - \tilde{F}_i.$$

Such a change leaves the expected utility for subscribing to any of the networks unaltered, and will thus lead to the same subscription decisions. The difference in profit for firm i is thus equal to

$$\begin{aligned} & \alpha_i [\tilde{F}_i - F_i - \alpha_i (p_i - c) q(p_i) - \alpha_j (\hat{p}_i - (c + m)) q(\hat{p}_i)] = \\ & \alpha_i [\alpha_i [v(c) - v(p_i) + v'(p_i)(p_i - c)] + \alpha_j [v(\hat{c}) - v(\hat{p}_i) + v'(\hat{p}_i)(\hat{p}_i - (c + m))]] > 0 \end{aligned}$$

where the equality follows from self-fulfilled expectations ($\beta_k = \alpha_k$) whereas the inequality follows from the fact that $v(\cdot)$ is a strictly convex and decreasing function. The deviation is thus profitable.

- (ii) Suppose that $p_i \neq \tilde{c}_i$. We claim that firm i can improve its profit by changing its tariff from (p_i, F_i) to $(\tilde{c}_i, \tilde{F}_i)$ where \tilde{F}_i is defined by

$$(\beta_i + \beta_j) v(p_i) - F_i = (\beta_i + \beta_j) v(\tilde{c}_i) - \tilde{F}_i.$$

Such a change leaves the utility for subscribing to any of the networks unaltered, and will thus lead to the same subscription decisions. Given self-fulfilled expectations ($\beta_k = \alpha_k$) the difference in profit for firm i is thus equal to

$$\alpha_i [\tilde{F}_i - F_i - (\alpha_i + \alpha_j) (p_i - \tilde{c}_i) q(p_i)] = \alpha_i (\alpha_i + \alpha_j) [v(\tilde{c}_i) - v(p_i) + v'(p_i)(p_i - \tilde{c}_i)] > 0$$

where the inequality follows from the fact that $v(\cdot)$ is a strictly convex and decreasing function. The deviation is thus profitable.

■

Proof of Lemma 3:

Let $m = 0$ and $\alpha \in (0, 1/2)$. Then

$$\frac{\partial F^{FOC}}{\partial \alpha} = \frac{\mu}{(1 - \alpha)^2} > 0,$$

while

$$\frac{\partial F^{RE}}{\partial \alpha} = 2v(c) - \frac{\mu}{\alpha(1 - 2\alpha)} < 0$$

whenever $\mu > v(c)/4$. So, for $m = 0$, the equilibrium curve intersects the rational expectations curve from below. By continuity, the same holds for $|m|$ small enough. Hence, there is exactly one solution. ■

Proof of Proposition 6:

Proof. Note that

$$\left. \frac{dF^{RE}}{dm} \right|_{m=0} = -\alpha^* q(c)$$

while

$$\left. \frac{dF^{FOC}}{dm} \right|_{m=0} = -\alpha^* q(c) \frac{1 - 2\alpha^*}{1 - \alpha^*}.$$

An increase in m thus lowers the rational expectations curve by more than the equilibrium curve. The intersection point thus shifts to the south-west, lowering subscription rate and fixed fee. ■

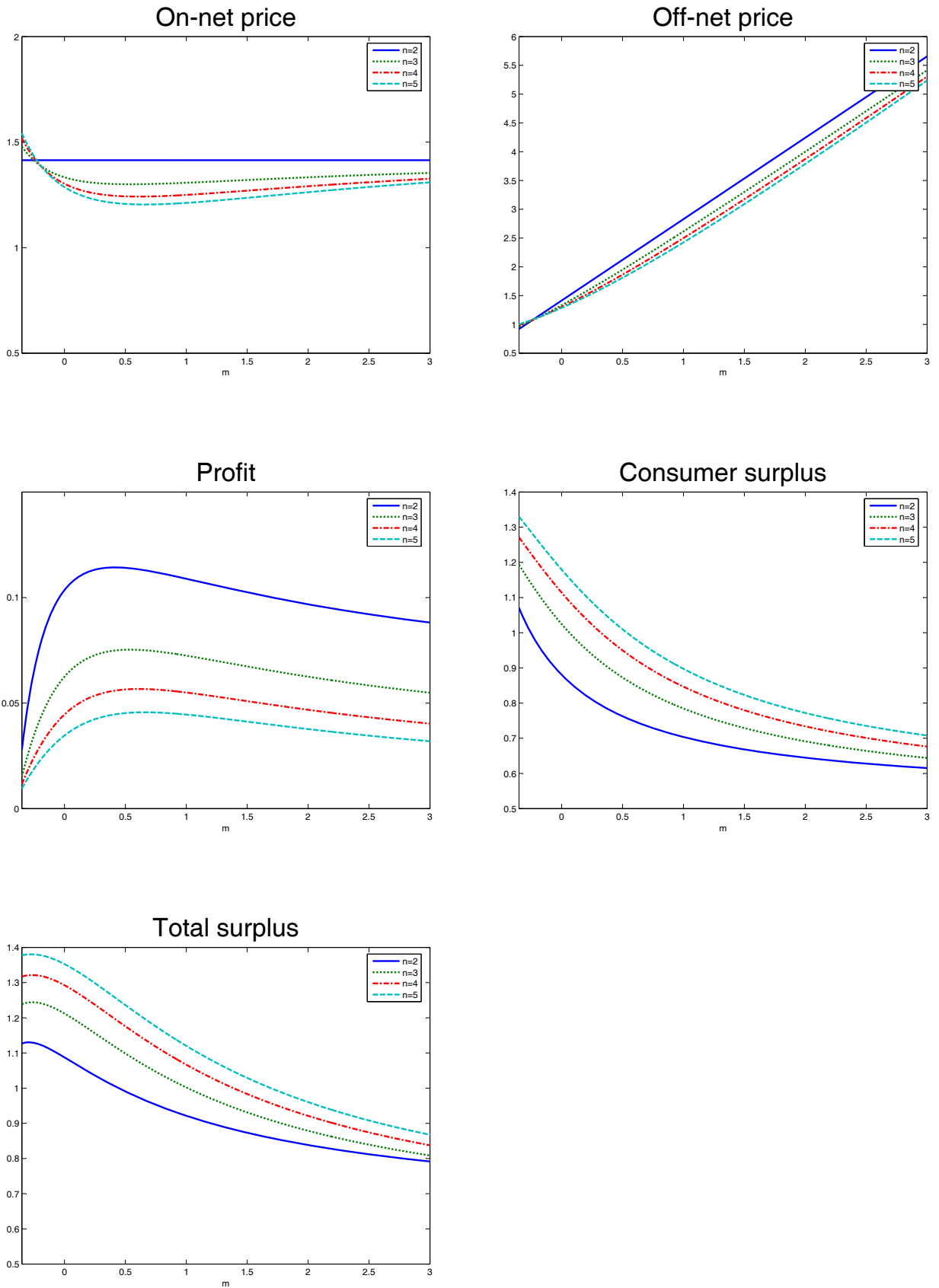


Figure 3: Effect of termination mark-up m on on-net price, off-net price, profit, consumer surplus and total welfare in the case of linear pricing.