Mobile Termination, Network Externalities, and Consumer Expectations[*]

short title: Mobile Termination and Consumer Expectations

Sjaak Hurkens and Ángel L. López

September 2013

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 when termination charges increase. This is puzzling since in reality regulators are pushing termination rates down while being opposed to do so by network operators. 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 responsive to non-equilibrium prices, as assumed until now.

This paper re-examines the effects of interconnection agreements, and in particular of termination charges, on competition and welfare in the mobile telephony market. Interconnection requires mobile operators to provide a wholesale service (called 'call termination'), whereby each network completes a call made to one of its subscribers by a caller from a different network (typically termed as 'off-net' call). In most countries, call termination is provided in exchange for a fee or termination charge to be paid by the originating operator to the terminating operator. Naturally, the termination charge affects the operators' cost of off-net calls and therefore has an impact on retail prices, competition and efficiency. Moreover, the termination charge affects the revenues accruing from providing termination services. This has then an impact on the competition for market share, which again affects retail prices and welfare.

There is a striking discrepancy between what real world practitioners (regulators and firms) believe the impact of termination charges on competition to be, and the insights provided by a vast theoretical reflection on the topic of termination charges. We believe a general re-examination of this archival wisdom is called for. We will examine the role of consumer expectations in this literature and show that a modification in modelling them is sufficient to change the theoretical results and bring them in line with real world practice.

Regulators around the world, and especially in the European Union, are concerned about inflated termination charges and have intervened in the markets of termination. For example, the European Commission recommended national regulators to diminish termination rates to reflect costs by the end of 2012 (EC 2009a). Mobile operators during the last decade have repeatedly opposed the cuts in termination rates imposed by the national regulatory authorities (NRAs), a clear indication that they expect a reduction

2

in profits when termination charges are decreased.[1] Some operators claim 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.[2] Other operators have even warned regulators that reducing termination charges would distort competition and hurt consumers because increased subscription fees would reduce mobile penetration.[3,4]

The existing theoretical literature supports the view that high termination charges can be used as a collusive device by firms and hurt consumer and total welfare, but only in the case of competition in linear prices, i.e. in the market segment of pre-paid cards. However, most clients are on contracts with monthly subscription fees, and these clients account for most of the total call volume. In this market segment firms compete in non-linear prices, and for this type of competition the theoretical literature has given the opposite predictions. When firms compete in two-part tariffs, marginal prices (i.e., on-net and off-net prices) will be set equal to perceived marginal cost, so that equilibrium profits accrue from both the collection of fixed fees and the provision of termination services. In this setting, and assuming subscription demand is inelastic, Laffont *et al.* (1998*b*) 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 as this softens competition.[5] The intuition behind this result is the following:

---

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

[2] See Ofcom, 2006, par. 7.7.

[3] Ofcom, 2007, par. 7.8.

[4] 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 subsidise marginal consumers to join a network, increase mobile penetration and thereby internalise the network externality. The UK regulator Ofcom calculated the identified externality surcharge to be positive, but very mildly, and took it into account when determining the termination charge (Ofcom, 2007). The European Commission, however, recommended against applying a surcharge and aims for termination charges equal to cost (EC 2009*b*, par 5.2.4.)

[5] Seminal models of network competition include Armstrong (1998) and Laffont *et al.* (1998*a,b*). For a complete review of the literature on access charges see Armstrong (2002), Vogelsang (2003) and Peitz

when the termination charge is above cost, off-net calls will be more expensive than on-net calls and 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 thus intensified. Firms prefer instead to soften competition and this can be attained by setting the termination charge below cost, which comes at the expense of reduced total welfare and consumer surplus.[6] This result also holds when the model is extended in various directions. For example, 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, 2012). Hurkens and Jeon (2012) show that the result continues to hold even when subscription demand is elastic.

The intuition for the Gans and King (2001) result reveals that consumer expectations play an important role. Namely, for the intuition to work it has to be the case that consumers realise and expect that when (starting from a situation with symmetric market shares) one firm lowers its price it will increase its market share and become the larger network. That is, consumers must be able to adjust their expectations in response to a price change. Indeed, the literature to date has taken for granted a sequence — that firms first compete in prices, then consumers form 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 call such expectations *responsive.* Consumers having responsive expectations means that any change by one firm of a price, no matter how tiny it may be, is assumed to lead to an instantaneous *rational* change in expectations of all consumers.

---

et al. (2004).

[6]Total welfare would be maximised by a termination charge equal to the cost of termination, whereas consumer surplus would be maximised by a termination charge strictly above the cost of termination.

It is presumed that given these changed expectations, optimal subscription decisions will lead realised and expected network sizes to coincide.

Our alternative proposal is to relax the assumption of responsive expectations and to replace it by one of fulfilled equilibrium expectations. This concept was first introduced by Katz and Shapiro (1985). They assumed that consumers first form expectations about network sizes, and firms then compete (in their Cournot model by setting quantities), and finally consumers make optimal subscription or purchasing decisions, given their expectations. These decisions then lead to actual market shares and network sizes. Katz and Shapiro contend that, *in equilibrium*, realised and expected network sizes are the same. We call such expectations *passive*, as they do not respond to out of equilibrium deviations by firms.

Our first set of findings concerns the case where subscription demand is assumed inelastic and firms may charge different prices for on- and off-net calls. When expectations are assumed passive, results about termination charges in mobile network industries are 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. We overturn the Gans and King (2001) result by showing that when firms compete in non-linear prices, they prefer a 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 case, there is no waterbed effect.[7] The 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

---

[7]The waterbed effect refers to the fact that the profit that a customer generates on fixed-to-mobile or mobile-to-mobile termination is (partially) competed away on the retail market. The term *waterbed* was coined by Prof. Paul Geroski during the investigations of the impact of fixed-to-mobile termination charges on competition.

above cost.[8] Total welfare maximising termination charges are equal to cost, whereas a termination charge below cost is optimal if maximising consumer surplus is the objective.

It turns out that characterising equilibrium prices by means of first-order conditions is easier when expectations are assumed passive rather than responsive. This allows us to take into account a variety of extensions of the baseline model. First, we consider the case of brand loyalty causing asymmetric networks and show that both networks will prefer above cost termination charges. This happens despite the fact that the smaller firm will typically compete more aggressively for market share when consumers come with termination profit. Second, 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.[9] Third, we re-examine Laffont *et al.* (1998*b*) where two symmetric 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 maximised by a termination charge above cost.

We also consider the possibility that subscription demand is elastic. When there are both direct and tariff-mediated network effects, we find that 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 internalise the network effect. Although *Bill 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 charges above cost, unless the direct network effect is sufficiently strong that firms would actually prefer to increase penetration rather than to increase

---

[8]Our 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 (2011).

[9]If the call externality is very strong, however, firms prefer a 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.

fixed fees. This means that in most European countries — with effective penetration rates now close to 100 % — firms prefer above cost termination charges. When there is only a direct network effect, because firms are not allowed to charge different prices for on- and off-net calls, the result is reversed: a termination charge above cost in that case increases participation, consumer surplus and total welfare.

Our aim to reconcile theory with real world practice is normally not shared with the related literature. However, a few attempts have recently been made in this direction. Armstrong and Wright (2009),[10] Jullien *et al.* (2013),[11] and Hoernig *et al.* (forthcoming),[12] all introduce additional realistic features of the telecommunication industry into the framework of Laffont *et al.* (1998*b*), and then show that for some parameter range joint profits are maximised at termination charges above cost. Moreover, these authors conclude that the need to regulate termination charges is reduced since the socially optimal termination charge would also be above cost. This paper offers a rather different solution to the puzzle, and draws a very different conclusion. First, instead of adding another realistic feature of telecommunication competition to their contribution, we confine our attention to the issue of how consumers form expectations. Second, while we also conclude that firms prefer above-cost termination charges, in contrast to the above papers we find that total welfare is maximised with termination charges at or below cost.

The paper proceeds as follows. In the next section we briefly discuss the different

---

[10]Armstrong and Wright (2009) argue that mobile-to-mobile (MTM) and fixed-to-mobile (FTM) termination 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 (this is the case if there is some room for mobile market expansion, and income from fixed lines is sufficiently important).

[11]Jullien *et al.* (2013) argue that the willingness to pay for subscription is related to the volume of calls. They consider two types, light and heavy users. The former only receive calls and are assumed to have an elastic subscription demand. There is full participation for the latter, who can place calls and obtain a fixed utility from receiving calls.

[12]Hoernig *et al.* (forthcoming) consider the existence of calling clubs so that the calling pattern is not uniform but skewed.

assumptions about consumer expectations. Section 2 introduces the general model of passive expectations. Section 3 deals with the models in which all consumers subscribe to one of the networks: Section 3.1 examines the symmetric duopoly case in which firms use non-linear prices and distinguish between on- and off-net calls, 3.2 symmetric oligopoly and 3.3 asymmetric duopoly. In 3.4 we extend the base model to allow for call externalities. Finally, in 3.5 we re-examine Laffont *et al.* (1998*b*) where firms compete in linear prices and distinguish between on- and off-net calls. Section 4 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-based price discrimination and the case where firms must set the same price for on- and off-net calls. Section 5 concludes. Proofs for sections 3 and 4 are collected in Appendix A and B, respectively.

1.   Passive versus Responsive Expectations

Expectations are important in any market with network effects, not just in the case of telecommunications. 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 are also present in financial markets. If the riskiness of a bank depends on the number of its depositors, depositors will care about the number of other people who will deposit in a given bank. (See Matutes and Vives, 1996.)

Many papers have been written on markets with network effects, and consumer expectations have been modeled as passive in some and as responsive in others.[13] Very

---

[13]Examples of the first include Katz and Shapiro (1985), Matutes and Vives (1996), Economides (1996), Lee and Mason (2001) and Malueg and Schwartz (2006). Examples of the latter include Crémer

few of these 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.[14] Lee and Mason (2001) point out in their pricing game that the results change dramatically if responsive beliefs are used instead of passive ones. Matutes and Vives (1996) characterise the equilibria under passive beliefs, but do point out that with responsive expectations any pair of deposit rates leading to non-negative profits can be sustained as an equilibrium. Griva and Vettas (2011) analyse price competition in a duopoly where products are 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 the case where firms can influence expectations through prices (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 explain why responsive expectations may intensify competition, let us look closely at a duopolistic industry with network effects.[15] 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 a network of size α equals $v_0 + \kappa\alpha$, where $\kappa > 0$ is a parameter that measures the strength of the positive network effect. Assume that networks 1 and 2 compete for consumers in flat fees, denoted by F_1 and F_2. Given these fees and expected market shares $(\alpha_0, 1 - \alpha_0)$, the consumer at location α_0 is exactly indifferent between the

_et al._ (2000), Caillaud and Jullien (2003) and Galeotti and Moraga-González (2009).

[14]Hermalin and Katz (2011) 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.'

[15]This model is a simplified version of the one introduced in Laffont _et al._ (1998_a_,_b_) which we will also employ in this paper.

two networks when

$$v_0 + \kappa \alpha_0 - t\alpha_0 - F_1 = v_0 + \kappa(1 - \alpha_0) - t(1 - \alpha_0) - F_2,$$

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

$$\alpha_0 = \frac{1}{2} + \frac{F_2 - F_1}{2(t - \kappa)}.$$

Now let us investigate what happens when suddenly firm 1 lowers its price to $F_1 - \Delta$. There are two fundamental questions: How will market shares respond? How will consumers react?

If consumers have passive expectations, they will take into account the *direct* pecuniary effect of the lower price but will not expect the size of networks to change. The result will be that some consumers will switch to network 1, namely the ones at locations $x \in (\alpha_0, \alpha_1)$, where α_1 is defined by

$$v_0 + \kappa \alpha_0 - t\alpha_1 - (F_1 - \Delta) = v_0 + \kappa(1 - \alpha_0) - t(1 - \alpha_1) - F_2.$$

That is,

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

If consumers actually expect the network sizes to change, their reaction may be different. For example, for a consumer located at $x > \alpha_1$ the direct pecuniary effect alone is insufficient to make him decide to switch, but if he realises that the size of network 1 will be at least α_1, the sum of the pecuniary effect and the network effect may be enough

10

to make him switch. More consumers will switch when they realise that the value of subscribing to network 1 has increased and the value of subscribing to network 2 has decreased. In particular, if consumers have responsive expectations and correctly anticipate how network sizes will change, those at locations in (α_0, α_2) will switch to network 1 where

$$\alpha_2 = \alpha_0 + \frac{\Delta}{2(t - \kappa)}.$$

A decrease of the fee by Δ increases the market share of network 1 under responsive expectations by $\Delta / [2(t - \kappa)]$, while under passive expectations its market share is only increased by $\Delta/(2t)$. It is thus reasonable to conclude that lowering the fee is a more effective competitive tool for gaining market share when expectations are responsive, whenever there are positive network externalities. Of course, in case of negative network externalities (*e.g.* congestion effects), $\kappa < 0$ and competition is more intense under passive expectations.[16]

We believe the assumption of passive expectations is at least as plausible as that of responsive expectations. One self-evident reason is that having responsive expectations implicitly assumes that consumers are sufficiently sophisticated to solve a fixed point problem. Passive expectations, on the other hand, may be generated by simply observing past market shares. Apart from this computational complexity argument, there is a more fundamental reason for emphasising the legitimate role of passive expectations. Namely, responsive expectations require consumers to expect that all other consumers respond, but that at the same time prices remain fixed.

---

[16]It is straightforward to calculate the symmetric equilibrium prices in the simplified model outlined here, both under passive and under responsive expectations. If f denotes the cost of serving a customer, and $\pi_i = \alpha_i(F_i - f)$ denotes network i's profit, then the symmetric equilibrium price is $F^* = f + t$ under passive expectations and $F^{**} = f + t - \kappa$ under responsive expectations. When positive network effects exist, profits are lower under responsive expectations than under passive expectations.

Assuming that all consumers will respond to a price cut may not be reasonable when price discounts are offered privately. For example, a consumer that is being poached by a competing network by the offer of a special deal may not know how many consumers are approached in this way. Moreover, their current network may also be poaching subscribers from rival networks with similar deals. It will not be easy to predict how the size of networks will change, and simply assuming consumers believe network sizes will not change does not seem unreasonable. Moreover, in the telecommunications industry many consumers are on long term contracts (18 to 24 months) and unable to switch even if they would want to when they receive a more appealing offer from the rival. Hence, market shares normally lag behind price cuts in the short to medium run. In this context, assuming that consumers believe market shares do not change at all may be more reasonable than assuming consumers believe they will respond as strongly as responsive expectations dictate.

Assuming that prices will remain fixed may also not always be plausible. 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 with the promise of free calls for the rest of their lives if they expected others to do so as well. That is, the entrant could take over the whole market based on the assumption of self-fulfilling responsive expectations. However, this does not and will not occur. Once the market share of the entrant passes a certain threshold, it will certainly withdraw the offer (or go bankrupt). In any case rival firms are certain to react before it gets too far. Notice that both responsive and self-fulfilling passive expectations are rational, so the difference between them does not occur for equilibrium prices (when all consumers have correct expectations in either case), but rather when prices are out of

equilibrium. The general point is that exactly when prices are out of equilibrium, at least one firm has an incentive to deviate, which makes its price commitment not credible.

2. The Model

We consider competition between two full-coverage networks, 1 and 2, indexed by $i \neq j \in \{1, 2\}$. Each network 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 p_i and off-net price is $\hat{p}_i$. 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.

We make the standard assumption of a balanced calling pattern, which means that the percentage of calls originating from a given network and completed on another (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:

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

The first term represents the profit made on consumer services (on-net and off-net calls, fixed fee and cost), and the second term 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 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 the networks' tariffs.

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

3. Full Participation

In this section we assume that the networks are differentiated à la Hotelling. Consumers are uniformly distributed on the segment $[0, 1]$, while the two networks are located at the two ends of this segment ($x_1 = 0$ and $x_2 = 1$). A consumer located at x and joining network i obtains a net utility given by

$$w_i - |x - x_i| / (2\sigma),$$

where $\sigma > 0$ measures the degree of substitutability between the two networks, and w_i is the value to a consumer subscribing to network i (as defined below). We assume full participation so that each consumer subscribes to the network that yields the highest net utility. We will focus our attention on the properties of shared market equilibria, where both firms have strictly positive market shares.[17]

---

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

3.1. Non-linear Pricing and Termination-based Price Discrimination

In this section 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 characterise the prices in a shared market equilibrium, and then 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 maximise

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

The first-order conditions are

$$(\alpha_i - \beta_i)q(p_i) + \alpha_i(p_i - c)q'(p_i) = 0 \tag{2}$$

and

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

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 $\hat{p}_i = c + m$. 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

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

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

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

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

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

Note that

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

This means that a necessary local second-order condition is that $1 + \sigma R(c + m) > 0$,

which we will assume to hold.[18] Solving the first-order condition for F_i we obtain the reaction function

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

Calculating the intersection of both reaction functions yields:

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

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

$$\alpha_i = \frac{1}{2} + 2\sigma \left[\frac{1 + 2\sigma R(c+m)}{3 + 4\sigma R(c+m)} - \frac{1}{2}\right](1 - 2\beta_i)\left[v(c) - v(c+m)\right].$$ (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.[19] There simply does not exist any asymmetric shared market equilibrium. It follows immediately that at the equilibrium

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

The preceding analysis 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 maximisers. In particular, one needs to verify that no firm wants to try to corner the market, given both the prices chosen by its competitor and the expectations of consumers.

---

[18]This condition holds for all $m \geq 0$ and also for $m < 0$, as long as $\sigma < -1/R(c+m)$.
[19]This assumption is relaxed in section 3.3.

Note that the firm that corners the market by lowering its fixed fee will also want to adjust the on- and off-net prices; that firm will want to set the off-net price at zero.[20] The following proposition establishes the conditions for the existence and uniqueness of the shared market equilibrium.[21]

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

Previous literature has suggested that a decrease of the access charges will result in the increase of (some) retail prices for mobile subscribers, commonly known as the waterbed effect. When consumer expectations are passive, the equilibrium 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.

To provide some intuition for the absence of a waterbed effect on the equilibrium fixed fee, imagine that both firms use perceived marginal cost pricing and that the fixed fee of firm j is held constant at F_j. The profit of firm i, because of perceived marginal cost pricing, stems only from the fixed fee and the termination profits. We are interested in knowing how the optimal reply of firm i changes as the termination mark-up varies. Let $F_i(m)$ denote the optimal reply. When m increases, users become more profitable, in the sense that they bring with them higher termination profits. This does not necessarily mean that firm i will compete more fiercely for market share. Namely, termination profits are only made on calls that originate from the rival network, and firm i terminates $n_i = \alpha_i(1 - \alpha_i)$ of such calls. If $F_i(m) > F_j$ (and thus $\alpha_i < 1/2$), firm i will indeed try to

---

[20] The firm will also increase the on-net price above cost.

[21] For the sake of a brief exposition we prove the existence of equilibrium only for this case. Our subsequent results will focus solely on the characterisation of equilibria.
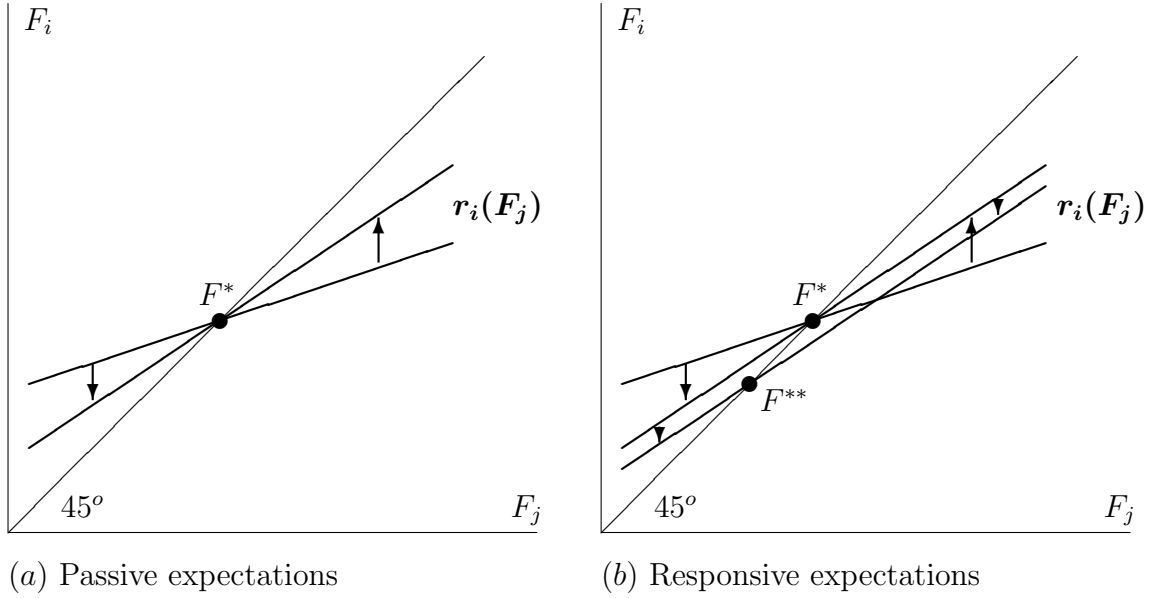
(a) Passive expectations   (b) Responsive expectations

Fig. 1: *The Effect of an Increase of m on Equilibrium Fixed Fees in Duopoly*

increase market share (and thus n_i) by lowering the fixed fee: $F_i'(m) < 0$. On the other hand, if $F_i(m) < F_j$ (and thus $\alpha_i > 1/2$), firm i will instead try to *decrease* market share (as this increases n_i) by *raising* the fixed fee: $F_i'(m) > 0$. Hence, an increase of m makes the reaction function of each firm rotate counterclockwise around the intersection point with the 45 degree line, but does not affect the equilibrium fixed fee. This is illustrated in Figure 1a.

Comparative statics. The symmetric equilibrium profit is

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

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 $c + m > p^M$. We thus have

COROLLARY 1. *Under non-linear pricing and termination-based price discrimination, shared-market equilibrium profits are maximised with the termination mark-up m^* that maximises the termination profit: $m^* = p^M - c > 0$. Total welfare is maximised at*

20

$m^W = 0$.

The results of Proposition 1 and Corollary 1 are in stark contrast with those obtained by Laffont *et al.* (1998*b*) and Gans and King (2001). Laffont *et al.* (1998*b*) show that the equilibrium fixed fees equal $F_1 = F_2 = f + 1/(2\sigma) - v(c) + v(c + m)$, which is strictly decreasing in m, so that a waterbed effect exists. Moreover, they show that the waterbed effect is so strong that equilibrium profits are strictly decreasing in m. Namely, $\pi = (1/2)(F - f) + (1/4)R(c + m)$ and

$$\frac{d\pi}{dm} = \frac{1}{2}v'(c + m) + \frac{1}{4}R'(c + m) = -\frac{1}{4}q(c + m) + \frac{1}{4}mq'(c + m) < 0,$$

for $m \geq 0$. Gans and King (2001) show that from the operators' viewpoint the optimal termination mark-up is negative. They even show that a zero termination charge (also known as *Bill and Keep*) is optimal if call demand is concave.

Since the only difference between our model and that of Laffont *et al.* (1998*b*) and Gans and King (2001) is that we assume that expectations are passive, it is not difficult to identify where the divergence in results originates. From Eq. (4) we see that with passive expectations $d\alpha_i/dF_i = -\sigma$ is independent of m, while in the models with responsive expectations

$$\alpha_i = \frac{1}{2} - \frac{\sigma(F_i - F_j)}{1 - 2\sigma\left[v(c) - v(c + m)\right]},$$

so that $d\alpha_i/dF_i$ is decreasing in m. Hence, only under responsive expectations does an increase of the termination mark-up make (an individual firm's) subscription demand more elastic.

Gans and King (2001) provide some intuition for their result: 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. In this scenario there are positive network effects and they become stronger when m increases. When consumers have responsive expectations, they anticipate that the firm with the lower price will be larger. It is exactly this aspect of consumers' expectations that makes it easier for firms to gain market share as the termination mark-up increases. Firms thus compete more aggressively for market share and reaction functions shift downward, which results in lower fixed fees in equilibrium. This intuition crucially hinges on the assumption of responsive expectations.

Figure 1 illustrates the above findings and intuitions. 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. An increase in the termination mark-up under both passive and under responsive expectations leads the smaller (larger) firm to compete more (less) aggressively, rotating the reaction function counterclockwise around the intersection point. (See Figure 1a and b.) On top of that, in the case of responsive expectations an increase in the termination mark-up shifts the reaction function downward. (See Figure 1b.) This explains why only in this case an increase in the termination mark-up reduces the equilibrium fixed fee (from F^* till F^{**}).

3.2. Oligopolistic Competition

The discussion thus far has focused exclusively on the rather special case of symmetric duopoly. It is important to know how the strength of the waterbed effect depends on market structure, particularly the number of competing networks. There are several widely applied and accepted models of oligopoly that could be used here: the circular city model of Salop (1979), the random utility Logit model of Anderson and de Palma

(1992), or the Spokes model of Chen and Riordan (2007).[22] In each of these models, usage prices will be equal to perceived marginal cost, both under passive and under responsive expectations. The profit function of firm i will thus be

$$\pi_i = \alpha_i \left[F_i - f + (1 - \alpha_i) R(c + m) \right].$$

The first-order condition thus reads

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

The fixed fee in a symmetric equilibrium in an n-firm oligopoly is thus characterised by

$$F = f - \frac{1}{n(d\alpha_i/dF_i)} - \left(1 - \frac{2}{n} \right) R(c + m).$$

This can be rewritten as follows:

$$\frac{F - f}{F} = \frac{1}{\varepsilon} - \frac{(1 - 2/n) R(c + m)}{F},$$

where $\varepsilon = -(d\alpha_i/dF_i)(F_i/\alpha_i)$ denotes the elasticity of subscription demand. The mark-up firms in equilibrium apply depends on two parts. The first corresponds to the standard inverse elasticity pricing rule and the second is an adjustment term that accounts for the profitability of consumers in terms of the termination profits they bring with them. In particular, when $n > 2$ the negative adjustment term is stronger when m is closer to $p^M - c$. That is, firms compete more fiercely in fixed fee when termination profits are higher

---

[22]Calzada and Valletti (2008) analysed oligopolistic competition in non-linear pricing with termination-based price discrimination using the Logit model.
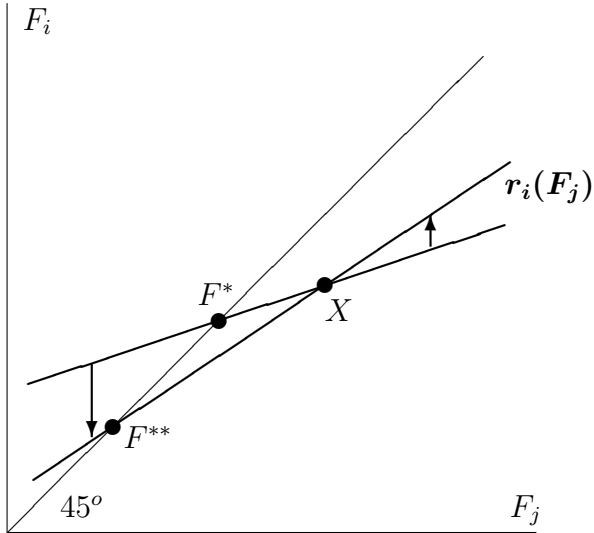
Fig. 2: *The Effect of an Increase of m on Equilibrium Fixed Fees in Oligopoly (Passive Expectations)*

(except for the special case of duopoly, for the reasons outlined above). The adjustment term is independent of whether expectations are assumed passive or responsive. The first term, however, depends crucially on how expectations are formed: Under passive expectations the elasticity does not depend on the termination mark-up, while under responsive expectations higher termination mark-ups make subscription demand more elastic.

Figure 2 explains why there is a 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 on the reaction function where firm i's market share would be 1/2. This is as well 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^{**}.

The total absence of a waterbed effect is thus restricted to the case of duopoly and passive expectations. In all other cases there is a waterbed effect, and it is always stronger under responsive expectations than under passive expectations. In order to evaluate this waterbed effect's strength, let us analyse the equilibrium profit:

$$\pi = \frac{1}{n}(F - f) + \frac{n-1}{n^2}R(c + m) = \frac{-(d\alpha_i/dF_i)^{-1} + R(c + m)}{n^2}.$$

When expectations are passive, $d\pi/dm = R'(c + m)/n^2 > 0$ for all $m < p^M - c$. We conclude that the waterbed effect is always less than full. The profit maximising termination mark-up equals $p^M - c > 0$, so that in equilibrium all off-net calls are priced at the monopoly price p^M. When expectations are responsive, $d\alpha_i/dF_i$ is decreasing in m. The exact formula depends on the model under consideration (Salop, Logit, or Spokes). In any case, we have

$$\frac{d\pi}{dm} = \frac{1}{n^2}\left[-\frac{d}{dm}\left(\frac{d\alpha_i}{dF_i}\right)^{-1} + R'(c + m) \right]$$

The first term between brackets is negative while the second term is positive for $m < p^M - c$. For any of the three models of oligopoly considered here, the first effect dominates and firms prefer termination charges below cost.[23]

3.3. Asymmetric Networks

In this section we analyse the competition between two asymmetric networks. Like Carter and Wright (1999; 2003) we model asymmetry by means of brand loyalty, but unlike them we allow firms to use termination-based price discrimination so that network effects

---

[23]In the Salop model, the first term between brackets equals $-3q(c + m)/2$; in the Logit and Spokes models, the first term equals $-nq(c + m)/(n - 1)$. The second term equals $mq'(c + m) + q(c + m)$ so that the total expression is strictly negative for all $m \geq 0$.

appear. Our model is therefore far closer to López and Rey (2012), apart from the fact that we assume expectations to be passive. They find that in the shared-market equilibrium a below-cost access charge generates higher equilibrium profits (for both the large and the small network) than any above-cost access charge. We will now show that under passive expectations this puzzling result is again reversed.

The net surplus from subscribing to network i is

$$w_i = \gamma_i/(2\sigma) + \beta_i v(p_i) + \beta_j v(\hat{p}_i) - F_i,$$

where γ_i denotes the brand loyalty parameter for firm i. Network i's profit is given by Eq. (1). Thus, as in section 3.1, 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

$$\alpha_1 = 1 - \alpha_2 = \frac{1+\gamma}{2} + \sigma\left(F_2 - F_1\right) + 2\sigma\left(\beta_1 - \frac{1}{2}\right)\left[v(c) - v(c+m)\right], \qquad (9)$$

where $\gamma = \gamma_1 - \gamma_2$ measures the degree of asymmetry between the networks. The first-order condition yields

$$F_i = f + \frac{\alpha_i}{\sigma} + 2\left(\alpha_i - \frac{1}{2}\right)R(c+m). \qquad (10)$$

The equilibrium profit of firm i is thus

$$\pi_i = \alpha_i^2\left[\frac{1}{\sigma} + R(c+m)\right], \qquad (11)$$

where α_i is given by Eqs. (9) and (10).

PROPOSITION 2. *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 the small network,

(iv) reduces total welfare.

The intuition for our results closely parallels the one given in the case of symmetric oligopolistic competition. An increase in termination mark-up means that consumers bring with them higher termination profits. This makes the large firm compete less fiercely for market share, because by reducing $\alpha_1 > 1/2$ it increases the number of calls to be terminated, $\alpha_1(1 - \alpha_1)$. The small firm will compete more fiercely for market share, because by increasing $\alpha_2 < 1/2$ it increases the number of calls to be terminated, $\alpha_2(1 - \alpha_2)$. This makes equilibrium market shares less asymmetric and reduces market concentration (as measured by the HHI index). The consequence is that more calls are off-net, and these are inefficiently high priced. This explains why firms obtain higher profits while consumer surplus is reduced. Moreover, taking into account that the larger network is the one that provides higher value to consumers, reducing asymmetry between the firms lowers total and consumer welfare even more.

3.4. 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, because otherwise nobody would bother to answer a call.

How strong such call externalities are is no doubt an empirical matter. Ofcom contends that 'call externalities - while they almost certainly exist - probably do not justify any adjustment to call prices. [...] these are likely to be effectively internalised 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.' (Ofcom, 2004, p.166). 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 private perspective. Berger (2005) even argues that regulatory intervention is superfluous when the social optimal termination regime is *Bill and Keep*, because firms will then always voluntarily agree on this regime. We agree that socially optimal termination charges are below cost when call externalities exist, and that they are equal to zero when call externalities are very strong. However, we show that firms will always want to set termination charges above the level that is socially optimal under passive expectations, and that for reasonable levels of the elasticity of call demand firms will prefer termination charges above cost.

We assume that consumers derive utility $\overline{u}(q)$ from receiving calls of volume q, with $\overline{u} = \lambda u$, 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

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

from subscribing to network i, for $i \neq j \in \{1, 2\}$. The actual market share, α_i, as a

function of consumer expectations and prices, is determined by the indifferent consumer:

$$\alpha_i \;=\; \frac{1}{2} + \sigma\beta_i \left\{ v(p_i) - v(\hat{p}_j) + \overline{u}\left[q(p_i)\right] - \overline{u}\left[q(\hat{p}_i)\right] \right\} \tag{12}$$

$$-\sigma\beta_j \left\{ v(p_j) - v(\hat{p}_i) + \overline{u}\left[q(p_j)\right] - \overline{u}\left[q(\hat{p}_j)\right] \right\} + \sigma(F_j - F_i).$$

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*

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

Therefore, in a symmetric equilibrium $p_i < c$ and $\hat{p}_i > c + m$, i.e., usage prices do not reflect the perceived marginal cost of calls. Networks find it optimal to internalise the call externality by setting the on-net price so as to maximise surplus from on-net calls, $(1+\lambda)u\left[q(p)\right] - cq(p)$, and then 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 and making the own network, in relative terms, more attractive. The off-net price is chosen to maximise the relative surplus $u\left[q(\hat{p})\right] - \lambda u\left[q(\hat{p})\right] - (c+m)q(\hat{p})$. When λ tends to 1 (which amounts to saying 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).

After substituting the usage prices from Lemma 1, the first-order condition with

respect to the fixed fee is

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

$$+ \sigma \left(\alpha_i - \alpha_j \right) m q(\hat{p}^*),$$

which defines i's reaction function. Hence, in a symmetric equilibrium $(\alpha_1 = \alpha_2 = 1/2)$ the first-order condition is satisfied at fixed fee

$$F^* = f + \frac{1}{2\sigma} - R(p^*).$$

The equilibrium profit is thus

$$\pi^* = \frac{1}{4\sigma} + \frac{1}{4} \left[R(\hat{p}^*) - R(p^*) \right].$$

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

PROPOSITION 3. *Under non-linear pricing, termination-based price discrimination and call externalities, symmetric equilibrium profits are maximised with the termination mark-up m^* that maximises the retail profit earned on the off-net calls (gross of termination payments): $m^* = \arg\max_{m \geq -c_T} R\left[(c+m)/(1-\lambda) \right] = \max\{(1-\lambda)\, p^M - c, -c_T\}$. Hence $m^* > 0$ if and only if $\lambda < \left(p^M - c \right)/p^M = 1/\varepsilon$, where ε denotes the elasticity of call demand at the monopoly price.*

In stark contrast with Jeon *et al.* (2004) and Berger (2005), the termination mark-up does not affect the fixed fee. The reason is as before: when expectations are passive and the two firms share the market equally, higher termination mark-ups will not lead to more

30

fierce competition in fixed fees. Then, networks maximise shared-market equilibrium profits by setting the termination mark-up m^* that maximises the retail profit from the off-net calls made by their subscribers. That is, m^* is such that $\hat{p}^* = (c + m^*)/(1 - \lambda)$ 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 a risk of connectivity breakdown because then $\hat{p} > p^M$, even if $m = 0$. In this extreme case a below cost access charge brings $\hat{p}^*$ down towards p^M and increases profits.[24] Note that when there are call externalities, firms do not prefer the termination mark-up that maximises the profits from termination, because they also make profits from off-net calls that are priced above their perceived marginal cost.[25]

Independent of whether expectations are passive or responsive, the welfare maximising termination mark-up m^W is such that $\hat{p}^*$ satisfies the condition $\hat{p}^* = p = c/(1 + \lambda)$.[26] Therefore, we have

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

This result contrasts with Berger (2005), who shows that in the presence of responsive expectations the best termination charge from the operators' perspective is lower than the socially optimal termination charge. This is formalised as $m^* \leq m^W = \max\{-2\lambda c/(1 + \lambda), -c_T\} < 0$, where the inequality binds when *Bill and Keep* is socially and privately

---

[24]Hurkens and López (2012) calibrate the call demand and cost functions for the Spanish market. Their calibration yields $\varepsilon = 1.35$ so that firms will prefer positive termination mark-ups as long as $\lambda < 0.74$.

[25]In the absence of call externalities, both termination profits and profits from off-net calls are equal to $mq(c + m)$.

[26]As 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 true cost of these calls. Given that $\overline{\mu} = \lambda\mu$, the socially optimal price maximises the expression $u[q(p)] + \lambda u[q(p)] - cq(p)$. Hence, this price coincides with the equilibrium on-net price.

optimal, i.e., when $m^* = m^W = -c_T$. This occurs when externalities are relatively strong $(\lambda \geq c_T/(2c_O + c_T))$. Berger (2005) even argues regulatory intervention is superfluous in this case since private and social incentives are then perfectly aligned. Our analysis confirms that *Bill and Keep* may be optimal from a social point of view, but shows firms will most likely prefer termination charges above cost. Firms would also prefer *Bill and Keep* only if the call externality is extremely high.[27] Such a level of call externality is arguably extreme, as it would imply that firms set off-net prices far above the monopoly price even when the termination charge is set at cost.

3.5. *Linear Pricing*

In this section we discuss how passive expectations affect the equilibrium results when firms compete in discriminatory linear prices − i.e., networks charge on- and off-net calls but not the fixed fee. One can think of this as competition in pre-paid tariffs. Laffont *et al.* (1998b) analysed competition in linear prices with price discrimination under the standard assumption of responsive expectations, assuming a constant elasticity call demand function. Their main message is that firms may use the termination charge as a collusion device by setting it above the cost of termination, while the socially optimal termination charge is below cost. This extends their result from Laffont *et al.* (1998a) where *uniform* linear prices are considered. Hence, contrary to the case of non-linear pricing, there seems to be no puzzle to be explained in this case. However, their result for discriminatory linear prices only holds when $\sigma > 0$ is very small. For larger σ, the result may well be reversed so that firms' profits are then maximised by a termination

---

[27]According to the calibration of the call demand function in Hurkens and López (2012), this occurs for call externality $\lambda > 0.87$.

charge below cost, while the welfare maximising termination charge is above cost.[28]

The intuition for this puzzling result is similar to that given for the case of non-linear pricing by Gans and King (2001). First, when call demand has constant elasticity, it can be shown that the equilibrium on- and off-net prices must satisfy a proportionality rule

$$\frac{\widehat{p}}{p} = \frac{c+m}{c}. \tag{13}$$

The off-net price thus exceeds the on-net price if and only if $m > 0$, as in the case of non-linear pricing. There exist tariff-mediated network externalities whenever $m \neq 0$. In particular, when the termination mark-up is positive consumers prefer to belong to the larger network. When consumers have responsive expectations, lowering the on-net price becomes a more effective competitive tool for gaining market share. As a consequence, firms end up setting lower on-net prices as the termination mark-up increases. One can interpret this again as a waterbed effect. The effects on the off-net price and profits are ambiguous because of the proportionality rule and turn out to crucially depend on the degree of substitutability σ.

We obtain under passive expectations the unambiguous and intuitive result that above cost termination charges serve as a collusive device and are welfare reducing.

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;

(ii) the shared-market equilibrium profits are maximised with the termination mark-up

$m^* > 0$ that maximises the retail profit earned on the off-net calls;

(iii) total welfare is maximised by a termination subsidy $m^W < 0$.

The driving force behind our results is simply that under passive expectations there is again no waterbed effect. That is, the on-net price is independent of the termination charge. The intuition for this result is similar to the one we gave for non-linear pricing. Namely, imagine that both firms set the same off-net price $\hat{p}$ (which may depend on m) and fix the on-net price of firm j at p_j. Then termination payments and revenues cancel each other out, and the profit of firm i is the sum of profits from on-net and off-net calls (gross termination payments). Recall that $n_i = \alpha_i(1 - \alpha_i)$ is the volume of off-net calls. How does the optimal on-net price of firm i, $p_i(m)$, vary with m? When $p_i(m) > p_j$ (and thus $\alpha_i < 1/2$), firm i will want to increase market share (and thus n_i) if and only if an increase in m increases profits from off-net calls, that is, when $\hat{p}$ gets closer to the monopoly price. On the other hand, when $p_i(m) < p_j$ (and thus $\alpha_i > 1/2$), firm i will want to decrease market share (and thus increase n_i) if and only if an increase in m increases profits from off-net calls. Hence, when an increase in m leads to higher off-net profits, $p_i'(m) < (>)0$ when $p_i(m) > (<)p_j$. When an increase in m leads to lower profits from off-net calls, the effect is reversed. In any case, the intersection point of firm i's optimal on-net price with the 45 degree line is unchanged. That is, the symmetric equilibrium on-net price is completely independent of the off-net price $\hat{p}$ and m. On the other hand, the proportionality rule (13) still holds. The off-net price is thus always strictly increasing in m. The profit maximising termination mark-up is the one that makes the off-net price equal to the monopoly price and is thus positive. In contrast, the welfare maximising termination mark-up is the one that makes the off-net price equal to cost and is thus negative.

4. 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 presence of termination-based price discrimination consumers will also 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 from fully internalising 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 internalisation of the externality. Dessein (2003) and Hurkens and Jeon (2012) show that when subscription demand is elastic, a surcharge may indeed increase penetration and improve total welfare. However, these models assume responsive expectations and predict again 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.[29] We therefore will use a Logit model in which consumers have random utility.[30]

We consider competition between two networks. Each firm i ($i = 1, 2$) charges a fixed fee F_i and may or may not be allowed to discriminate between on-net and off-net calls. For ease of exposition we continue to use the notation p_i and $\hat{p}_i$ for on- and off-net call

---

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

[30]See Anderson and de Palma (1992) and Anderson *et al.* (1992) for more details about the Logit model.

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 i equals

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

while not subscribing at all yields utility w_0.

We now add a random noise term and define $U_k = w_k + \mu \varepsilon_k$ for $k = 0, 1, 2$. 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 distributed. They reflect the 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 $U_2 > U_0$; otherwise he will not subscribe to any network. The probability of subscribing to network i is denoted by α_i where α_0 represents the probability of remaining unsubscribed. The probabilities are given by

$$\alpha_i = \frac{\exp\left(w_i/\mu\right)}{\sum_{k=0}^{2} \exp\left(w_k/\mu\right)} \tag{14}$$

Note that for $i = 1, 2$

$$\frac{\partial \alpha_i}{\partial F_i} = -\frac{\alpha_i(1 - \alpha_i)}{\mu}, \tag{15}$$

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

$$\frac{\partial \alpha_j}{\partial F_i} = \frac{\alpha_i \alpha_j}{\mu}. \tag{16}$$

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

$$CS = \mu \ln \left[\sum_{k=0}^{2} \exp(w_k/\mu) \right] = w_0 - \mu \ln(\alpha_0), \tag{17}$$

where the right-hand side follows from (14). Clearly, consumer surplus is increasing in market penetration $1 - \alpha_0$.

4.1. Equilibrium

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

$$\pi_i^{UNI} = \alpha_i \left[F_i - f + (\alpha_i + \alpha_j)(p_i - \tilde{c}_i) q(p_i) \right] + \alpha_i \alpha_j m q(p_j).$$

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 $\tilde{c}_i = c + m/2$.

Next we will characterise fixed fees in equilibrium. It is important to treat the case of on-net/off-net price discrimination separately from the case where firms set a uniform usage price.

Given the perceived marginal cost pricing result, in the case of termination-based price discrimination, profits stem only from the fixed fee and termination services:

$$\pi_i^{PD} = \alpha_i(F_i - f) + \alpha_i\alpha_j mq(c + m).$$

The necessary first-order condition with respect to the fixed fee thus gives

$$0 = \frac{\partial \pi_i^{PD}}{\partial F_i} = \frac{\partial \alpha_i}{\partial F_i}(F_i - f) + \alpha_i + \left(\alpha_i \frac{\partial \alpha_j}{\partial F_i} + \alpha_j \frac{\partial \alpha_i}{\partial F_i} \right) mq(c + m).$$

Substituting (15) and (16) and re-arranging yields

$$F_i = f + \frac{\mu}{1 - \alpha_i} - mq(c + m)\frac{\alpha_j(1 - 2\alpha_i)}{1 - \alpha_i}.$$

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:

$$F^{PD} = f + \frac{\mu}{1-\alpha} - mq(c+m)\frac{\alpha(1-2\alpha)}{1-\alpha}. \tag{18}$$

We will denote the right-hand side of equation (18) by $F_{PD}^{FOC}(\alpha, m)$ and refer to this curve as the equilibrium curve (when termination-based price discrimination is allowed).

When termination-based price discrimination is not allowed, profits can be rewritten as

$$\pi_i^{UNI} = \alpha_i \left[F_i - f + (\alpha_i + \alpha_j)(p_i - c)q(p_i) \right] + \alpha_i \alpha_j m [q(p_j) - q(p_i)].$$

Keeping both call prices fixed at $c + m/2$, this expression further simplifies to:

$$\pi_i^{UNI} = \alpha_i \left[F_i - f + (\alpha_i + \alpha_j)\frac{m}{2}q(c+m/2) \right].$$

Profits stem from the fixed fee and the fact that all calls are charged at $c + m/2$, while termination payments and revenues cancel each other out. The necessary first-order condition with respect to the fixed fee thus gives

$$0 = \frac{\partial \pi_i^{UNI}}{\partial F_i} = \frac{\partial \alpha_i}{\partial F_i} \left[F_i - f + (\alpha_i + \alpha_j)\frac{m}{2}q(c+m/2) \right] + \alpha_i \left[1 + \left(\frac{\partial \alpha_i}{\partial F_i} + \frac{\partial \alpha_j}{\partial F_i} \right)\frac{m}{2}q(c+m/2) \right].$$

Substituting (15) and (16) and re-arranging yields

$$F_i = f + \frac{\mu}{1-\alpha_i} + \frac{m}{2}q(c+m/2)\left(-2\alpha_i - \alpha_j + \frac{\alpha_i \alpha_j}{1-\alpha_i} \right).$$

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:

$$F^{UNI} = f + \frac{\mu}{1-\alpha} + \frac{m}{2}q(c+m/2)\left(\frac{4\alpha^2 - 3\alpha}{1-\alpha}\right). \tag{19}$$

We will denote the right-hand side of equation (19) by $F_{UNI}^{FOC}(\alpha, m)$ and refer to this curve

as the equilibrium curve (when termination-based price discrimination is not allowed).

From (14) 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

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

This can be rewritten as

$$F = \alpha v(p) + \alpha v(\hat{p}) - w_0 - \mu \log\left(\frac{\alpha}{1-2\alpha}\right). \tag{20}$$

We will denote the right-hand side of equation (20) by $F^{RE}(\alpha, m)$ and refer to this curve

as the rational expectations curve.

A symmetric equilibrium with fulfilled expectations with (respectively, without) ter-

mination-based price discrimination is thus found by solving the system of equations (18)

(respectively, (19)) and (20). 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 $\alpha \downarrow 0$, and it approaches $-\infty$ as $\alpha \uparrow 1/2$. The following lemma gives a sufficient

condition for the uniqueness of such a solution.

LEMMA 3.

(i) For $|m|$ small enough and $\mu > v(c)/4$, the system of equations (18) and (20) has a

unique solution.

(ii) *For $|m|$ small enough and $\mu > v(c)/4$, the system of equations (19) and (20) has a unique solution.*

4.2. Comparative Statics

We now investigate how the equilibrium behaves in a neighbourhood of $m = 0$. We first establish that an increase in the termination mark-up above 0 always reduces equilibrium fixed fees, but increases the number of subscribers if and only if on-net/off-net price discrimination is not allowed.

PROPOSITION 5.

(i) *In the case of termination-based price discrimination, a marginal increase in the termination mark-up above 0 lowers overall subscription and lowers equilibrium fixed fees.*

(ii) *In the case of no termination-based price discrimination, a marginal increase in the termination mark-up above 0 increases overall subscription and lowers equilibrium fixed fees.*

Proposition 5 states that in both cases a waterbed effect exists, that in the case of no price discrimination it is so strong that the number of subscribers increases, while with price discrimination it is weak and the number of subscribers decreases. To understand these results, observe that in both cases equilibrium profit equals[31]

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

---

[31]We omit indices to the case at hand when it does not matter. Recall that in the case of no on-net/off-net price discrimination $\hat{p} = p = c + m/2$.

Hence, in equilibrium profits stem from the fixed fee and termination services. When m increases, consumers bring with them higher termination profits. Competition for customers becomes fiercer and this leads firms to charge lower fixed fees in equilibrium. In contrast to the case of inelastic subscription demand, here there is a waterbed effect in the symmetric duopoly equilibrium. The reason is that when one firm attracts more subscribers by lowering its fixed fee, only a portion of these additional subscribers come from the other firm. Therefore the total number of off-net calls increases.[32]

The waterbed effect is stronger when on-net/off-net price discrimination is not allowed. This is because the volume (or duration in minutes) of off-net calls in this case, $q(c+m/2)$, responds less to increases in m than the volume of off-net calls when on/off-net price discrimination is allowed, $q(c+m)$.

It is not obvious how profits and welfare are affected by an increase in termination mark-up. Namely, an increase in the termination charge improves termination profits, lowers fixed fees and affects penetration. The total effect on profits and welfare will depend on the strength of the waterbed effect.

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

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

---

[32]For example, suppose that at the symmetric equilibrium each firm attracts 40 out of 100 potential consumers (and 20 consumers remain unsubscribed). In total there will be $40 \times 40 = 1600$ off-net calls. When the termination mark-up increases, a firm is willing to reduce the fixed fee so as to attract, for instance, 3 additional consumers. He can do so because only 2 of these come from the other firm while the other consumer was before unsubscribed. (The point is that a change in the fixed fee of firm 1 will not affect the ratio between subscribers of firm 2 and unsubscribed consumers in the Logit model with passive expectations.) This results in $43 \times 38 = 1634 > 1600$ off-net calls. With inelastic subscription demand, instead, the number of off-net calls would be $43 \times 37 = 1591 < 1600$.

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

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

The sign of $\partial \pi / \partial \alpha$ is negative for mature markets (when $\alpha \approx 1/2$) and positive for $\alpha \approx 0$ and $\mu < 2v(c)$. If the sign is negative, colluding networks would prefer to reduce market penetration (and thus increase fixed fees). We will refer to this case as one of effective competition. If the sign is positive, colluding networks would prefer to increase penetration (and thus reduce fixed fees). This would be the case if externalities are very important and not well internalised under competition.

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 each other out.

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

where the second equality follows from Eq. (20) and $v'(c) = -q(c)$. Putting the two effects together shows that

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

43

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.

Suppose first the case that profits increase along the rational expectations curve (low μ and low market penetration, competition is not effective). If on-net/off-net price discrimination is allowed, then an increase of the termination charge lowers profits, since $\partial\alpha/\partial m < 0$ from Proposition 5(i). When price discrimination is not allowed, then an increase of the termination charge increases profits, since $\partial\alpha/\partial m > 0$ from Proposition 5(ii).

Suppose next the case that profits decrease along the rational expectations curve, that is, competition is effective.[33] If on-net/off-net price discrimination is allowed, then an increase in the termination charge increases profits. When price discrimination is not allowed, then an increase of the termination charge lowers profits.

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 (17) we know that $dCS/dm = (\partial\alpha/\partial m)\left[2\mu/(1-2\alpha)\right]$. 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. We thus have

---

[33]This situation is the more likely scenario, especially for European countries where penetration rates are close to 100 %.

PROPOSITION 6.

(i) *Suppose on-net/off-net price discrimination is allowed. In order to maximise either consumer surplus or total surplus, the termination charge has to be set strictly below the cost of termination. Firms' profits are maximised by a termination charge above the cost of termination if and only if $\mu > 2v(c)(1 - \alpha^*)(1 - 2\alpha^*)$ (i.e., if and only if competition is effective).*

(ii) *Suppose on-net/off-net price discrimination is not allowed. In order to maximise either consumer surplus or total surplus, the termination charge has to be set strictly above the cost of termination. Firms' profits are maximised by a termination charge below the cost of termination if and only if $\mu > 2v(c)(1 - \alpha^*)(1 - 2\alpha^*)$ (i.e., if and only if competition is effective).*

This result is in stark contrast with Dessein (2003) and Hurkens and Jeon (2012) who implicitly use responsive expectations. Dessein (2003) does not allow for termination-based price discrimination while Hurkens and Jeon (2012) do. Both papers find that firms always prefer termination charges below the cost of termination. Moreover, they both find that only in the (plausible) case of effective competition, consumer surplus and total welfare are maximised when the termination charge is above cost. However, we find instead that an externality surcharge only improves penetration when termination-based price discrimination is not allowed.

One might expect that firms, for marketing purposes, will never charge on-net prices above off-net prices, even if that is optimal from a theoretical point of view when termination charges are below cost. Under this assumption, termination charges below cost will result in no on-net/off-net price discrimination, while termination charges above cost will

result in on-net/off-net price discrimination. The socially optimal termination charge would then be exactly equal to cost while, in the case of effective competition, firms would see their profits increase in both circumstances: when the termination charge is reduced below cost and when it is increased above cost. This could potentially explain two points. On the one hand, that regulators are right when they propose to set termination charges at cost, and on the other hand why operators complained when termination charges were reduced (but were still far above cost). Moreover, it is consistent with the recent campaign by smaller operators to fully adapt *Bill and Keep*.

5. Conclusion

This article has explored how consumer expectations affect retail competition when network externalities exist. Like Katz and Shapiro (1985), we assume that consumers first 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 *et al.*, 1998*b*), rational expectations are imposed on an interim basis: any change of a price by one firm leads to a *rational* change in consumer expectations. 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. The present paper has surveyed a number of relevant theoretical models (linear and non-linear pricing, duopoly and oligopoly, symmetric and asymmetric firms, elastic and inelastic subscription demand, and call externalities) and shown that a waterbed effect often does exist, but that it is always less than full: consumer welfare is improved and

networks' profits are diminished when termination charges are reduced toward or even below cost. 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 (2011). They also provide formal support to the relatively commonly held view of 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 have opposed cuts in termination charges over the past decade, and continue doing so. The arguments they employ to defend their opposition sometimes make reference to the existence of a waterbed effect. They warn regulators that cutting termination rates may lead to higher prices that would hurt consumers. Regulators have been relatively unpersuaded by this argument and sometimes even denied the existence of a waterbed effect. For example, the Australian Competition and Consumer Commission wrote 'The Commission considers that these trends of lower average retail prices [...] demonstrate that the converse of the 'waterbed' effect has been in operation.' (ACCC, 2007, p.24). The New Zealand Commerce Commission (NZCC, 2006) initially discarded the existence of a waterbed effect, and later noted that to the extent that there is a waterbed effect, 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 an externality surcharge to promote subscription was appropriate. Our model shows that this conclusion is not warranted and that the recommendation by the European Commission to not allow for such a mark-up is correct (unless termination-based price discrimination would

be prohibited).

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 significant monopolistic bottlenecks. In fact, regulation may be even more important in this case 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 this normal background, we also believe it is important to consider a truly dynamic model where consumers face switching costs and where expectations are formed endogenously. A key question would be whether and to what extent 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.

An alternative way of assessing the plausibility and validity of the two different assumptions about expectations would make use of the rich data sets that national regulators often possess. For example, one could first use data from or until 2010 about termination rates, (on- and off-net) prices and call volumes, profits, market shares and penetration rates in order to calibrate all relevant parameters of the model. Then one could make predictions of prices, profits and market shares for 2011 and 2012 (based on new and actually imposed termination rates). One should do this exercise both for the model with passive and for the model with responsive expectations. Finally one could examine which of the two models gives more accurate predictions (by comparing with the actual data from 2011 and 2012).[34]

_____

[34]Hurkens and López (2012) and Harbord and Hoernig (2012) already go some way in performing this

Although the current paper addressed a wide range of models, a number of issues that deserve further investigation remained unexplored and worthy of attention. For example, how do passive expectations affect equilibrium outcomes when (i) different types of consumers are taken into account?; (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?

Institute for Economic Analysis (CSIC), Barcelona GSE, and Public-Private Sector Research Center, IESE Business School.

Departament d'Economia Aplicada, Universitat Autònoma de Barcelona, and Public-Private Sector Research Center, IESE Business School

---

exercise. They calibrate parameters using the models of passive and responsive expectations, respectively. However, they both use publicly available data (from Spain and the UK, respectively) that are aggregated at the industry level and make predictions under hypothetical future termination rates.

References

ACCC. (2007). 'MTAS Pricing Principles Determination: 1 July 2007 to 31 December 2008', November 2007.

Anderson, S.P., and de Palma, A. (1992). 'The Logit as a model of product differentiation', *Oxford Economic Papers*, vol. 44(1), pp. 51-67.

Anderson, S.P., de Palma, A. and Thisse, J.-F. (1992). *Discrete Choice Theory of Product Differentiation,* The MIT Press, Cambridge (MA).

Armstrong, M. (1998). 'Network interconnection in telecommunications', ECONOMIC JOURNAL, vol. 108(448), pp. 545-64.

Armstrong, M. (2002). 'The theory of access pricing and interconnection', in (M. Cave, S. Majumdar and I. Vogelsang, eds.), *Handbook of Telecommunications Economics*, pp. 295-384. Amsterdam: North Holland.

Armstrong, M., and Wright, J. (2009). 'Mobile call termination', ECONOMIC JOURNAL, vol. 119(538), F270-F307.

Berger, U. (2005). 'Bill-and-keep vs. cost-based access pricing', *Economics Letters,* vol. 86(1), pp. 107-12.

Caillaud, B., and Jullien, B. (2003). 'Chicken & egg: competition among intermediation service providers', *RAND Journal of Economics*, vol. 34(2), pp. 309-28.

Calzada, J., and Valletti, T. (2008). 'Network competition and entry deterrence', ECONOMIC JOURNAL, vol. 118(531), pp. 1223-44.

Carter, M., and Wright, J. (1999). 'Interconnection in network industries', *Review of Industrial Organization*, vol. 14(1), pp. 1-25.

Carter, M., and Wright, J. (2003). 'Asymmetric network interconnection', *Review of Industrial Organization*, vol. 22(1), pp. 27-46.

Chen, Y., and Riordan, M.H. (2007). 'Price and variety in the spokes model', ECONOMIC JOURNAL, vol. 117(522), pp. 897-921.

Crémer, J., Rey, P. and Tirole, J. (2000). 'Connectivity in the commercial internet', *Journal of Industrial Economics*, vol. 48(4), pp. 433-72.

Dessein, W. (2003). 'Network competition in nonlinear pricing', *Rand Journal of Economics*, vol. 34(4), pp. 593-611.

EC. (2009*a*). 'Commission recommendation on the regulatory treatment of fixed and mobile termination rates in the EU', *Official Journal of the European Union* (May 20, 2009), L124/67-L124/74.

EC. (2009*b*). *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.

Economides, N. (1996). 'Network externalities, complementarities, and invitations to enter', *European Journal of Political Economy,* vol. 12(2), pp. 211-33.

Galeotti, A., and Moraga-González, J.L. (2009). 'Platform intermediation in a market for differentiated products', *European Economic Review,* vol. 53(4), pp. 417-28.

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

413-20.

Genakos, C., and Valletti, T. (2011). 'Testing the 'waterbed' effect in mobile telephony', *Journal of the European Economic Association,* vol. 9(6), pp. 1114-42.

Griva, K., and Vettas, N. (2011). 'Price competition in a differentiated products duopoly under network effects', *Information Economics and Policy,* vol. 23(1), pp. 85-97.

Harbord, D., and Hoernig, S. (2012).'Welfare analysis of regulating mobile termination rates in the UK with an application to the Orange/T-Mobile merger'. Working Paper, FEUNL.

Harbord, D., and Pagnozzi, M. (2010). 'Network-based price discrimination and 'bill and keep' vs. 'cost-based' regulation of mobile termination rates', *Review of Network Economics,* vol. 9(1), Article 1.

Hermalin, B.E., and Katz, M.L. (2011). 'Customer or complementor? Intercarrier compensation with two-sided benefits', *Journal of Economics & Management Strategy,* vol. 20(2), pp. 379-408.

Hoernig, S., Inderst, R. and Valletti, T. (forthcoming). 'Calling circles: network competition with non-uniform calling patterns', *Rand Journal of Economics.*

Hurkens, S., and Jeon, D.S. (2012). 'Promoting network competition by regulating termination charges', *International Journal of Industrial Organization,* vol. 30(6), pp. 541-52.

Hurkens, S., and López, Á.L. (2012). 'The welfare effects of mobile ter-

mination rate regulation in asymmetric oligopolies: The case of Spain', *Telecommunications Policy*, vol. 36(5), pp. 369-81.

Jeon, D.S., Laffont, J.J. and Tirole, J. (2004). 'On the 'receiver-pays' principle', *Rand Journal of Economics,* vol. 35(1), pp. 85-110.

Jullien, B., Rey, P. and Sand-Zantman, W. (2013). 'Termination fees revisited', *International Journal of Industrial Organization,*

http://dx.doi.org/10.1016/j.ijindorg.2013.05.006.

Katz, M.L., and Shapiro, C. S. (1985). 'Network externalities, competition, and compatibility', *American Economic Review,* vol. 75(3), pp. 424-40.

Laffont, J.J., Rey, P. and Tirole, J. (1998*a*). 'Network competition I: overview and nondiscriminatory pricing', *Rand Journal of Economics*, vol. 29(1), pp. 1-37.

Laffont, J.J., Rey, P. and Tirole, J. (1998*b*). 'Network competition II: price discrimination', *Rand Journal of Economics,* vol. 29(1), pp. 38-56.

Lee, I.H., and Mason, R. (2001). 'Market structure in congestible markets', *European Economic Review,* vol. 45(4-6), pp. 809-18.

López, A.L., and Rey, P. (2012). 'Foreclosing competition through access charges and price discrimination', Working Paper, Institut d'Economie Industrielle.

Malueg, D.A., and Schwartz, M. (2006). 'Compatibility incentives of a large network facing multiple rivals', *Journal of Industrial Economics,* vol. 54(4), pp. 527-67.

Matutes, C., and Vives, X. (1996). 'Competition for deposits, fragility, and

insurance', *Journal of Financial Intermediation,* vol. 5(2), pp. 184-216.

NZCC. (2006). 'Telecommunications Act 2001: Schedule 3 investigation into regulation of mobile termination. Reconsideration Final Report'.

Ofcom. (2004). *Wholesale Mobile Voice Call Termination: Statement.*

Ofcom. (2006). *Mobile Call Termination. Proposals for consultation.*

Ofcom. (2007). *Mobile Call Termination: Statement.*

Peitz, M., Valletti, T.M., and Wright, J. (2004). 'Competition in telecommunications: an introduction', *Information Economics and Policy,* vol. 16(3), pp. 315-21.

Salop, S.C. (1979). 'Monopolistic competition with outside goods', *Bell Journal of Economics,* vol. 10(1), pp. 141-56.

Small, K.A., and Rosen, H.S. (1981). 'Applied welfare economics with discrete choice models', *Econometrica,* vol. 49(1), pp. 105-30.

Vogelsang, I. (2003). 'Price regulation of access to telecommunications networks', *Journal of Economic Literature,* vol. 41(3), pp. 830-862.

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 maximise 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 $0 < \alpha_1 < 1$

$$R'[p_1(\alpha_1)] = \frac{q[p_1(\alpha_1)]}{2\alpha_1} > 0 \qquad (A.1)$$

and

$$\hat{R}'[\hat{p}_1(\alpha_1)] = \frac{q[\hat{p}_1(\alpha_1)]}{2(1-\alpha_1)} > 0. \qquad (A.2)$$

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,

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

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

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

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 maximised 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

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

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 (A.1) and (A.2) we can write the second-order derivative as

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

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, maximising π_i with respect to p_i, while adapting F_i so as to keep market shares constant, yields

$$\alpha_i\left\{\alpha_i\left[q(p_i) + (p_i - c)\,q'(p_i)\right] + \frac{dF_i}{dp_i}\right\} = 0. \tag{A.3}$$

For a constant α_i, differentiating Eq. (12) with respect to p_i yields

$$\sigma\left(\beta_i\left\{q(p_i) - \bar{u}'\left[q(p_i)\right]q'(p_i)\right\} + \frac{dF_i}{dp_i}\right) = 0. \tag{A.4}$$

In equilibrium, expectations are fulfilled ($\beta_i = \alpha_i$), then from Eqs. (A.3) and (A.4) we

have that $c - p_i = \overline{u}'\left[q(p_i)\right]$. Since $\overline{u}(q) = \lambda u(q)$ and $u'(q) = p$, it follows that

$$p_i = p^* \equiv \frac{c}{1 + \lambda}. \tag{A.5}$$

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

$$\alpha_i \left\{ \alpha_j \left[q(\hat{p}_i) + (\hat{p}_i - c)q'(\hat{p}_i)\right] - \alpha_j m q'(\hat{p}_i) + \frac{dF_i}{d\hat{p}_i} \right\} = 0. \tag{A.6}$$

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

$$-\sigma \beta_i \overline{u}'\left[q(\hat{p}_i)\right] q'(\hat{p}_i) - \sigma \beta_j q(\hat{p}_i) - \sigma \frac{dF_i}{d\hat{p}_i} = 0. \tag{A.7}$$

Comparing Eqs. (A.6) and (A.7), we have that $\beta_i \overline{u}'\left[q(\hat{p}_i)\right] q'(\hat{p}_i) = \alpha_j(\hat{p}_i - c - m)q'(\hat{p}_i)$. Using $\overline{u}'(q) = \lambda \hat{p}_i$, we obtain $\beta_i \lambda \hat{p}_i = \alpha_j(\hat{p}_i - c - m)$, where $\beta_i = \alpha_i$. Hence

$$\hat{p}_i = \hat{p}^*(\alpha_i) \equiv \frac{(1 - \alpha_i)(c + m)}{1 - \alpha_i(1 + \lambda)}. \tag{A.8}$$

Proof of Proposition 2.

Profits in equilibrium are given by (11). Totally differentiating with respect to m gives

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

Evaluating this derivative at $m = 0$ yields

$$\left.\frac{d\pi_i}{dm}\right|_{m=0} = \frac{2\alpha_i}{\sigma} \left.\frac{d\alpha_i}{dm}\right|_{m=0} + \alpha_i^2 q(c). \tag{A.9}$$

Totally differentiating (9) and (10), 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

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

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

$$\left. \frac{d\alpha_i}{dm} \right|_{m=0} = -\frac{2\sigma}{3} \left(\alpha_i - \frac{1}{2} \right) q(c) \quad i = 1, 2, \tag{A.10}$$

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. (A.10) into Eq. (A.9) 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

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

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

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

Proof of Proposition 4.

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

$$
w(p_i, \hat{p}_i) \equiv \beta_i v(p_i) + (1-\beta_i)v(\hat{p}_i). \tag{A.11}
$$

Market shares are determined by the indifferent customer:

$$
\begin{aligned}
\alpha_1 \;=\; & \frac{1}{2} + \sigma\left[w(p_1, \hat{p}_1) - w(p_2, \hat{p}_2) \right] \tag{A.12} \\
=\; & \frac{1}{2} + \sigma \left\{ \beta_1 \left[v(p_1) - v(\hat{p}_2) \right] - \beta_2 \left[v(p_2) - v(\hat{p}_1) \right] \right\}.
\end{aligned}
$$

Differentiating Eq. (1) $-$ where α_i is given by Eq. (A.12) and $F_i = 0$ $-$ with respect to p_i and $\hat{p}_i$, we have that at a symmetric equilibrium ($p_1 = p_2 = p$, $\hat{p}_1 = \hat{p}_2 = \hat{p}$, $\alpha_i = \beta_i = 1/2$):

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

$$[R(p) - f] - \frac{\hat{R}'(\hat{p})}{2\sigma q(\hat{p})} = 0. \tag{A.14}$$

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

Using Eqs. (A.13) and (A.14), we obtain

$$\frac{q(p) + (p - c)q'(p)}{q(p)} = \frac{q(\hat{p}) + (\hat{p} - c - m)q'(\hat{p})}{q(\hat{p})}. \tag{A.15}$$

Assuming a constant elasticity demand function $(\eta \equiv -q'(p)(p/q))$,[35] we can rewrite Eq. (A.15) as Eq. (13), which is the proportionality rule derived in Laffont *et al.* (Lemma 1, 1998*b*). The off-net price is increasing in the termination mark-up: $d\hat{p}/dm = p^*/c > 0$ (since $dp^*/dm = 0$). This proves (i).

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

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

where $\hat{p}^* = p^D(c + m)/c$. The profit maximising termination profit is thus the one that makes $\widehat{p}^*$ equal to the monopoly price, that is $m^* = (p^M/p^D)c - c$. This proves (ii). Finally, the welfare maximising termination mark-up is the one that makes the off-net price equal to marginal cost c. Hence, $m^W = c(c/p^D - 1) < 0$. This proves (iii).

---

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

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

$$\alpha_i \left\{ \tilde{F}_i - F_i - \alpha_i(p_i - c)q(p_i) - \alpha_j \left[\hat{p}_i - (c + m) \right] q(\hat{p}_i) \right\} =$$

$$\alpha_i(\alpha_i[v(c) - v(p_i) + v'(p_i)(p_i - c)] + \alpha_j \left\{ v(\hat{c}) - v(\hat{p}_i) + v'(\hat{p}_i) \left[\hat{p}_i - (c + m) \right] \right\}) > 0$$

where the equality follows from self-fulfilling 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-fulfilling 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 $\alpha \in (0, 1/2)$. Let $K \in \{PD, UNI\}$. Then

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

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 5.

Note that

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

while

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

and

$$\left.\frac{dF_{UNI}^{FOC}}{dm}\right|_{m=0} = -\alpha^* q(c)\frac{3 - 4\alpha^*}{2(1 - \alpha^*)}.$$

In the case of on-net/off-net price discrimination, an increase in m lowers the rational expectations curve by more than the equilibrium curve, since $0 < (1-2\alpha^*)/(1-\alpha^*) < 1$. The intersection point thus shifts to the south-west, lowering the subscription rate and fixed fee.

In the case of no on-net/off-net price discrimination, an increase in m lowers the rational expectations curve by less than the equilibrium curve, since $1 < (3-4\alpha^*)/[2(1-\alpha^*)]$. The intersection point thus shifts to the south-east, lowering the fixed fee and increasing subscription rate.