TSE working paper series

10-198

Research Group: *Industrial Organization*

April, 2009

Revised August, 2010

Mobile Call Termination Revisited

BRUNO JULLIEN, PATRICK REY AND

WILFRIED SAND-ZANTMAN

Ecole d'économie de Toulouse (TSE) - Manufacture des Tabacs
Aile Jean-Jacques Laffont - 21, allée de Brienne - 31000 TOULOUSE
Tél : +33 (0) 5 61 12 85 89 - Fax : +33 (0) 5 61 12 86 37 - **www.TSE-fr.eu** - contact@TSE-fr.eu

Toulouse School of Economics

Mobile call termination revisited*

Bruno Jullien[†], Patrick Rey[‡] and Wilfried Sand-Zantman[§]

30 August 2010

Abstract

We propose an explanation of the reluctance of mobile operators to move toward low levels of mobile to mobile (MTM) termination rates, that is based on the heterogeneity of calling patterns and demand elasticities among users of the service. We show that when the elasticity of participation and the intensity of usage are negatively correlated, the following conclusions hold: i) The profit maximizing MTM reciprocal termination rate is above the marginal termination cost; ii) The welfare maximizing termination rate is also above cost, but below the former. We extend the analysis to on-net / off-net pricing, and also discuss the impact of fixed to mobile termination.

1 Introduction

In this paper we revisit the analysis of the effect of mobile-to-mobile (MTM) call termination rates on the market for mobile telephony by considering the effect of heterogeneous demands for calls and subscriptions. We show that the following conclusions hold when those who call less have also a more elastic demand for subscription:

i) The profit maximizing MTM reciprocal termination rate is above the marginal termination cost;

ii) The welfare maximizing MTM reciprocal termination rate is also above cost, but below the profit maximizing level.

*We thank Stefan Behringer, Doh-Shin Jeon and seminar participants at IESE (Barcelona), Leuven, ARCEP, CREST-LEI, Ecole Polytechnique and at ESWC (Shanghai) for helpful comments.

[†]Toulouse School of Economics (GREMAQ and IDEI).

[‡]Toulouse School of Economics (GREMAQ and IDEI).

[§]Toulouse School of Economics (GREMAQ and IDEI).

The analysis and the conclusions are consistent with casual observation of the European markets for mobile telephony. A first feature of these markets is the so-called "calling party pays" (CPP) principle, according to which users do not pay for receiving calls; instead, the operator of the calling party pays a *termination rate* to the network that completes the call. A second feature is that operators offer complex non-linear tariffs; they also offer handset subsidies to attract new customers or keep their current customers. Over the years, tariffs differentiating on-net calls (within the operator's network) from off-net calls (terminating on another network) have developed, although in different ways across countries.[1] As a first approximation, we may thus view Europeans markets as markets with CPP and non-linear tariffs, and some but limited extent of on-net / off-net pricing.

Under CPP, the market is perceived to provide little discipline on the level of termination rates, since the customer of a given network is not necessarily sensitive to the price paid by those who call him. While there seems to be a general consensus that there is scope for some form of regulation, there is more disagreement on the nature of this regulation and on the adequate level of termination rates. Currently rates are regulated at the national level and, despite the European Commission's attempt in 2002 at harmonizing the practices by defining rules for the regulation of telecommunication markets, as of 2010 there is still a large disparity among the regulated levels. Yet European regulators have steadily reduced termination rates over the last ten years, and mobile operators have consistently resisted this move, arguing that reducing termination rates would impede the development of the market to the detriment of mobile customers.

Starting with the work of Laffont, Rey and Tirole (1998a,b) and Armstrong (1998), researchers have developed a body of theoretical work modelling the competition between mobile operators and analyzing the determination of termination rates.[2] One key to understanding the effect of termination on retail prices is the so-called "waterbed" effect:[3] the profit that a customer may generate on fixed-to-mobile (FTM) or mobile-to-mobile termination[4] will be at least partially competed away through retail competition, since mobile operators will then fight more fiercely to attract customers. This

[1]For instance, the practice is common in UK since the 90s, while it started recently in France. Moreover, some operators simply charge different prices for (all) off-net calls, while others propose "friends and family" packages (that is, a special low price for calls directed to a small set of numbers, chosen by the subscriber) restricted to on-net numbers.

[2]See Hoernig (2010) for a recent and flexible model of the mobile sector.

[3]The term was coined by Paul Geroski.

[4]While the term is usually used for FTM termination, the same insight applies to MTM termination.

can take the form of reduced subscription fees but could also translate into increased advertising, larger handset subsidies or reduced fees on particular services.[5] This has recently been empirically studied by Genakos and Valletti (2007), who find a significant although not full waterbed effect. While the existing literature predicts that a partial waterbed effect leads the operators to favour large FTM termination rates, this is not so for MTM termination. This is partly due to the fact that, in the latter case, termination revenues operate transfers among mobile operators rather than from other networks, and thus affect both calls emitted and received. As pointed out recently by Armstrong and Wright (2008), one of the main conclusions of the literature is that network operators should collectively favour low MTM termination rates, which is somewhat at odds with the observation that, in practice, mobile operators resist reducing MTM rates. To reconcile theory and practice, Armstrong and Wright (2009) stress that arbitrage possibilities between fixed and mobile origination tend to link FTM to MTM rates, so that network operators may favour high rates if FTM revenues are large enough and the waterbed effect is only partial; they also contrast the case of bilateral negotiations over (reciprocal) termination rates with the situation where each operator unilaterally sets its own rates. In this paper we propose an alternative explanation for above-cost MTM rates.

We start by noting that there is considerable heterogeneity in usage patterns among users. This heterogeneity is reflected to some extent in the large variety of post-pay contracts, as well as in the differences between pre-pay and post-pay users.[6] It is a source of traffic imbalance at the customer level, since some customers call more than they receive while others receive more than they call. This makes customers more or less attractive for an operator, to an extent that depends on termination rates. The heterogeneity of calling patterns has been studied by Dessein (2003) and Hahn (2004) in contexts where total subscription demand is inelastic; they show that the waterbed effect remains full and, as a result, the profit remains unaffected by the level of the termination rate. However, as mentioned by Dessein (2003), it is not clear how this conclusion extends to situations where the subscription demand is elastic. We follow this route by allowing for elastic demand for subscription.[7]

[5]See Schiff (2008).

[6]Post-pay contracts include a monthly subscription fee as well as usage fees; pre-pay (or pay-as-you-go) contracts allow instead customers to buy minutes of calls in advance. It is worth noting here that operators are willing to maintain receiving services on pre-pay contracts even after the contracted volume of calls is exhausted.

[7]Dessein (2004) shows that the profit neutrality results break if the intensity of competition differs for different types of users. Poletti and Wright (2004) reach a similar con-

Our model is based on the observation that the willingness to pay for subscription is related to the volume of calls. Customers with very large volumes of calls are infra-marginal customers, who may switch between operators but do not renounce to the service when prices increase; marginal customers are instead those who also call less. We thus introduce two types of customers: heavy users and light users; the latter not only call less often, but their demand for subscription is moreover more elastic.

As an illustration, we use the change in data issued by the French regulator during the year 2005 to obtain some idea of the difference between pre-pay and post-pay clients. France moved away from bill-and-keep for MTM termination in January 2005,[8] which triggered a change in the statistics collected by the *Observatoire des mobiles*. During the first semester, volumes included the minutes of calls emitted, along with FTM termination and roaming. Afterwards, the volumes also included the number of minutes of off-net MTM termination. Using these data, for each quarter of 2005 we computed average volumes, for pre-pay and post-pay customers; we present the results in the following Table:

Volume per subscriber 2005 (mn)	$Q1$	$Q2$	$Q3$	$Q4$
Post-pay	786	798	837	867
Pre-pay	156	159	205	201

Table 1

The Table confirms that pre-pay customers call much less than post-pay ones (see e.g. the first two quarters). Assuming that usage is relatively stable across quarters, the difference between the third and second quarters provides moreover a rough estimate of off-net incoming calls from mobile. This amounts to 39 mn for post-pay customers and 47 mn for pre-pay clients. Therefore, adding off-net MTM termination raises volume by 35% for pre-pay and only 5% for post-pay; this confirms that, compared with post-pay customers, pre-pay customers tend to receive relatively more than they call.[9]

Formally, we use the framework of Laffont, Rey and Tirole (1998a) (hereafter LRT), in which we introduce light users; in particular, as in LRT and Armstrong (1998), we assume full participation for heavy users (that is, their aggregate demand for subscription is inelastic). To keep things simple, in the main part of the paper we assume that light users actually only receive calls; we then show that the analysis is robust when light users also call. We also

clusion by introducing a participation constraint on usage. Both papers however maintain the assumption of a fixed participation.

[8]The decision was to align MTM and FTM termination rate.

[9]Note that this 5% figure is probably an over-estimate since there was an positive trend on post-pay during this period.

consider first the case of explicit (third-degree) price discrimination, where the operators can thus offer different contracts for heavy and light users, and show later that the analysis carries over to the case of implicit (second-degree) price discrimination through a menu of tariffs (pre-pay and post-pay contracts, say).

Operators offer a menu of contracts, each including a subscription fee and a unit price for calls. We first focus on MTM termination and on customer prices that are uniform across networks. We then consider termination-based price discrimination, referred to simply as on-net pricing, by allowing different prices for off-net and on-net calls. We finally introduce FTM termination. In each situation, we analyze the impact of (reciprocal) termination rates on subscription and usage prices, as well as on profits and welfare. In equilibrium, usage prices are equal to perceived costs and there is no profit from origination; network operators' profit is thus driven by termination revenue (or deficit) and by subscription fees. We identify two new effects:

Raising termination revenue weakens competition for heavy users: Introducing light users reduces competition for heavy ones when the termination charge is above cost. The reason is that the operators then obtain more profit from terminating off-net calls than on-net calls. Therefore, losing a heavy user to the competitor raises the termination profit on light users – without generating an equivalent cost, since light users call less than they are called.

Raising termination revenue intensifies competition for light users: This is a variant of the waterbed effect. Since light users generate a positive termination balance, they become more profitable when the termination markup increases, hence a reduction in the equilibrium price. This waterbed effect is however modified here, due to the fact that losing light users to the competing network generates a termination deficit, since light users are mainly receivers. This additional cost further intensifies competition for light users.

We show that, in the absence of on-net pricing, the former effect dominates for profit while the latter dominates for welfare. As a result, both profit and welfare are maximal for termination rates that are above cost: adopting a positive termination markup increases welfare because it generates a market expansion that benefits all customers – in contrast, in the absence of any scope for demand expansion, welfare would be maximized for cost-based termination charges. The operators also prefer a positive markup because the extra revenue from termination by heavy users is not fully competed away anymore through subscription fees. A conflict arises, however, since network operators favor excessively high termination rates.

The analysis is more complex with on-net pricing, because the market then exhibits tariff-mediated network effects: with a positive termination markup, the off-net price is above the on-net price; a customer is thus better off joining a larger network, as a larger proportion of calls will then remain on-net. As pointed out by Laffont, Rey and Tirole (1998b) and Gans and King (2001), these network effects intensify competition. While these network effects mostly concern heavy users, the operators then compete more fiercely for both heavy and light users. We show that welfare is still maximized for a termination rate that lies above cost. The conclusion for the operators' profit depends on the characteristics of the market: the operators prefer a termination rate above cost to a cost-based termination rate if the size of the demand from light users is not too small nor too inelastic.

Introducing FTM termination revenues reduces all subscription prices. While this waterbed effect is stronger for heavy users (because their demand is inelastic), we show that the overall effect of FTM termination revenue on prices may actually be larger for either heavy or light users. The reason is that increasing the subscription fee for light users further weakens the competition for heavy users, which in turn limits the reduction of their subscription fees. Finally we show that the network operators collectively favor a positive FTM rate while they would be indifferent if they were no light users.

The paper is organized as follows. Sections 2 to 4 present the analysis for the basis model. Section 5 considers on-net pricing. Section 6 extends the model to allow calls from light users. Section 7 discusses FTM termination while section 8 concludes.

2 The model

Two mobile operators 1 and 2 compete for heterogenous customers. To keep the exposition simple, we assume that there are only two categories of users: *heavy users* wishing to call as well as to receive calls, and *light users* who are only interested in being reached. The case where light users also call is discussed in section 6.

Providing the service involves a fixed cost per customer, which we allow to differ across the two categories of users; we denote by f the cost for heavy users and by $\tilde{f}$ the cost for light users. Each call moreover generates an origination cost c_O of calls and a termination cost c_T. The total cost of a call is therefore $c = c_O + c_T$. In the case of an off-net call, the calling network pays a termination charge a to the receiving network, which is assumed to be reciprocal and non-negative.[10] The terminating network thus receives the termination markup $m \equiv a - c_T$, while the originating network bears a cost

[10] Negative termination charges could generate abuses.

$c+m$. To study the impact of the access charge on the competition between the two operators, we consider the following timing:

- first, the reciprocal access markup m is set (more on this below);
- second, the two operators compete in retail prices.

We moreover assume that the two types of users are sufficiently different that each network $i = 1, 2$ can discriminate between them by offering two contracts: a two-part tariff for heavy users, which consists of a subscription fee F_i and a unit price for calls p_i, and a simple fixed fee $\tilde{F}_i$ for light users (together with a high usage price, say). We show later on that the main insights obtained for such explicit (or third-degree) discrimination extends to the case of implicit (or second-degree) discrimination.

Networks are differentiated and face symmetric demands. More precisely, the demand from light users for network i is given by $\tilde{\alpha}_i = D(\tilde{F}_i, \tilde{F}_j)$. We will assume that demand for network i is bounded, twice continuously differentiable with bounded derivatives, decreasing with its own price[11] ($D_1 < 0$) and increasing with its rival's price ($D_2 > 0$), and that the aggregate demand, $\tilde{\alpha}_T(\tilde{F}) \equiv 2D(\tilde{F}, \tilde{F})$, is decreasing in $\tilde{F}$; the "replacement ratio", $\gamma(\tilde{F}) \equiv -D_2(\tilde{F}, \tilde{F})/D_1(\tilde{F}, \tilde{F})$, is thus such that $0 \leq \gamma(\tilde{F}) < 1$.[12]

Heavy users represent a mass 1 and are uniformly distributed on an Hotelling line of length 1, whereas the networks are located at the two ends of the segment. Heavy users moreover have a balanced calling pattern and thus call all subscribers (heavy and light) with equal probability; a volume of calls q gives them a utility $u(q)$ per subscriber. Assuming that all heavy users subscribe, so that the total number of subscribers is $1 + \tilde{\alpha}_T$, subscribing to network i gives a user located at a distance x a net utility given by

$$u_0 + (1 + \tilde{\alpha}_T)(u(q) - p_i q) - F_i - \frac{x}{2\sigma},$$

where σ measures the degree of substitution between the two networks and u_0 denotes the fixed utility from receiving calls. The volume of calls is then given by

$$q_i = q(p_i) \equiv \arg\max_{q \geq 0} \{u(q) - p_i q\},$$

where we assume that $q(p)$ is differentiable. Letting $v(p_i) \equiv \max_{q \geq 0} \{u(q) - p_i q\}$ denote the surplus so achieved, heavy users' overall variable surplus from the service is thus:

$$w_i \equiv (1 + \tilde{\alpha}_T) v(p_i) - F_i. \tag{1}$$

[11] In the following, D_i denotes the partial derivative of the demand function D with respect to it i^{th} argument.

[12] The limit case $\gamma = 1$ would correspond to the case of full (fixed) participation.

Operator i's market share among heavy users is then given by:

$$\alpha_i = \frac{1}{2} + \sigma \left(w_i - w_j \right). \tag{2}$$

Departing from cost-based termination charges (i.e., $m \neq 0$) may introduce non-concavity problems. However, building on the analysis of LRT, it can be checked that a unique symmetric equilibrium indeed exists as long as the termination markup is not too large and/or networks are sufficiently differentiated (i.e., σ and $\gamma(.)$ small). Throughout the paper, we will assume the following:

Assumption A

1. Heavy users' utility from receiving calls, u_0, is large enough to ensure that their entire segment is always covered;

2. $v(0)$ and $q(0)$ are bounded and the two networks are sufficiently differentiated (i.e., σ and $\gamma(.)$ small enough) that there always exists a unique, pure strategy equilibrium, in which the two networks moreover share the market equally.[13]

3 Retail market equilibrium

Assumption A ensures the existence of a unique equilibrium for a given access markup m, and this equilibrium is moreover characterized by the first-order conditions. For given prices $\left(p_i, F_i, \tilde{F}_i\right)_{i=1,2}$, and given subscription demands from heavy and light users $(\alpha_i, \tilde{\alpha}_i,)_{i=1,2}$, network i's profit is equal to, for $i \neq j = 1, 2$:

$$\begin{aligned} \Pi_i &= \alpha_i \left[(1 + \tilde{\alpha}_T)(p_i - c) q(p_i) - (\alpha_j + \tilde{\alpha}_j) m q(p_i) + F_i - f \right] \\ &\quad + (\alpha_i + \tilde{\alpha}_i) \alpha_j m q(p_j) + \tilde{\alpha}_i \left(\tilde{F}_i - \tilde{f} \right). \end{aligned}$$

A first and by now standard step consists in optimizing with respect to the usage price p_i, while adjusting the fee F_i so as to keep constant the variable

[13] Laffont, Rey and Tirole (1998a) show that, in the case of homogenous users, a symmetric shared-market equilibrium exists when m and/or σ is small enough.The argument can easily be extended here (as in Dessein (2003), who considers the case of implicit discrimination among heterogenous users); in particular, the bound on $v(0)$ puts a limit on non-concave terms in profit expressions for $m > 0$, while the restriction to non-negative termination charges puts a similar limit for $m < 0$. Lopez and Rey (2008) provide a detailed analysis of the existence of shared-market and cornered market equilibria.

surplus $w_i = (1+\tilde{\alpha}_T)\, v(p_i) - F_i$. Indeed a marginal change in prices dp_i and $dF_i = -(1+\tilde{\alpha}_T)\, q(p_i)\, dp_i$ allows the firm to maintain all the market shares constant and yields a marginal gain:

$$\begin{aligned}
\left.\frac{\partial \Pi_i}{\partial p_i}\right|_{w_i,\tilde{F}_i} &= \alpha_i \left[(1+\tilde{\alpha}_T)\left((p_i - c)\, q'(p_i) + q(p_i)\right) - (\alpha_j + \tilde{\alpha}_j)\, mq'(p_i) - (1+\tilde{\alpha}_T)\, q(p_i)\right] \\
&= \alpha_i q'(p_i) \left[(1+\tilde{\alpha}_T)(p_i - c) - (\alpha_j + \tilde{\alpha}_j)\, m\right],
\end{aligned}$$

which, evaluated at a symmetric equilibrium $(\alpha_i = \alpha_j = 1/2, \tilde{\alpha}_j = \tilde{\alpha}_T/2)$, leads to:

$$p_1 = p_2 = p^* \equiv c + \frac{m}{2}. \tag{3}$$

As in the previous literature, the networks thus price usage at the average perceived marginal cost. Given these equilibrium prices, network i's profit is equal to:

$$\begin{aligned}
\Pi_i &= \alpha_i \left[(\alpha_i + \tilde{\alpha}_i)\frac{mq^*}{2} - (\alpha_j + \tilde{\alpha}_j)\frac{mq^*}{2} + F_i - f\right] \\
&\quad + (\alpha_i + \tilde{\alpha}_i)\, \alpha_j mq^* + \tilde{\alpha}_i \left(\tilde{F}_i - \tilde{f}\right),
\end{aligned}$$

where $q^* \equiv q(p^*)$ denotes the equilibrium volume of calls per subscriber and:

$$\alpha_i = 1 - \alpha_j = \frac{1}{2} - \sigma (F_i - F_j).$$

Differentiating with respect to the subscription fee F_i yields, at a symmetric equilibrium:

$$\begin{aligned}
\left.\frac{\partial \Pi_i}{\partial F_i}\right|_{F_1=F_2=F, \tilde{\alpha}_1=\tilde{\alpha}_2=\frac{\tilde{\alpha}_T}{2},} &= -\sigma (F - f) + \frac{1}{2}\left[-\sigma mq^* + 1\right] + \left(-\sigma \frac{1}{2} + \sigma \frac{1+\tilde{\alpha}_T}{2}\right) mq^* \\
&= \frac{1}{2} - \sigma (F - f) + \sigma (\tilde{\alpha}_T - 1)\frac{mq^*}{2}.
\end{aligned}$$

Therefore, the equilibrium fixed fee F^* is given by

$$F^* = f + \frac{1}{2\sigma} + (\tilde{\alpha}_T - 1)\frac{mq^*}{2}, \tag{4}$$

and heavy users' net variable surplus is equal to:

$$(1+\tilde{\alpha}_T)\, v(p^*) - f - \frac{1}{2\sigma} - (\tilde{\alpha}_T - 1)\frac{mq^*}{2}.$$

Condition (4) is similar to that obtained by Laffont, Rey and Tirole (1998a), except for the term in $\tilde{\alpha}_T$. To understand this condition, consider first the revenues earned by network i on the calls made or received by a heavy user. If the user subscribes to network i, his own calls generate no revenue since the usage price reflects the average variable cost, taking into account the termination markup paid on the proportion of off-net calls; the calls received from the same network generate however a retail revenue equal to $\alpha_i (p^* - c) q^* = mq^*/4$, while the calls received from the rival network generate an access revenue equal to $\alpha_j mq^* = mq^*/2$. If the user subscribes instead to network j, his calls to network i's heavy users generate a termination revenue of $\alpha_i mq^* = mq^*/2$, while the calls he receives from network i generate a net revenue $\alpha_i (p^* - c - m) q^* = -mq^*/4$, due to the difference between the price and the cost of an off-net call. On the whole, attracting the user generates a net gain equal to $mq^*/2$, as in Laffont, Rey and Tirole (1998a), and the fixed fee is reduced by this amount.

The existence of light users mitigates this first impact. While the calls placed by network i's subscribers to light users still generate no revenue, calls from network j's subscribers generate an access revenue equal to $\tilde{\alpha}_i mq^* = \tilde{\alpha}_T mq^*/2$. Losing a heavy user to the rival thus generates an additional net gain of $\tilde{\alpha}_T mq^*/2$, which is reflected in the equilibrium fixed fees.

Conditions (3) and (4) characterize the retail equilibrium prices for heavy users, for a given mass of light users. Setting heavy users' prices to their equilibrium values (which yields $\alpha_1 = \alpha_2 = 1/2$), network i's profit becomes:

$$\begin{aligned}
\Pi_i &= \frac{1}{2}\left[\left(\frac{1}{2} + \tilde{\alpha}_i\right)\frac{mq^*}{2} - \left(\frac{1}{2} + \tilde{\alpha}_j\right)\frac{mq^*}{2} + F^* - f\right] \\
&\quad + \left(\frac{1}{2} + \tilde{\alpha}_i\right)\frac{1}{2} mq^* + \tilde{\alpha}_i(\tilde{F}_i - \tilde{f}) \\
&= \frac{F^* - f}{2} + (1 + 3\tilde{\alpha}_i - \tilde{\alpha}_j)\frac{mq^*}{4} + \tilde{\alpha}_i(\tilde{F}_i - \tilde{f}).
\end{aligned}$$

Optimizing this profit with respect to $\tilde{F}_i$ amounts to maximizing

$$G\left(\tilde{F}_i, \tilde{F}_j\right) \equiv (\tilde{F}_i - C)D(\tilde{F}_i, \tilde{F}_j) - \hat{C}D(\tilde{F}_j, \tilde{F}_i),$$

where

$$C \equiv \tilde{f} - \frac{3mq^*}{4} \text{ and } \hat{C} \equiv \frac{mq^*}{4}.$$

C represents the direct opportunity cost of attracting additional light users, taking into account that each new subscriber generates a retail revenue $\alpha_i (p_i - c) q(p_i) = \frac{mq^*}{4}$ from on-net calls and a termination revenue $\alpha_j mq(p_j) = \frac{mq^*}{2}$ from incoming off-net calls. $\hat{C}$ represents the indirect opportunity cost generated by the rival's customers and corresponds to the termination deficit $\alpha_i (p_i - c - m) q(p_i) = -\frac{mq^*}{4}$.

We will assume that the corresponding game is well-behaved, namely:

Assumption B

1. The game with payoff functions $G(\tilde{F}_1, \tilde{F}_2)$ and $G(\tilde{F}_2, \tilde{F}_1)$ has a unique equilibrium, $\tilde{F}_1 = \tilde{F}_2 = \tilde{F}^e(C, \hat{C})$, which is symmetric, continuously differentiable and the unique solution to the first-order condition:

$$\frac{\tilde{F}^e - C + \gamma\left(\tilde{F}^e\right)\hat{C}}{\tilde{F}^e} = \frac{1}{\varepsilon(\tilde{F}^e)}, \tag{5}$$

where $\varepsilon(\tilde{F}) \equiv -\tilde{F}D_1(\tilde{F}, \tilde{F})/D(\tilde{F}, \tilde{F})$ denotes the own price elasticity of the demand from light users.

2. This equilibrium price verifies

$$0 < \frac{\partial \tilde{F}^e}{\partial C} \leq 1, \qquad \frac{\partial \tilde{F}^e}{\partial \hat{C}} \leq 0 < \frac{\partial \tilde{F}^e}{\partial C} + \frac{\partial \tilde{F}^e}{\partial \hat{C}}.$$

Assumption $B.1$ simply ensures that first-order conditions uniquely characterize a unique, symmetric equilibrium. The first condition in Assumption $B.2$ is fairly reasonable and simply supposes that an increase in the direct cost of the operators is at least partially passed through to users.[14] The second condition states that an increase in the indirect cost attached to rival's customers results instead in lower prices, although this effect is less important than the impact of direct costs.

Under Assumption $B.1$, the equilibrium is unique, symmetric, and characterized by the first-order condition:

$$\frac{\tilde{F}^* - (\tilde{f} - \frac{3+\gamma\left(\tilde{F}^*\right)}{4} mq^*)}{\tilde{F}^*} = \frac{1}{\varepsilon(\tilde{F}^*)}. \tag{6}$$

4 Choosing the termination rate

Let us now derive the private and social optimal values for the termination markup m. The equilibrium profit of each operator is equal to:

$$\Pi^* = \frac{1}{4\sigma} + \frac{\tilde{\alpha}_T^*}{2}\left(\tilde{F}^* - \tilde{f} + mq^*\right), \tag{7}$$

[14]In standard linear models with inelastic demands, there is full pass-through, although this is not necessarily implied by the assumption of a fixed demand. Conversely, our assumption of an elastic demand is not incompatible with a full pass-through.

where $\tilde{\alpha}_T^* \equiv \tilde{\alpha}_T(\tilde{F}^*)$ denote the equilibrium number of light users. The term $\frac{\tilde{\alpha}_T^* mq^*}{2}$ captures two effects. First, the presence of light users reduces the intensity of competition for heavy users. As a result, the profit on heavy users increases by $\frac{\tilde{\alpha}_T mq^*}{4}$. Moreover, light users generate a termination revenue $\frac{\tilde{\alpha}_T mq^*}{4}$ (which is however partially granted back through a reduction in $\tilde{F}^*$).

The equilibrium price $\tilde{F}^*$ defined by (6) depends on the termination markup m only through the access revenue

$$r(m) \equiv mq^*(p^*) = mq\left(c + \frac{m}{2}\right).$$

We have moreover:

Lemma 1 $-1 < \frac{\partial \tilde{F}^*}{\partial r} < 0$.

Proof. Using Assumption $B.1$, $\tilde{F}^* = \tilde{F}^e\left(\tilde{f} - \frac{3r}{4}, \frac{r}{4}\right)$. The conclusion then follows from $B.2$, since:

$$\frac{\partial \tilde{F}^*}{\partial r} = -\frac{3}{4}\frac{\partial \tilde{F}^e}{\partial C} + \frac{1}{4}\frac{\partial \tilde{F}^e}{\partial \hat{C}},$$

where

$$-1 \leq -\frac{\partial \tilde{F}^e}{\partial C} < \frac{\partial \tilde{F}^e}{\partial \hat{C}} \leq 0.$$

■

Since $\tilde{F}^*$ depends on m solely through the revenue r, the profit Π^* given by (7) can also be expressed as a function of r:

$$\Pi^* = \frac{1}{4\sigma} + \frac{\tilde{\alpha}_T^*}{2}\left(\tilde{F}^* - \tilde{f} + r\right). \tag{8}$$

Similarly, the equilibrium surplus of the light users, S^L, depends only on $\tilde{F}^*$ and thus on r. As for the heavy users, their surplus can be written as:

$$\begin{aligned} S^H &= (1 + \tilde{\alpha}_T^*)\, v(p^*) - F^* - \frac{t}{4} \\ &= (1 + \tilde{\alpha}_T^*)\, v(p^*) + \frac{r}{2}(1 - \tilde{\alpha}_T^*) - \frac{5}{8\sigma}. \end{aligned}$$

The termination markup m thus affects this surplus both through the access revenue r and through the equilibrium price $p^* = c + m/2$.

Let us now define the "monopoly" termination markup:

$$m^R \equiv \arg\max_m r(m),$$

which, for the sake of exposition, is assumed to be unique.[15] We can note a useful preliminary result:

Proposition 1 *For any $m > m^R$, there exists $\tilde{m} < m^R$ that Pareto dominates m.*

Proof. Take any candidate $m > m^R$. Since $r(0) = 0$ and $r(.)$ is continuous (by the continuity of demand), there exists $\tilde{m} \in [0, m^R]$ such that $r(\tilde{m}) = r(m)$. Then,

1. The profit is the same for m and $\tilde{m}$ since it only depends on r.
2. The surplus on light users is the same for m and $\tilde{m}$ for the same reason as above.
3. The surplus of heavy users is higher with $\tilde{m}$ than with m since p^* is lower for $\tilde{m}$.

■

Therefore, any access markup above m^R is Pareto dominated by an alternative markup below this threshold. We now show that the monopoly rate maximizes networks' equilibrium profit:

Proposition 2 *The profit maximizing termination markup is positive and equal to the monopoly termination markup m^R.*

Proof. The impact of m on total profits is given by:

$$\frac{\partial\left(2\Pi^*\right)}{\partial m} = \left[\tilde{\alpha}_T'(\tilde{F}^*)\frac{\partial \tilde{F}^*}{\partial r}\left(\tilde{F}^* - \tilde{f} + r\right) + \tilde{\alpha}_T^*(\frac{\partial \tilde{F}^*}{\partial r} + 1)\right]\frac{\partial r}{\partial m}. \qquad (9)$$

Consider first the case where $m \geq 0$ (and thus $r \geq 0$). Since $\gamma^* \equiv \gamma\left(\tilde{F}^*\right) < 1$, (6) then yields:

$$\tilde{F}^* - \tilde{f} + r > \tilde{F}^* - \tilde{f} + \frac{3+\gamma^*}{4}r = \frac{D}{-D_1}(\tilde{F}^*, \tilde{F}^*) \geq 0.$$

Since $\tilde{\alpha}_T' < 0$ and $-1 < \frac{\partial \tilde{F}^*}{\partial r} < 0$, the bracket term in (9) is therefore positive. Hence, in the range $m \geq 0$, the profits are maximal for $m = m^R$.

[15] If the monopoly rate m^R is not uniquely defined, the same analysis applies to its lowest value.

Now consider the case $r < 0$. A similar reasoning applies as long $\tilde{F}^* - \tilde{f} + r \geq 0$, since $\frac{\partial r}{\partial m} = \frac{mq'}{2} + q > 0$ for $m < 0$. If instead $\tilde{F}^* - \tilde{f} + r < 0$, then from (8) the equilibrium profit is lower than $\frac{1}{4\sigma}$; m is therefore dominated by a zero markup ($r = 0$), for which $\tilde{F}^* - \tilde{f} + r = \tilde{F}^* - \tilde{f} > 0$ from (6), and thus $\Pi^* > \frac{1}{4\sigma}$. Therefore, the profit maximizing termination markup is m^R. ■

Let us now turn to users. Light users' surplus is of the form $S^L\left(\tilde{F}^*, \tilde{F}^*\right)$, where $S^L\left(\tilde{F}_1, \tilde{F}_2\right)$ is such that $\frac{\partial S^L}{\partial \tilde{F}_i} = -\tilde{\alpha}_i$. Therefore:

$$\frac{\partial S^L}{\partial m} = \left(\frac{\partial S^L}{\partial \tilde{F}_1} + \frac{\partial S^L}{\partial \tilde{F}_2}\right)\frac{\partial \tilde{F}^*}{\partial m} = -\tilde{\alpha}_T^* \frac{\partial \tilde{F}^*}{\partial m}.$$

Thus, as long as $m < m^R$, light users' surplus increases with the termination markup. As for heavy users, we show in the appendix that, at $m = 0$:

$$\left.\frac{\partial S^H}{\partial m}\right|_{m=0} = \tilde{\alpha}_T^* v\left(c\right)\left(\frac{\tilde{\alpha}_T'}{\tilde{\alpha}_T}(\tilde{F}^*)\frac{\partial \tilde{F}^*}{\partial m} + \frac{v'\left(c\right)}{v\left(c\right)}\right). \tag{10}$$

Therefore:

Proposition 3 *For m small, increasing m raises heavy users' surplus if light users' subscription demand is very elastic or if heavy users' usage surplus is not very elastic.*

Proof. See the appendix. ■

The effect on heavy users is two-fold. First, raising the termination markup reduces the net surplus from usage, which in the presence of light users is no longer fully compensated by a reduction in subscription fees. Second, heavy users benefit from the increased participation of light users, due to intensified competition on this customer segment. The latter effect dominates if the subscription demand of light users is sufficiently elastic.

Total welfare can be written as follows:

$$W^* = \left[(1 + \tilde{\alpha}_T^*)\left(v^* + \frac{mq^*}{2}\right) - f - \frac{t}{4}\right] + \left[S^L(\tilde{F}^*, \tilde{F}^*) + \tilde{\alpha}_T^*\left(\tilde{F}^* - \tilde{f}\right)\right]. \tag{11}$$

The first term within bracket represents the joint surplus generated with heavy users, including call termination revenues. The second term represents the joint surplus generated with light users (excluding termination revenues). We then obtain:

Proposition 4 *The welfare maximizing termination markup is positive and strictly less than m^R.*

Proof. Using $p^* = c + m/2$ and $\frac{\partial S^L}{\partial m} = -\tilde{\alpha}_T^* \frac{\partial \tilde{F}^*}{\partial m}$, we have:

$$\frac{\partial W^*}{\partial m} = (1 + \tilde{\alpha}_T^*) \frac{mq'(p^*)}{4} + \left(v^* + \frac{mq^*}{2} + \tilde{F}^* - \tilde{f} \right) \tilde{\alpha}_T' \left(\tilde{F}^* \right) \frac{\partial \tilde{F}^*}{\partial m}. \tag{12}$$

Consider first the case $m \leq 0$. We then have $\tilde{\alpha}_T' \left(\tilde{F}^* \right) \frac{\partial \tilde{F}^*}{\partial m} > 0$, since $\frac{\partial \tilde{F}^*}{\partial m} = \frac{\partial \tilde{F}^*}{\partial r} \left(q^* + \frac{mq'(p^*)}{2} \right) < 0$. In addition, (6) yields $\tilde{F}^* \geq \tilde{f} - \frac{3+\gamma^*}{4} mq^*$, and thus:

$$\frac{\partial W^*}{\partial m} \geq (1 + \tilde{\alpha}_T^*) \frac{mq'(p^*)}{4} + \left(v^* - \frac{1+\gamma^*}{4} mq^* \right) \tilde{\alpha}_T' \left(\tilde{F}^* \right) \frac{\partial \tilde{F}^*}{\partial m}.$$

It follows that $\frac{\partial W^*}{\partial m}$ is positive for $m \leq 0$. From Proposition 1, the socially optimal termination markup thus lies in the range $\left]0, m^R\right]$. To conclude the proof, it suffices to note that, at $m = m^R \, (> 0)$, $\frac{\partial \tilde{F}^*}{\partial m} = \frac{\partial \tilde{F}^*}{\partial r} \frac{\partial r}{\partial m} = 0$ and thus:

$$\left. \frac{\partial W^*}{\partial m} \right|_{m = m^R} = (1 + \tilde{\alpha}_T^*) \frac{mq'}{4} < 0.$$

■

Therefore, the presence of light users, whose participation is elastic,[16] leads to favoring a positive termination markup. Note that the above analysis puts the same weight on both categories of users. If a regulator wanted to promote the participation of light users, thus placing a higher weight on those users, the optimal termination markup would be even higher. Note moreover that raising the termination charge above cost may benefit here *all* categories of agents. In particular, if the participation of light users is quite elastic, heavy users are better off with a positive markup, which increases their calling opportunities.

5 On-net Pricing

We now allow networks to set different prices for on-net and off-net calls. We keep the same notation as before except that p_i and $\hat{p}_i$ now denote the prices charged by network i for on-net and off-net calls respectively. Network i's profit becomes, for $i \neq j = 1, 2$:

$$\begin{aligned} \Pi_i &= \alpha_i \left[(\alpha_i + \tilde{\alpha}_i) (p_i - c) \, q(p_i) + (\alpha_j + \tilde{\alpha}_j) (\hat{p}_i - c - m) \, q(\hat{p}_i) + F_i - f \right] \\ &\quad + (\alpha_i + \tilde{\alpha}_i) \, \alpha_j mq(\hat{p}_j) + \tilde{\alpha}_i (\tilde{F}_i - \tilde{f}), \end{aligned}$$

[16] In the case of a fixed participation (i.e., β_T constant), $\frac{\partial W^*}{\partial m} = (1 + \beta_T^*) \frac{mq'(p^*)}{4}$ and thus welfare is maximal for $m = 0$.

where:

$$\begin{aligned}\alpha_i &= \frac{1}{2}+\sigma(w_i-w_j),\\ w_i &= (\alpha_i+\tilde{\alpha}_i)v(p_i)+(\alpha_j+\tilde{\alpha}_j)v(\hat{p}_i)-F_i.\end{aligned}$$

This profit can also be written as a function of w_i rather than F_i:

$$\begin{aligned}\Pi_i &= \alpha_i[(\alpha_i+\tilde{\alpha}_i)\left((p_i-c)\,q(p_i)+v(p_i)\right)\\ &\quad (\alpha_j+\tilde{\alpha}_j)\left((\hat{p}_i-c-m)\,q\left(\hat{p}_i\right)+v(\hat{p}_i)\right)-w_i-f]\\ &\quad +(\alpha_i+\tilde{\alpha}_i)\,\alpha_j m q(\hat{p}_j)+\tilde{\alpha}_i(\tilde{F}_i-\tilde{f})\end{aligned}$$

Differentiating with respect to usage prices p_i and $\hat{p}_i$ while adjusting the subscription fee F_i so as to keep constant the net surplus w_i (and thus the market shares) yields:

$$p_1=p_2=c \text{ and } \hat{p}_1=\hat{p}_2=\hat{p}=c+m.$$

Using the notation $\hat{q}=q(c+m)$, $v=v(c)$ and $\hat{v}=v(c+m)$, network i's profit can be written as:

$$\Pi_i=\alpha_i(F_i-f)+\alpha_j(\alpha_i+\tilde{\alpha}_i)m\hat{q}+\tilde{\alpha}_i(\tilde{F}_i-\tilde{f}), \tag{13}$$

where the market shares can be expressed as a function of the fixed fees:

$$\begin{aligned}\alpha_i &= \frac{1}{2}+\sigma(w_i-w_j)\\ &= \frac{1}{2}+\sigma[(2\alpha_i-1+\tilde{\alpha}_i-\tilde{\alpha}_j)(v-\hat{v})-(F_i-F_j)],\end{aligned}$$

and thus:

$$\alpha_i-\frac{1}{2}=\sigma\frac{(\tilde{\alpha}_i-\tilde{\alpha}_j)(v-\hat{v})-(F_i-F_j)}{1-2\sigma(v-\hat{v})}. \tag{14}$$

Differentiating (13) with respect to F_i then yields, at a symmetric equilibrium:

$$\left.\frac{\partial\Pi_i}{\partial F_i}\right|_{\tilde{\alpha}_1=\tilde{\alpha}_2=\frac{\tilde{\alpha}_T}{2},F_1=F_2=F^{**}}=\frac{1}{2}-\frac{\sigma}{1-2\sigma(v-\hat{v})}\left(F_1-f-\frac{1+\tilde{\alpha}_T}{2}m\hat{q}+\frac{m\hat{q}}{2}\right),$$

which leads to:

$$F_1=F_2=F^{**}=f+\frac{1}{2\sigma}+\tilde{\alpha}_T\frac{m\hat{q}}{2}-(v-\hat{v}). \tag{15}$$

To understand this determination of the equilibrium fees, it is useful to decompose again the revenues generated by the calls made or received by a heavy user. As in the absence of on-net pricing, the calls made by one of network 1's subscribers generate no revenue, since usage prices reflect again marginal costs, including the termination markup in the case of off-net calls. As for the calls received, those that originate off-net still generate an access revenue $\frac{m\hat{q}}{2}$, but on-net calls no longer generate any revenue since the price of these calls now reflect their actual cost. If the user switches to network 2 then his off-net calls generate an access revenue $(1+\tilde{\alpha}_T)\frac{m\hat{q}}{2}$ but the calls received from network 1 no longer generate any net revenue, since the price of off-net calls now also reflects their actual cost. On the whole, the net gain of attracting this user is reduced by $\frac{mq}{2}$ compared to the situation without on-net pricing, which induces an increase in the fixed fee by the same amount.

This first effect is mitigated by a tariff-mediated network effect. As in LRT (1998b), on-net pricing increases competition between networks: since attracting an additional user raises the value of a network by $v - \hat{v}$, networks compete more fiercely for subscribers; and the higher the difference between the utilities generated by on-net and off-net calls, the more intense the competition and the lower the fixed fee.

On-net pricing thus generates two conflicting effects. On the one hand, the opportunity cost of losing a heavy user is reduced, since there is less cross-subsidy between different types of calls; this first effect tends to decrease competition. On the other hand, network effects tend to increase competition. The following proposition shows that, for small termination markups, the second effect actually dominates and on-net pricing therefore benefits heavy users:

Proposition 5 *For termination rates close to the marginal cost, on-net pricing leads to lower (resp., higher) subscription fees for heavy users when m is positive (resp., negative).*

Proof. The equilibrium fixed fees without and with on-net pricing are respectively defined by (4) and (15). Therefore:

$$F^* - F^{**} = \frac{\tilde{\alpha}_T^* - 1}{2} m q^* - \frac{\tilde{\alpha}_T^{**}}{2} m \hat{q} + v - \hat{v},$$

where $\tilde{\alpha}_T^{**}$ denotes the equilibrium number of users under price discrimination. At $m = 0$, there is no difference in prices (since $p^* = \hat{p} = c$) and thus $F^{**} = F^*$ and $\tilde{\alpha}_T^{**} = \tilde{\alpha}_T^*$, which implies:

$$\left.\frac{\partial(F^* - F^{**})}{\partial m}\right|_{m=0} = \frac{\tilde{\alpha}_T^* - 1}{2} q(c) - \frac{\tilde{\alpha}_T^*}{2} q(c) + q(c) = \frac{q(c)}{2} > 0.$$

■

We now characterize the equilibrium subscription price $\tilde{F}$ for light users. Setting heavy users' prices to their equilibrium values, network i's profit becomes:

$$\Pi_i = \alpha_i(F^{**} - f) + \alpha_j(\alpha_i + \tilde{\alpha}_i)m\hat{q} + \tilde{\alpha}_i(\tilde{F}_i - \tilde{f}), \tag{16}$$

where $\tilde{F}_1$ and $\tilde{F}_2$ affect here market shares in both segments: $\tilde{\alpha}_i = D\left(\tilde{F}_i, \tilde{F}_j\right)$ and

$$\alpha_i = \frac{1}{2} + \frac{\sigma(\tilde{\alpha}_i - \tilde{\alpha}_j)(v - \hat{v})}{1 - 2\sigma(v - \hat{v})}.$$

Differentiating (16) with respect to $\tilde{F}_i$ then yields, at a symmetric equilibrium:

$$\begin{aligned}\left.\frac{\partial \Pi_i}{\partial \tilde{F}_i}\right|_{\tilde{F}_1 = \tilde{F}_2 = \tilde{F}^{**}} &= \frac{\sigma\left(v - \hat{v}\right)\left(D_1 - D_2\right)}{1 - 2\sigma\left(v - \hat{v}\right)}\left(F^{**} - f - \tilde{\alpha}_T\frac{m\hat{q}}{2}\right) \\ &\quad + D_1\left(\frac{m\hat{q}}{2} + \tilde{F}^{**} - \tilde{f}\right) + \frac{\tilde{\alpha}_T}{2},\end{aligned}$$

which, using $D_2 = -\gamma D_1$ and (15), can be rewritten as:

$$\frac{\tilde{F}^{**} - (\tilde{f} - \dfrac{\left(1 + \gamma(\tilde{F}^{**})\right)\left(v - \hat{v}\right) + m\hat{q}}{2})}{\tilde{F}^{**}} = \frac{1}{\varepsilon(\tilde{F}^{**})}. \tag{17}$$

$\tilde{F}^{**}$ satisfies the first-order condition (5) of the duopoly game for $C = \tilde{f} - \frac{m\hat{q} + v - \hat{v}}{2}$ and $\hat{C} = \frac{v - \hat{v}}{2}$; therefore, from Assumption $B.1$:

$$\tilde{F}^{**} = \tilde{F}^{e}\left(\tilde{f} - \frac{m\hat{q} + v - \hat{v}}{2}, \frac{v - \hat{v}}{2}\right).$$

Assumption $B.2$ then ensures that $\tilde{F}^{**}$ decreases with $\hat{r}$ and increases with $\hat{v}$. Therefore, an increase in the termination rate benefits light users at least as long as it also raises the termination profit. Building on this, we now characterize the impact of on-net pricing on the price offered to light users:

Proposition 6 *For termination rates close to the marginal cost, on-net pricing leads to lower (resp., higher) tariffs for light users when m is positive (resp., negative).*

Proof. In the absence of termination-based price discrimination, the price for light users satisfies $\tilde{F}^{*} = \tilde{F}^{e}(\tilde{f} - \frac{3mq^{*}}{4}, \frac{mq^{*}}{4})$. Therefore, $\tilde{F}^{*} = \tilde{F}^{**}(=$

$\tilde{F}^e\left(\tilde{f},0\right)$) for $m=0$ and:

$$\left.\frac{\partial\left(\tilde{F}^{*}-\tilde{F}^{**}\right)}{\partial m}\right|_{m=0} = \frac{\partial\tilde{F}^e}{\partial C}\left(-\frac{3q\left(c\right)}{4}+q\left(c\right)\right)+\frac{\partial\tilde{F}^e}{\partial\hat{C}}\left(\frac{q\left(c\right)}{4}-\frac{q\left(c\right)}{2}\right)$$
$$= \left(\frac{\partial\tilde{F}^e}{\partial C}-\frac{\partial\tilde{F}^e}{\partial\hat{C}}\right)\frac{q\left(c\right)}{4}>0.$$

■

On-net pricing thus induces a decrease in the price for light users when the termination charge is raised above cost. This is again partly driven by network effects: adding an additional light user renders a network comparatively more attractive for heavy users, which encourages networks to compete more fiercely for light users. In addition, while on-net calls to light users no longer generate any net revenue, off-net incoming calls still generate an access revenue equal to $\frac{m\hat{q}}{2}$, which contributes again to reduce prices.

Using the same decomposition as before, total welfare now becomes:

$$W^{**}=\left[(1+\tilde{\alpha}_T^{**})\frac{v+\hat{v}+m\hat{q}}{2}-f-\frac{t}{4}\right]+\left[S^L(\tilde{F}^{**},\tilde{F}^{**})+\tilde{\alpha}_T^{**}\left(\tilde{F}^{**}-\tilde{f}\right)\right].$$

It is then again socially desirable to raise the termination charge above cost:

Proposition 7 *With on-net pricing, the welfare maximizing termination markup is positive.*

Proof. Using $\frac{\partial S^L(\tilde{F}^{**},\tilde{F}^{**})}{\partial\tilde{F}^{**}}=-\tilde{\alpha}_T^{**}$, we have

$$\frac{\partial W^{**}}{\partial m}=(1+\tilde{\alpha}_T^{**})\frac{mq'\left(\hat{p}\right)}{2}+\frac{\partial\tilde{\alpha}_T^{**}}{\partial m}\left(\frac{v+\hat{v}+m\hat{q}}{2}+\tilde{F}^{**}-\tilde{f}\right),$$

where, from (17):

$$\frac{v+\hat{v}+m\hat{q}}{2}+\tilde{F}^{**}-\tilde{f}=\frac{D}{-D_1}\left(\tilde{F}^{**},\tilde{F}^{**}\right)+\hat{v}-\gamma\left(\tilde{F}^{**}\right)\frac{v-\hat{v}}{2}.$$

Since $\hat{v}\geq v$ for $m\leq 0$, this implies $\frac{\partial W^{**}}{\partial m}>0$ for $m\leq 0$. ■

Unsurprisingly, the result on profit is more ambiguous. Indeed, while the competition weakening effect described in the case without on-net pricing is still present, it is now mitigated by the impact of tariff-mediated network effects. Total profit can be written as:

$$2\Pi^{**}=\left(\frac{1}{2\sigma}+\frac{m\hat{q}}{2}+\hat{v}-v\right)+\tilde{\alpha}_T^{**}\left(\tilde{F}^{**}-\tilde{f}+m\hat{q}\right).$$

The first term is maximal at some negative value of m, as shown by Gans and King (2001) and Dessein (2003). The second term is more complex. Still, we can establish:

Proposition 8 *If at* $m = 0$*,*

$$\tilde{\alpha}_T^{**}\left(1-\gamma^{**}\left(1+\frac{\gamma^{**}}{2}\right)\frac{\partial \tilde{F}^e}{\partial C}(.,.)\right) > \frac{1}{2}, \tag{18}$$

then the profit is increasing with m *for* m *close to zero.*

Proof. See the appendix. ■

The result does not extend easily to larger departures from cost-based termination charges, due to the impact on the volume of traffic. We can however extend it when the latter is not too sensitive to usage prices. To get some intuition, suppose that the individual demand is inelastic:

$$q(p) = \begin{cases} \bar{q} \text{ if } p \leq \bar{p}, \\ 0 \text{ if } p > \bar{p}, \end{cases}$$

where $\bar{p} > c$ and $\bar{q} > 0$. Then, as long as $c + m < \bar{p}$:

$$2\Pi^{**} = \frac{1}{2\sigma} - \frac{m\bar{q}}{2} + \tilde{\alpha}_T^{**}\left(\tilde{F}^{**} - \tilde{f} + m\bar{q}\right).$$

In that case we have:

Corollary 1 *If individual usage is inelastic and condition* (18) *holds for any* $m < 0$*, then the profit maximizing termination charge is above cost.*

Proof. See the appendix. ■

Condition (18) is such that, as long as light users' participation is elastic (γ small) and/or pass-through rate of the duopoly price is not too high ($\frac{\partial \tilde{F}^e}{\partial C}$ small), the profit maximizing termination margin is positive when the population of light users is large enough. Note that, since scaling the demands $\tilde{\alpha}_i$ by a multiplicative factor λ does not affect the equilibrium prices, the condition is indeed easier to satisfy when the demand from light users is large. When for example the operators have a local monopoly over their own clientele of light users ($D_2 = 0$, which implies $\gamma = 0$), the profit maximizing termination markup is positive if the equilibrium proportion of light users exceeds one third of the total customer base. In contrast, the condition is unlikely to be satisfied when the participation of light users is quite inelastic (γ close to 1) and/or the pass-through rate is large ($\frac{\partial \tilde{F}^e}{\partial C}$ close to 1).[17]

[17] As γ and $\frac{\partial \tilde{F}^e}{\partial C}$ are between 0 and 1, the term $\left(1-\gamma\left(1+\frac{\gamma}{2}\right)\frac{\partial \tilde{F}^e}{\partial C}\right)$ lies in the range $\left[-\frac{1}{2}, 1\right]$. It is positive for instance if $\frac{\partial \tilde{F}^e}{\partial \hat{C}} < 2/3$ or if $\gamma < \sqrt{3} - 1 \approx 0.73$.

6 Robustness

So far we have neglected light users' demand for calls; the analysis however still applies as long as light users call sufficiently less than heavy users. We have also ignored self-selection constraints, which is fine if for example the operators can explicitly discriminate light users from heavy ones, or if the equilibrium tariffs are incentive compatible; the operators must otherwise take into account self-selection constraints. We address these two issues in turn.

6.1 Demand for calls from light users

We show here that our basic analysis applies when light users call as well, although to a lesser extent than heavy users. To see this, we now suppose that light users derive a utility $\theta u(q/\theta)$ from calls (with the same balanced calling pattern as for heavy users), where θ is significantly lower than 1. For the sake of exposition, we assume that explicit, third-price discrimination is possible, but the reasoning extends as well to implicit, second-degree price discrimination.

Letting $\tilde{p}_i$ denote the unit price of calls for light users, the per subscriber demand for calls is then $\theta q(\tilde{p}_i)$, and the surplus from calls is $(1+\tilde{\alpha}_T)\theta v(\tilde{p}_i)$. We extend the model by assuming that light users' subscription demand now relies on net surpluses: $\tilde{\alpha}_i = D(-\tilde{w}_i, -\tilde{w}_j)$, where

$$\tilde{w}_i = (1+\tilde{\alpha}_T)\theta v(\tilde{p}_i) - \tilde{F}_i.$$

The operators' customer bases are then solution to the system:

$$\begin{aligned}
\tilde{\alpha}_1 &= D\left(\tilde{F}_1 - (1+\tilde{\alpha}_T)\theta v(\tilde{p}_1), \tilde{F}_2 - (1+\tilde{\alpha}_T)\theta v(\tilde{p}_2)\right), \\
\tilde{\alpha}_2 &= D\left(\tilde{F}_2 - (1+\tilde{\alpha}_T)\theta v(\tilde{p}_2), \tilde{F}_1 - (1+\tilde{\alpha}_T)\theta v(\tilde{p}_1)\right), \\
\tilde{\alpha}_T &= \tilde{\alpha}_1 + \tilde{\alpha}_2,
\end{aligned} \tag{19}$$

and are uniquely defined for θ small enough. Operator i's profit is now:

$$\begin{aligned}
\Pi_i &= \alpha_i\left[(1+\tilde{\alpha}_T)(p_i - c)q(p_i) - (\alpha_j + \tilde{\alpha}_j)mq(p_i) + F_i - f\right] \\
&\quad + (\alpha_i + \tilde{\alpha}_i)\alpha_j mq(p_j) + \tilde{\alpha}_i\left(\tilde{F}_i - \tilde{f}\right) \\
&\quad + (\alpha_i + \tilde{\alpha}_i)\tilde{\alpha}_j m\theta q(\tilde{p}_j) + \tilde{\alpha}_i\theta\left[(1+\tilde{\alpha}_T)(\tilde{p}_i - c) - (\alpha_j + \tilde{\alpha}_j)m\right]q(\tilde{p}_i).
\end{aligned}$$

It differs from the benchmark profit by two terms (the last line in the above equation), representing the termination and retail revenues generated by light-users' calls.

We follow the same steps as before to derive equilibrium conditions. First, keeping net surpluses w_i and $\tilde{w}_i$ (and thus subscription demands) unchanged, the effect of usage prices is the same as before; at a symmetric equilibrium, we thus have again:

$$\tilde{p}_i = p_i = p^* = c + \frac{m}{2}.$$

Second, for $p_i = \tilde{p}_i = p^*$, optimizing with respect to F_i yields, at a symmetric equilibrium:

$$\left.\frac{\partial \Pi_i}{\partial F_i}\right|_{\tilde{\alpha}_1=\tilde{\alpha}_2=\tilde{\alpha}_T/2, F_1=F_2=F} = \frac{1}{2} - \sigma\left(F - f\right) + \sigma\left(\tilde{\alpha}_T - 1\right)\frac{mq^*}{2} - \sigma\theta\tilde{\alpha}_T mq^*.$$

Therefore, the equilibrium fixed fee $F^*\left(\theta\right)$ is given by

$$F^*\left(\theta\right) = f + \frac{1}{2\sigma} + \left(\tilde{\alpha}_T - 1\right)\frac{mq^*}{2} - \theta\tilde{\alpha}_T mq^*. \tag{20}$$

Since light users now call heavy users, the competition weakening effect is smaller; this is reflected in the term $\theta\tilde{\alpha}_T mq^*$, which includes the termination revenue $\theta mq^*\tilde{\alpha}_T/2$ from light users calling the marginal heavy user, as well as the cost saving $\theta mq^*\tilde{\alpha}_T/2$ on the calls from light users. *Thus, for given participation levels, accounting for light users' calls intensifies competition for heavy users.*

Finally, it is shown in the appendix that the equilibrium is continuous at $\theta = 0$:

Proposition 9 *When θ tends to 0, the equilibrium tariffs (and thus participation as well as call volumes) converge to the benchmark equilibrium values ($\theta = 0$).*

Proof. See appendix. ■

The conclusions of our analysis thus extend to the more general case where the demand for calls of light users is positive but small. It is moreover shown in the appendix that light users' net surplus satisfies:

$$\tilde{w}^*\left(\theta\right) = -\tilde{F}^e\left(\tilde{f} - \frac{3}{4}mq^* - \theta\Xi, \frac{mq^*}{4}\right),$$

where Ξ is positive for θ and m close to 0, implying that light users' participation is higher when they have a positive demand for calls: *when the termination charge is close to the termination cost, accounting for light users' calls raises the participation of light users.*

6.2 Second-degree price discrimination

When operators cannot discriminate users explicitly ("third-degree" price discrimination), they can still do so implicitly by offering a menu of options that induce different customers to adopt different tariffs ("second-degree" discrimination). In this case, however, simple two-part tariffs are not necessarily optimal anymore, and more sophisticated tariffs can be desirable.[18] For example, when light users have no demand for calls, offering them an option that only allows for receiving calls does not lower their utility, and makes the option less appealing for heavy users. As a result, in the absence of on-net pricing the above analysis remains relevant as long as: (i) light users prefer paying $\tilde{F}^*$ than F^*; and (ii) heavy users favour instead the two-part tariff (F^*, p^*) to being able to receive calls for a flat fee $\tilde{F}^*$. In other words, the above tariffs (together with a quota of zero calls for light users or, equivalently, with a prohibitively high usage price) are incentive compatible whenever:

$$F^* > \tilde{F}^* > F^* - (1 + \tilde{\alpha}_T^*)\, v(p^*) \tag{21}$$

Similarly, when on-net pricing is allowed, heavy users must prefer the tariff $(F^*, p = c, \hat{p} = c + m)$, and thus the analysis holds whenever:

$$F^* > \tilde{F}^* > F^* - (1 + \tilde{\alpha}_T)\, \frac{v + \hat{v}}{2} \tag{22}$$

Conversely, under these conditions the equilibria are the same, whether the operators can discriminate explicitly or must induce self-selection.

The situation is different when the above condition does not hold. For the sake of exposition, we will focus here on the case where: (i) operators cannot engage in on-net pricing, and: (ii) heavy users are the ones that may be tempted to choose the contract designed for the other category of users (that is, $\tilde{F}^* < F^* - (1 + \tilde{\alpha}_T)\, v\,(p^*)$).

Without loss of generality, the operators could restrict attention to options granting a fixed volume of calls, q_i, for a given fee, F_i. It is moreover clearly optimal here to set light users' quotas to zero since this relaxes heavy users' incentive constraints, without any adverse effect on light users. As for heavy users, there is no need to distort their usage (that is, $q_i = q^*$), since

[18] Without loss of generality, operators could restrict attention here to "quantity forcing" contracts specifying a volume of calls in exchange of a flat fee. In more general contexts, in which users' demand for calls may for example be uncertain, smoother nonlinear contract menus might be preferable. In practice, some contracts impose a cap, while others include free minutes of calls which can be interpreted as a minimum level of usage.

light users are not tempted by heavy users' options anyway.[19] Therefore, we can interpret the operators' offers as presenting two options, a two-part tariff $(F_i, p_i = p^*)$ for heavy users and a flat rate $\tilde{F}_i$ for light users, which only allows them to receive call. When designing their offers, the operators must now adjust the fixed fees so as to accommodate heavy users' self-selection constraints, namely:

$$(1 + \tilde{\alpha}_T)\, v(p^*) - F_i \geq -\tilde{F}_i.$$

The shadow cost of this constraint (which is binding when (21) is violated) leads the operators to increase the fee for light users and decrease that for heavy users:[20]

$$\begin{aligned} F &< f + \frac{1}{2\sigma} + (\tilde{\alpha}_T - 1)\, \frac{mq^*}{2}, \\ \tilde{F} &> \tilde{F}^* = \tilde{F}^e \left(\tilde{f} - \frac{3}{4} mq^*, \frac{mq^*}{4} \right). \end{aligned}$$

The second inequality implies that light users' participation $\tilde{\alpha}_T$ is reduced: $\tilde{\alpha}_T < \tilde{\alpha}_T^*$. This, along with the first inequality, implies in turn that heavy users face a lower subscription fee, $F < F^*$. Notice that while light users' welfare is reduced, the effect on heavy users' welfare is unclear, since the contraction of the market reduces their utility from calls.

7 Fixed to mobile termination charges

So far we have focused on MTM termination and ignored calls from/to fixed networks. We now turn to the potential incentives to raise termination charges for fixed to mobile calls. As for the calls to the fixed networks, for the sake of exposition we will simply assume that all mobile users derive a fixed utility from them, which is moreover the same on both networks.[21]

[19] As usual, there is thus "no distortion at the top" (that is, for the largest customers). When light users have a demand for calls, their usage may instead be distorted downwards: starting from a symmetric situation where $\tilde{q}_i = q^*$, reducing slightly the volume of calls below q^* (while maintaining light users' net surplus $\tilde{w}_i$) does not affect light users but relaxes heavy users' incentive constraint and generates only a second-order loss of efficiency.

[20] A similar reasoning as above shows that the equilibrium conditions differ only by a first-order term in θ; the equilibrium allocations thus remain in the neighborhood when θ is small.

[21] In Europe, the termination charge of fixed networks is regulated and is the same for all mobile operators. Introducing this feature in our framework, and allowing mobile operators to discriminate between calls terminating on fixed and mobile networks, as is

Thus, mobile users' choice of network and usage are unaffected by the presence of fixed networks. The only difference is that now the network receives an additional FTM termination revenue, ρ, which depends on the FTM termination markup and on the volume of calls that a mobile customer receives from fixed networks, but is independent from mobile operators' strategies: it simply constitutes a "windfall gain" per customer, which amounts to reduce the per customer fixed costs f or $\tilde{f}$. In the absence of on-net pricing, the equilibrium prices will thus be:

$$\begin{aligned}
p^* &= c + \frac{m}{2}, \\
F^* &= f - \rho + \frac{1}{2\sigma} + (\tilde{\alpha}_T - 1)\frac{mq^*}{2}, \\
\tilde{F}^* &= \tilde{F}^e\left(\tilde{f} - \rho - \frac{3mq^*}{4}, \frac{mq^*}{4}\right).
\end{aligned}$$

In this context it is interesting to compare the effect of FTM termination revenues on the subscription prices and the total bills of the two user categories. In the absence of light users, the FTM termination revenue would be entirely passed through to customers: it is indeed immediate that $\frac{\partial F^*}{\partial \rho} = -1$ for $\tilde{\alpha}_T = 0$. But with the presence of light users, the pass-through to heavy users is only partial. The reason is that the prices for light users decrease:

$$\frac{\partial \tilde{F}^*}{\partial \rho} = -\frac{\partial \tilde{F}^e}{\partial C} < 0.$$

Their participation therefore increases, which in turn tends to weaken competition for heavy users; as a result:

$$\frac{\partial F^*}{\partial \rho} = -1 + \frac{mq^*}{2}\tilde{\alpha}_T' \frac{\partial \tilde{F}^*}{\partial \rho} > -1.$$

Thus, while in the absence of light users, the FTM termination revenue would be fully absorbed by a reduction in heavy users' subscription fees, the presence of light users limits this waterbed effect. The comparison between the impact of the termination revenue on the two user categories depends on the pass-through rate for light users:

Proposition 10 *Increasing the FTM per customer revenue ρ reduces more the subscription fee for heavy users than for light users if and only if* $\frac{\partial \tilde{F}^e}{\partial C} < 1/\left(1 - \frac{mq^*}{2}\tilde{\alpha}_T'\right)$.

the case in practice, they then adopt the same price, reflecting the regulated cost of these calls. Thus the utility derived from calls to fixed networks is indeed independent of the network, and these calls moreover generate no profit for the mobile operators.

Proof. It follows directly from the above analysis, since $\frac{\partial F^*}{\partial \rho} - \frac{\partial \tilde{F}^*}{\partial \rho} = \left(1 - \frac{mq^*}{2}\tilde{\alpha}_T'\right)\frac{\partial \tilde{F}^e}{\partial C} - 1$. ■

A second question concerns the preferred level of termination rates. In the absence of light users, the mobile operators are indifferent to the (common) level of FTM termination rate, which does not affect their profits due to a full waterbed effect. When light users are present, each mobile operator's profit becomes (replacing $\tilde{f}$ with $\tilde{f} - \rho$ in the previous expressions):

$$2\Pi^* = \frac{1}{2\sigma} + \tilde{\alpha}_T^*\left(\tilde{F}^* - \tilde{f} + \rho + mq^*\right).$$

Thus:

Proposition 11 *Whenever $\tilde{\alpha}_T^* > 0$, as long as $m \geq 0$ (or more generally if $\tilde{F}^* > \tilde{f} - \rho - mq^*$), an increase in FTM termination revenue ρ raises the equilibrium profit (while it has no effect on profit without light users).*

Proof. We have:

$$\frac{\partial\left(2\Pi^*\right)}{\partial \rho} = \tilde{\alpha}_T^* + \left[\tilde{\alpha}_T^* + \tilde{\alpha}_T'\left(\tilde{F}^* - \tilde{f} + \rho + mq^*\right)\right]\frac{\partial \tilde{F}^*}{\partial \rho}$$

Since $\frac{\partial \tilde{F}^*}{\partial \rho} < 0$, this is positive if $\tilde{\alpha}_T^* + \tilde{\alpha}_T'\left(\tilde{F}^* - \tilde{f} + \rho + mq^*\right) < 0$. Suppose instead that $\tilde{\alpha}_T^* + \tilde{\alpha}_T'\left(\tilde{F}^* - \tilde{f} + \rho + mq^*\right) > 0$. Since $\frac{\partial \tilde{F}^*}{\partial \rho} > -1$, we then have:

$$\frac{\partial\left(2\Pi^*\right)}{\partial \rho} > -\tilde{\alpha}_T'\left(\tilde{F}^* - \tilde{f} + \rho + mq^*\right),$$

and is thus positive if $\tilde{F}^* > \tilde{f} - \rho - mq^*$. In particular, from (6), replacing $\tilde{f}$ with $\tilde{f} - \rho$:

$$\tilde{F}^* - \tilde{f} + \rho + mq^* = \frac{D}{-D_1} + \left(1 - \gamma^*\right)\frac{mq^*}{4},$$

and thus $\partial\left(2\Pi^*\right)/\partial \rho > 0$ whenever $m \geq 0$. ■

Therefore, under fairly general conditions the mobile operators would prefer to coordinate on a positive FTM termination markup.

8 Conclusion

This paper proposes an explanation of mobile operators' reluctance to reduce MTM termination rates. We show that the insights of the existing literature, which suggest profit-maximizing rates at or below cost, rely critically on the related assumptions of fixed participation and full pass-through. Accounting for the heterogeneity among users, we show that when the elasticity of subscription and the intensity of usage are negatively correlated across users, then the profit maximizing MTM (reciprocal) termination rate is instead always above cost in the absence of on-net pricing, and can still be so with on-net pricing; in addition, the welfare maximizing termination rate is also above cost, although it is below the former one in the absence of termination-based price discrimination. We also study the robustness of these insights when taking fixed to mobile termination revenues into consideration.

Our results thus imply that while some cap on termination rates is desirable, the regulated cap should be above termination costs. This optimal rate depends on factors such as the proportion of light users and their demand elasticity. Thus local market conditions matter, suggesting that, at least in Europe, there should be some discretion left to national regulators in defining these rates.

Our model has been motivated by casual observation of the mobile markets and of business practices. The analysis shows that demand heterogeneity is a key element that needs to be accounted for in the regulatory debate. It points to the need for better empirical facts on the composition of the demand for mobile services and on the participation elasticities of the various categories of users.

References

[1] Armstrong, M. (1998), "Network Interconnection in Telecommunications", *Economic Journal*, 108:545-564.

[2] Armstrong, M. (2002), "The Theory of Access Pricing and Interconnection", in *Handbook of Telecommunications Economics*, eds. M. Cave, S. Majumdar and I. Vogelsang, North-Holland, 295-384.

[3] Armstrong, M., and J. Wright (2007), "Two-sided Markets, Competitive Bottlenecks and Exclusive Contracts", *Economic Theory*, 32(2):353-380.

[4] Armstrong, M., and J. Wright (2009), "Mobile Call Termination", *Economic Journal*, 119(538): 270-307.

[5] Dessein, W. (2003), "Network Competition in Nonlinear Pricing", *Rand Journal of Economics*, 34(4):593-611.

[6] Dessein, W. (2004), "Network Competition with Heterogeneous Customers and Calling Patterns", *Information Economics and Policy*, 16:323-345.

[7] Gans, J., and S. King (2001), "Using 'Bill and Keep' Interconnect Agreements to Soften Network Competition", *Economics Letters*, 71(3):413-420.

[8] Genakos, C., and T. Valletti (2007), "Testing the "Waterbed" Effect in Mobile Telephony", CEP Discussion Papers dp0827.

[9] Hahn, J.-H. (2004), "Network competition and interconnection with heterogeneous subscribers", *International Journal of Industrial Organization*, 22(5):611-631.

[10] Hoernig S. (2010), "Competition between multiple asymmetric networks: A toolkit and applications", mimeo Universidade Nova de Lisboa.

[11] Laffont, J.-J., P. Rey and J. Tirole (1998a), "Network Competition: I. Overview and Nondiscriminatory pricing", *Rand Journal of Economics*, 29(1):1-37.

[12] Laffont, J.-J., P. Rey and J. Tirole (1998b), "Network Competition: II. Price Discrimination", *Rand Journal of Economics*, 29(1):37-56.

[13] Lopez, A. and P. Rey (2008), "Foreclosing Competition through Access Charges and Price Discrimination", mimeo IDEI.

[14] Poletti, S. and J. Wright (2004), "Network interconnection with participation constraints", *Information Economics and Policy*, 16(3):347-373.

[15] Schiff, A. (2008), "The Waterbed Effect and Price Regulation", *Review of Network Economics*, 7(3):392-414.

A Proof of proposition 3

We have:

$$\frac{\partial S^H}{\partial m} = -\left(1+\tilde{\alpha}_T^*\right)\frac{q^*}{2} + v^*\tilde{\alpha}_T'(\tilde{F}^*)\frac{\partial \tilde{F}^*}{\partial m} + \left(1-\tilde{\alpha}_T^*\right)\frac{q^*+mq'(p^*)/2}{2} - \frac{mq^*}{2}\tilde{\alpha}_T'(\tilde{F}^*)\frac{\partial \tilde{F}^*}{\partial m}.$$

Therefore, at $m=0$:

$$\frac{\partial S^H}{\partial m} = -\tilde{\alpha}_T^* q\left(c\right) + v\left(c\right)\tilde{\alpha}_T'(\tilde{F}^*)\frac{\partial \tilde{F}^*}{\partial m} = \tilde{\alpha}_T^* v\left(c\right)\left(\frac{\tilde{\alpha}_T'}{\tilde{\alpha}_T}(\tilde{F}^*)\frac{\partial \tilde{F}^*}{\partial m} - \frac{q\left(c\right)}{v\left(c\right)}\right),$$

from which the expression (10) is obtained using $v'\left(c\right) = -q\left(c\right)$.

B Proof of proposition 8

We have:

$$\left.\frac{\partial\left(2\Pi^{**}\right)}{\partial m}\right|_{m=0} = -\frac{q}{2} + \left(\tilde{\alpha}_T + \tilde{\alpha}_T'\left(\tilde{F}-\tilde{f}\right)\right)\frac{\partial \tilde{F}^{**}}{\partial m} + \tilde{\alpha}_T q,$$

where $q = q\left(c\right)$, whereas $\tilde{\alpha}_T$ and $\tilde{\alpha}_T'$ are evaluated at $\tilde{F} = \tilde{F}^{**}\left(0\right)$, characterized by (17):

$$\tilde{F} - \tilde{f} = \frac{\tilde{F}}{\varepsilon\left(\tilde{F}\right)} = -\frac{\tilde{\alpha}_T}{2D_1(\tilde{F},\tilde{F})}.$$

Together with $\tilde{\alpha}_T' = 2\left(1-\tilde{\gamma}\right)D_1(\tilde{F},\tilde{F})$, where $\tilde{\gamma} = \gamma(\tilde{F})$, this yields:

$$\left.\frac{\partial\left(2\Pi^{**}\right)}{\partial m}\right|_{m=0} = \frac{-q}{2} + \tilde{\alpha}_T q + \tilde{\gamma}\tilde{\alpha}_T\frac{\partial \tilde{F}^{**}}{\partial m}.$$

In addition, from $\tilde{F}^{**} = \tilde{F}^e\left(\tilde{f} - \frac{\hat{r}+v-\hat{v}}{2}, \frac{v-\hat{v}}{2}\right)$:

$$\left.\frac{\partial \tilde{F}^{**}}{\partial m}\right|_{m=0} = -\frac{\partial \tilde{F}^e}{\partial C}q + \frac{\partial \tilde{F}^e}{\partial \hat{C}}\frac{q}{2}.$$

Since $\tilde{F}^e\left(C,\hat{C}\right)$ is implicitly defined by

$$\left(\tilde{F}-C\right)D_1\left(\tilde{F},\tilde{F}\right) - \hat{C}D_2\left(\tilde{F},\tilde{F}\right) + D\left(\tilde{F},\tilde{F}\right) = 0,$$

we have:

$$\frac{\frac{\partial \tilde{F}^e}{\partial \hat{C}}}{\frac{\partial \tilde{F}^e}{\partial C}} = \frac{D_2}{D_1}\left(\tilde{F}^e, \tilde{F}^e\right) = -\gamma(\tilde{F}^e),$$

and thus:

$$\frac{1}{q}\left.\frac{\partial\left(2\Pi^{**}\right)}{\partial m}\right|_{m=0} = -\frac{1}{2} + \tilde{\alpha}_T\left(1 - \tilde{\gamma}\left(1+\frac{\tilde{\gamma}}{2}\right)\frac{\partial \tilde{F}^e}{\partial C}\right).$$

C Proof of corollary 1

Using $q = \bar{q}$ and $q' = 0$, we have:

$$\frac{\partial\left(2\Pi^{**}\right)}{\partial m} = -\frac{\bar{q}}{2} + \frac{\partial \tilde{F}^{**}}{\partial m}\left(\tilde{\alpha}_T^{**} + \tilde{\alpha}_T'\left(\tilde{F}^{**} - \tilde{f} + m\bar{q}\right)\right) + \tilde{\alpha}_T^{**}\bar{q},$$

where

$$\begin{aligned}
\tilde{F}^{**} - \tilde{f} &= -\frac{\tilde{\alpha}_T^{**}}{2D_1\left(\tilde{F}^{**}, \tilde{F}^{**}\right)} - \left(1+\frac{\gamma^{**}}{2}\right)m\bar{q},\\
\tilde{\alpha}_T' &= \left(1-\gamma^{**}\right)2D_1\left(\tilde{F}^{**}, \tilde{F}^{**}\right).
\end{aligned}$$

Thus:

$$\begin{aligned}
\frac{1}{\bar{q}}\frac{\partial\left(2\Pi^{**}\right)}{\partial m} &= -\frac{1}{2} + \tilde{\alpha}_T^{**}\\
&\quad + \frac{1}{\bar{q}}\frac{\partial \tilde{F}^{**}}{\partial m}\left(\tilde{\alpha}_T^{**} + \left(1-\gamma^{**}\right)2D_1\left(\tilde{F}^{**}, \tilde{F}^{**}\right)\left(-\frac{\tilde{\alpha}_T^{**}}{2D_1\left(\tilde{F}^{**}, \tilde{F}^{**}\right)} - \frac{\gamma^{**}}{2}m\bar{q}\right)\right)\\
&= -\frac{1}{2} + \tilde{\alpha}_T^{**} + \frac{1}{\bar{q}}\frac{\partial \tilde{F}^{**}}{\partial m}\left(\gamma^{**}\tilde{\alpha}_T^{**} - \left(1-\gamma^{**}\right)\gamma^{**}D_1\left(\tilde{F}^{**}, \tilde{F}^{**}\right)m\bar{q}\right).
\end{aligned}$$

Since

$$\frac{\partial \tilde{F}^{**}}{\partial m} = -\bar{q}\left(1+\frac{\gamma^{**}}{2}\right)\frac{\partial \tilde{F}^e}{\partial C}(.,.) < 0,$$

for $m < 0$ the previous expression is larger than:

$$-\frac{1}{2} + \tilde{\alpha}_T^{**}\left(1 + \frac{\gamma^{**}}{\bar{q}}\frac{\partial \tilde{F}^{**}}{\partial m}\right) = -\frac{1}{2} + \tilde{\alpha}_T^{**}\left(1 - \gamma^{**}\left(1+\frac{\gamma^{**}}{2}\right)\frac{\partial \tilde{F}^e}{\partial C}(.,.)\right).$$

D Proof of proposition 9

Using $p_i = \tilde{p}_i = c + m/2$ and $F_i = F^*(\theta)$, operator i's profit is equal to:

$$\begin{aligned}
\Pi_i &= \frac{1}{2}\left[(\tilde{\alpha}_i - \tilde{\alpha}_j)\frac{mq^*}{2} + F^*(\theta) - f\right] + \tilde{\alpha}_i\left[(\tilde{\alpha}_i - \tilde{\alpha}_j)\frac{m\theta q^*}{2} + \tilde{F}_i - \tilde{f}\right] \\
&\quad + \left(\frac{1}{2} + \tilde{\alpha}_i\right)\left(\frac{1}{2} + \tilde{\alpha}_j\theta\right) mq^* \\
&= \frac{1}{2}\left(F^*(\theta) - f + \frac{mq^*}{2}\right) + \tilde{\alpha}_i\left(\tilde{F}_i - \tilde{f} + \frac{3mq^*}{4}\right) - \tilde{\alpha}_j\frac{mq^*}{4} + \frac{\theta mq^*}{2}\left(\tilde{\alpha}_i^2 + \tilde{\alpha}_j + \tilde{\alpha}_i\tilde{\alpha}_j\right).
\end{aligned}$$

Optimizing with respect to $\tilde{F}_i$ then yields, at a symmetric equilibrium $\tilde{F}_i = \tilde{F}^*(\theta)$:

$$\frac{\partial\tilde{\alpha}_i}{\partial\tilde{F}_i}\left[(1 + \theta\tilde{\alpha}_T)\frac{3mq^*}{4} + \tilde{F}^*(\theta) - \tilde{f}\right] - \frac{\partial\tilde{\alpha}_j}{\partial\tilde{F}_i}\left(1 - \theta\left(2 + \tilde{\alpha}_T^*\right)\right)\frac{mq^*}{4} + \frac{\tilde{\alpha}_T^*}{2} = 0, \tag{23}$$

where $\tilde{\alpha}_T^* = \tilde{\alpha}_T\left(\tilde{F}^*(\theta)\right)$ and, using (19), the partial derivatives of the demand can derived from $\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}$:

$$\frac{\partial\tilde{\alpha}_i}{\partial\tilde{F}_i} = \left(1 - \theta v^*(1 - \gamma^*)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}\right)D_1 \text{ and } \frac{\partial\tilde{\alpha}_j}{\partial\tilde{F}_i} = -\left(\gamma^* + \theta v^*(1 - \gamma^*)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}\right)D_1,$$

where $\gamma^* = \gamma\left(\tilde{F}^*(\theta)\right)$. Using these expressions and dividing (23) by $\frac{\partial\tilde{\alpha}_i}{\partial\tilde{F}_i}$ yields:

$$\begin{aligned}
&(1 + \theta\tilde{\alpha}_T^*)\frac{3mq^*}{4} + \tilde{F}^*(\theta) - \tilde{f} + \left(1 - \theta\left(2 + \tilde{\alpha}_T^*\right)\right)\frac{mq^*}{4}\frac{\left(\gamma^* + \theta v^*(1 - \gamma^*)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}\right)}{\left(1 - \theta v^*(1 - \gamma^*)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}\right)} \\
&= -\frac{D}{D_1}\frac{1}{1 - \theta v^*(1 - \gamma^*)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}} \\
&= -\frac{D}{D_1} - \frac{D}{D_1}\frac{\theta v^*(1 - \gamma^*)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}}{1 - \theta v^*(1 - \gamma^*)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}},
\end{aligned}$$

which can be rearranged as:

$$\begin{aligned}
-\frac{D}{D_1} &= (1 + \theta\tilde{\alpha}_T^*)\frac{3mq^*}{4} + \tilde{F}^*(\theta) - \tilde{f} \\
&\quad + \left(1 - \theta\left(2 + \tilde{\alpha}_T^*\right)\right)\frac{mq^*}{4}\frac{\left(\gamma^* + \theta v^*(1 - \gamma^*)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}\right)}{\left(1 - \theta v^*(1 - \gamma^*)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}\right)} + \frac{D}{D_1}\frac{\theta v^*(1 - \gamma^*)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}}{1 - \theta v^*(1 - \gamma^*)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}},
\end{aligned}$$

or:

$$
\begin{aligned}
-\frac{D}{D_1} &= \tilde{F}^*(\theta) - \theta\left(1+\tilde{\alpha}_T^*\right)v^* - \left(\tilde{f} - \frac{3mq^*}{4}\right) + \gamma^*\frac{mq^*}{4} \\
&\quad + \left(\left(1-\theta\left(2+\tilde{\alpha}_T^*\right)\right)\frac{\left(\gamma^* + \theta v^*\left(1-\gamma^*\right)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}\right)}{\left(1-\theta v^*\left(1-\gamma^*\right)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}\right)} - \gamma^*\right)\frac{mq^*}{4} \\
&\quad + \theta\left(1+\tilde{\alpha}_T^*\right)v^* + \theta\tilde{\alpha}_T^*\frac{3mq^*}{4} + \frac{D}{D_1}\frac{\theta v^*\left(1-\gamma^*\right)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}}{1-\theta v^*\left(1-\gamma^*\right)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}} \\
&= \tilde{F}^*(\theta) - \theta\left(1+\tilde{\alpha}_T^*\right)v^* - \left(\tilde{f} - \frac{3mq^*}{4}\right) + \gamma^*\frac{mq^*}{4} + \theta\Xi,
\end{aligned}
$$

where:

$$
\begin{aligned}
\Xi &= \frac{\left(1+\gamma^*\right)v^*\left(1-\gamma^*\right)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i} - \left(2+\tilde{\alpha}_T^*\right)\left(\gamma^* + \theta v^*\left(1-\gamma^*\right)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}\right)}{\left(1-\theta v^*\left(1-\gamma^*\right)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}\right)}\frac{mq^*}{4} \\
&\quad + \left(1+\tilde{\alpha}_T^*\right)v^* + \tilde{\alpha}_T^*\frac{3mq^*}{4} + \frac{D}{D_1}\frac{v^*\left(1-\gamma^*\right)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}}{1-\theta v^*\left(1-\gamma^*\right)\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i}}.
\end{aligned}
$$

Since $D(.)$ is bounded with bounded derivatives, $\tilde{\alpha}_T$ and

$$\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i} = \frac{(1-\gamma)D_1}{1+2\theta v^*(1-\gamma)D_1}$$

are well defined and uniformly bounded in a neighborhood of $\theta = 0$. Moreover $\frac{D}{D_1}\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i} = \frac{(1-\gamma^*)D}{1+2\theta v^*(1-\gamma^*)D_1}$ is bounded. Therefore Ξ is bounded in such a neighborhood. Since $\theta\Xi$ converges to 0, assumption B implies that light users' net surplus, $\tilde{w}^*(\theta) = -\left(\tilde{F}^*(\theta) - \theta\left(1+\tilde{\alpha}_T^*(\theta)\right)v^*\right)$, converges to $-\tilde{F}^*$. The participation levels thus converge. Likewise, light users' volume of calls, $\left(1+\tilde{\alpha}_T^*\right)\theta q^*$, converges to zero.

Finally, light users' net surplus, $\tilde{w}^*(\theta)$, satisfies:

$$-\tilde{w}^*(\theta) = \tilde{F}^*(\theta) - \theta\left(1+\tilde{\alpha}_T^*\right)v^* = \tilde{F}^e\left(\tilde{f} - \frac{3}{4}mq^* - \theta\Xi, \frac{mq^*}{4}\right).$$

Compared to the benchmark case $\theta = 0$, the sign of Ξ determines whether the participation level of light users increases or decreases with their demand for calls. While in general the sign of Ξ is ambiguous, for $\theta = 0$ we have:

$$\frac{\partial\tilde{\alpha}_T}{\partial\tilde{F}_i} = \left(1-\gamma^*\right)D_1,$$

and thus:

$$\begin{aligned}\Xi &= \left(v^*\left(1-\gamma^*\right)^2\left(1+\gamma^*\right)D_1+3-\left(2+\tilde{\alpha}_T^*\right)\gamma^*\right)\frac{mq^*}{4}\\&\quad+\frac{\tilde{\alpha}_T^*}{2}\left(v^*\left(1-\gamma^*\right)^2\right)+\left(1+\tilde{\alpha}_T^*\right)v^*,\end{aligned}$$

which is positive for m small enough.