

Mobile call termination revisited

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

April 2009

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 price discrimination, 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.

*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 European markets as markets with CPP and non-linear tariffs, and some but limited extent of on-net / off-net price discrimination.

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 2009 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. One key to understanding the effect of termination on retail prices is the so-called waterbed effect: the profit that a customer may generate on fixed-to-mobile (FTM) or mobile-to-mobile termination² will be at least partially competed away through retail competition, since mobile operators will then fight more fiercely to attract customers. This can take the form of reduced subscription fees but could also translate into increased advertising, larger handset subsidies or reduced fees on particular

¹For 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.

²While the term is usually used for FTM termination, the effect is at work also in MTM.

services.³ 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 from other networks, and thus affect both calls emitted and received. As pointed 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 (2008) 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 discuss the unilateral setting of (possibly non-reciprocal) termination 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 in the large variety of post-pay contracts.⁴ Similarly, pre-pay contracts clients include occasional users as well as users who mainly receive calls.⁵ Overall, heterogeneity is a source of traffic imbalance at the customer level, meaning that 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 participation.⁶

Our model is based on the idea that the willingness to pay for subscription

³See Schiff (2008).

⁴Post-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.

⁵It is worth noting here that operators are willing to maintain the receiving service on pre-pay contracts even after the contracted volume of calls is exhausted.

⁶Dessein (2004) shows that the profit neutrality results break if the intensity of competition differs for different types of users. Polletti and Wright (2004) reach a similar conclusion by introducing a participation constraint on usage. Both papers however maintain the assumption of a fixed participation.

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. Heavy users have a larger demand for calls than light users; to simplify things we assume furthermore that light users actually only receive calls. Heavy and light users also differ in their demands for subscription; for the sake of exposition, we assume that the aggregate subscription demand of heavy users is inelastic, whereas that of light users is elastic. Formally the model is an extension of the model in Laffont, Rey and Tirole (1998a) (hereafter LRT) where we add the presence of light users (pure receivers) with elastic subscription demand.

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 allow for off-net / on-net discrimination, before introducing FTM termination as well. 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. They will therefore account for the fact that, if a heavy user leaves them for the competitor, this 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 of losing light users further intensifies competition for them.

We show that, in the absence of termination-based price discrimination, 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 / off-net price discrimination, because the market then exhibits tariff-mediated network effects: with positive termination markups, the off-net price is above the on-net price, and thus a customer is better off joining a larger network as this means more on-net calls. As pointed out by Laffont, Rey and Tirole (1998b) and Gans and King (2001), these network effects intensify competition. While these network effects directly affect heavy users only, 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 than for small 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 favour a positive FTM rate while they would be indifferent if they were no light users.

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. We sometimes refer to the first category of users as “callers”, by contrast with the second category, referred to as “pure receivers”. We assume that users derive a fixed utility from being called, independent of the volume of calls received, but variable among users. We moreover assume that the two types of users are sufficiently different that each network $i = 1, 2$ can perfectly discriminate between them by offering two contracts: a simple fixed fee P_i for light users (together with a high

usage price, say), and a two-part tariff for heavy users, which consists of a subscription fee F_i and a unit price for calls p_i .⁷

Networks are differentiated and face symmetric demands. More precisely, the demand from light users for network i is given by $\beta_i(P_1, P_2)$, where $\beta_1(P_1, P_2) = \beta_2(P_2, P_1)$. We will assume that demand for network i decreases with its own price ($\partial\beta_1/\partial P_1 < 0$) and increases with its rival's price ($\partial\beta_2/\partial P_1 < 0$), and that the aggregate demand, $\beta_T(P) \equiv 2\beta_i(P, P)$, is decreasing in P ; the "replacement ratio", $\gamma(P) \equiv -\frac{\partial\beta_2(P, P)}{\partial P_1} / \frac{\partial\beta_1(P, P)}{\partial P_1}$, is thus such that $0 \leq \gamma(P) < 1$.⁸

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 $1 + \beta_T$, subscribing to network i gives a user located at a distance x a net utility given by

$$u_0 + (1 + \beta_T)(u(q) - p_i q) - F_i - tx,$$

where t denotes the Hotelling differentiation parameter and u_0 denotes the fixed utility from receiving calls. The volume of calls is then given by

$$q_i = q(p_i) \equiv \arg \max_q (u(q) - p_i q),$$

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

$$w_i \equiv (1 + \beta_T)v(p_i) - F_i. \quad (1)$$

Operator i 's market share is then given by:

$$\alpha_i = \frac{1}{2} + \sigma(w_i - w_j), \quad (2)$$

where

$$\sigma \equiv \frac{1}{2t}$$

⁷If only second-degree price discrimination is feasible, tariffs must be incentive compatible: each category must prefer "its" tariff to the one designed for the other category. If for example heavy users are mainly interested in calling rather than receiving, and light users face high usage prices, these constraints boil down to $F_i \geq P_i$, which we assume to hold in what follows.

⁸The limit case $\gamma = 1$ would correspond to the case of full (fixed) participation.

represents the degree of substitution between the two networks.

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 φ 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.⁹ The terminating network thus receives the termination markup $m \equiv a - c_T$, while the originating network bears a cost $c + m$. 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, of the form (p_i, F_i) and P_i .

As usual, departing from cost-based termination charges (i.e., $m \neq 0$) introduces 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 $\gamma(\cdot)$ 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. $v(0)$ is bounded and the two networks are sufficiently differentiated (i.e., σ and $\gamma(\cdot)$ small enough) that there always exists a unique, pure strategy equilibrium, in which the two networks moreover share the market equally.¹⁰

⁹Negative termination charges could generate abuses.

¹⁰Laffont, Rey and Tirole (1998a) show that, in the case of homogenous users, a symmetric shared-market equilibrium exists when the access markup m is small enough and/or the differentiation parameter t is large enough. The argument can easily be extended here (as in Dessein (2003), who considers the case of implicit discrimination among heterogeneous 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.

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 $(p_i, F_i, P_i)_{i=1,2}$, market shares (α_1, α_2) in the heavy user segment and demands (β_1, β_2) from light users, network i 's profit is equal to, for $i \neq j = 1, 2$:

$$\begin{aligned} \Pi_i &= \alpha_i [(1 + \beta_T) (p_i - c) q(p_i) - (\alpha_j + \beta_j) m q(p_i) + F_i - f] \\ &\quad + (\alpha_i + \beta_i) \alpha_j m q(p_j) + \beta_i (P_i - \varphi). \end{aligned}$$

Optimizing with respect to the usage price p_i , while adjusting the fee F_i so as to keep constant the variable surplus $w_i = (1 + \beta_T) v(p_i) - F_i$, and thus the market shares, yields:

$$\begin{aligned} \left. \frac{\partial \Pi_i}{\partial p_i} \right|_{w_i, P_i} &= \alpha_i [(1 + \beta_T) ((p_i - c) q'(p_i) + q(p_i)) - (\alpha_j + \beta_j) m q'(p_i) - (1 + \beta_T) q(p_i)] \\ &= \alpha_i q'(p_i) [(1 + \beta_T) (p_i - c) - (\alpha_j + \beta_j) m], \end{aligned}$$

which, evaluated at a symmetric equilibrium ($\alpha_1 = \alpha_2 = 1/2, \beta_1 = \beta_2 = \beta_T/2$), leads to:

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

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:

$$\begin{aligned} \Pi_i &= \alpha_i \left[(\alpha_i + \beta_i) \frac{m q^*}{2} - (\alpha_j + \beta_j) \frac{m q^*}{2} + F_i - f \right] \\ &\quad + (\alpha_i + \beta_i) \alpha_j m q^* + \beta_i (P_i - \varphi), \end{aligned}$$

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:

$$\begin{aligned} \left. \frac{\partial \Pi_i}{\partial F_i} \right|_{\beta_1 = \beta_2 = \beta_T/2, F_1 = F_2 = F} &= -\sigma (F - f) + \frac{1}{2} [-\sigma m q^* + 1] + \left(-\sigma \frac{1}{2} + \sigma \frac{1 + \beta_T}{2} \right) m q^* \\ &= \frac{1}{2} - \sigma (F - f) + \sigma (\beta_T - 1) \frac{m q^*}{2}. \end{aligned}$$

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

$$F^* = f + \frac{1}{2\sigma} + (\beta_T - 1) \frac{mq^*}{2}, \quad (4)$$

and heavy users' net variable surplus is equal to:

$$(1 + \beta_T) v(p^*) - f - \frac{1}{2\sigma} - (\beta_T - 1) \frac{mq^*}{2}.$$

Condition 4 is similar to that obtained by Laffont, Rey and Tirole (1998a), except for the term in β_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 $\alpha_i(p^* - c)q^* = mq^*/4$, and the calls received from the rival network generate an access revenue equal to $\alpha_jmq^* = mq^*/2$. If the user subscribes instead to network j , his calls to network i 's heavy users generate a termination revenue of $\alpha_imq^* = mq^*/2$, 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 $\beta_imq^* = \beta_Tmq^*/2$. Attracting a heavy user away from the rival network thus generates an additional net loss equal to $\beta_Tmq^*/2$, which is again 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 the heavy users' prices to their equilibrium values (which yields $\alpha_1 = \alpha_2 = \alpha^* = 1/2$), network i 's profit writes as:

$$\begin{aligned} \Pi_i &= \frac{1}{2} \left[\left(\frac{1}{2} + \beta_i \right) \frac{mq^*}{2} - \left(\frac{1}{2} + \beta_j \right) \frac{mq^*}{2} + F^* - f \right] \\ &\quad + \left(\frac{1}{2} + \beta_i \right) \frac{1}{2} mq^* + \beta_i (P_i - \varphi) \\ &= \frac{F^* - f}{2} + (1 + 3\beta_i - \beta_j) \frac{mq^*}{4} + \beta_i (P_i - \varphi). \end{aligned}$$

Optimizing this profit with respect to (P_i, P_j) amounts to maximize

$$G(P_i, P_j) \equiv (P_i - C) \beta_i(P_1, P_2) - \hat{C} \beta_j(P_1, P_2),$$

where

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

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 $\alpha_jmq(p_j) = \frac{mq^*}{2}$ 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 $G(P_1, P_2)$ and $G(P_2, P_1)$ has a unique equilibrium, $P_1 = P_2 = P^D(C, \hat{C})$, which is symmetric and the unique solution to the first-order conditions.
2. This equilibrium price verifies

$$0 < \frac{\partial P^D}{\partial C} \leq 1, \quad \frac{\partial P^D}{\partial \hat{C}} \leq 0 < \frac{\partial P^D}{\partial C} + \frac{\partial P^D}{\partial \hat{C}}.$$

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 amounts to say that an increase in the direct cost of the operators is at least partially passed through to users.¹¹ 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 characterized by the first-order conditions:

$$\left. \frac{\partial \Pi_i}{\partial P_i} \right|_{p_1=p_2=p^*, F_1=F_2=F^*} = \frac{3mq^*}{4} \frac{\partial \beta_i}{\partial P_i} - \frac{mq^*}{4} \frac{\partial \beta_j}{\partial P_i} + \beta_i + (P_i - \varphi) \frac{\partial \beta_i}{\partial P_i} = 0$$

¹¹In standard 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.

which, evaluated at a symmetric equilibrium, yields:

$$\frac{P^* - \left(\varphi - \frac{3+\gamma(P^*)}{4}mq^* \right)}{P^*} = \frac{1}{\varepsilon(P^*)}, \quad (5)$$

where $\varepsilon(P) \equiv -P \frac{\partial \beta_i(P,P)}{\partial P_i} / \beta_i(P,P)$ denotes the own price elasticity of the demand from light users.

4 Choosing the termination rate

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

$$\Pi^* = \frac{1}{4\sigma} + \frac{\beta_T(P^*)}{2} (P^* - \varphi + mq^*) \quad (6)$$

The term $\frac{\beta_T}{2}mq^*$ 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 $\frac{\beta_T}{4}mq^*$. Moreover, light users generate a termination revenue $\frac{\beta_T}{4}mq^*$ (which is however partially granted back through a lower price P^*).

The equilibrium price P^* defined by (5) depends on the termination markup m only through the access revenue $r(m) \equiv mq^* = mq(c + m/2)$; we have moreover:

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

Proof. Using Assumption B.1, $P^* = P^D(\varphi - \frac{3r}{4}, \frac{r}{4})$. The conclusion then follows from B.2, since:

$$\frac{\partial P^*}{\partial r} = -\frac{3}{4} \frac{\partial P^D}{\partial C} + \frac{1}{4} \frac{\partial P^D}{\partial \hat{C}},$$

where

$$-1 \leq -\frac{\partial P^D}{\partial C} < \frac{\partial P^D}{\partial \hat{C}} \leq 0.$$

■

Since P^* depends on m solely through the revenue r , the profit Π^* given by (6) can also be expressed as a function of r :

$$\Pi^* = \frac{1}{4\sigma} + \frac{\beta_T(P^*)}{2} (P^* - \varphi + r). \quad (7)$$

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

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

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:

$$m^R \equiv \arg \max_m m q \left(c + \frac{m}{2} \right),$$

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

Proposition 1 *For any $m > m^R$ such that $\Pi(m) > 0$, there exists $\tilde{m} < m^R$ that Pareto dominates m .*

Proof. Take any candidate $m > m^R$ such that $\Pi(m) > 0$. Since $r(0) = 0$ and $r(\cdot)$ 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.¹² 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 .*

¹²If the monopoly rate m^R is not uniquely defined, then the same applies to its lowest value, $\hat{m} \equiv \min \left\{ \hat{M} \equiv \arg \max_m m q (c + m/2) \right\}$. That value Pareto dominates any other monopoly value in \hat{M} , and the reasoning of the above proof applies to any other rate m above of \hat{m} .

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

$$\frac{\partial (2\Pi^*)}{\partial m} = \left[\beta'_T \frac{\partial P^*}{\partial r} (P^* - \varphi + r) + \beta_T \left(\frac{\partial P^*}{\partial r} + 1 \right) \right] \frac{\partial r}{\partial m} \quad (8)$$

Consider first the case where $m \geq 0$ (and thus $r \geq 0$). Since $\gamma^* = \gamma(P^*) < 1$, (5) then yields:

$$P^* - \varphi + r \geq P^* - \varphi + \frac{3 + \gamma^*}{4} r > 0.$$

Since $\beta'_T < 0$ and $-1 < \frac{\partial P^*}{\partial r} < 0$, the bracket term in (8) is therefore strictly positive. Hence, in the range $m \geq 0$, the profits are maximal for $m = m^R$.

Now consider the case $r < 0$. A similar reasoning applies as long $P^* - \varphi + r \geq 0$, since $\frac{\partial r}{\partial m} = \frac{mq'}{2} + q > 0$ for $m < 0$. If instead $P^* - \varphi + r < 0$, then from (7) the equilibrium profit is lower than $\frac{1}{4\sigma}$; m is therefore dominated by a zero markup ($r = 0$), for which $P^* - \varphi + r = P^* - \varphi > 0$ from (5), and thus $\Pi^* > \frac{1}{4\sigma}$. Therefore, the profit maximizing termination markup is m^R . ■

Let us now turn to users. The surplus of light users is of the form $S^L(P^*, P^*)$, where $S^L(P_1, P_2)$ is such that $\frac{\partial S^L}{\partial P_i} = -\beta_i$. Therefore:

$$\frac{\partial S^L}{\partial m} = \left(\frac{\partial S^L}{\partial P_1} \frac{\partial P^*}{\partial r} + \frac{\partial S^L}{\partial P_2} \frac{\partial P^*}{\partial r} \right) \frac{\partial r}{\partial m} = -\beta_T \frac{\partial P^*}{\partial r} \frac{\partial r}{\partial m}.$$

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

$$\left. \frac{\partial S^H}{\partial m} \right|_{m=0} = \beta_T(P^*) v(c) \left(\frac{\beta'_T}{\beta_T}(P^*) \frac{\partial P^*}{\partial m} + \frac{v'(c)}{v(c)} \right).$$

Therefore:

Proposition 3 *For m small, increasing m raises heavy users' surplus if the participation of light users is very elastic or if the 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 participation of light users is sufficiently elastic.

Total welfare can be written as follows:

$$W^* = \left[(1 + \beta_T^*) \left(v^* + \frac{mq^*}{2} \right) - f - \frac{t}{4} \right] + [S^L(P^*, P^*) + \beta_T^*(P^* - \varphi)] \quad (9)$$

The first term into 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 $\frac{\partial S^L}{\partial m} = -\beta_T^* \frac{\partial P^*}{\partial m}$, we have:

$$\frac{\partial W^*}{\partial m} = (1 + \beta_T^*) \frac{mq'(p^*)}{4} + \left(v^* + \frac{mq^*}{2} + P^* - \varphi \right) \beta_T'(P^*) \frac{\partial P^*}{\partial m}. \quad (10)$$

Consider first the case $m \leq 0$. We then have $\beta_T'(P^*) \frac{\partial P^*}{\partial m} > 0$, since $\frac{\partial P^*}{\partial m} = \frac{\partial P^*}{\partial r} \left(q^* + \frac{mq'(p^*)}{2} \right) < 0$. In addition, (5) yields $P^* > \varphi - \frac{3+\gamma^*}{4}mq^*$, and thus:

$$\frac{\partial W^*}{\partial m} > (1 + \beta_T^*) \frac{mq'(p^*)}{4} + \left(v^* - \frac{1 + \gamma^*}{4}mq^* \right) \beta_T'(P^*) \frac{\partial P^*}{\partial m}.$$

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 $]0, m^R]$. To conclude the proof, it suffices to note that, at $m = m^R (> 0)$, $\frac{\partial P^*}{\partial m} = \frac{\partial P^*}{\partial r} \frac{\partial r}{\partial m} = 0$ and thus:

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

■

Therefore, the presence of light users, who furthermore have an elastic participation,¹³ 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 by raising the termination markup to increase their calling opportunities.

¹³In 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 $m = 0$.

5 Price Discrimination

We now allow networks to discriminate between on-net and off-net calls. We keep the same notations 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. With this termination-based price discrimination, network i 's profit writes as, for $i \neq j = 1, 2$:

$$\begin{aligned} \Pi_i = & \alpha_i [(\alpha_i + \beta_i)(p_i - c)q(p_i) + (\alpha_j + \beta_j)(\hat{p}_i - c - m)q(\hat{p}_i) + F_i - f] \\ & + (\alpha_i + \beta_i)\alpha_j m q(\hat{p}_j) + \beta_i(P_i - \varphi), \end{aligned}$$

where:

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

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

$$\begin{aligned} \Pi_1 = & \alpha_1 [(\alpha_1 + \beta_1)((p_1 - c)q(p_1) + v(p_1)) \\ & (\alpha_j + \beta_j)((\hat{p}_1 - c - m)q(\hat{p}_1) + v(\hat{p}_1)) - w_1 - f] \\ & + (\alpha_1 + \beta_1)\alpha_j m q(\hat{p}_j) + \beta_1(P_1 - \varphi) \end{aligned}$$

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 $\hat{v} = v(c + m)$, the profit of network 1, say, can be written as:

$$\Pi_1 = \alpha_1(F_1 - f) + \alpha_2(\alpha_1 + \beta_1)m\hat{q} + \beta_1(P_1 - \varphi), \quad (11)$$

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

$$\begin{aligned} \alpha_i &= \frac{1}{2} + \sigma(w_i - w_j) \\ &= \frac{1}{2} + \sigma[(2\alpha_i - 1 + \beta_i - \beta_j)(v - \hat{v}) - (F_i - F_j)], \end{aligned} \quad (12)$$

and thus:

$$\alpha_i - \frac{1}{2} = \sigma \frac{(\beta_i - \beta_j)(v - \hat{v}) - (F_i - F_j)}{1 - 2\sigma(v - \hat{v})}.$$

Differentiating (11) with respect to F_1 then yields, at a symmetric equilibrium:

$$\frac{\partial \Pi_1}{\partial F_1} \Big|_{\beta_1=\beta_2=\hat{\beta}_T/2, F_1=F_2=\hat{F}} = \frac{1}{2} - \frac{\sigma}{1 - 2\sigma(v - \hat{v})} \left(F_1 - f - \frac{1 + \hat{\beta}_T}{2} m\hat{q} + \frac{m\hat{q}}{2} \right),$$

which leads to:

$$F_1 = F_2 = \hat{F} = f + \frac{1}{2\sigma} + \hat{\beta}_T \frac{m\hat{q}}{2} - (v - \hat{v}).$$

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 discrimination, 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, as in the absence of discrimination, his off-net calls generate an access revenue $(1 + \hat{\beta}_T) \frac{m\hat{q}}{2}$ 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 $\frac{m\hat{q}}{2}$ compared to the no discrimination case, 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), price discrimination 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.

Price discrimination thus generate two conflicting effects. On the one hand, the opportunity cost of losing a caller 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 price-discrimination therefore benefits heavy users:

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

Proof. The equilibrium fixed fees without and with price discrimination are respectively defined by:

$$F^* = f + \frac{1}{2\sigma} + \frac{\beta_T^* - 1}{2}mq^*, \hat{F} = f + \frac{1}{2\sigma} + \frac{\hat{\beta}_T}{2}m\hat{q} - (v - \hat{v}).$$

Therefore:

$$F - \hat{F} = \frac{\beta_T^* - 1}{2}mq^* - \frac{\hat{\beta}_T}{2}m\hat{q} + v - \hat{v}.$$

At $m = 0$, $F^* = \hat{F}$, $p^* = \hat{p} = c$ and $\beta_T^* = \hat{\beta}_T = \beta_T$, which implies:

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

■

We now characterize the equilibrium subscription price P for light users. Setting $F_1 = F_2 = \hat{F}$ in (12) shows how the prices P_1 and P_2 , by determining the participation of light users, affect the market shares for heavy users as well; differentiating (11) with respect to P_1 then yields, at a symmetric equilibrium:

$$\begin{aligned} \frac{\partial \Pi_1}{\partial P_1} &= \frac{\sigma(v - \hat{v})}{1 - 2\sigma(v - \hat{v})} \left(\frac{\partial \beta_1}{\partial P_1} - \frac{\partial \beta_2}{\partial P_1} \right) \left(\hat{F} - f + (\alpha_2 - \beta_1 - \alpha_1)m\hat{q} \right) \\ &\quad + \frac{\partial \beta_1}{\partial P_1} \frac{m\hat{q}}{2} + \frac{\partial \beta_1}{\partial P_1} (P_1 - \varphi) + \beta_1 \end{aligned}$$

Using $F_1 = \hat{F}$ and evaluating the corresponding first-order condition at a symmetric equilibrium $P_1 = P_2 = \hat{P}$ then yields, using $\hat{r} = m\hat{q}$:

$$\frac{\hat{P} - \left(\varphi - \frac{(1 + \gamma(\hat{P}))(v - \hat{v}) + m\hat{q}}{2} \right)}{\hat{P}} = \frac{1}{\varepsilon(\hat{P})}. \quad (13)$$

\hat{P} satisfies the first-order condition of the above duopoly game – namely, $(P - C + \gamma(P)\hat{C})/P = 1/\varepsilon(P)$ – for $C = \varphi - \frac{\hat{r} + v - \hat{v}}{2}$ and $\hat{C} = \frac{v - \hat{v}}{2}$; therefore, from Assumption B.1:

$$\hat{P} = P^D \left(\varphi - \frac{\hat{r} + v - \hat{v}}{2}, \frac{v - \hat{v}}{2} \right).$$

Assumption B.2 then ensures that \hat{P} 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 price discrimination on the price offered to light users:

Proposition 6 *For termination rates close to the marginal cost, price discrimination 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 $P^* = P^D(\varphi - \frac{3r^*}{4}, \frac{r^*}{4})$. Therefore, $P^* = \hat{P}$ for $m = 0$ and:

$$\begin{aligned} \left. \frac{\partial (P^* - \hat{P})}{\partial m} \right|_{m=0} &= \frac{\partial P^D}{\partial C} \left(-\frac{3q(c)}{4} + q(c) \right) + \frac{\partial P^D}{\partial \hat{C}} \left(\frac{q(c)}{4} - \frac{q(c)}{2} \right) \\ &= \left(\frac{\partial P^D}{\partial C} - \frac{\partial P^D}{\partial \hat{C}} \right) \frac{q(c)}{4} > 0. \end{aligned}$$

■

Price discrimination 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 writes as:

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

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

Proposition 7 *Under price discrimination, the total welfare maximizing termination markup is positive.*

Proof. We have

$$\frac{\partial \hat{W}}{\partial m} = \left(1 + \hat{\beta}_T \right) \frac{mq'(\hat{p})}{2} + \frac{\partial \hat{\beta}_T}{\partial m} \left(\frac{v + \hat{v} + m\hat{q}}{2} + \hat{P} - \varphi \right).$$

But, defining $\hat{\gamma} = \gamma(\hat{P})$:

$$\frac{v + \hat{v} + m\hat{q}}{2} + \hat{P} - \varphi = \frac{\hat{P}}{\varepsilon(\hat{P})} + \hat{v} - \hat{\gamma} \frac{v - \hat{v}}{2}.$$

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

Unsurprisingly, the result on profit is more ambiguous. Indeed, while the competition weakening effect described in the case of no-discrimination is still present, it is now mitigated by the impact of tariff-mediated network effects. Total profit writes as:

$$2\hat{\Pi} = \left(\frac{1}{2\sigma} + \frac{1}{2}\hat{r} + \hat{v} - v \right) + \hat{\beta}_T \left(\hat{P} - \varphi + \hat{r} \right).$$

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 $m = 0$,*

$$\hat{\beta}_T \left(1 + \left(1 + \frac{\hat{\gamma}}{2} \right) \frac{\partial P^D}{\partial \hat{C}} \right) > \frac{1}{2},$$

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 given by, for some $\bar{p} > c$:

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

Then, as long as $c + m < \bar{p}$:

$$2\hat{\Pi} = \frac{1}{2\sigma} - \frac{m\bar{q}}{2} + \hat{\beta}_T \left(\hat{P} - \varphi + m\bar{q} \right).$$

In that case we have:

Corollary 1 *Suppose that individual usage is inelastic at \bar{q} ; then if*

$$\hat{\beta}_T \left(1 + \left(1 + \frac{\hat{\gamma}}{2} \right) \frac{\partial P^D}{\partial \hat{C}} \right) > \frac{1}{2}$$

for any $m < 0$, the profit maximizing termination charge is above cost.

Proof. See the appendix. ■

The condition states that provided that the duopoly price is not too much affected by the opportunity cost \hat{C} of letting a light user join the competing network (that is, $\frac{\partial P^D}{\partial \hat{C}}$ not too negative), the profit maximizing termination margin is positive when the population of light users is large enough.¹⁴ Note that, since scaling the demands β_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 ($\frac{\partial \beta_2}{\partial P_1} = 0$, which implies $\frac{\partial P^D}{\partial \hat{C}} = \gamma = 0$), 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 there is a large pass-through rate ($\frac{\partial P^D}{\partial \hat{C}}$ close to 1).

6 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. Let μ denote the FTM termination markup and $Q(\mu)$ represent the volume of calls that a mobile customer receives from fixed networks. 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.¹⁵ Thus, mobile users' participation and usage are unaffected by the presence of fixed networks. The only difference is that now the network receives an additional FTM termination revenue $\rho = \mu Q(\mu)$, which 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 φ . In the absence

¹⁴The term $1 + \frac{\partial P^D}{\partial \hat{C}} \left(1 - \frac{1}{2} \left(\frac{\partial P^D}{\partial \hat{C}} / \frac{\partial P^D}{\partial C}\right)\right)$ varies from 1 to $-\frac{1}{2}$. It is positive for instance if $\frac{\partial P^D}{\partial \hat{C}} > 1 - \sqrt{3}$ or $\frac{\partial P^D}{\partial \hat{C}} < \frac{2}{3}$.

¹⁵In Europe, the termination charge of fixed networks is regulated and 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 the case in practice, the two mobile operators 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.

of on-net / off-net price discrimination, the equilibrium prices will thus be:

$$\begin{aligned} p^* &= c + \frac{m}{2}, \\ F^* &= f + \frac{1}{2\sigma} + (\beta_T - 1) \frac{r^*}{2} - \rho, \\ P^* &= P^D \left(\varphi - \rho - \frac{3r^*}{4}, \frac{r^*}{4} \right). \end{aligned}$$

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 $\beta_T = 0$. 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:

$$\frac{\partial P^*}{\partial \rho} = -\frac{\partial P^D}{\partial C} < 0.$$

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

$$\frac{\partial F^*}{\partial \rho} = \frac{\beta'_T}{2} \frac{\partial P^*}{\partial \rho} - 1 > -1.$$

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 9 *Increasing the FTM per customer revenue ρ reduces more the subscription fee for heavy users than for light users if and only if $\frac{\partial P^D}{\partial C} < 1 / \left(1 - \frac{\beta'_T}{2}\right)$.*

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

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 φ with $\varphi - \rho$ in the previous expressions):

$$2\Pi^* = \frac{1}{2\sigma} + \beta_T^* (P^* - \varphi + \rho + r^*).$$

Thus:

Proposition 10 *Whenever $\beta_T^* > 0$, as long as $m \geq 0$ (or more generally if $P^* > \varphi - \rho - r$), an increase in FTM termination revenue ρ raises the equilibrium profit (while it has no effect on profit without light users).*

Proof. We have:

$$\frac{\partial (2\Pi^*)}{\partial \rho} = \beta_T^* + [\beta_T^* + \beta_T' (P^* - \varphi + \rho + r^*)] \frac{\partial P^*}{\partial \rho}$$

Since $\frac{\partial P^*}{\partial \rho} < 0$, this is positive if $\beta_T^* + \beta_T' (P^* - \varphi + \rho + r^*) < 0$. Suppose instead that $\beta_T^* + \beta_T' (P^* - \varphi + \rho + r^*) > 0$. Since $\frac{\partial P}{\partial \rho} > -1$, we then have:

$$\frac{\partial (2\Pi^*)}{\partial \rho} > -\beta_T' (P^* - \varphi + \rho + r^*),$$

But from (5), adapted by replacing φ with $\varphi - \rho$:

$$P^* - \varphi + \rho + r^* = -\frac{\beta_T^*}{2 \frac{\partial \beta_1}{\partial P_1}} + (1 - \gamma^*) \frac{r^*}{4},$$

and is thus positive when $m \geq 0$. ■

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

7 Conclusion

This paper proposes an explanation of the reluctance of mobile operators 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 a fixed participation and full pass-through rates. Using a framework based on the heterogeneity of users, we show that when the elasticity of participation 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 termination-based price discrimination, 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 of obtaining better empirical facts on the composition of the demand for mobile services and participation elasticities of 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 (2008), “Mobile Call Termination”, mimeo MPRA.
- [4] Dessein, W. (2003), “Network Competition in Nonlinear Pricing”, *Rand Journal of Economics*, Vol 34 (4), 593-611.
- [5] Dessein, W. (2004), “Network Competition with Heterogeneous Customers and Calling Patterns”, *Information Economics and Policy*, Vol 16, 323-345.
- [6] Gans, J., and S. King (2001), “Using ‘Bill and Keep’ Interconnect Agreements to Soften Network Competition”, *Economics Letters*, 71(3), 413—420.
- [7] Genakos, C., and T. Valletti (2007), “Testing the ”Waterbed” Effect in Mobile Telephony”, CEP Discussion Papers dp0827.
- [8] Hahn, J.-H. (2004), “Network competition and interconnection with heterogeneous subscribers”, *International Journal of Industrial Organization*, vol. 22(5), 611-631.
- [9] Laffont J.J., Rey P. and J. Tirole (1998a), “Network Competition: I. Overview and Nondiscriminatory pricing”, *Rand Journal of Economics*, 29(1), 1-37.
- [10] Laffont J.J., Rey P. and J. Tirole (1998b), “Network Competition: II. Price Discrimination”, *Rand Journal of Economics*, 29(1), 37-56.
- [11] Lopez A. and P. Rey (2008), “Foreclosing Competition through Access Charges and Price Discrimination”, mimeo IDEI.
- [12] Poletti S. and J. Wright (2004), “Network interconnection with participation constraints”, *Information Economics and Policy*, vol. 16(3), 347-373.
- [13] Schiff, A. (2008), “The Waterbed Effect and Price Regulation”, *Review of Network Economics*, 7(3), 392-414.

Proof of proposition 3.

$$\frac{\partial S^H}{\partial m} = -(1 + \beta_T) \frac{q^*}{2} + v^* \beta'_T \frac{\partial P^*}{\partial m} + (1 - \beta_T) \frac{q^* + mq'(p^*)/2}{2} - \frac{mq^*}{2} \beta'_T \frac{\partial P^*}{\partial m}.$$

Therefore, at $m = 0$:

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

The conclusion follows using $v'(c) = -q(c)$. ■

Proof of proposition 8.

$$\frac{\partial (2\hat{\Pi})}{\partial m} = \frac{1}{2} (mq'(c+m) - \hat{q}) + \hat{\beta}'_T \frac{\partial \hat{P}}{\partial m} (\hat{P} - \varphi + \hat{r}) + \hat{\beta}_T \left(\frac{\partial \hat{P}}{\partial m} + mq'(c+m) + \hat{q} \right),$$

which evaluated at $m = 0$, yields:

$$\left. \frac{\partial (2\hat{\Pi})}{\partial m} \right|_{m=0} = \frac{-q}{2} + \hat{\beta}'_T \frac{\partial \hat{P}}{\partial m} (\hat{P} - \varphi) + \hat{\beta}_T \left(\frac{\partial \hat{P}}{\partial m} + q \right),$$

where $q = q(c)$, $\hat{\beta}'_T = 2(1 - \gamma) \frac{\partial \beta_i}{\partial P_i}$ and, from (13) at $m = 0$:

$$\hat{P} - \varphi = \frac{\hat{P}}{\varepsilon(\hat{P})} = -\frac{\hat{\beta}_T}{2 \frac{\partial \beta_i}{\partial P_i}}.$$

In addition, from $\hat{P} = P^D(\varphi - \frac{\hat{r} + v - \hat{v}}{2}, \frac{v - \hat{v}}{2})$:

$$\left. \frac{\partial \hat{P}}{\partial m} \right|_{m=0} = -\frac{\partial P^D}{\partial C} q + \frac{\partial P^D}{\partial \hat{C}} \frac{q}{2}.$$

Hence:

$$\begin{aligned} \left. \frac{1}{q} \frac{\partial (2\hat{\Pi})}{\partial m} \right|_{m=0} &= -\frac{1}{2} + 2(1 - \gamma) \frac{\partial \beta_i}{\partial P_i} \left(-\frac{\partial P^D}{\partial C} + \frac{1}{2} \frac{\partial P^D}{\partial \hat{C}} \right) \left(-\frac{\hat{\beta}_T}{2 \frac{\partial \beta_i}{\partial P_i}} \right) \\ &\quad + \hat{\beta}_T \left(\left(-\frac{\partial P^D}{\partial C} + \frac{1}{2} \frac{\partial P^D}{\partial \hat{C}} \right) + 1 \right) \\ &= -\frac{1}{2} - \hat{\beta}_T (1 - \gamma) \left(-\frac{\partial P^D}{\partial C} + \frac{1}{2} \frac{\partial P^D}{\partial \hat{C}} \right) + \hat{\beta}_T \left(\left(-\frac{\partial P^D}{\partial C} + \frac{1}{2} \frac{\partial P^D}{\partial \hat{C}} \right) + 1 \right) \\ &= -\frac{1}{2} + \hat{\beta}_T \left(1 - \gamma \left(\frac{\partial P^D}{\partial C} - \frac{1}{2} \frac{\partial P^D}{\partial \hat{C}} \right) \right). \end{aligned}$$

Finally, since $P^D(C, \hat{C})$ is implicitly defined by

$$(P - C) \frac{\partial \beta_1}{\partial P_1}(P, P) - \hat{C} \frac{\partial \beta_2}{\partial P_1}(P, P) + \beta_1(P, P) = 0,$$

we have:

$$\frac{\frac{\partial P^D}{\partial \hat{C}}}{\frac{\partial P^D}{\partial C}} = \frac{\frac{\partial \beta_2}{\partial P_1}}{\frac{\partial \beta_1}{\partial P_1}} = -\gamma.$$

Therefore:

$$\frac{1}{q} \frac{\partial (2\hat{\Pi})}{\partial m} \Big|_{m=0} = -\frac{1}{2} + \hat{\beta}_T \left(1 + \left(1 + \frac{\gamma}{2} \right) \frac{\partial P^D}{\partial \hat{C}} \right).$$

■

Proof of corollary 1.

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

$$\frac{\partial (2\hat{\Pi})}{\partial m} = -\frac{\bar{q}}{2} + \hat{\beta}'_T \frac{\partial \hat{P}}{\partial m} (\hat{P} - \varphi + m\bar{q}) + \hat{\beta}_T \left(\frac{\partial \hat{P}}{\partial m} + \bar{q} \right),$$

Using

$$\begin{aligned} \hat{P} - \varphi &= -\frac{\beta_T}{2 \frac{\partial \beta_i}{\partial P_i}} - \left(1 + \frac{\hat{\gamma}}{2} \right) m\bar{q}, \\ \frac{\partial \hat{P}}{\partial m} &= -\frac{\partial P^D}{\partial C} \bar{q} + \frac{\partial P^D}{\partial \hat{C}} \frac{\bar{q}}{2}, \\ \hat{\beta}'_T &= 2(1 - \gamma) \frac{\partial \beta_i}{\partial P_i} \end{aligned}$$

we obtain:

$$\begin{aligned} \frac{1}{\bar{q}} \frac{\partial (2\hat{\Pi})}{\partial m} &= -\frac{1}{2} + \hat{\beta}'_T \left(-\frac{\partial P^D}{\partial C} + \frac{1}{2} \frac{\partial P^D}{\partial \hat{C}} \right) \left(-\frac{\beta_T}{2 \frac{\partial \beta_i}{\partial P_i}} - \left(1 + \frac{\hat{\gamma}}{2} \right) m\bar{q} + m\bar{q} \right) \\ &\quad + \hat{\beta}_T \left(\left(-\frac{\partial P^D}{\partial C} + \frac{1}{2} \frac{\partial P^D}{\partial \hat{C}} \right) + 1 \right) \\ &= -\frac{1}{2} + \hat{\beta}_T - \left(1 + \frac{\gamma}{2} \right) \frac{\partial P^D}{\partial C} \left(\hat{\beta}_T \left(1 - \frac{\hat{\beta}'_T}{2 \frac{\partial \beta_i}{\partial P_i}} \right) - \frac{\hat{\gamma}}{2} \hat{\beta}'_T m\bar{q} \right). \end{aligned}$$

For $m < 0$ this is larger than:

$$\begin{aligned}
& -\frac{1}{2} + \hat{\beta}_T - \left(1 + \frac{\gamma}{2}\right) \frac{\partial P^D}{\partial C} \left(\hat{\beta}_T \left(1 - \frac{2(1-\gamma) \frac{\partial \beta_i}{\partial P_i}}{2 \frac{\partial \beta_i}{\partial P_i}}\right) \right) \\
&= -\frac{1}{2} + \hat{\beta}_T - \hat{\beta}_T \left(1 + \frac{\gamma}{2}\right) \frac{\partial P^D}{\partial C} \gamma \\
&= -\frac{1}{2} + \hat{\beta}_T \left(1 - \left(1 + \frac{\gamma}{2}\right) \frac{\partial P^D}{\partial \hat{C}}\right).
\end{aligned}$$

■