Using on-net / off-net price differential to measure the size of call externalities and its implications for setting efficient MTRs

Jonathan Sandbach,[1] Luke van Hooft [2] [3]

ABSTRACT

This paper discusses the significance of call externalities on efficient mobile termination pricing in the two-sided mobile market. It shows by intuitive argument that call externalities should be largely internalised by subscribers and hence should not influence efficient platform pricing. This proposition is then tested by confronting the economic theory of on-net/off-net price discrimination with observed pricing behaviour of mobile network operators (MNOs). Economic theory suggests, to the extent that there exists a positive residual call externality (the un-internalised benefit accruing to subscribers from the calls they receive), MNOs will price on-net calls below marginal costs and off-net calls to competing networks above marginal costs. This allows the perceived strength of the residual call externality to be observed from MNO tariffs. If this theory is correct, we confirm that the residual call externality is small and that below-cost termination rates are not welfare maximising.

INTRODUCTION

It has long been accepted that the efficient level of mobile termination rates (MTRs) is in part influenced by the relative size of the access externality and call externality. The literature has traditionally assumed an access externality when examining the efficient level of MTRs. However, recent papers have assumed the presence of a call externality and no access externality.[4] Not surprisingly, papers in the two different camps come to different conclusions about the efficient level of MTRs. Papers assuming an access externality conclude that MTRs should be above cost, and those assuming a call externality conclude MTRs should be below cost.

The results from papers assuming a call externality have been interpreted as suggesting that efficient MTRs need to be set below costs whenever the receiving party benefits from receiving a call.[5] However, such a conclusion misinterprets the meaning of the externality factors within the theoretical models, and has been made without reference to empirical evidence on the existence or size of any call externality.

This paper examines the significance of call externalities on pricing in the two-sided mobile market in two ways: first, by clarifying the interpretation of the externality factor in economic models in the existing literature and, second, by measuring the size of any residual call externality using observed retail pricing by MNOs. We seek to show that the residual call externality is sufficiently small such that regulators are correct to continue with their existing practice of ignoring it for the purposes of setting MTRs.

This paper is structured as follows: Section 2 looks at the roll of call externalities in pricing mobile termination; Section 3 examines the conceptual considerations when assessing the level of residual call externality; Section 4 highlights existing regulatory views on the size of call externality; Section 5 analyses retail pricing of MNOs using an on-net/off-net framework to estimate the observed size of the call externality; and Section 6 provides concluding remarks.

ROLE OF EXTERNALITIES IN PRICING MOBILE TERMINTION WITHIN A TWO-SIDED MARKET

Efficient pricing in two-sided markets takes into account the effect pricing on one side has on the other side of the market. This has two aspects: first, the elasticity of demand on one side of the market influences the price to be charged on the other side, so that the more elastic one side of the market is, the higher the price the other side will pay. Second, platform pricing also internalises the inter-group externalities, so that the price faced by one side is influenced by the benefit the other side gains from the first side's presence.[6]

This implies that the side of the market which is (i) more competitive (has higher elasticity) and (ii) causes a larger benefit to the other group than vice versa, will face the lowest price. This may result in that side being subsidised to participate. The classic example is the night club market, where females receive free entry and males face the full cost of providing services to both sides.[7]

The economic literature addressing the issue of welfare maximising termination charges, using a two-sided market framework, shows that the relationship of these rates to the average incremental cost is centrally determined by the existence and size of the access externality and the call externality, amongst other factors.[8]

The literature initially focused on estimating the efficient termination rate in the presence of an access externality – or where the access externality is larger than the call externality. In such a case, the literature agrees that the efficient level of termination rate is above the cost of providing the service. In effect, the calling side of the market assists in the subsidisation of the acquisition and retention of mobile subscribers.[9]

Some of these papers have noted that when call externalities are introduced into the analysis, the welfare maximising level of the MTR falls back towards cost. Recently, there has been a growing number of papers focusing on the efficient level of termination rates in the presence of call externalities, assuming that access externalities are zero[10]. These papers conclude that the efficient termination charge is below the cost of termination in the presence of call externalities and put into question the rationale of regulating mobile termination rates[11].

Some have interpreted the literature as implying that if the receiving party gets a benefit from answering a call, the efficient MTR is necessarily below cost.[12] However, this view misunderstands the role of a platform and the concept of externalities in two-sided markets.

The role of platforms in two-sided markets is to internalise the cross-group externalities that the members of each group are not able to internalise themselves (residual externalities).[13] At one extreme, where all cross-group externalities are internalised by parties, and the parties can agree on an optimal price structure which maximises joint benefit, the market is one-sided. But where this does not occur, there is a role for the intermediary platform to bring the two sides together and to assist in facilitating efficient trading.

The purpose of a platform in a two-sided market is to act as an intermediary between the two sides and coordinate their transactions. In essence, the platform seeks to internalise the external value that the parties cannot do so themselves and sets the price structure so as to maximise the joint welfare of the two sides of the market. If one side receives a large benefit from the participation of the other side, the platform will recover more fees from the first side.

We call the external value which the parties themselves cannot internalise the residual external value. It is the relative size of the residual externality effects which determines whether there is a net access or call externality effect.

CONCEPTUAL CONSIERATIONS WHEN ASSESSING LEVEL OF RESIDUAL EXTERNALITY

Before progressing onto an empirical assessment of on-net pricing to see whether it shows the presence of call externalities, we first need to look at the relevant conceptual factors that determine the extent that any call externality will be internalised, and as a result, the size of the residual call externality.

A sub-optimal level of network usage (here assumed to be minutes) will only occur where the parties cannot negotiate to internalise the call externality themselves. The necessary conditions for this to occur are shown in Figures 1 and 2.

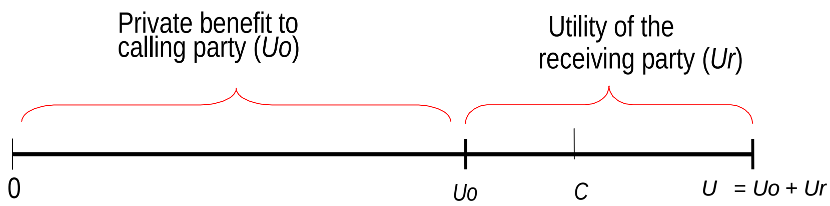

Figure 1 Utility and cost of a marginal minute – no need for internalisation

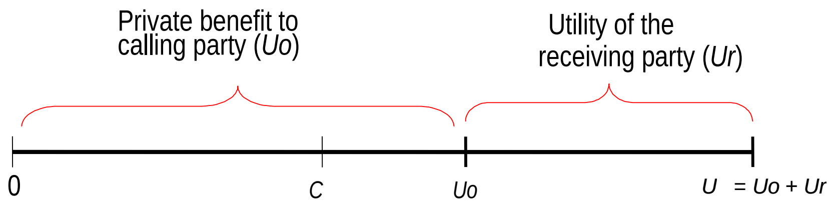

Figure 2 Utility and cost of a marginal minute – need for internalisation

Figures 1 and 2 show the utility to both the calling party (*Uo*) and to the receiving party (*Ur*) of an extra minute of a call. The total utility of an extra minute is the combination of the utility of the calling and receiving parties ($U=Uo+Ur$). Where the cost of the extra minute falls below the private benefit of the calling party ($C \leq Uo$) there is an optimal level of call minutes and no need to internalise the receiver's call externality. This is the case shown in Figure 1. However, Figure 2 shows the case where the cost of the extra minute is greater than the private benefit to the calling party ($C>Uo$).

When this occurs the calling party (who incurs the cost of the extra minute) will end the call, even though the joint utility of an extra minute is greater than the cost ($U>C$). It is this foregone utility which gives rise to an inefficient consumption of minutes and a welfare loss.

However, the parties themselves could negotiate to internalise this externality. This will occur when the utility of the extra minute to the receiving party is greater than or equal to the cost differential including the transaction costs[14] (*Ct*), that is $Ur \geq C+Ct-Uo$[15]. Only when such negotiated internalisation cannot occur is there any impact on efficient platform pricing.

There are, therefore, two necessary conditions that must hold for there to be a residual call externality:

(i) $C > Uo$, i.e .the cost of an extra minute exceeds the utility to the caller.

(ii) $C + Ct > U > C$, i.e. the total utility of an extra minute is bounded above by the call cost including transaction cost, and below by the call cost excluding transaction cost.

When these conditions hold, the value of the residual call externality is given by $C + Ct - U$.

An assessment of the likelihood of internalising call externalities is assisted by separating calls into three types. The first type is calls within a closed-user group with a single bill payer. This typically is an immediate family unit (e.g., parent paying the phone bill of children) or company phone (where the company pays the bill of its employees). The second type is calls within a closed-user group (CUG), with frequent calling patterns. Examples include groups of friends, or colleagues, or categories of users (such as students). The third call type is calls made outside CUGs.

The likelihood of internalisation of the call externality for these three call types is summarised in Table 1.

Table 1 Likelihood of internalisation of call externalities

	Size of group	Call frequency	Likelihood of internalising the call externality
CUG with single bill payer	Very small	Very frequent	Very likely
CUG	Small to large	Frequent	Likely
Non-CUG	Very large	Infrequent	Unlikely

The first consideration to note for all call types is that the cost to the calling party of a marginal minute is likely to be quite low for most calls made. The marginal cost could be zero, for minutes within a bundle, or for calls to a selected number of friends. The cost of a minute outside of a bundle is typically quite low as well – less than 20 €c/minute.

Call externalities for calls made within CUGs with a single bill payer are most likely to be internalised by the parties to the call. This is because:

- The utility of the calling party is likely to be high. This implies that there would be few minutes where the cost is higher than the utility gained by the calling party. So condition (1) will not be satisfied.
- For minutes where calling party utility is less than the cost, it is likely the utility gained by the receiving party would also be very high and the transaction costs would be very low. Parties within sole-payer CUGs have reciprocal, close and repeat relationships, not requiring any search costs to be incurred. Repeat and reciprocal relationship also facilitates easy negotiation between parties – the calling party can be compensated in many ways, the simplest being an arrangement whereby one party pays the cost of calls of the other party; for instance, a business providing a work phone to an employee or parents paying the mobile bills of their children. So condition (2) will not be satisfied.

Calls externalities within CUGs are also likely to be internalised by the parties to the call, because:

- The utility of the calling party and the receiving party would be high, meaning that the total utility is large. Only when the cost of the call and the transaction costs of internalisation (*C+Ct*) is greater than the total utility of the call (*U*) can intervention be justified.
- Transaction costs (*Ct*) would be low. Subscribers within CUGs have repeat and reciprocal calling patterns. Search costs are low, as are negotiation costs. The repeat and reciprocal nature of relationships enable a variety of simple yet effective ways through which a receiving party can 'compensate' the calling party for incurring the total cost of the marginal minute. The simplest example being that the parties agree to call each other half of the time, ensuring that they share the total cost of calls. While enforcement costs would be higher than within a sole-payer CUG, the costs would be still be relatively low – constant offenders risk social exclusion from a group of friends, and business colleagues/suppliers/clients may choose not to deal with constant offenders.

Call externalities generated from calls received from subscribers outside of the receiving party's CUGs would not likely be internalised. Subscribers have infrequent and often one-way calling relationships with other callers outside their CUG, and so the search and negotiation costs would be relatively high. There are also no effective enforcement mechanisms through which callers could enforce any deal. In the absence of informal social enforcement mechanisms, partiers have only the option to rely on formal dispute mechanisms (e.g. court proceedings) which are an unrealistic option given the value of the bargain. However this by itself is not necessarily an issue of concern. Many calls made to and received from callers outside one's CUG may not derive a high level of total utility. Thus, while the cost of the marginal minute may be greater than the utility of the caller (and hence the marginal minute will not take place), it may also be above total utility of both consumers. Hence it is efficient for the marginal minute not to be consumed.

Of course, a negative call externality is also a possibility in individual cases, especially outside of a CUG, e.g. a call from a telesales agent who interrupts dinner, but is unlikely to be a universal phenomenon.

Call externalities are likely to be internalised by the parties to a call for calls made between members of frequent call circles, or CUGs. This represents the majority of call volumes carried over mobile networks. However, for calls received from parties outside subscribers' CUGs, there is limited ability to internalise any call externality. Consequently, one may expect to see a small residual call externality effect.

However, it does not follow that the possible existence of a residual call externality for some subscriber relationships results in an overall residual call externality effect, or that efficient MTRs are below cost. The access externality effect also needs to be taken into account, and ultimately though, it is an empirical question to determine which effect will dominate.

EXISTING REGULATORY VIEWS ON SIZE OF CALL EXTERNALITY

It is commonly recognised that the UK Competition Commission and the telecommunications regulator (Oftel and now Ofcom) analysed the extent of the access externality in the context of setting welfare optimising termination charges in 2003, 2004 and 2007.[16] This was done using a model developed by Dr Jeffrey Rohlfs (the Rohlfs model). It is less commonly recognised, however, that the Rohlfs model acknowledged the existence of, and accounts for a small call externality, through the derivation of the externality factor.[17]

The Rohlfs model dealt with externalities through the use of a gross network externality factor (gross Rohlfs-Griffen or R-G factor). The gross R-G factor is the 'ratio of the total social value of subscription to the private value that accrues to the marginal subscriber'.[18] The gross R-G factor includes both usage (call) and membership (access) externality effects. Rohlf's commenting on Oftel's assumed range of 1.3 to 1.7, stated:

> This range of the externality factor should be interpreted to include cross-elastic, as well as other, externalities. The logic that justifies this range relates to total

consumer benefit and does not distinguish between cross-elastic and other externalities.[19]

In deciding to use a range for the gross R-G factor of 1.3 to 1.7, Oftel was of the view that while users can internalise most of the call externality, uninternalised call externalities may still be significant – especially for calls received from callers outside of the receivers' closed user groups.[20] The UKCC approved the use of an R-G factor of 1.5.[21] The same value was used by Ofcom in its 2007 decision on mobile termination rates[22]. It should also be noted that the analyses undertaken by the UK Competition Commission in 2003 and Ofcom in 2007 focused on residual externality effects. The issue of internalisation was analysed in depth by the UKCC and its conclusions have been adopted since by Ofcom[23].

As such, the surcharge should therefore be interpreted as being the residual externality surcharge taking into account the level of residual call externalities due to calls received from parties outside the receivers' CUGs. The approach adopted by UKCC and Ofcom is consistent with the theoretical approach outlined in this paper. Namely, that the overall externality effect is determined by the relative size of the residual access externality and the residual call externality.

EMPRICAL ESTIMATION OF CALL EXTERNALITIES

The theory of on-net/off-net price discrimination in telecommunication networks has been analysed in a series of papers starting with Laffont, Rey and Tirole (1998b), and further developed by Gans and King (2000b), using a Bertrand model of competition, with a Hotelling-type differentiation between two competing networks. These papers include a basic result: in a two-part tariff structure (i.e. fixed monthly fees plus usage related charges) mobile network operators (MNOs) will price usage at perceived marginal cost. Therefore, on-net calls will be priced at the marginal cost of originating and terminating calls on the MNO's own network, whilst off-net calls will be priced at the marginal cost of call origination plus an interconnection charge for call termination on another network.

The basic model used in these papers has subsequently been extended to include call externalities, whereby subscribers gain utility not just from calls they make, but also from calls they receive. The analysis for the two-part tariff structure is very well developed by Berger (2004), Hoernig (2007) and Armstrong and Wright (2007). These papers predict that if an MNO believes a positive call externality exists, on-net calls will be priced below marginal cost and off-net calls to competing networks priced above marginal cost. The intuition for this prediction is simply that the receiver's benefit of on-net calls is internalised within the same network, and so MNOs will charge a lower call price for on-net calls, but also will be able to charge a higher subscription price to reflect the benefit from receiving on-net calls. In the case of off-net calls, however, the externality benefit accrues to the subscribers of the receiving MNO, and so enhances its ability to compete. Thus, the MNO originating off-net calls will want to increase the price to mitigate this loss. For example, Vodafone Ireland's pre-pay "Advantage Plus" tariff charges 29c/minute for calls within the Vodafone network or to fixed networks, but 39c/minute for calls to other mobile networks – a difference of 10c. The mobile termination rate is about 7.8c (while the fixed termination rate is about 0.6c[24]), suggesting that factors other than pure costs may be at work encouraging MNOs to discount on-net prices and/or charge a premium for off-net calls (e.g. the call externality). In contrast, however, the pre-pay packages of Vodafone UK make no distinction between on-net and off-net calls, with simply a 20p/minute charge for calls to all networks, despite a mobile termination rates of around 4.7p on mobile networks (and 0.2p on fixed networks[25]). The explanation for these differences must lie in the different market positions between the two countries.

It is clear, therefore, that the existence and magnitude of a call externality should have an important role in pricing and, consequently, we should be able to observe MNOs' perceptions of the strength of the call externality in actual pricing plans. However, the complexity of mobile pricing plans mean that direct observations of the differences between on-net and off-net tariffs will often be contradictory, and a richer model and more sophisticated methodology is necessary. This is the objective of this section.

We take as a starting point the call externality specification of Armstrong and Wright (2007), where the call externality is modelled as a fixed amount per call received irrespective of volume, in contrast to the models of Berger (2004) and Hoernig (2007) where the call externality is modelled as a fixed proportion of the utility of making calls which diminishes with volume. Armstrong and Wright provide justification for their approach by arguing that generally subscribers have little control over the calls they receive, and so each received call can be considered to have a random value drawn from a distribution of fixed mean.

We also introduce a simple modification to the model whereby our pricing equations holds not only in the case of evenly distributed calling patterns between networks (as is usually analysed in the literature), but also for the case where a higher proportion of traffic is on-net, consistent with the existence of limited calling circles[26] where consumers in each circle cluster on the same network. In such a situation a network can attract the full membership of a calling circle, especially where it offers on-net discounts. When combined with retail competition, such a strategy has been used successfully by smaller operators to attract subscribers.[27] Other network models capture exogenous customer groups and market segmentation, e.g. Jullien (2001), Banerji and Dutta (2005), and Ambrus and Argenziano (2004).

The derivation of the pricing equations we use is shown in Appendix 1, for the case of calling party pays with two-part tariffs.[28] The pricing equations for on-net and off-net calls respectively are:

$$p_{iit} = rc_t + 2nc_t - \gamma \tag{1a}$$

$$p_{ijt} = rc_t + nc_t + a_t + \gamma \frac{\phi(1-\phi) + s_{it}}{(1-\phi)^2 - s_{it}} \tag{1b}$$

where:

- p_{iit} is the price of an on-net call on mobile network i in country t;
- p_{ijt} is the price of an off-net call from mobile network i to another mobile network in country t;
- rc_t is the marginal retail cost of call origination on a mobile network in country t;
- nc_t is the marginal network cost of call origination or termination on a mobile network (assumed to be the same[29]) in country t;
- ϕ is the proportion of calls that are on-net irrespective of market share (as may happen with limited calling circles or CUGs);
- a_t is the interconnection charge for termination on mobile networks in country t;
- γ is the residual call externality[30] in country t;
- s_{it} is the market share of network i in country t.

Note that equations (1) split out retail and network costs. However, since there is no reason to suppose retail costs will differ between on-net and off-net calls, we can eliminate the need to consider these costs by taking the differential price:

$$p_{ijt} - p_{iit} - a_t = -nc_t + \frac{(1-\phi)\gamma}{(1-\phi)^2 - s_{it}} \tag{2}$$

Empirical estimation of equation (2) is best done by introducing a residual term to capture variations in network costs between countries, and other unaccounted factors that influence pricing. We write:

$$p_{ijt} - p_{iit} - a_t = -nc + \frac{(1-\phi)\gamma}{(1-\phi)^2 - s_{it}} + u_{it} \qquad (3)$$

where we assume $u_{it} \sim N(0,\sigma)$. Equation (3) can be estimated by non-linear least squares. The parameters to be estimated are nc (the network cost), ϕ (the proportion of traffic on-net irrespective of market shares), and γ (the call externality). All three should be interpreted as the average level across all networks and subscribers in the same countries. In the case of the network cost national variations are explicitly allowed for through the residual term, u_{it}. We would expect the call externality and the proportion of calls within a CUG to be different for each subscriber. What we are attempting to measure is the average levels across all customers of the network operator.[31]

Data

The dataset analysed consisted of prices for two MNOs in each of 22 European countries, giving a total of 44 MNOs in total. For each MNO we calculated on-net and off-net (to other mobile networks) call charges using each of pre-pay and contract tariffs, giving a total of 88 sets of prices in all.

The data used is described in more detail in Appendix 2. Network operators have numerous different tariff packages available to subscribers and on-net and off-net price differentials will vary accordingly. In order to get typical prices we took a basket of 1,000 calls split equally between on-net calls, off-net calls to other mobiles, and calls to fixed lines. This split is a good rough approximation to the calling pattern in most European countries. We calculated the incremental bill saving if each category of call were individually subtracted from the basket (e.g. removing only the on-net calls from the basket), and divided this saving by the number of subtracted call minutes. This allows calculation of the effective price per minute for each of (a) on-net calls, and (b) off-net calls to other mobile networks, on the assumption that subscribers choose the most appropriate tariff. In this way we are able to measure the typical difference in price between on-net and off-net calls, reflecting the network operators' assessment of the "average" call externality.

Mobile termination rates were taken from the European Regulators Group (ERG) benchmarks. The marginal network cost of termination on a mobile network was assumed to be 5c, although lower estimates of 1-2c have been suggested. We present a sensitivity using a marginal cost of 3c and 1.5c.

Generally, the lower the estimate of the marginal network cost, the lower the implied estimate of the call externality.

Analysis

Equation (3) was estimated by the non-linear least squares algorithms of LIMDEP. Three models were estimated:

- Model A: contract tariffs only (allowing 44 observations);
- Model B: all tariffs (allowing 88 observations), estimating separate coefficient values for ϕ (the proportion on traffic on-net irrespective of market shares) and γ (the call externality) for each of contract and pre-pay subscribers. The network cost, nc, will be the same;
- Model C: all tariffs (allowing 88 observations), constraining ϕ and γ to be the same for both contract and pre-pay subscribers.

Results are shown in Table 2. Although overall the models explain only a small proportion of the variation in on-net and off-net price differentiation – indicating that other localised factors are important in pricing - many of the estimated coefficients are statistically significant at the 1% level.

Although only Model A (contract tariffs) is strictly consistent with the two-part tariff, doubling the sample size does have some empirical benefits and, as argued above, the model is likely to also provide a good approximation to pre-pay tariffs.

The constraints imposed on Model C are not statistically significant at the 5% level,[32] and so we have a preference for accepting Model C compared to model B.

Table 2: Model results

Model A: Contract tariffs only

Numbers of observations	44
Degrees of freedom	41
Standard error of residuals	0.0475
Adjusted R^2	0.0838

	Coefficient	**Standard Error**	**Coefficient Standard Error**	**P- value**
nc	0.0794	0.0201	3.940	0.0001
ϕ	0.1932	0.0392	4.928	0.0000
γ	0.0032	0.0065	0.495	0.6204

Model B: Contract and pre-pay tariffs – unconstrained

Numbers of observations	88
Degrees of freedom	83
Standard error of residuals	0.0634
Adjusted R^2	0.0552

	Coefficient	**Standard Error**	**Coefficient Standard Error**	**P- value**
nc	0.0821	0.0269	3.048	0.0023
ϕ contract	0.1889	0.0559	3.377	0.0007
γ contract	0.0041	0.0095	0.430	0.6674
ϕ pre-pay	0.1382	0.0931	1.485	0.1376
γ pre-pay	0.0148	0.0164	0.905	0.3655

Model C: Contract and pre-pay tariffs – constrained

Numbers of observations	88
Degrees of freedom	85
Standard error of residuals	0.0634
Adjusted R^2	0.0551

	Coefficient	**Standard Error**	**Coefficient Standard Error**	**P- value**
nc	0.0783	0.0294	2.664	0.0077
ϕ	0.1725	0.0687	2.512	0.0120
γ	0.0072	0.0125	0.578	0.5636

There is a large degree of similarity between all three models. In all cases the network cost is estimated to be around 8€c/minute, and is statistically significant at the 1% level. Similarly the proportion of traffic on-net irrespective of market share is estimated to be around 17-19%.[33] In all cases, however, this parameter is statistically significant at the 1% level.

Crucially for this paper, although all models estimate a positive residual call externality, in no cases is this statistically significant even at the 10% level. We must conclude, therefore, that the magnitude of any call externality is small.

CONCLUSIONS

Economic theory predicts that MNOs will price on-net and off-net calls taking account of marginal costs (including interconnection costs), and any perceived residual externality effects. The model examined in this paper focus on the residual call externality effect. In so far as a residual call externality is present, this will lead to lower margins on on-net calls, but higher margins on off-net calls to competing networks. The empirical evidence of this paper suggests that this differential is not as large as we might expect, and so we conclude that the residual call externality is generally low.

Of course, a negative call externality is a possibility in individual cases, e.g. a call from a telesales agent who interrupts a dinner party, but is unlikely to be a universal phenomenon. Therefore, it is more likely that the call externality is often internalised (as discussed in this paper), and although the average residual call externality is positive, it is very small.

These empirical results have implications for the application of the theoretical literature surrounding efficient MTRs. The literature shows that efficient MTRs can be above or below cost – depending on whether there is a net residual access or call externality effect - ultimately an empirical question. Based on the empirical results in this paper, it is not appropriate to apply the findings of papers assuming a net residual call externality effect, such as Hoernig (2007), Jeon, Laffont, Tirole (2004), Berger (2005), to the regulation of MTRs. Setting MTRs below cost is likely to lead to welfare-reducing outcomes.

APPENDIX 1 SERIVATION OF PRICING EQUATION

For simplicity we assume two mobile networks compete within a Bertrand pricing framework, with Hotelling differentiation. The basic theory behind this model is laid out in the paper by Laffont, Rey and Tirole (1998a). We also assume a one-period model, in which networks maximise profits with respect to prices (in a two-part pricing structure).

Subscribers are assumed to be distributed uniformly along a segment [0,1], differentiating the two networks. We represent the consumer surplus that an individual subscriber would receive from network *i* as:

$$w_i - tx \tag{A1}$$

where t is the loss of utility to the subscriber for each unit of distance from network i.

The market share of network i, s_i, is found by determining the value of x at which subscribers are indifferent between the fixed and mobile networks, thus:

$$s_i = \frac{1}{2} + \frac{w_i - w_j}{2t} = \frac{1}{2} + \sigma(w_i - w_j) \tag{A2}$$

where $\sigma = 1/(2t)$ is an index of substitutability between the two networks.

We can further write:

$$w_i = \alpha_i[v(p_{ii}) + \gamma q_{ii}] + \alpha_j[v(p_{ij}) + \gamma q_{ji}] - F_i \tag{A3}$$

where:

α_i is the proportion of on-net calls;

q_{ij} is the number of calls from network i to network j;

$v(p_{ij})$ is the variable consumer surplus from these calls, where p_{ij} is their price;

γ is the utility the subscriber gains form each call received; and

F_i is the fixed monthly payment network i.

Furthermore, we will model the percentage of on-net calls as the sum of a fixed proportion of calls irrespective of network size (perhaps due to limited calling circles or CUGs), ϕ, with the remainder evenly distributed according to market share. Thus

$$\alpha_i = \phi + \frac{s_i}{1-\phi} \tag{A4}$$

Substituting (A3) and (A4) into (A2) gives:

$$(1-\phi)\alpha_i - \phi(1-\phi) = \frac{1}{2}$$
$$+\sigma\{\alpha_i[v(p_{ii})+\gamma q_{ii}]+\alpha_j\lfloor v(p_{ij})+\gamma q_{ji}\rfloor-\alpha_j\lfloor v(p_{jj})+\gamma q_{jj}\rfloor-\alpha_i\lfloor v(p_{ji})+\gamma q_{ij}\rfloor\}$$
$$-F_i+F_j \tag{A5}$$

This is essentially a relationship between the proportion of calls that are on-net (α_i), closely related to market share through equation (A4), and prices (both p_{ii}, p_{ij} and F_i). The network would normally choose both p_{ii}, p_{ij} and F_i, with the later determining market share and hence α_i. However, for algebraic convenience we equivalently assume that the network chooses p_{ii}, p_{ij} and α_i, consequentially determining the implied level of F_i. Therefore, differentiating with respect to p_{ii} and p_{ij} whilst holding constant α_i we have:

$$\frac{\partial F_i}{\partial p_{ii}} = -\alpha_i q_{ii}\left(1+\frac{\gamma\varepsilon}{p_{ii}}\right) \tag{A6a}$$

$$\frac{\partial F_i}{\partial p_{ij}} = -\alpha_j q_{ij} + \alpha_i \frac{\gamma\varepsilon q_{ii}}{p_{ij}} \tag{A6b}$$

where ε is the price elasticity of calls defined as $\varepsilon = -\frac{\partial q}{\partial p}\frac{p}{q}$.

We write network i 's profit function as:

$$\pi = s_i\lfloor\alpha_i(p_{ii}-rc-2nc)q_{ii}+\alpha_j(p_{ij}-rc-nc-a)q_{ij}++\alpha_j(a-nc)q_{ij}+F_i-f\rfloor \tag{A7}$$

where:

rc is the marginal retail cost of a call on a network;

nc is the marginal network cost of a call originated or terminated on a network;[34]

a is the interconnection charge for terminating calls on a network; and

f is the marginal cost of a network subscription (excluding calls).

Maximising π_i with respect to p_{ii} and p_{ij} whilst holding constant α_i gives first order conditions:

$$p_{ii} = nr + 2nc - \gamma \tag{A8a}$$

$$p_{ij} = nr + nc + \frac{\gamma\alpha_i}{1-\alpha_i} = nr + nc + \gamma \frac{\phi(1-\phi)+s_i}{(1-\phi)^2 - s_i} \tag{A8b}$$

APPENDIX 2 DATA AND SOURCES

Monthly subscriptions Price of call minute: • On-net; • Off-net to other mobile networks	We used the Teligen T-Basket database to estimate call prices. MNOs offer numerous tariff options and bundles, and subscribers will choose a tariff (or migrate between tariffs) according to which is cheapest for their particular usage. Although MNO tariffs are published, the number of subscribers on each tariff is commercially confidential information. It is necessary, therefore, to make an assumption about how subscribers choose between tariff options. The Teligen T-Basket software calculates, for a pre-defined call basket, the cheapest tariff available. For the purposes of this analysis we assumed a bundle of 1,000 outgoing calls/year at 1.8 minutes each, split equally between on-net calls, off-net calls to other mobiles, and calls to fixed lines. This split is a good rough approximation to the calling pattern in most European countries. We calculated the incremental bill saving if each category of call were individually subtracted from the basket (e.g. removing only the on-net calls from the basket), and divided this saving by the number of subtracted call minutes (33.33% x 1,000 x 1.8 minutes). This allows calculation of the effective price per minute for each of (a) on-net calls, and (b) off-net calls to other mobile networks, all on the assumption that subscribers choose the most appropriate tariff. The Teligen data relates to May 2008.
Interconnection charges	For mobile termination rates we took the European Regulators Group's (ERG) snapshot benchmarks for January 2008.
Market share	Subscriber market share of each operator at Q2 2008. Source: Wireless Intelligence.

REFERENCES

Ambrus A and Argenziano "Network Markets and Consumer Coordination", N 1481, Cowles Foundation Discussion Paper, Yale University, September 2004.

Armstrong M "The theory of access pricing and interconnection", chap. 8, in Cave M, et al., Handbook of Telecommunications Economics, vol. 1, North Holland Elsevier (2002)

Armstrong M (2006) "Competition in Two-sided Markets", Rand Journal of Economics, 37(3), 668-91 (2006).

Armstrong M and Wright J "Mobile Call Termination", Economic Journal forthcoming, October 2008, accessible at http://profile.nus.edi.sg/fass/ecsjkdw/.

Banerji A and Dutta B "Local Network Externalities and Market Segmentation", No 725, Warwick Economic Research Papers, Department of Economics, University of Warwick (2005).

Baranes E and Flochel L, "Competition in telecommunication networks with call externalities," Journal of Regulatory Economics , 34(1) (2008)

Berger U "Access Charhes in the Presence of Call Externalities", Contributions to Economic Analysis and Policy, 3(1), article 21.

Berger U "Bill-and-Keep vs. Cost-Based Access Pricing Revisited", Economics Letters, 86(1), 107-12 (2005)

Coase R "The Problem of Social Cost," Journal of Law and Economics 3, 1-44 (1960).

Competition Commission "Vodafone, O2, Orange and T-Mobile: Reports on References Under Section 13 of Telecommunications Act 1984 on Charges Made by Vodafone, Orange, O2 and T-Mobile for Terminating Calls Made by Fixed and Mobile Networks", HMSO, London (2003)

Dahlman, Carl J. (1979). "The Problem of Externality," Journal of Law and Economics 21 (2): 141–162

Gans J and King S "Mobile Network Competition, Consumer Ignorance and Fixed-to-mobile Call Prices," Information Econ. and Policy, vol. 12 (2000a)

Gans J and King S "Using Bill and Keep Interconnect Arrangements to Soften Network Competition", Journal of Economic Literature (2000b).

Hausman J and Wright J, "Two Sided Markets with Substitution: Mobile Termination Revisited," ms. (2006)

Hermalin B and Katz M, "Customer or Complementor? Intercarrier Compensation with Two-Sided Benefits," ms (2006)

Hoernig S "On-net and off-net pricing on asymmetric telecommunications networks", Information Economics and Policy, 19 (2007) pp171-188.

Jeon, D, Laffont, J-J., and Tirole J. "On the Receiver Party Pays Principles", Rand Journal of Economics, 29(1), 1-37 (2004)

Jullien B “Competing in Network Industries: Divide and Conquer”, Working Paper, 2001.

Laffont J-J, Rey P and Tirole J “Network competition: I. Overview and Non-discriminatory Pricing”, RAND Journal of Economics, Vol 29, No 1, Spring (1998a).

Laffont J-J, Rey P and Tirole J “Network competition: II. Price Discrimination”, RAND Journal of Economics, Vol 29, No 1, Spring (1998b).

Ofcom “Wholesale Mobile Voice Call Termination: Statement”, Office of Communications, London (2004)

Ofcom “Mobile Call Termination: Statement”, Office of Communications, London (2007)

Ordover J “Recovering fixed and common costs for mobile networks in Europe”, Compass Lexecon (2008)

Richot J and Tirole J “Platform Competition in Two-sided Markets”, Journal of the European Economic Association, 1(4), 990-1029 (2003)

Rohlfs J “A Model of Prices and Costs of Mobile Network Operators,” Report prepared at the request of Oftel (2002a)

Rohlfs J, “Network Externalities and their Internalization With Respect to the UK Mobile Market”, (2002b)

Rochet J and Tirole J ”Two-sided Markets: a Progress Report”, Rand Journal of Economics, 37(3) (2006)

Williamson, Oliver E. (1981). “The Economics of Organization: The Transaction Cost Approach,” The American Journal of Sociology, 87(3), pp, 548-577

Wright J, “Access Pricing under Competition: An Application to Cellular Networks,” J. Industrial Econ., vol. 50 (2002)

Vodafone, “On-net Pricing in Mobile”, Vodafone Public Policy Series, No.8, April 2008

NOTES

[1] Head of Regulatory Economics, Vodafone Group, e-mail: jonathan.sandbach@vodafone.com
[2] Economist, Vodafone Group, e-mail: luke.vanHooft2@vodafone.com
[3] The views expressed in this paper are those of the author, and should not necessarily be attributed to Vodafone.
[4] E.g. Berger (2004), Hoernig (2007).
[5] See, for example, EC *Draft Recommendation on the Regulatory Treatment of Fixed and Mobile Termination Rates in the EU*; and ARCEP *Les référentiels de coûts des opérateurs mobiles en 2008.*
[6] Armstrong (2006); Richot and Tirole (2003, 2006).
[7] See Armstrong (2006).
[8] For a full discussion of the literature surrounding efficient termination rates, see Ordover (2008).
[9] See Armstrong, (2002), Wright (2002), Gans and King (2000a), and Hausman and Wright (2006).
[10] Hoernig (2007), Jeon, Laffont, Tirole (2004), Berger (2005). See also Hermalin and Katz (2006), who model benefits to both callers and calling parties in a one-way access setting.
[11] See Baranes and Flochel (2008).
[12] See, for example, ARCEP (2008) and EC (2008).
[13] Richot and Tirole (2006) note that the failure of the Coase Theorem is a necessary condition for a two-sided market. In addition, the structure of prices must also matter for a market to be two-sided.
[14] There are three broad categories of transaction costs: search, negotiation, and enforcement costs. See Dahlman (1979) and Williamson (1981).
[15] Coase (1960) demonstrated that where transaction costs are zero all externalities will be internalised by the parties. Transaction costs do not equal zero in the real world, but the key insight from this seminal work was that the size of the transaction costs determines the ability to internalise externalities.
[16] See Ofcom (2004, 2007) and UK Competition Commission (2003).
[17] See Rohlfs (2002a, 2002b).
[18] Rohlfs (2002a) *ibid.*, p.3.
[19] Rohlfs (2002b), *supra*, p.7.
[20] *Ibid.*, p.2-3.
[21] UKCC (2003), *supra* note 14, p.88.
[22] Ofcom (2007), *supra* note 14, p. 342.
[23] See, UKCC (2003), *supra* note 14, chapter 8.
[24] Approximate average for single transit termination of a 2 minute call on the Eircom network.
[25] Approximate average for single transit termination on the BT network.
[26] The limited calling circle is also likely to be a CUG, described in Section 3 of this paper.
[27] For a more complete discussion of TMNE and on-net discounts, see Vodafone (2008). This coefficient measures the effect on network traffic of adopting such pricing strategies.

[28] An equivalent mathematical formula for linear tariffs (per minute charges) is more difficult to specify, and depends on the price elasticity of calls. Berger (2004) and Hoernig (2007) give analytical results that link Lerner Index margins to the call externality, market shares and price elasticities, and provide numerical simulations of the relationship. Broadly, however, we would expect similar results, but with prices exceeding their respective marginal costs in order to cover subscriber specific costs. In any event, the two-part tariff model probably provides a better overall approximation to the actual price structure whereby the average price will be decreasing in usage.

[29] In actual fact termination has a higher cost than origination due to the need to locate the subscriber on the network and transport the call to that location (compared to origination where the call is simply transported to the nearest point on interconnect).

[30] This coefficient will be measuring only the residual externality, i.e. the call externality that is not internalised by the parties to a call. As shown above, it is the un-internalised value which influences efficient platform pricing.

[31] The model used explains any observed differences in retail prices in terms of changes in interconnection and network costs, with a call externality factor accounting for the remaining differences. However, real world pricing decisions account for many other factors. For example, mobile number portability, whereby callers' knowledge of whether a call is on-net or off-net may become inaccurate, would reduce the rationale for an MNO to offer reduced on-net pricing, consistent with the findings of this paper. That is, where subscribers are unable to tell which network another subscriber belongs to, MNOs are less able to use reduced on-net pricing to internalise any residual call externality or tariff-mediated network effect. Also it is possible that a large operator may attempt to use on-net pricing as an anti-competitive predatory tool to "tip" the market in its favour. This would result in a larger than predicated on-net/off-net price differential. Hoernig (2007), however, shows that this is an inefficient strategy, and also counter to the finding in this paper of on-net/off-net price differentials being lower than would predicted by competitive models.

[32] $F_{2,83} = 1.5528$.

[33] Although this drops to 14% when separately estimated for pre-pay customers alone, this difference is not statistically significant, and more generally the equality constraints on contract and pre-pay coefficients imposed in model C are statistically acceptable.

[34] In practise the network cost of terminating a call is slightly higher than that of originating a call, but we ignore this difference for the sake of simplicity.