ERG Consultation on Fixed/Mobile Termination Rates

Response from Horrocks Technology Limited

Horrocks Technology Limited is a UK-based consultancy on telecommunications regulations and technology. Its principal John Horrocks was Deputy Technical Director of OFTEL from 1986-1990 and since 1990 has worked for many assignments for regulators, operators and the European Commission both directly and through larger consultancies such as Ovum, PricewaterhouseCoopers, Indepen, Mott McDonald, A-Cing and TERA Consultants. He currently chairs the CEPT ECC TRIS project team on behalf of UK BERR and was an ETSI Technical Committee chairman for more than eight years.

This response does not in any way represent the views of BERR or other ECC TRIS members or any other client. BERR is aware that this independent response is being submitted.

---

0 Summary

The ERG Common Position states on p6 :

"Economic principles tend to recommend a unique and uniform TR, determined with reference to costs incurred by an hypothetic efficient operator, i.e. a termination rate which does not depend on costs effectively incurred by the operators or on their market shares."

The main purpose of this response is to comment on this point, which underlies the rest of the consultation.

Our view is that all termination rates should be regulated to a symmetrical rate that is not cost based but is below cost and low, possibly, but not necessarily zero. The rate should be applied to all networks that use the same part of the numbering scheme. The rates for fixed and mobile should not necessarily be the same because there may be reasons to retain a distance element in fixed if there is geographic numbering and this element is not relevant to mobile.

We believe that prices are held artificially high by cost based mobile termination rates and that if the mobile termination rates are reduced substantially then retail prices will drop and the market will expand and more than compensate the operators.

Thus this whole response is largely a response to Question G1.

Since mobile termination rates are much higher than fixed, the practical application of the arguments presented here has wider implications and urgency for mobile than for fixed. In practice fixed termination rates are already quite low because the line rental is paid by the subscriber.

1 A global and historical perspective

Cost based termination has been accepted as a principle in Europe since the early days of liberalisation. Although it is accepted in some other parts of the world, it is not accepted in many significant countries and the evidence that will be presented later in this paper is that in the case of mobile these other markets operate more efficiently.

Cost based termination developed in the early days of liberalisation from the payment mechanism used internationally between countries where accounting rates were paid to the destination country. Within the international telecommunications community it was recognised that many poorer countries would set accounting rates above cost as a means of obtaining revenue from richer countries. This imbalance still exists in some places. For example the cost of calls from the UK to a particular Gulf State is more than twice the cost from that state to the UK. These were early examples of the exercise of monopoly power in termination.

When liberalisation started in the UK in the mid 1980s, BT proposed to copy the international system, which was not unreasonable as there was no other system to copy. Since the regulator needed to control the rate it seemed obvious to control it on the basis of costs. Our argument is that although this seemed reasonable it was an undesirable model to adopt and there is now much greater understanding of network economics that demands a different approach.

Other countries outside Europe to adopt a different approach as is shown by the following slide produced by Tim Kelly of ITU.

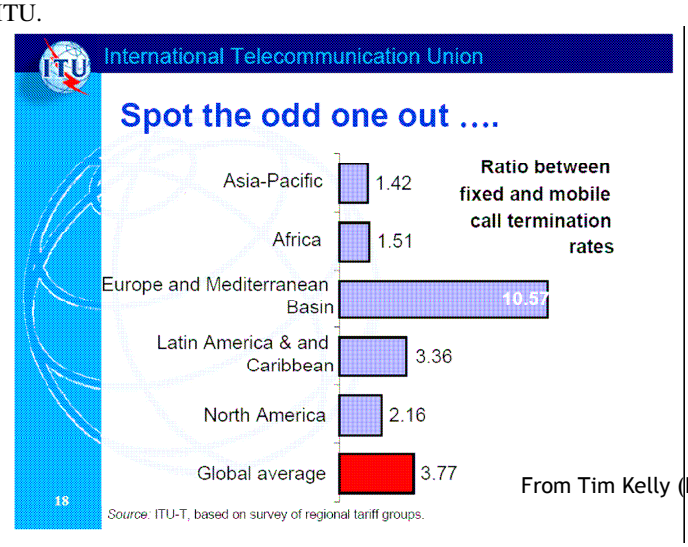


Figure 1: Ratios between fixed and mobile termination rates

These differences alone should provoke a more radical re-assessment of the position in Europe.

2 Whose service?

The underlying concept behind cost based termination is that the terminating operator is providing a service to the originating operator. The demand for this service arises from the attempt by the caller to make the call. It follows then that if the termination price needs to be regulated it should be regulated to cost. If this is the correct or most appropriate concept, then cost based termination would be the correct approach. We believe that this service concept is not appropriate and that this inappropriate concept is the underlying issue.

We believe that the appropriate service concept is that both operators are providing jointly a service to both the calling and called parties. If this is the appropriate concept then the case for cost based termination is destroyed and other totally different payment models need to be considered.

In the paper "Do mobile operators have a dominant position in a market for the wholesale termination of calls from fixed to mobile?"[1] Dr Evans argues that regulators are treating termination as a single sided market and that this is incorrect, since it is a double-sided market. This is another way of saying that the service concept is wrong and we agree with this conclusion.

Dr Evans then argues that two-sided markets do not need regulating, and he gives many examples of such markets that operate satisfactorily and where the price paid between the two sides is far different from cost. However the markets that he quotes are more of a client-host nature (eg a shopping mall), and so intrinsically different from the telecommunications market for client-client services. In the market that he quotes, there is normally adequate external competition, which is largely absent in telecommunications. For example, he quotes shopping malls where the shops and the mall owner cooperate successfully together without regulation to attract subscribers. These analogies are poor because shoppers can normally use different malls. In contrast, in telecommunications the only way to reach a given caller party is via the terminating operator that that party has chosen. Telecommunications is a special and possibly unique form of a two-sided market that is dominated by the value of connectivity.

[1] Published by Vodafone, Policy Paper Series No 7, Nov 2007

Thus we agree with Dr Evans that call termination is clearly not a single sided market and so there is no reason why any needed regulation of termination rates should be based on cost rather than any other approach, but we do not agree that there is no case for regulation.

3 The case for regulation

Regulation is a remedy for market failure. The sign of market failure is the ability to raise prices without losing market share. We believe that the situation of unregulated termination rates is even worse - the prices are unstable and there is strong incentive to raise rates to gain market share and some operators are adopting this strategy.

What is the instability effect? The effect is that an operator can raise its termination rate and thus generate more revenue from other operators without affecting its own customers. It can then use the revenue to offer deeper on-net discounts to attract more customers. The increase in customers generates more incoming call minutes and more revenue. Thus in control systems terms there is positive feedback and instability. This is shown in the following diagram:

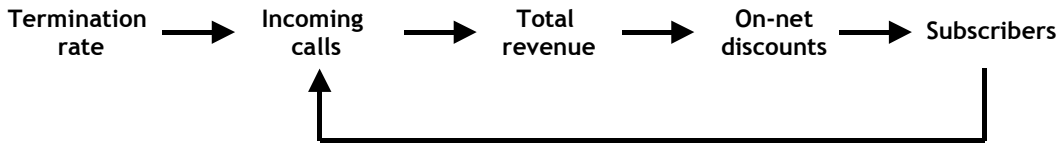


Figure 2: The instability loop

If unregulated termination prices are unstable, as we argue they are, then one might expect complete chaos and rates rising towards infinity. This has not happened because there are some other effects and non-linearities, and in most cases at least the rates of the operators that are dominant in the retail market are regulated.

The existence of the instability effect is commonly observed. Many new entrants adopt the policy of on-net discounts and high termination rates, and use it as much as they dare without attracting regulatory intervention. Where new entrants have experimented with lower termination rates as a means to attract subscribers they have nearly always found it ineffective and changed to a high termination rate strategy. This has demonstrated that customers are not sufficiently sensitive to the termination rates of their own operator to counteract the instability effect.

Thus the case for regulating all termination rates is based on market failure, which is manifested in the incentive and ability to raise termination rates to attract more subscribers through on-net discounts.

If mobile operators had voluntarily adopted a system where raising termination rates tended to reduce rather than increase market share then there would not be market failure; but they have not done so. Possibly they are now unable to do so since the necessary discussions could attract accusations of collusion and so the operators are locked into the current system.

4 Technology neutrality

Many regulators adhere to the principle of technology neutrality in regulation. Cost based termination is in conflict with this principle since costs are technology dependent. If termination rates are regulated to cost then the incentive to adopt a cheaper technology is reduced.

5 Which economic principle?

When economics are applied to telecommunications prices, as for example in the treatment of charges for number portability and carrier pre-selection, there is normally a discussion about which economic principle to apply. The principles are:

- Cost minimisation
- Cost causation
- Distribution of benefits
- Promotion of competition
- Reciprocity
- Practicability

Cost causation and a misapplication of distribution of benefits have been accepted as the basic principles that justify cost based termination. The application of the other principles such as cost minimisation, promotion of competition, practicability and a proper application of distribution of benefits point to a different approach.

The argument for cost-based termination is based on the following:

- Distribution of benefits: The main beneficiary of the call is the caller, and so the caller should pay and this payment flows through to the terminating end;
- Cost causation: The prime cause of the call termination is the caller and therefore the principle of cost causation demands call termination charges.

The application of distribution of benefits is obviously flawed and the principle is not applied correctly. The benefits of a call depend on the contents of the call and vary from call to call. They may lie 99% at one end or 99% at the other or be distributed between the ends.

The application of cost causation is also flawed if the wider context of the call is taken into account as the call may be made in response to a request to be called.

Furthermore, call termination charges are time based and yet either party may terminate the call. Therefore the total cost of the call after initial connection may be determined by either end, and so the principle of cost causation does not point wholly to the call being paid by the caller.

Thus we conclude that:

- neither the distribution of benefits nor cost causation justify payment by the caller for the whole of the call, and so they do not justify cost based termination
- there is no economic absolute for having duration-related cost-based call termination charges - other economic principles can apply and point to other solutions.

Consequently a new debate is needed to achieve better charging arrangements for the future and ones that better reflects all the relevant economic principles.

6 Inconsistency between fixed and mobile

Furthermore, it is worth noting that cost based termination is not applied consistently between fixed and mobile. In fixed networks the called subscriber pays for most or all of the local access for incoming as well as outgoing calls, and so the called subscriber pays for part of the termination costs. In contrast in mobile networks the calling party pays for all the termination.

This difference has arisen largely because of the separation between fixed costs and traffic dependent costs, where regulators aimed to have the fixed costs of a fixed network paid by subscription once the tariffs were fully rebalanced. In contrast, all the costs of a mobile network are traffic dependent.

This difference is highly significant because a large part of the cost of modern networks is the access system.

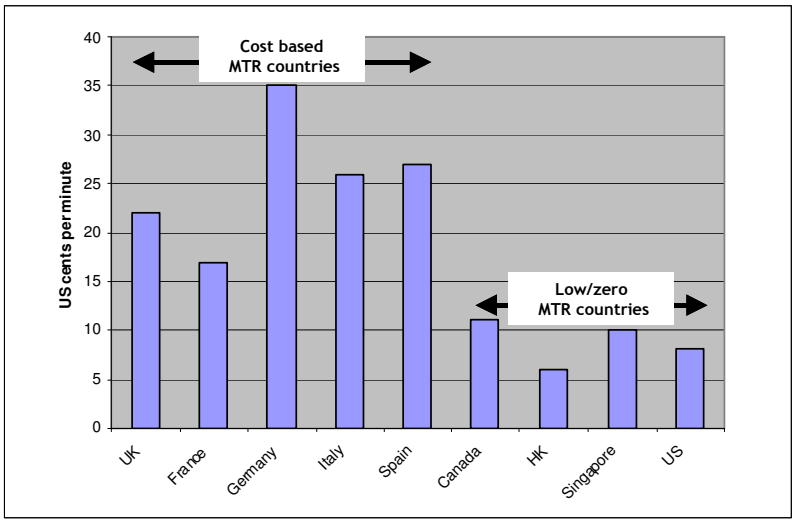
If mobile networks were treated in the same way as fixed networks with the mobile access element of call termination paid by the called party out of their subscription or out of the higher outgoing charges then the termination rates would fall significantly even on a cost basis.

7 What should be the basis for termination rate regulation?

We have argued that cost should not be the basis, then what should be the basis? We are unable to propose a theoretically justifiable answer. From theory one can only argue that the rate should be fixed and symmetrical.

However experience has shown that a low to zero rate has proved more effective in other parts of the world and a zero rate has proved highly effective for services on the Internet such as email. Low symmetrical rates also avoid many other problems that regulators have encountered. There is therefore a strong pragmatic argument for moving towards low symmetrical rates and this approach is consistent with some of the economic principles mentioned earlier.

Figure 3 compares the mobile revenue per minute of use between groups of countries with (high) cost based mobile termination and ones with low/zero rate termination.



Source: Merrill Lynch Wireless Matrix Q404 and Regulation at Ovum for MTRs

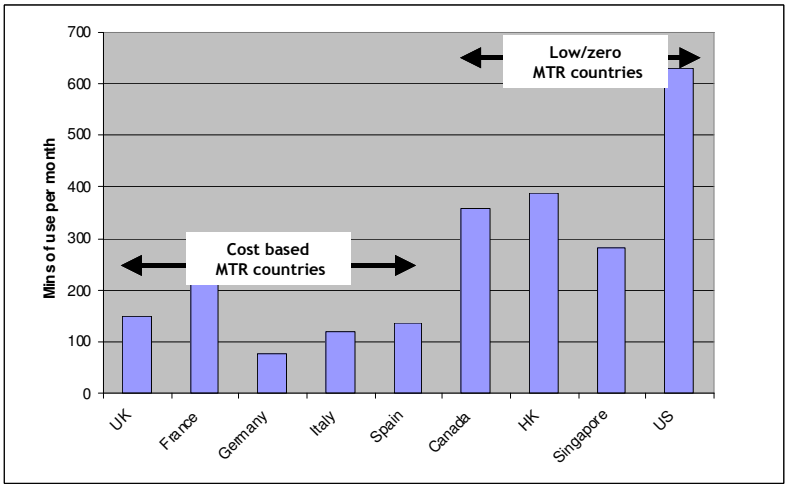
Figure 3: Average mobile revenue per minute of use

In Figure 3:

- The revenue includes both retail revenues received from the mobile subscriber and call termination revenues received on inbound calls
- The minutes of use include both outbound and inbound minutes.

The average revenue per minute is a good measure of the average prices faced by both mobile subscribers and fixed subscribers calling to mobiles (given that the price component for the fixed element of a call is typically only a small part of the total price paid).

Figure 4 shows the minutes of use per month in the same countries.



Source: Merrill Lynch Wireless Matrix Q404 and Regulation at Ovum for MTRs

Figure 4: Mobile minutes of use per month

The figures show that countries with low/zero mobile termination rates have lower prices involving mobile and consequently greater volumes of calls. The strong correlation suggests that this is because of the low/zero termination rates.

This comparison raises an important question - if mobile termination rates drop, will mobile prices have to rise? If demand were inelastic then the answer would be yes, although this would not necessarily be a problem since most of the benefit of mobility goes to the mobile party. However the figures from the low termination rate countries suggest that thus will not be the case. Operators in these countries are profitable with higher volumes and lower rates. This suggests that the reduction in termination rates, which will be reflected in prices, will lead to a significant growth in call volumes and this growth, because of economies of scale, will make up for the lost revenue and potential reduction in profitability. Thus low mobile termination rates may well lead to a win-win situation.

8 Symmetry

If the linkage to costs is broken, then the case for asymmetry is largely eliminated. The argument for asymmetry is normally based on the differences in cost of different technologies or the additional cost of the early years of operation of a network.

The benefit of symmetry is that calls to similar networks are likely to have similar prices and so number portability does not cause problems of tariff transparency. Furthermore the inter-operator billing is simplified.

9 Low rates

The simplest arrangement would be to move to very low termination charges that would be the same in principle for all networks.

This would give the following advantages:

- All revenue for termination would have to come from the subscribers of the terminating network and so be fully competitive. Low termination rates would increase competitiveness;
- Regulators would no longer have to determine rates and could withdraw from this part of the market;
- The regime would become technology neutral;
- The main obstacle to fixed-mobile convergence would be removed;
- Incentives to invest in new technology would increase.
- Prices for calls to fixed and mobile would converge so number portability between fixed and mobile networks would become practicable and would no longer be prevented by the need for tariff transparency.

We believe that the lowering of prices will generate more income through increased volume due to price elasticity and that since there are economies of scale the potential loss in revenue may be more than offset by expansion of the market. The operators in the low termination rate countries are profitable at higher volumes and lower prices. There is the prospect of a win-win situation for mobile.

For fixed networks the practical effects would be minor since the cost of the access is paid for normally by a line rental and termination rates are already quite low in some countries. Flat rate retail tariffs are becoming popular and fixed operators would benefit from becoming able to bundle calls to mobiles in their flat rate packages.

10 Is this “Bill and Keep” or “Sender keeps all”? - who pays transit?

The interconnection charging regimes known as “Bill and Keep” or “Sender keeps all” are used in the Internet, eg for email, and are used with subtle modifications in some countries. Taken literally, these schemes are NOT exactly what we are advocating. There is also confusion that they imply the US system of the called mobile party paying to receive calls on a per minute basis. We are not advocating this.

The existence of a termination rate acts as an incentive for the originating operator to carry the call as far as is cost effective. A pure “Bill and keep” scheme inverts this incentive resulting in “hot potato” routing

where calls are passed to other operators as close to the origination as possible. It would be hugely disruptive to change the direction of the incentives and this is not proposed. Consequently we think that any transit costs should still be paid by the originating operator.

For fixed networks with geographic number structures, the local call termination rate would be reduced to a low level and probably best expressed as a flat rate port charge. Single and double transit termination would take account of the costs of transit for the typical distances concerned. In many countries transit is now a competitive market and so the double transit termination charge need not be regulated. Single transit might remain an area of significant market power and continue to need to be regulated.

Where the numbering does not follow a geographic structure then an average distance element would need to be added to take this into account. This situation would apply for example to mobile networks where the location of the subscriber is not known by the calling network.

11 Carrier selection and pre-selection

Carrier selection is a technique for creating competition in call handling. Local loop operators are required to provide call origination services at cost plus some mark-ups. The rationale is quite different from that of termination because the issue is about creating scope for competition whereas termination is about connectivity. Call origination would still need to be regulated, and a cost basis would continue to be justified. Since the revenue system of the operator would change as the termination rate reduces, call origination charges would increase.

12 A practical migration path

If it is accepted that there is no absolute theoretical justification for cost based termination, then a practical approach can be taken to establishing a migration path to a new regime.

Concern about the reduction of revenues from fixed to mobile operators probably means that mobile operators will not accept the changes on a voluntary basis as they will lose the revenue that they have enjoyed from fixed subscribers for so long. Consequently regulators will need to start to regulate the mobile termination rates down on a glide path towards the level of the fixed operators. In order to avoid sudden changes, this could take place over several years. Once the mobile charges have reached the level of national fixed termination charges, fixed termination can be reduced if they are not already sufficiently low.

The gradual migration path will enable regulators to monitor the effects of these changes and see to what extent lower prices lead to increased volumes and create a win-win situation. If the win-win does not occur then the regulators are able to stop the process or reverse it. This is a low-risk not a high-risk strategy.

The following would be a logical end of the process:

- Local fixed termination: A capacity based port charge, lower than current termination.
- Regional fixed termination (single transit): A capacity based port charge plus a cost based regional transit charge set by reference to costs and the market levels, ie no detailed modelling
- Nationwide fixed termination (double transit): Unregulated if the transit market is reasonably competitive
- Mobile termination: A charge based on the average cost of transit across the country

13 Recommendation

We recommend that the ERG should look even more radically than the current consultation at the fundamental characteristics of the telecommunications call market, and consider moving away from the cost based termination concept to one of low symmetrical termination rates. We think that low termination rates, especially for mobile, could expand the market substantially to everyone's benefit resulting in a win-win for all.

John Horrocks
24 January 2008
+44 1483 797807
john@horrocks.co.uk