

EN

EN

EN



COMMISSION OF THE EUROPEAN COMMUNITIES

Brussels,
(2008)

Draft

COMMISSION STAFF WORKING DOCUMENT

EXPLANATORY NOTE

Accompanying document to the

COMMISSION RECOMMENDATION

of [...]

on the Regulatory Treatment of Fixed and Mobile Termination Rates in the EU

This is a draft document which does not necessarily represent the official position of the Commission.

TABLE OF CONTENTS

1.	Introduction.....	4
2.	Rationale for regulating fixed and mobile call termination markets.....	5
2.1.	General competition issues in fixed and mobile termination markets.....	5
2.2.	Rationale for cost-based pricing.....	7
3.	Commission decisional practice/European Regulators Group (ERG) experience.....	7
3.1.	Key insights from the Article 7 procedure to-date.....	7
3.1.1.	Type of price control.....	7
3.1.2.	Cost models used.....	8
3.1.3.	Symmetry of remedies	9
3.2.	Some practical experience to-date as reported by the ERG.....	10
4.	Common principles for regulating termination markets	12
4.1.	Common principles in relation to cost determination.....	12
4.2.	Common principles for symmetry/asymmetry of termination rates	16
5.	The application of cost-based remedies	18
5.1.	Fixed networks	18
5.1.1.	Choice of technology	18
5.1.2.	Definition of the increment	19
5.1.3.	Efficient scale of operators.....	20
5.2.	Mobile networks.....	20
5.2.1.	Choice of technology	20
5.2.2.	Definition of the increment	21
5.2.3.	Efficient scale of operators.....	22
5.2.4.	Externalities.....	22
5.2.5.	Implication of recommended approach for mobile termination rates.....	23
6.	Forward-looking consideration	24
6.1.	Possible alternative approaches.....	24
6.1.1.	Introduction.....	24
6.1.2.	Bill and Keep.....	24
6.1.3.	Reciprocity	25

6.1.4.	RPP.....	25
6.1.5.	Conclusion.....	26
6.2.	Migration to IP Interconnection.....	26
7.	Transition period.....	27
	ANNEX.....	29

EXPLANATORY NOTE

1. INTRODUCTION

This document provides the background to the Recommendation on the Regulatory Treatment of Fixed and Mobile Termination Rates in the EU. A key observation during the assessment of approximately 770 notifications under Article 7 of the Framework Directive¹ concerns inconsistencies in the application of remedies to voice call termination markets². Although some form of cost orientation is provided for in most Member States, a considerable divergence between average termination rates, particularly as regards mobile termination rates, still exists across Member States³.

Additionally, National Regulatory Authorities (NRAs) have, in a number of cases, authorised higher termination rates for smaller fixed or mobile operators on the grounds that these operators are new entrants into the market and have not benefited from economies of scale and/or are subject to differing cost conditions. These asymmetries still exist, although they are slowly decreasing⁴. Furthermore, the absolute level of termination rates remains high in a number of Member States, thus continuing to translate into high, albeit decreasing, prices for end-consumers.

A number of inconsistencies in the regulation of mobile call termination rates have also been identified by the European Regulators Group (ERG)⁵, in particular in relation to the form of price regulation, treatment of asymmetries and the implementation of glide paths.

The above indicates significant differences in the regulatory treatment of terminating operators both within and across national boundaries. The distinct approach taken in different Member States as regards market players operating in similar conditions is difficult to justify.

The lack of harmonisation in the application of cost-accounting principles to termination markets to-date demonstrates a need for common guidelines and a common approach which will provide greater legal certainty and the right incentives for potential investors. It will also reduce the regulatory burden on existing operators

¹ Article 7 of Directive 2002/21/EC of the European Parliament and the Council of 7 March 2002 on a common regulatory framework for electronic communication networks and services (“the Framework Directive”), OJ L 108, 24.4.2002, p. 33.

² See the Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on market reviews under the EU Regulatory Framework (2nd report), COM(2007) 401 final of 11.7.2007.

³ These differences are illustrated in Annex 1 below.

⁴ According to the European Regulators Group Common Position on symmetry of fixed call termination rates and symmetry of mobile call termination rates (ERG (07) 83 final 080312) (“ERG Common Position on symmetry”), average asymmetry of mobile termination rates (within individual countries) decreased from 1.4 €-cents in January 2004 to 0.9 €-cents in January 2007. The ERG has recognised in its Common Position on symmetry that termination rates should normally be symmetric and that asymmetry requires an adequate justification.

⁵ See the ERG Common Position on symmetry.

that are currently active on a pan-European basis. The objective of coherent regulation in termination markets is clear and recognised by the NRAs.

This common approach builds on the decisional practice of the Commission to-date and is set out in the Recommendation. The objective of the Recommendation is to define and set out clear common principles on:

- (a) the regulation of cost-oriented fixed and mobile termination rates in the EU, including common principles on the concepts of an efficient operator and symmetric regulation; and
- (b) the identification and calculation of efficient costs consistent with those incurred in a competitive market.

The Recommendation also considers how termination rates might be regulated in a changing technological environment, e.g. in the presence of Next Generation Networks (NGNs).

The Recommendation furthermore considers approaches other than cost-based regulation of termination rates. These alternative approaches may help alleviate the competitive and regulatory issues inherent in the Calling Party Pays (CPP) convention.

The remainder of this document is structured as follows:

- Chapter 2: Rationale for regulating fixed and mobile call termination markets
- Chapter 3: Commission decisional practice/ERG experience
- Chapter 4: Common principles for regulating termination markets
- Chapter 5: The application of cost-based remedies
- Chapter 6: Forward-looking considerations
- Chapter 7: Transition period.

2. RATIONALE FOR REGULATING FIXED AND MOBILE CALL TERMINATION MARKETS⁶

2.1. General competition issues in fixed and mobile termination markets

Call termination can only be supplied by the network provider to which the called party is connected. There are currently no demand- or supply-side substitutes for call termination on an individual network. Therefore, each network constitutes a separate relevant market and each network operator has a monopolistic position on the market for terminating calls on its own network.

⁶ This assessment of competition problems in call termination markets is largely based on the prevailing interconnection arrangements in the fixed (PSTN) and mobile telephone networks.

Moreover, under the prevailing CPP principle in the EU, the calling party pays entirely for the call, and the wholesale termination rate paid by the originating operator is normally passed on to its end customer. As the called party is not billed for incoming calls, it is generally indifferent to the termination charge set by its network provider (i.e. the terminating operator) and has little or no incentive to change its own network provider in the event that those charges are raised⁷.

Consequently, in the absence of other factors such as countervailing buyer power, the criteria necessary to merit *ex-ante* regulation are normally met, and the terminating operator is designated as having significant market power (SMP).

The main potential competition concern common to both fixed and mobile termination markets is that of *excessive pricing*, implying that operators may extract excessive profits at the wholesale level. Moreover, fixed and mobile terminating operators are vertically integrated into retail calls markets and compete with their wholesale customers on those markets. Consequently, terminating operators have incentives to raise rivals' costs by setting termination prices at a level that impedes their ability to compete in downstream retail markets.

Termination has been analysed as a situation of “two-way” interconnection whereby two wholesale prices have to be negotiated and each operator could potentially use the price charged for termination on its own network as leverage in the relevant negotiations. This may lead to efficient rates being negotiated, particularly among symmetrically sized networks, which is more likely in mobile markets. This type of interaction may, however, still facilitate anti-competitive behaviour in the form of excessive pricing. High termination charges may be used to foreclose a new entrant network, where the majority of originated calls are off-net. High termination rates may also facilitate collusive behaviour between two or more terminating operators.

In the past, negotiations between fixed and mobile operators typically evolved differently because mobile operators could raise the initially unregulated mobile termination rates without experiencing a reciprocal increase in the often tightly regulated fixed termination rate (FTR). This raised allocative-efficiency concerns where there is an implicit cross-subsidy from fixed network operators and their customers to mobile operators and mobile customers.

Furthermore, with the evolution of fixed–mobile hybrid services and a move towards convergence, a different regulatory treatment of fixed and mobile termination rates raises a possible inconsistency issue. The regulatory model underlying the FTR regulation assumes that operators will recover the cost of the local loop via retail subscription charges, and that these costs are not included in the FTR paid by other operators, including mobile operators. This is not the case in mobile networks where the access network costs are largely recovered via the termination rate. This needs to be considered in order to ensure that competitive distortions do not arise and that allocative-efficiency concerns as described above are addressed.

⁷ A different rationale applies to numbers used by Service Providers (SP). A called SP is sensitive to the level of termination charges — which directly affect its revenues — and may therefore switch between providers of termination services.

2.2. Rationale for cost-based pricing

In the light of the ability and incentives of terminating operators to raise prices substantially above cost, cost orientation is the most appropriate intervention to address this concern over the medium term. Cost orientation addresses both productive- and allocative-efficiency concerns. From a productive-efficiency perspective, low termination rates facilitate low retail call charges and higher consumption. It is also important that the relevant price is based on the costs of an *efficient* operator. If the regulation of termination charges was based on the *actual* costs of the operator, this would not provide the right incentives for operators to innovate and increase efficiency, as their inefficiency would be covered by their competitors. This will also give rise to allocative-efficiency concerns as customers of other operators would ultimately bear the costs of the inefficient operators.

Allocative efficiency suggests that one group of customers should not subsidise another group of customers. Apart from the fixed-to-mobile cross-subsidisation outlined above, this is also relevant within markets (e.g. in mobile markets). Late entrants argue that due to large traffic imbalances and on-net/off-net price differentiation they cannot compete effectively at the retail level. A large proportion of calls originated on late entrant networks is terminated on other networks, i.e. off-net. If new entrants pay a regulated termination charge in excess of actual costs they effectively give a transfer to the large network. As a result, their ability to offer retail rates comparable to the retail rates of an established operator, which terminates a majority of its calls on-net, is impeded.

3. COMMISSION DECISIONAL PRACTICE/EUROPEAN REGULATORS GROUP (ERG) EXPERIENCE

3.1. Key insights from the Article 7 procedure to-date

Any Recommendation regarding greater harmonisation of regulation in the EU must be guided by regulatory experience as well as by Commission decisional practice. In line with the Commission Recommendation on relevant markets, all NRAs of the EU25 have notified the markets for fixed and mobile call termination⁸ and imposed *ex ante* obligations on all SMP operators. Regulatory practice has demonstrated, however, that NRAs do not employ a consistent set of remedies in these markets. Differences exist in regulating different operators within a Member State, and across Member States. This has led the Commission to comment *inter alia* on three principal sources of inconsistencies: the type of price control, the cost model used, as well as the issue of asymmetric termination rates.

3.1.1. Type of price control

One NRA decided not to impose price regulation on alternative operators, citing their higher cost of call termination, their significantly smaller scope of operation than that of the incumbent operator, the decreasing termination rates of alternative operators,

⁸ At the time of writing, Romania and Bulgaria have not yet notified all of their wholesale termination markets.

absent regulation and their limited asymmetry in comparison with the incumbent⁹. In this instance, the Commission invited the NRA to impose effective price regulation also on the alternative operators if the downward trend of unregulated fixed termination rates did not continue, or if the asymmetry with the incumbent's rates increased. Similarly, when an NRA decided not to impose obligations of cost orientation, cost accounting and accounting separation¹⁰, the Commission stated that some form of cost control, such as benchmarking against a larger operator which is under a cost-orientation obligation, should also be imposed on smaller operators.

Where an NRA decided that the imposition of cost-orientation and cost-accounting obligations may be disproportionate¹¹, the Commission reiterated the need for a cost control to be imposed on smaller operators, e.g. by benchmarking against a larger operator whose termination rates are cost-oriented. The Commission has also noted that a glide path towards an efficient rate should be established without delay as any grace period could remove the incentive to become cost-effective as quickly as possible¹². The Commission also encourages an NRA to complement the imposed cost-orientation remedies by an appropriate *ex ante* price control obligation supported by an appropriate cost-accounting methodology¹³.

3.1.2. Cost models used

In its responses under the Article 7 procedure, the Commission has noted the importance of regulating termination rates based on the costs of an efficient operator. The Commission has also encouraged NRAs to develop cost models which take into account the necessity for alternative operators to become efficient over time¹⁴. At the same time, the Commission has acknowledged that these models could reflect different cost structures of different operators caused by objective cost differences¹⁵.

In several cases the Commission indicated the necessity of also imposing an obligation of accounting separation, which would allow internal transfers to be visible. The Commission found that imposition of accounting separation as a separate measure would facilitate effective price control, increase transparency and decrease the risk of cross-subsidisation¹⁶.

As regards the selection of the appropriate type of cost model, the Commission has encouraged an NRA to impose a cost-calculation obligation and to assess whether a forward-looking long-run incremental cost model (LRIC) would not be the most appropriate model for calculating termination rates, notably in terms of tariffs, and potential excessive costs and inefficiencies of the mobile operators¹⁷. In addition, the Commission indicated that it is important that LRIC models use current costs and not historical costs which risk over-estimating the appropriate costs¹⁸.

⁹ Cases PL/2006/0502, PL/2007/0633, PL/2007/0641.

¹⁰ Case FI/2003/0029.

¹¹ Case FR/2005/0228.

¹² Case IE/2008/0746.

¹³ Case FI/2006/0403.

¹⁴ Cases IT/2006/0384, PL/2006/0502, EE/2007/0598.

¹⁵ Cases IT/2006/0384, EE/2007/0598, LV/2006/0464.

¹⁶ Cases PL/2006/0379, DE/2005/0234 and DE/2006/0421.

¹⁷ Case PL/2006/0379.

¹⁸ Case UK/2006/0498.

In terms of costs included in the relevant cost model, the Commission has noted that the value of 3G licences should be calculated at current value on a forward-looking basis and not on the basis of spectrum values which approximate past levels. In that respect, termination rates should be set at the cost which would be faced by an efficient operator providing the relevant service. The Commission stated that this consideration was particularly relevant for spectrum fees which had been written off by operators since the relevant frequencies had been auctioned, and for which there may be an overstatement of the opportunity cost of 3G spectrum. Therefore, with a view to allowing end-users to obtain the benefits of regulation, the Commission invited the NRA concerned to reconsider the valuation of 3G licences¹⁹. The broader issue of how spectrum should be treated within the cost model was not addressed.

3.1.3. *Symmetry of remedies*

In several cases, the Commission has stated that in circumstances where a NRA intends to impose different remedies on different operators within similarly defined markets, such differential treatment should be adequately reasoned²⁰.

More specifically, the Commission considered that termination rates should normally be symmetric and that asymmetry requires an adequate justification. The Commission recognised that in certain exceptional cases asymmetry might be justified by objective cost differences outside the control of the operators concerned. Such possible justifications could be objective network cost differences, for instance owing to cost differences between the operation of a GSM900 network and a DCS1800 network²¹, or substantial differences in the date of market entry²². However, the Commission has also emphasised that the fact that an operator entered the market later and that it therefore has a smaller market share can only justify higher termination rates for a limited transitory period. The persistence of a higher termination rate would not be justified after a period long enough for the operator to adapt to market conditions and become efficient over time, and could even discourage smaller operators from seeking to expand their market share²³.

The Commission has also commented upon traffic imbalances in the context of mobile termination markets by stating that such traffic imbalances may in fact be caused by the current asymmetric level of mobile termination rates, as well as by an on-net/off-net retail price differentiation which is within the control of the operators. The Commission also stressed the importance of reducing termination rates to the level of costs of an efficient operator, which would take into account objective cost differences²⁴.

¹⁹ Case UK/2006/0498.

²⁰ Cases FI/2003/0029, FI/2003/0031, HU/2005/0152, DK/2005/0207, FR/2005/0228, IT/2006/0384, PL/2006/0502, EE/2007/0598.

²¹ However, in cases BE/2006/0433 and LV/2006/0464 the Commission stated that it expects the differences related to technology to be small.

²² Cases DE/2006/0421, BE/2006/0433, FR/2006/0461, LV/2006/0464, LV/2007/0574, FR/2007/0596, ES/2007/0654, IT/2007/0659.

²³ Cases DE/2006/0421, AT/2006/0544, LV/2007/0574, FR/2007/0596, EE/2007/0598, BE/2006/0433, FR/2006/0461, FR/2007/0669.

²⁴ Case FR/2007/0669.

Finally, the Commission, indicating the EU-wide importance of regulating mobile termination rates effectively and in a consistent manner, has in multiple cases encouraged the NRAs to work in close cooperation with the European Regulators Group in order to arrive at a coherent EU approach, as well as to revisit the NRAs' analysis in the light of a common approach as soon as this has been established. In this respect, relevant aspects of the work of the ERG on a common position on the regulation of both fixed and mobile call termination — as reported to-date — are also presented here.

3.2. Some practical experience to-date as reported by the ERG

The ERG Common Position on symmetry of fixed call termination rates and symmetry of mobile call termination rates adopted on 28 February 2008 (ERG Common Position on symmetry) helps to illustrate some inconsistencies observed in the NRAs' implementation of remedies in fixed and mobile termination markets to date.

A number of inconsistencies in the regulation of mobile call termination rates have been identified. According to ERG data, 21 out of 28 countries that provided information in response to an ERG questionnaire²⁵ indicated that they imposed a cost orientation obligation on at least the first mobile operator having entered the market. For later entrants, the price control obligation could sometimes take the form of a “non-excessive” or “fair and reasonable” price rule. A wholesale price cap was imposed in some countries, although not necessarily on all mobile operators.

In addition, significant variety was noted in respect of the cost models already in place. According to the ERG Common Position on symmetry:

- top-down accounting data was used by eleven NRAs as the main tool and by two NRAs as a complementary tool;
- a bottom-up model was used by two NRAs as the main tool while one NRA was developing it;
- a hybrid model (bottom-up model calibrated with data provided by Mobile Network Operators (MNOs)) was used by seven NRAs as the main tool and by one NRA as a complementary tool, while three NRAs were developing it; and
- international benchmarking was used by eight NRAs as the main tool and by five NRAs as a complementary tool²⁶.

Furthermore, even where NRAs chose the same costing tool, the ERG noted differing practices in implementing those models. For example, in relation to top-down models, the ERG observed large disparities in the way top-down accounting data are first produced and then how they are checked/verified. With regard to bottom-up modelling, the ERG also noted a large disparity with regard to the way depreciation is implemented in the model.

²⁵ The ERG questionnaire was also sent to non-EU Member States.

²⁶ Two NRAs, Hungary and Poland, have two main tools.

Finally, with regard to the definition of an “efficient” operator (whose charges are used as a reference target for the model, especially those models whose remit spans a number of years), the ERG acknowledged a large variety of definitions chosen by the NRAs (including the lowest cost of all the MNOs, the highest costs of the MNOs, an average or a weighted average of the costs of all the MNOs, the cost reference of an efficient operator, the actual costs of each operator as well as a benchmark).

As regards fixed call termination, ERG noted that a different, more stringent set of remedies is usually imposed on the incumbent operators as compared with the remedies imposed on alternative operators.

The NRAs have imposed the full scope of remedies set out in the Access Directive on the fixed incumbent operators²⁷. However, differences between Member States in implementing cost orientation are observable²⁸: although in most cases the termination rates are regulated on the basis of a LRIC model, a Fully Allocated Cost (FAC) model or other means of regulation are also applied. Moreover, Current Cost Accounting (CCA) is most commonly, but not exclusively, used for calculating FTRs. As a result, the different application of the same regulatory tool produces diverse results.

The diversity of methods is also apparent in the regulation of termination rates for fixed alternative operators. One of the following regulatory approaches is usually applied:

- requiring reasonable prices or forbidding excessive prices;
- adding a mark-up to the incumbent’s fixed termination rates;
- benchmarking the termination rates of alternative operators against the charges of the incumbent operator (higher rates may be approved on the basis of cost calculation);
- imposing symmetry gradually, after a “glide path” — i.e. the difference (asymmetry) between the termination rates of the incumbent and of an alternative operator is progressively decreased, so that both become equal (symmetric) at a given point in time;
- imposing delayed reciprocity where alternative operators’ termination rates are set equal to the incumbent’s termination rates but lagged by a specified number of years.

Theoretically, symmetry may also be achieved in the latter case if the incumbent’s termination rates do not change over several years.

²⁷ However, the following exceptions were noted by the ERG: one NRA has not imposed an obligation of transparency, but transparency followed, however, from the obligation to publish a reference offer; two NRAs have not imposed an obligation of accounting separation, but in one case it is stipulated by national law in the event an *ex ante* price control obligation is imposed.

²⁸ See, for example, ERG Report — Regulatory Accounting in Practice, 2007 (ERG (07)22).

Finally, in some cases no *ex ante* price control was imposed on alternative operators²⁹.

In conclusion, as a consequence of the diverse approaches taken on regulating both mobile and fixed termination rates, these rates differ more between Member States and between operators than may be justified by different national circumstances or by exogenous cost factors.

4. COMMON PRINCIPLES FOR REGULATING TERMINATION MARKETS

4.1. Common principles in relation to cost determination

Article 8(2) of the Framework Directive requires NRAs to promote competition by *inter alia* ensuring that all users derive maximum benefit in terms of choice, price and quality, and that there is no distortion or restriction of competition. In relation to these obligations the Commission has previously emphasised that termination rates should be brought down to the costs of an efficient operator as soon as possible. As outlined in section 3.1 above, the Commission has also encouraged NRAs to develop cost models which take into account the necessity for alternative operators to become efficient over time and which take into account the costs of an efficient operator. In addition, the Commission has encouraged NRAs to assess whether a forward-looking LRIC model would not be the most appropriate model for calculating termination rates, notably in terms of potentially excessive tariffs and inefficiencies of operators.

A key regulatory decision relates to the appropriate cost base for calculating an efficient operator's costs, and the question arises as to which cost base is more in line with the above-stated regulatory objectives. Today regulators may use either the costs actually incurred by the regulated company (historic costs) or the costs that would be incurred if a notional network would be built today (current costs). While both approaches can, in principle, be used to satisfy the efficiency objective, the current-costs approach is more compatible with the competitive standard of efficiency, since in a competitive market prices would be set on the basis of the prevailing technology. In a competitive environment, operators would compete on the basis of current costs and would not be compensated for costs incurred through inefficiency; neither should high-cost operators be allowed through regulation to pass on their inefficiencies to final consumers. Operators that are compensated for actual costs incurred have few incentives to increase efficiency. In these circumstances, the operator that was able to terminate calls more cheaply would not be the operator to benefit from the efficiency gains. On the contrary, it would be the less efficient (competing) operator that would pay the lower termination charge and thereby gain an undue competitive advantage.

Final consumers also stand to gain from the use of current costs. Termination charges are expected to be lower using a current-cost base due to the impact of technological improvements in relation to the core network, where most of the termination costs are incurred. This gain in consumer surplus is unlikely to be outweighed by the fact that assets already depreciated in the past may under a current-cost methodology be

²⁹ For example Poland, Denmark.

included again. These costs primarily concern the access network which is less relevant for the calculation of termination charges.

The choice of the appropriate cost base is also related to the choice of cost model, i.e. whether a top-down (TD), bottom-up (BU) or hybrid model is used. In a TD model the starting source of information is the cost actually incurred by the operator. Consequently TD models are said to avoid disincentives to invest, since incurred costs are usually allowed to be recovered, even if this does not necessarily promote efficiency.

BU models use demand data as a starting point and determine an efficient network capable of serving that demand by using economic, engineering and accounting principles. BU models give more flexibility regarding network efficiency considerations and reduce the dependence on the regulated operator for data. A BU model is synonymous with the theoretical concept of developing the network of an efficient operator because it reflects the equipment quantity needed rather than actually provided and the model ignores legacy costs. A BU model does not ensure that all actually incurred costs are eventually recovered from the regulated service.

Also, BU models may understate the costs where technologies are rapidly changing and operators cannot instantaneously change their technologies.

Although BU models are generally developed by NRAs, operators can contribute to the model inputs and assumptions. This will increase the transparency and objectivity of BU models, although it carries the risk that ‘negotiated’ figures, as opposed to more accurate figures, will be used in the model.

Given the fact that a bottom-up model is based largely on derived data, e.g. network costs are computed using information from equipment vendors, regulators may wish to reconcile the results of a BU model with the results of a TD model in order to produce as robust results as possible and to avoid large discrepancies in operating cost, capital cost and cost allocation between a hypothetical and a real operator. The purpose of the reconciliation is to show and to quantify the sources of differences between both models, thus assisting in the verification of the BU model. This may be appropriate, for example, where there is an information asymmetry or a risk of certain cost categories being erroneously omitted. However, any modification of the BU model must take into account the necessity of showing the costs of an efficient operator; it should not be done merely to bring the results of both models closer.

Concerning cost standards, the Commission has stated³⁰ that the long-run incremental cost (LRIC)³¹ methodology is consistent with cost orientation. LRIC is normally based on forward-looking cost (FL-LRIC). “Forward-looking” is a term which is used interchangeably with current cost.

Standard economic theory determines that prices be set equal to marginal costs. However, should this be applied in the telecommunications sector, operators would

³⁰ Recommendation of 8 January 1998 on interconnection in a liberalised telecommunications market (Part 1 — Interconnection pricing), (98/195/EC), OJ L 73, 12.3.1998, p. 42.

³¹ The forward-looking long-run incremental cost provides an analytical framework which can be used to obtain an estimate of the cost that would be found in a competitive market.

not be in a position to fully recover their joint costs (that can be directly attributed to more than one specific service) and common costs (which are not directly attributable to specific services). These cost categories constitute the majority of costs in fixed telecommunications.

LRIC is conceptually between marginal cost and stand-alone cost. This is achieved by considering the long run (as opposed to the short run for marginal costing) and rendering all assets variable, i.e. assuming that they can vary in response to demand. Additionally, instead of taking into account an additional unit of output, LRIC considers an additional increment. The increment can be defined narrowly, as a small change in the volume of a particular service, or broadly, as the addition of a whole group of services, with many possible increments of different size. By considering the long run, LRIC also facilitates efficient recovery of costs relevant to the defined increment which in the short run are regarded as fixed.

Depending on the size of the increment, only costs associated with the services included in the increment would be allocated to that increment. If, for example, there was only one increment including all services provided by an operator, then LRIC would cover all costs and, in fact, be equivalent to Fully Allocated Cost (FAC). If smaller increments are chosen, a LRIC model requires the allocation of joint costs and a decision on an appropriate cost-allocation mechanism. LRIC does not include common costs but regulators often apply a mark-up to account for common costs.

In this respect, it is important to note that LRIC provides scope for discretion as to how certain regulatory objectives are most effectively met. Under the current regulatory framework, the primary mechanism for ensuring that users derive maximum benefit in terms of choice, price and quality is competition.

The main advantage of low interconnection charges based on an incremental-cost approach which allocates only efficiently incurred costs is that it promotes efficient production and consumption and minimises potential competitive distortions. It sends correct signals to originating operators as to the costs generated by their activities and they can therefore adjust their behaviour in the most efficient manner. For example, allowing network costs to be recovered from the wholesale termination rate which do not result directly from the provision of that service can lead to distorted signals and higher prices for the originating operators and, consequently, their consumers. In effect, this results in them cross-subsidising the investment costs of other operators' networks and may also result in a sub-optimal number of calls being made. Currently, there are net transfers of wholesale termination revenues from:

- fixed network operators to mobile network operators, creating an effective cross-subsidy between fixed and mobile markets and consumers; and
- net senders to net receivers of voice traffic, which can reinforce network effects and increase barriers to smaller operators expanding within markets.

Furthermore, this can promote inefficiency on the part of the terminating operator.

Basing the regulated wholesale termination charge on the incremental cost of providing this service alleviates such distortions, for example, by reducing the

magnitude of cross-subsidies between groups of customers and by reducing the impact of any financial disadvantages arising from traffic imbalances between smaller and larger operators. Thus, termination rates which approximate the long-run incremental cost of providing the service can be expected to lead to enhanced competition and lower retail tariffs across the range of consumers while still facilitating efficient cost recovery and appropriate investment incentives. Furthermore, it may be claimed that high termination rates charged on a per-minute price basis create pressure on operators to adopt per-minute retail tariffs, thereby limiting the possible emergence of more innovative offers such as those based on flat-rate tariff structures which could in turn promote greater retail consumption.

When deciding on the correct level of the regulated wholesale termination rate, it is essential to ensure that the methodology adopted promotes efficient production and consumption decisions, and minimises any artificial transfers and distortions between competitors and consumers. Therefore, regulators should construct models which set wholesale termination charges as close to marginal cost as possible. The closer the termination price of all operators is to the marginal cost, the more likely it is that this will lead to the most efficient and least distortionary use of call termination services, and minimise the risk of problems such as cross-subsidisation between operators and customers and inefficient pricing and investment behaviour. Therefore, it is justified to apply a pure LRIC approach where the relevant increment is the wholesale call termination service and which includes only those costs that would not be incurred if that service were no longer produced (i.e. avoidable costs).

A pure LRIC approach implies the exclusion of costs which do not vary in response to an increase in wholesale termination traffic. For example, where costs arise which are common across all services and which do not increase in response to an increase in wholesale termination traffic, these common costs should not be allocated to the regulated voice call termination service.

It should also be noted that the existing system of cost allocation used for cost orientation of wholesale termination rates in the EU, i.e. Calling Party Network Pays, assumes that the calling party is the only party causing costs to arise. However, it is important to recognise that both calling and called parties jointly cause a call to be made and jointly benefit from that call. If the receiver did not receive a benefit, they would not accept the call. In that respect, call termination differs from other markets where the creation of costs and attribution of benefits can be ascribed to one side only. The use of traditional cost-causation principles in setting cost-oriented prices suggests that the creator of the costs should bear those costs. Given the two-sided nature of call termination, not all related termination costs must necessarily be recovered from the wholesale charge levied on the originating operator. Even if wholesale termination rates were set at zero, terminating operators would still have the ability to recover their costs from non-regulated retail services. Rather it is a question of how these financial transfers are distributed across operators in a way that best promotes economic efficiency to the benefit of consumers. However, for the purposes of this Recommendation, it is proposed that all of the incremental (avoidable) costs of providing the wholesale termination service to third parties may be recovered via the regulated wholesale termination charge.

4.2. Common principles for symmetry/asymmetry of termination rates

As the relevant cost standard for setting termination rates should be BU LRIC which reflects the cost of an efficient operator, there should in principle be no asymmetries between the rate of the established operator(s) and the rates of later entrants to the market. This is broadly consistent with the ERG Common Position on symmetry which states that termination rates should normally be symmetric and that asymmetry requires adequate justification.

A key argument frequently used in support of the authorisation of temporary asymmetric rates in favour of later entrants is that it forms part of an overall *entry assistance policy* which is aimed at promoting new entry and longer-term competition in fixed and mobile markets. The rationale is that allowing higher post-entry profits will encourage entry and investment and lead to more intense competition in the long run. However, it is generally accepted that such a policy may also attract inefficient entry. It may also be expected that consumers will end up paying higher retail prices than would otherwise be the case in a situation of cost-based symmetric termination rates. In addition, providing a mark-up for new entrants while regulating incumbents at cost effectively creates a cross-subsidy and can simultaneously reduce the incumbents' investment incentives.

In the light of the above, it is questionable whether asymmetric termination rates should be used as a form of entry assistance. On the contrary, it may be argued that symmetric price control based on an efficient-cost benchmark, rather than on the costs actually incurred by an operator, gives efficient investment incentives to firms. These considerations apply to both fixed and mobile markets.

Arguments relating to *economies of scale* and the higher unit costs initially incurred by new entrants have in particular been raised as possible justification for transitory asymmetry in termination rates. As regards the extent to which new entrants might be expected to have higher unit costs than incumbents, it has been argued that this consideration is more relevant for mobile than for fixed operators. Fixed operators have the opportunity to build their networks in a particular geographic area and focus on higher density routes. Furthermore, they can lease relevant network services from the incumbent to reduce the fixed costs of network build and thereby reduce the impact of economies of scale. It has been argued, however, that scale economies play a bigger role in mobile networks: due to coverage requirements new entrants initially incur higher per-unit costs arising from their smaller customer base, although there is some debate regarding the magnitude of the costs involved.

However, rewarding an operator for its smaller size can give inappropriate investment signals and risks promoting inefficient entry. Such a policy may, for example, act as a disincentive to smaller operators to innovate and expand. Furthermore, the fact that many fixed and mobile markets in the EU have been liberalised for some time now raises a question as to the efficiency of the operators in question if they still cannot survive in the market without some form of entry assistance. In that respect, the Commission has previously stated that the persistence of a higher termination rate would not be justified after a period long enough for an operator to adapt to market conditions and become efficient, and could even discourage smaller operators from seeking to expand their market share.

A further (albeit related) argument cited in support of temporary asymmetry is the existence of *traffic imbalances* between larger incumbent operators and smaller new entrants. Where a new market entrant initially has lower traffic volumes than the more established incumbents, this can result in net payments to the incumbents which are typically net receivers of traffic. It is further argued that the financial disadvantages which new entrants face as a result of their lower traffic volumes can be exacerbated by *differential on-net/off-net pricing* policies pursued by the incumbent operators.

It is difficult to see how arguments regarding financial imbalances resulting from differences in traffic volumes and differential on-net/off-net pricing would justify setting asymmetric termination rates. This is because asymmetric wholesale pricing is likely to reinforce the asymmetric pricing observed at retail level. That is, the off-net retail prices of the incumbents will likely rise to compensate for the increased cost of off-net wholesale termination to the new entrants. As long as traffic imbalances persist, asymmetric pricing will likely only contribute to perpetuating any resulting financial imbalances.

Moreover, it has been argued that on mobile markets there may be exogenous cost factors associated with *uneven spectrum assignments* which result in per-unit cost differences between mobile operators. Exogenous cost differences may arise where spectrum assignments have not taken place using market-based mechanisms but on the basis of a sequential licensing process where, for example, later entrants mainly receive 1800 MHz frequencies, thus facing higher unit costs in certain areas than operators with a 900 MHz allocation. Due to the better propagation characteristics of the 900 MHz spectrum, it is argued, for example, that in urban areas fewer base stations are needed to ensure indoor coverage than is the case with the 1800 MHz spectrum.

The extent of this cost disadvantage depends on a number of factors, including the regulatory situation, the nature of the demand for coverage and the geography and topology of the country. It appears that this relative cost disadvantage decreases as the market shares of the later entrants grow, increasing their capacity needs. In addition, where the spectrum assignment takes place through a market-based mechanism such as an auction or where there is a secondary market in place, any frequency-induced cost differences become more endogenously determined and are likely to be significantly reduced or eliminated. Moreover, with further spectrum liberalisation taking place, it needs to be examined whether on a forward-looking basis additional spectrum is likely to be made available through market-based assignment processes which might erode any cost differences arising from existing assignments. For example, the digital dividend is leading to the release of spectrum that is being freed up as a result of the switchover from analogue to all-digital television. The spectrum that will be released by the digital switchover is in the prime Ultra High Frequency (UHF) band. Since these bands are located in the lower spectrum range they can cover large geographical areas with relatively few base stations, offering nationwide network rollout at lower costs when compared to services delivered at higher frequencies, offering greater capacity but at shorter range.

An important argument for symmetric termination rates at the level of efficient cost is that asymmetric pricing can foster inefficient behaviour and generate productive

inefficiencies. Productive efficiency takes place when a good is produced at the lowest cost possible. Rewarding an operator with a price above an efficient or cost-based level can reduce its incentives to innovate and minimise costs. For example, asymmetries based on differences in dates of market entry and scale may discourage innovation and cost efficiency on the part of the later entrant/smaller operator, and may give rise to inappropriate investment incentives and inefficient entry.

Consequently, consumers may end up paying higher prices than would otherwise be the case in a situation of cost-based symmetric termination rates. This is because the higher termination rates have to be recovered by the originating operators and will presumably be passed onto consumers in the form of higher retail prices. This effectively creates a cross-subsidy from lower-cost operators and their consumers to their less efficient rivals, thereby generating allocative-efficiency concerns. Meanwhile, the less efficient operator benefits from the lower termination rates of its rivals, thus enabling it to lower its retail prices and win customers. As the subsidised operators expand, the negative impact on retail prices and consumer welfare is even greater. Given that the stated purpose of the regulation of wholesale termination charges is to prevent excessive pricing and its negative impact on consumer welfare, it is arguably counter-intuitive to apply a remedy that also generates allocative and productive inefficiencies.

5. THE APPLICATION OF COST-BASED REMEDIES

Following the choice of the appropriate cost base, cost model and cost standard (i.e. a BU LRIC model based on current costs), this section deals with the implementation of the chosen model in practice. The first question deals with the choice of technology, from that follows the definition of the increment. Since both fixed and mobile termination rates are generally subject to regulation, and since fixed and mobile networks are to an increasing extent in competition with each other, attention must be paid to consistent treatment of both network types.

5.1. Fixed networks

5.1.1. Choice of technology

From a forward-looking perspective, a new operator would choose a packet-switched network with all services delivered over an IP core network. Hence, a BU model built today should assume that the *core* network is firmly NGN-based. Specifically, this implies that all existing PSTN switches should be assumed to be replaced with NGN equivalents and that Synchronous Digital Hierarchy/Asynchronous Transfer Mode (SDH/ATM) transmission equipment becomes redundant. It also implies that voice traffic needs to be converted to/from IP packets at the edges of the network. Whilst connecting operators still interconnect via Time Division Multiplexing (TDM) technologies, there will be a need to include Media Gateways in the bottom-up model in order to interconnect with operators using PSTN-based equipment.

Technology developments in the *access* network may focus around the shortening of the local loop by partial replacement with fibre to the kerb or street cabinet (using Very-High-Rate Digital Subscriber Line (VDSL) technology), the replacement of the copper local loop with Fibre-To-The-Building (FTTB), or the replacement of the copper local loop with Fixed Wireless Access (FWA) equipment.

In principle, the concept of forward-looking costs would value all assets at the cost of a modern equivalent asset (MEA), which is the lowest cost asset with the latest available and proven technology. While it can be argued that an operator constructing a brand-new, nationwide access network today would install fibre directly to the home, the level of investment necessary to replace the existing copper-based access network with fibre on a nationwide basis precludes it from being a MEA for the twisted copper pair. Furthermore, FWA technologies are not likely to form a long-term replacement for the twisted copper pair since they appear to be deployed only in specialist cases. The question of how the cost of deploying fibre to the street cabinet should be treated in the LRIC model will largely depend on the increment chosen and is addressed in the next section.

5.1.2. *Definition of the increment*

As stated in section 4.1., LRIC models include only those costs which are caused by the provision of a defined increment. This increment can contain single or multiple services or network components. Incremental costs can also be considered as the costs that would be saved if those services were no longer produced. Costs that span more than one increment are either joint or common costs. The smaller the increment, the larger the proportion of costs which are joint or common and *vice versa*.

The relevant incremental cost (i.e. avoidable costs) of the wholesale call termination increment is the difference between the total long-run costs of an operator providing its full range of services and the total long-run costs of that operator not providing wholesale call termination service to third parties.

In this context a distinction needs to be made between traffic- and non-traffic-related costs to ensure the appropriate attribution of those costs. The relevant costs for the calculation of the regulated wholesale termination charges are the traffic-related costs which are only attributable to wholesale voice call termination services. Other costs, i.e. those traffic-related costs attributed to other services (e.g. call origination, data services, IPTV, etc.) and non-traffic-related costs are not to be taken into account.

The default demarcation point between traffic- and non-traffic-related costs is typically where the first point of traffic concentration occurs. In a PSTN network this is normally deemed to be the upstream side of the line card in the (remote) concentrator. The broadband NGN equivalent is the line card in the Digital Subscriber Line Access Multiplexer/Multi-Service Access Node (DSLAM/MSAN). The deployment of fibre to the street cabinets and the installation of active devices (DSLAM or MSAN) at that network level might be seen as an extension of the traffic-sensitive part of the network. The logic behind this is that under current technology the loop is customer-specific and not traffic-dependent. In a Next Generation Access (NGA) network the former loop between the cabinet and the exchange/Main Distribution Frame (MDF) becomes a shared medium and might not be treated as being subscriber-driven, but rather as being traffic-sensitive.

The demarcation point between traffic- and non-traffic-sensitive costs in an NGA context is subject to considerable uncertainty. Certain regulators have decided that traffic is *amalgamated* at the cabinet and not concentrated, and that the fibre capacity between the cabinet and MDF is dedicated to each subscriber for the purposes of

voice traffic. The existing demarcation point (i.e. at the line card) also remains unchanged if operators dedicate sufficient capacity for voice traffic to ensure there is never congestion. However, if operators would opt to prioritise traffic rather than dedicate capacity, then in a fibre-to-the-building scenario the traffic-sensitive part of the network could move closer to the final consumer.

For the time being it can be assumed that an efficient forward-looking network will allocate dedicated capacity to the voice channel irrespective of the technology deployed. Hence, the existing demarcation remains unchanged, unless there are significant NGA developments inducing an observable general trend towards using shared capacity, which would be reflected in the regulated access regime.

To facilitate accurate identification of avoidable costs that should be attributed to the wholesale call termination service, it may be appropriate to allocate operators' costs in the first instance to business segments/services other than wholesale voice call termination, with only the residual cost being allocated to the wholesale call termination increment. Given that wholesale call termination is a traffic-related service, non-traffic-related costs should not be taken into account when calculating wholesale termination rates. Then, it may be appropriate to first identify among the traffic-related costs (discussed above) those that are related to other services, such as data traffic (e.g. broadband, leased lines, IPTV, etc.) and other relevant voice services (e.g. call origination), and to develop cost-volume relationships according to which costs can be allocated to those services. When the costs for all other services have been calculated and attributed, only then should the remaining costs be allocated to voice call termination service. As a consequence, the incremental costs of call termination are only those costs that can be avoided if the call termination service were no longer provided (avoidable costs).

5.1.3. Efficient scale of operators

It is particularly difficult to determine minimum efficient scale for fixed networks due to a number of factors. Firstly, in fixed networks operators have the ability to rent infrastructure and to purchase interconnection. Secondly, fixed operators have the opportunity to build their networks in a particular geographic area and focus on higher-density routes. Consequently, fixed operators can potentially achieve low unit costs at low levels of output and thereby reduce the impact of economies of scale.

When deciding on the appropriate efficient scale of the modelled operator, NRAs should take into account the need to promote efficient entry, while also recognising that under certain conditions smaller operators can produce at low unit costs by operating in smaller geographic areas. Furthermore, smaller operators which cannot match the largest operators' scale advantages over broader geographic areas can be assumed to purchase wholesale inputs rather than self-provide termination services.

5.2. Mobile networks

5.2.1. Choice of technology

Just as in fixed networks, a forward-looking perspective would imply that all services will be delivered over an IP core network. A BU model built today should assume that the core network is firmly NGN-based. Similar issues arise in relation to the mobile access network as compared to the fixed access network. In the same way as

fibre to the node or to the home is replacing copper, so too are 3G- or UMTS-based technologies gradually replacing 2G. Some very important differences remain. In mobile networks economic conditions driven by demand concentration and geographic characteristics influence the selection of a range of spectrum-based technologies to match those conditions. It can be expected that 2G and 3G networks are likely to co-exist for a number of years. Hence, the model should be based on both 2G and 3G employed in the access part of the network to reflect the actually anticipated situation facing operators, while the core part is firmly NGN-based.

5.2.2. *Definition of the increment*

As in fixed networks, for the purposes of calculating the incremental costs of wholesale call termination in mobile networks, it is necessary to identify only those costs that would not be incurred if the wholesale termination services were no longer provided to third-party operators (i.e. the avoidable costs only). The avoidable costs of the wholesale call termination increment may be calculated by identifying the total long-run cost of an operator providing its full range of services and then identifying the long-run costs of this same operator in the absence of the wholesale call termination service being provided to third parties. This may then be subtracted from the total long-run costs of the business to derive the defined increment.

Furthermore, as for the fixed network, a broad distinction will need to be made between traffic- and non-traffic-related costs to ensure the appropriate attribution of those costs.

The costs of the handset and the SIM card are not traffic-related and should be excluded from any costing model for wholesale voice call termination services.

Although there is no equivalent to the local loop or the line card as in the fixed access model, there is a requirement to provide certain minimum coverage to subscribers, and this will cause certain subscriber-related costs to be incurred to meet that requirement. Those costs should consequently be treated as non-traffic-sensitive costs and should not be attributed to the wholesale call termination increment. Coverage can be best described as the capability or option to make a single call from any point of the network at a point in time. Coverage costs would, for example, include site preparation costs, the base station cost and the first transceiver cost of coverage sites. Investments in mature mobile markets are largely driven by capacity increases. Capacity represents the additional network costs which are necessary to carry increasing levels of traffic (above the minimum necessary to offer a retail service to subscribers). These capacity costs can be regarded as traffic-related and may fall within the wholesale call termination service increment to the extent that those capacity requirements are driven by the provision of a wholesale call termination service. The incremental cost of wholesale voice call termination services should therefore exclude coverage costs, but should include additional capacity costs to the extent that they are caused by the provision of wholesale voice call termination services.

As for the fixed network, an appropriate way of accurately identifying only those residual costs which may be attributed to the wholesale call termination service may be to first allocate costs to services other than wholesale voice call termination, with only the residual being allocated to the wholesale call termination increment. Given

that wholesale call termination is a traffic-related service, it may be appropriate to identify from the traffic-related cost category those other services (e.g. data, SMS, MMS, call origination, etc.) which also fall within that broader cost category, and to develop cost-volume relationships according to which costs can be allocated to those services. When the costs for all other services have been calculated and attributed, only then should the remaining costs be allocated to voice call termination services. As a consequence, the incremental costs of call termination are only those costs that can be avoided if the call termination service were no longer provided (avoidable costs).

The costs of spectrum usage (the authorisation to retain and use spectrum frequencies) incurred in providing retail services to network subscribers are initially driven by the number of subscribers, and thus are not traffic-driven and should not be calculated as part of the wholesale call termination service increment. The costs of acquiring additional spectrum to increase capacity (above the minimum necessary to provide retail services to subscribers) for the purposes of carrying additional traffic resulting from the provision of a wholesale voice call termination service should be included on the basis of forward-looking opportunity costs, where possible.

Following the approach outlined above, only the additional network capacity needed to transport additional traffic deriving from the provision of the wholesale call termination service to third-party operators is included in the increment (e.g. additional network infrastructure to the extent that it is driven by the need to increase capacity for the purposes of carrying the additional wholesale traffic), as well as the additional spectrum costs and wholesale commercial costs directly related to the provision of the wholesale termination service to third parties. This implies that coverage costs, general business overhead costs and retail commercial costs are not included.

5.2.3. *Efficient scale of operators*

To determine the efficient scale for the purposes of the cost model, the recommended approach is to set that scale as follows: $1/\text{Number of Mobile Infrastructure Operators}$. The rationale for this is that the number of market players is to a large extent determined by regulatory policy, such as spectrum assignment, as well as by competitive factors. It may be expected that operators, having entered the market, would strive to maximise efficiency and revenues and thus be motivated to achieve greater market share. Thus, one could reasonably expect that an efficient operator would or could achieve a minimum market share of $1/\text{Number of Mobile Network Operators}$.

5.2.4. *Externalities*

It is argued that in the presence of *network externalities*, the addition of a marginal subscriber to a mobile network may also be of value to other subscribers. For example, other fixed and mobile subscribers derive a benefit from being able to contact and be contacted by this additional subscriber. The externality arises because the benefit to other subscribers is not taken into account when the decision of whether or not to join a network is made. Thus a sub-optimal number of customers may choose to become network subscribers. Consequently, it is argued that it may be appropriate for wholesale termination charges to include an externality mark-up

above cost which may then be used by the operators to subsidise the addition of marginal subscribers to their networks with associated benefits for all consumers calling those networks. However, this argument relies on a number of assumptions. The first is that the network operators are not in a position to internalise this externality in the absence of a mark-up. It also assumes that there is a direct and complete pass-through of the wholesale termination profits to marginal subscribers at retail level rather than being retained by the relevant operator as excess profits, i.e. that there is a complete waterbed effect. Furthermore, this argument in favour of a network externality mark-up is also based on the premise that customer penetration levels are not yet near saturation levels, as otherwise network externalities would be largely exhausted. There is evidence that where call externalities have not been applied, network operators have still acted to bring and maintain marginal subscribers onto the network. There is also insufficient evidence as regards the magnitude of the waterbed effect and both current and foreseeable market developments in terms of penetration levels. Therefore, a network externality does not seem sufficient justification to allow for an increase in termination rates.

5.2.5. *Implication of recommended approach for mobile termination rates*

The recommended approach for setting termination rates constitutes a significant departure from the currently applied cost-accounting principles where a majority of joint and common costs have traditionally been allocated to termination services. This methodological shift will have an impact on the level of termination rates, in particular in mobile networks.

When examining the cost structure of mobile operations, it can be noted that on average around 75% of the costs of mobile call termination are currently network-related, slightly more than half of which are generated by the radio access network. According to the recommended approach, only those costs which are capacity-driven and incremental to the provision of a wholesale call termination service would be taken into account. Furthermore, the remaining 25% of the total cost of mobile call termination is typically accounted for by spectrum costs, business overheads and wholesale commercial costs. Under the recommended approach only that portion of spectrum costs driven by the need to increase capacity, above the spectrum necessary to provide retail services to subscribers, for the purposes of carrying wholesale voice call termination traffic for third parties would be included. Cost models currently applied by NRAs treat a large proportion of the radio access network as traffic-driven and therefore a sizeable proportion of the radio access network costs are taken into account in calculating the costs of providing termination services. Under the recommended approach only a reduced proportion of the radio access network costs would be allocated to the overall mobile termination cost. Business overhead costs would be regarded as a common cost and therefore excluded. Furthermore, wholesale commercial costs which are directly related to the provision of a wholesale call termination service would be included; however of the latter cost categories, wholesale commercial costs are typically the smallest in magnitude.

6. FORWARD-LOOKING CONSIDERATION

6.1. Possible alternative approaches

6.1.1. Introduction

As noted above, call termination services are two-sided, with the network(s) being the platform and the caller and receiver being on either side of that platform. The demand elasticities on either side of the platform mean that the *structure* of prices impacts on the levels of consumption; therefore, it often plays a crucial role in bringing the two sides of the market together³².

It has also been noted that in a call there are benefits to both the calling and to the called party, which in turn suggests that both parties have a part in the creation of costs. Currently, the benefit of the called party is largely ignored when regulating termination rates³³. The consideration of call externalities raises issues about how costs ought to be recovered in a forward-looking context.

In addition, there are non-trivial costs associated with developing cost models for the setting of wholesale termination rates. These costs need to be considered in the context of possible alternative mechanisms for remunerating termination services. A number of alternative arrangements for the exchange of termination traffic are considered below.

6.1.2. Bill and Keep

A few countries³⁴ use alternative arrangements, under which network operators negotiate termination fees, subject to an obligation to interconnect and usually subject to the requirement that rates received by both networks that are parties to the same agreement are reciprocal. These operators often choose to set termination rates at zero. That system, where traffic is exchanged without financial settlements, is known as *Bill and Keep*. Bill and Keep may be related to Receiving Party Pays (see 6.1.4.), as it allows operators to directly charge their customers for received calls without resorting to wholesale charges from other operators.

There is no record of Bill and Keep being imposed by a regulatory authority. It generally results from voluntary agreement between interested parties, which in certain circumstances choose to set these fees at zero, particularly where the net financial settlements are equal to or close to zero.

It is argued that Bill and Keep obviates the need for regulatory intervention and resolves the termination bottleneck. Moreover, it is further argued that Bill and Keep leads to lower retail prices for call origination and appears to increase usage due to the price elasticity of demand. Furthermore, proponents of Bill and Keep consider that it facilitates development of innovative offers, e.g. flat-rate offers promoting increased usage. It also brings immediate benefits by decreasing transaction and measurement costs. Finally, Bill and Keep takes account of the call externality.

³² See Rochet/Tirole (2004).

³³ Unlike in the regulation of international roaming calls, Regulation (EC) No 717/2007.

³⁴ The USA, Canada and Singapore.

Nevertheless, one should note that setting the price of any service at zero may cause distortionary behaviour, bring arbitrage opportunities, lead to inefficient traffic routing and inefficient network utilisation. For instance, a potentially problematic issue might be inefficient routing of traffic from operators not participating in the Bill and Keep scheme.

When assessing the possible introduction of the Bill and Keep system, potential merits and drawbacks of such an approach would have to be carefully considered. Given the high current level of termination rates under the prevailing CPP system in the EU, the full effects of switching to a Bill and Keep system may not be reliably foreseen.

However, a significant reduction of termination rates from current levels might create appropriate incentives for voluntary inter-operator agreements³⁵ and consequently Bill and Keep type arrangements could evolve naturally.

6.1.3. *Reciprocity*

One of the possible regulatory approaches is to require that interconnecting operators negotiate termination rates among themselves, subject only to the requirement that these rates are reciprocal. Bill and Keep is a particular example of a reciprocal arrangement, where the termination rates are set at zero. However, other levels of reciprocal termination rates may be applied.

Potentially, this could require limited regulation, assuming that prices are negotiated at an efficient level. Nevertheless, reciprocity may lead to collusion between operators aimed at maintaining high wholesale and retail prices. Moreover, it may be expected that the outcome of reciprocal arrangements would depend on the level of traffic flows between two interconnecting networks. A net recipient of traffic would likely prefer a higher termination rate and *vice versa*. Thus, efficient termination rates do not necessarily have to result from the imposition of reciprocity. Consequently, additional regulatory intervention may be needed.

6.1.4. *RPP*

Some countries (e.g. Canada, Singapore, Hong Kong, the United States) use Receiving Party Pays (RPP) as an alternative arrangement to the CPP system at retail level. Under RPP the receiving network terminates calls without charging the originating operator the full cost of that termination service, leading the operator to potentially recover part of the termination costs from their own retail customers. Since this charge is now noticeable to the consumer, there is an incentive for the consumer to respond to that charge where more competitive alternatives exist. Thus, both incoming and outgoing call charges are subject to competition. Such a settlement system is consistent with an argument that while the calling party causes a cost originating the call, the called party causes a cost by accepting it and thus it recognises the existence of a positive call externality to the receiving party.

RPP avoids the deficiencies of the CPP system, e.g. high termination rates resulting from the monopoly on termination markets and which thus produce negative

³⁵ Voluntary agreements are always subject to competition law.

competitive consequences both at the wholesale and retail level. If subscribers are charged for incoming calls, they can be expected to be more sensitive to the price charged for them. Thus, competition between operators for mobile subscribers could be expected to exert a constraint on the setting of wholesale termination charges with associated implications for retail prices.

Nevertheless, it could possibly meet resistance from customers unwilling to meet the termination charge. RPP might not be efficient if the calling party values the call highly but the called party does not and, as a result, an efficient call might not be completed³⁶. The reverse issue may arise in the CPP system, where an efficient call may not be initiated even if the called party values it highly but the calling party does not. In addition, if the originating operator continues to cover part of the termination cost, RPP would still require a degree of regulatory oversight, as otherwise RPP would likely revert to a CPP arrangement.

As noted above, there are potential merits in an RPP system given that it recognises that both calling and receiving parties benefit from a call and contribute towards its cost. However, it is difficult to envisage such a situation at present given the current high level of termination rates in the EU. Nevertheless, it may not be excluded that RPP will emerge if operators decide to recover part of the termination charge directly from the called party, in particular RPP may evolve after a reduction of the regulated termination charge or as a response to a Bill and Keep system.

6.1.5. *Conclusion*

Further to the above, a number of possible alternative approaches may be implemented over the longer term to the extent that they may promote efficiency and decrease the need for regulation. However, in view of the current high level of termination rates, particularly in the mobile sector, it is difficult to see how these alternative systems may be introduced in the short to medium term. Reducing termination rates to an efficient level is an appropriate first step before other potential approaches may be introduced.

6.2. **Migration to IP Interconnection**

Another issue which needs to be considered with respect to forward-looking termination rates is the effect that termination arrangements are likely to have on investment and network evolution in the context of IP developments.

Currently, the most noteworthy driver of change in networks is the convergence of the network, with a single integrated IP-based network delivering a combination of data, voice and video services. This evolution makes it possible for different underlying platforms (for example, fixed telecommunications and cable television) to offer equivalent services, potentially benefiting competition. This same evolution enables bundled offers of multiple services to the end-user, thus changing the character of competition.

³⁶ However, operators may employ certain measures aimed at counteracting possible sub-optimal usage, e.g. flat rate offers with free incoming calls.

Although migration to an IP network enables a direct decrease of network costs, in addition to increasing the economies of scope resulting from an ability to offer a wider range of services, in the transition to IP certain inefficiencies may occur. Such short-term inefficiencies, resulting from the operators' own policy, should not serve as a justification for higher termination rates even for a limited period. For some operators, high termination rates represent an important source of revenue. They may therefore perceive this evolution as a threat, and possibly resist the emergence of these new forms of interconnection since they may undermine the current charging mechanisms.

Generally, IP-based interconnection (data traffic) is currently implemented by a mixture of peering and transit³⁷. With *peering*, two Internet Service Providers (ISPs) agree to exchange traffic solely among their respective customers, sometimes without payment; with *transit*, one ISP agrees to carry the traffic of a customer (possibly also an ISP) to third parties for a fee. These freely negotiated arrangements result in a globally interconnected Internet, and do not (in most cases) depend on any regulatory obligations.

If call termination fees remain at current levels, it might be that many mobile operators and some fixed operators might choose not to evolve their networks to IP-based interconnection. They might perceive the migration as a risk of losing termination revenues. This suggests that waiting for the migration to IP-based NGNs to be implemented by operators in the presence of high termination charges might be a self-defeating strategy.

Furthermore, even in the event of a move to IP-based interconnection of voice calls, the inevitable question remains as to whether interconnection of future NGNs should or could be based on the Internet economic model, on the switched network model, or some third model (possibly a blend of the two). In the presence of high termination rates, any spontaneous move from the existing charging mechanisms for voice traffic exchange seems unlikely. The conclusion must be that given the current level of termination rates, the evolution of IP interconnection is likely to be slower and that any transition to alternative charging mechanisms is likely to be significantly impeded.

7. TRANSITION PERIOD

The transition to the Recommendation raises issues for all stakeholders. Article 16(1) of the Framework Directive states that NRAs shall carry out an analysis of the relevant markets as soon as possible after adoption of the Recommendation or any updating thereof. That implies that "as soon as possible" in Article 16(1) is interpreted as respecting regulatory measures that have already been notified and agreed.

A period of transition can therefore be anticipated. Such a time period should, on the one hand, be long enough to allow regulators to put the cost model in place and for operators to adapt their business plans accordingly while, on the other hand, ensuring that consumers derive maximum benefits in terms of efficient cost-based termination

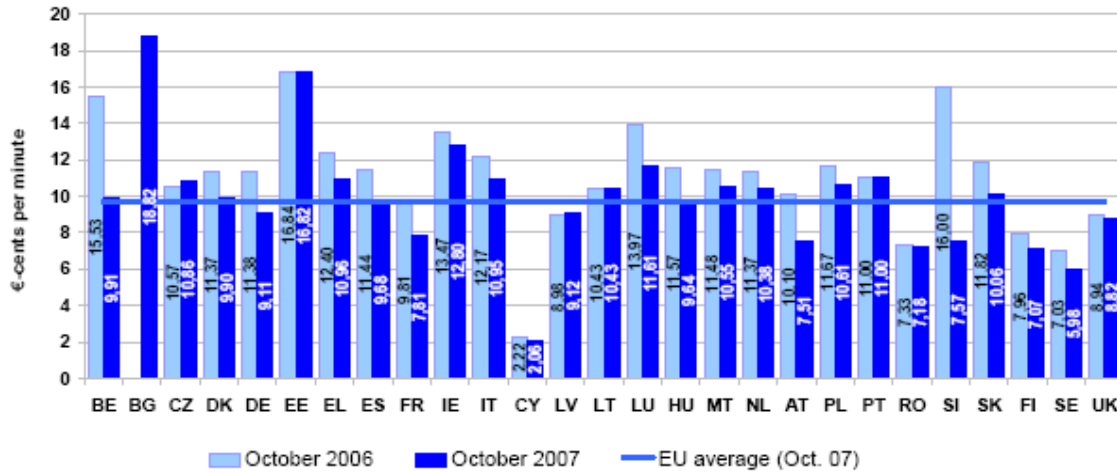
³⁷ Other IP interconnection arrangements exist, such as mutual transit, but they are less frequently used.

rates. Such a period should be limited to 31.12.2011 as of which date the NRAs should ensure that the termination rates are set in accordance with this Recommendation.

ANNEX

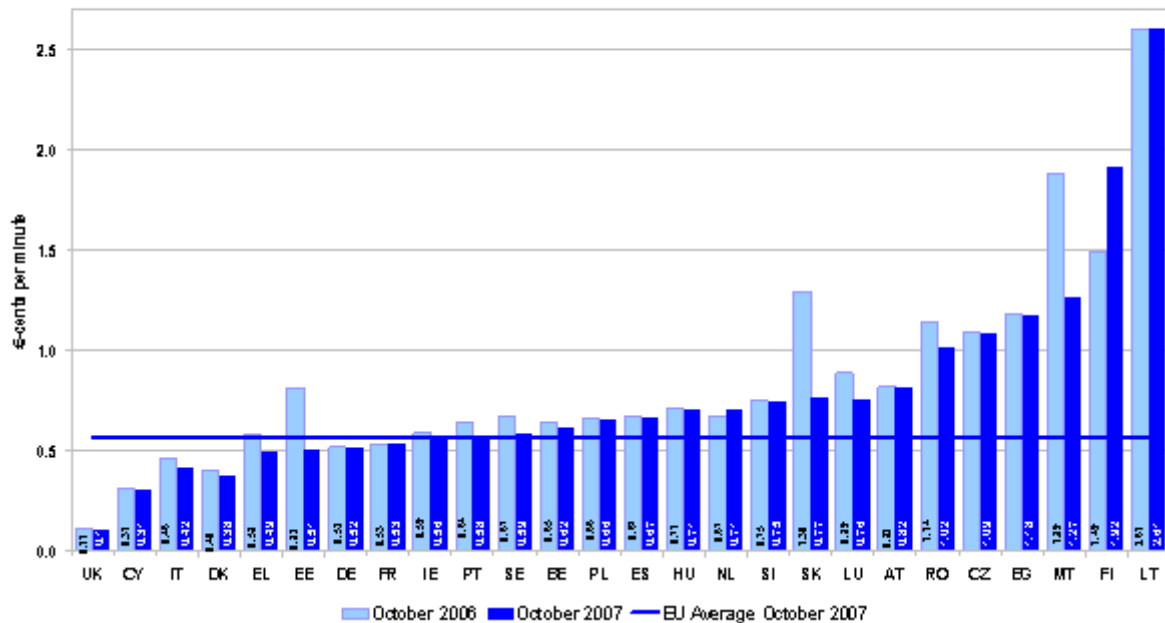
Interconnection charges for call termination on mobile networks
(national average on the basis of subscribers)

EU average Oct. 2007: 9,67 €-cents



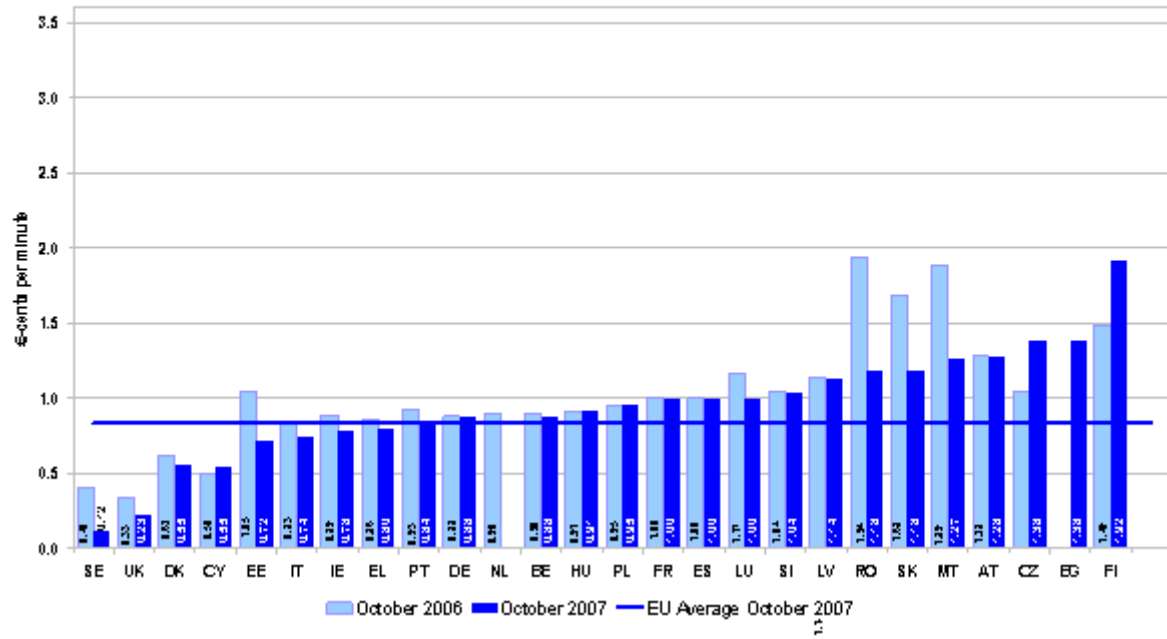
Source: European Electronic Communications Regulation and Markets, 13th Report

Interconnection charges for terminating calls on INCUMBENT'S FIXED NETWORK
(at 1/10/2007)(peak time)
Local level - EU average: 0,57 €-cents



Source: European Electronic Communications Regulation and Markets, 13th Report

Interconnection charges for terminating calls on INCUMBENT'S FIXED NETWORK
 (at 1/10/2007)
 Single transit - EU average: 0,83 €-cents



Source: European Electronic Communications Regulation and Markets, 13th Report