



Regulating mobile call termination

Foreword

I hope you enjoy our first Vodafone Policy Paper. Our aim in these papers is to provide a platform for leading experts to write on issues in public policy that are important to us at Vodafone. These are the people that we listen to, even if we do not always agree with them. These are their views, not ours. We think that they have important things to say that should be of interest to anybody concerned with good public policy.

Arun Sarin, Chief Executive, Vodafone

Contents

	Page
Foreword	00
– Arun Sarin, Chief Executive, Vodafone	
Introduction	01
– Stewart White, Group Public Policy Director, Vodafone	
About the essays	02
– Richard Feasey, Director of Public Policy, Vodafone	
The size of fixed common costs in mobile networks: empirical evidence from Europe	03
– Alastair Macpherson, Partner, PricewaterhouseCoopers	
The importance of price elasticities in the regulation of mobile call termination	06
– Dan Elliott, Director, Frontier Economics	
Application of Ramsey pricing for regulating mobile call termination charges	11
– Professor David Newbery, Professor of Applied Economics, University of Cambridge	
Mobile to mobile call termination	19
– Professor Patrick Rey, Professor of Economics, University of Toulouse & Dr. Bruno Jullien, Director of Research, CNRS	
Price controls on mobile termination charges	25
– Professor Stephen C Littlechild, Honorary Professor, University of Birmingham Business School and Principal Research Fellow, Judge Institute of Management Studies, University of Cambridge	

Introduction

You will find that in this first Paper on the regulation of mobile call termination we have asked the authors to give their views on difficult and often complex issues. Many of these were debated by Vodafone and others before the UK Competition Commission during 2002 and some have emerged since then.

This Paper is intended to try to move the debate on call termination forward. The issues covered by the authors are very much alive today. In Europe those regulators who have not already done so must review the mobile call termination market under the new EU legislative framework. In Australia the ACCC is revisiting mobile. In the US the FCC has investigated charges levied on US carriers by foreign mobile operators.

One of the themes that emerges in these essays is that intervening in a competitive market is far more complex and challenging than the traditional utility regulation of the kind normally applied to monopolies in gas, electricity and fixed line telecommunications. Yet many regulators have done nothing but utility regulation in the past. With mobile, every action is more finely calibrated. The benefits of intervention are more ambiguous and the error costs larger. That does not mean it should not or cannot be done. But it means that regulators must be acutely aware of their fallibility, should be as sophisticated in their collection and handling of data as they possibly can – and should entertain the possibility that the application of regulation can get better over time. We hope this Policy Paper may contribute to that.

Stewart White, Group Public Policy Director, Vodafone

About the essays

This document begins where the UK Competition Commission report on call termination left off, so it assumes a general familiarity with the issues raised in the report. We asked a number of internationally respected academics and advisers to reflect further upon them. Although many authors focus heavily on the UK experience, we were keen to explore these issues given their relevance to debates elsewhere in the world.

We first had two empirical matters to resolve. Professor David Newbery had told us that the size of fixed costs ought not to be a matter of opinion but a matter capable of being resolved empirically one way or the other. We therefore asked PricewaterhouseCoopers, who had advised Vodafone during the UK enquiry, whether we could do better in quantifying the size of fixed and common costs in a mobile business. Whether these were large – as we held – or small, as the Commission concluded – clearly matters in any debate about how costs should be properly recovered and accounted for in setting regulated prices.

If fixed and common costs are modest, the choice of mark up methodology to recover them is unimportant. But if they are large and if the elasticities of different mobile services are different then some form of Ramsey pricing is a better way to mark up costs. We therefore asked Frontier Economics, who had also advised Vodafone during the UK enquiry, to reconsider what the evidence we have today tells us about the elasticities of mobile services.

Some commentators have accepted that Ramsey pricing is conceptually the right way to recover common costs, but argue that it is simply impractical to do so. We asked Professor David Newbery to consider that.

Vodafone have argued that mobile to mobile termination rate setting had characteristics that distinguished it from rate setting for fixed to mobile. We thought this important because we expect mobile to mobile communications to grow in future. So we concluded that if we could find a path which led to the deregulation of mobile to mobile then that would be a significant step towards less regulation in the mobile industry.

The Commission declined to take that step, although other regulators have done so in the past. We asked Professor Patrick Rey, who has written extensively on these matters, to consider further some of the objections that the Commission had made to our proposals.

Finally, some regulators in Europe are prepared to tolerate large differences in the rates they allow different operators to recover. We think this ignores the positive incentive properties of price caps. We asked Professor Stephen Littlechild, the original advocate of price controls in telecommunications in the 1980s, for his views.

These papers aim to stimulate debate and thought. The views expressed are those of the authors and cannot be attributed to Vodafone. However, we would welcome and encourage comments on the issues discussed. These should be sent to me at richard.feasey@vodafone.com

Richard Feasey, Director of Public Policy, Vodafone



Alastair Macpherson

Partner, PricewaterhouseCoopers

Alastair Macpherson is a Partner in PwC's Telecoms, Media and Technology Team, Valuation and Strategy and has extensive experience of the telecommunications industry. He is particularly skilled in the areas of regulatory strategy and economics, network and service costing, interconnect costing and pricing, accounting separation and regulatory reporting. Prior to joining PricewaterhouseCoopers, Alastair worked for British Telecommunications Plc for eleven years in a variety of roles.

The size of fixed common costs in mobile networks: empirical evidence from Europe¹

Background

In the recent regulatory UK Competition Commission inquiry into calls to mobiles it was clear that identifying the level of fixed common costs is important. For these costs to be recovered mobile service prices need to be marked up above their marginal cost.

The current view of Ofcom is that fixed common costs are small, i.e. that mobile networks enjoy few economies of scale and/or scope. The implication of this contention and the absence of other significant economies of scale or scope is that the size of fixed common and joint costs are not material and therefore prices could be set close to marginal cost.

It has been suggested that the issue of the magnitude of fixed and common costs is a matter of empirical investigation and PricewaterhouseCoopers have been asked to investigate further the actual level of these costs.

Methodology²

Using standard econometric techniques and drawing on data supplied by three Vodafone operating companies, we set out the size of the fixed common costs in mobile networks that we have estimated.

Our approach has been to estimate a statistically robust relationship between mobile equipment quantities and mobile network usage. This allows us to estimate the level of marginal

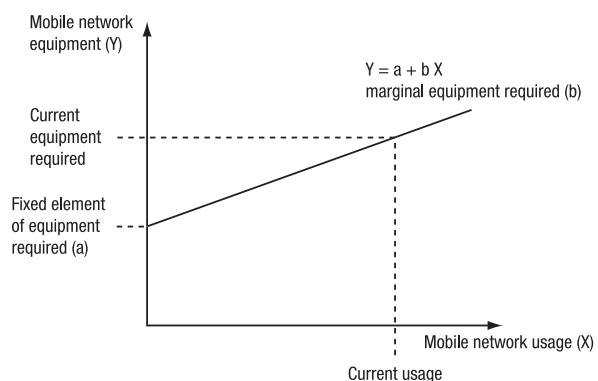
equipment required to meet an increase in demand and therefore we can estimate the level of fixed common costs.

We have focussed on the quantities of the key components of the mobile network rather than their actual costs. This avoids complications introduced by any differences in cost accounting, input prices and reporting policies both across jurisdictions and over time.

Ofcom's view is that as usage grows equipment quantities grow at the same rate. This is a testable assertion. If it can be demonstrated that the quantity of equipment grows at a slower rate than the mobile usage then this is empirical evidence that mobile networks display economies of scale as a result of fixed costs.

Results

In the diagram below we illustrate the relationship we found for each equipment type in each country:



This demonstrates the proportion of current equipment installed that is fixed and common. To turn this into a proportion of total costs that are fixed we weighted each equipment type by an indicative proportion of total costs it accounts for. In the table below we report the estimates for the three countries:

Country	% of fixed common costs in the mobile network
Greece	14%
Netherlands	23%
Spain	45%

Conclusion

The results show that there are significant fixed common costs in contrast to Ofcom's contention that these are very small. However, there is a difference between their levels across the three countries.

This is not surprising, as we would not expect the proportions of fixed common costs in each country to be the same. The main reasons for this are firstly, each country has a different geographical setting. Secondly, the distribution of demand across the country will impact the mix of coverage- and capacity-related equipment to meet total demand. Thirdly, the operational process of network rollout is unlikely to be the same across different countries with different strategies around how much capacity to build incrementally given the forecast availability of network equipment at different times and forecasts of likely demand.

Appendix 1: Methodology

To estimate the level of fixed common costs in the mobile network, it is necessary to explore the relationship between service volumes and mobile equipment quantities through time.

We have focused on mobile network equipment quantities with the intention of avoiding the complications, which stem from different cost accounting and reporting policies through time³.

Vodafone Group coordinated a data collection exercise, which included the operating companies in Greece, Netherlands and Spain. We requested and were provided with:

- Mobile service volumes in Erlangs for total usage and usage in the busiest hour of the day. We have chosen the busy hour Erlang ("BH Erlang") as the measure of output because it is the closest proxy for the actual demand on capital that may drive network rollout; and
- The number of key network components: i.e. Sites, Cells, Transceivers, Base station controllers ("BSC"), Mobile switching centres ("MSC") and Home location registers ("HLR").

In the table below we summarise the data we were provided with:

Operating company	Mobile usage	Network equipment
Greece	1994 to 2003 quarterly	1993 to 2003 quarterly
Netherlands ⁴	1996 to 2003 quarterly	1996 to 2003 quarterly
Spain	2000 to 2003 quarterly	2000 to 2003 quarterly

It is worth noting that what we are interested in is how equipment deployed has responded to demand for additional capacity. In the early periods of network rollout the quantity of equipment deployed is responding to both demand for capacity and the need to attain a given level of coverage as part of a licence condition:

- In Greece we understand that this coverage commitment was to reach 85% population coverage by September 1998, 6 years from the issue of the licence in September 1992. This obligation was met in June 1995. Our data starts in 1994 and so includes an element of coverage rollout that could reduce our overall estimate of fixed common costs.
- In the Netherlands we understand that this coverage commitment was to reach 98% coverage by July 1996. We understand this commitment was met. The data in our analysis starts in Q3 1997.
- In Spain we understand that the license commitment was to have coverage of every town of over 500k people by end of 1999. This was achieved. The data in our analysis starts in Q1 2000.

Since the data in our analysis begins in each country after the license coverage commitments were met, our analysis should reflect the actual marginal equipment required to meet demand on capacity rather than an element of coverage.

We have plotted mobile equipment quantities and busy hour (BH) Erlangs for the Netherlands, Greece and Spain respectively. These demonstrate visually the relationships between equipment and output. To estimate the level of fixed capital from the data requires fitting a relationship between mobile equipment quantities and usage.

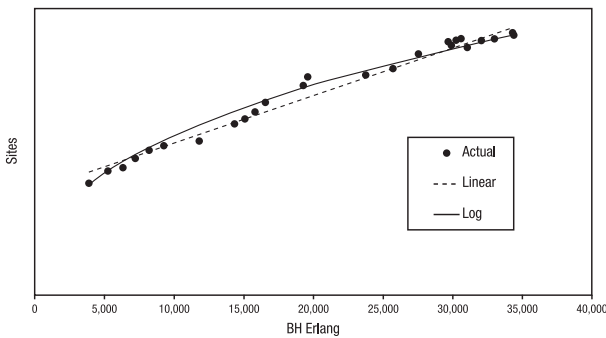
Our analysis has found that there are two basic functional forms that fit the data:

- Linear: $y = a + b x$; and
- Log: $y = a x^b$ or equivalently $\log(y) = A + b \log(x)$; $a = \exp(A)$.

Where: y = Quantity of equipment and x = BH Erlang.

To conduct the econometric analysis we have used the statistical software package Stata (v.8) to investigate the relationship between each equipment type and BH Erlang in each country. We have not presented the results from these regressions due to their commercially sensitive nature. However, we note that all final estimates from each equipment regression that has been included in the analysis were statistically significant to the 5% level⁵ and other issues around spurious regression, heteroskedasticity and serial correlation have been tested for and ruled out (or in the case of serial correlation corrected for with an AR(1) specification).⁶

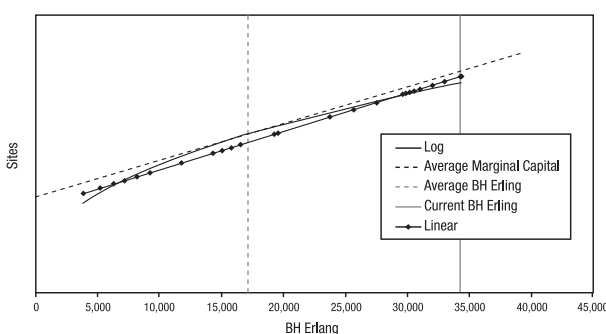
As an example, in the chart below we show the results of the econometric analysis for the relationship between the number of Sites and the BH Erlang in the Netherlands. (Left hand scale removed due to commercial sensitivity).



The estimated coefficients can be used to infer the proportion of fixed capital:

- In the Linear model, the proportion of fixed capital is the intercept (coefficient a) divided by the total current capital; and
- In the Log model the same approach is taken with reference to a tangent line at the average BH Erlang (the mid point from zero to the current BH Erlang). This line represents the marginal capital at the average output level.

In the chart below we show the Linear model, Log model and the tangent line (described above) for the number of sites in the Netherlands:



It is not coincidence that the tangent line drawn at the average output level is similar to the equation estimated in the Linear model. The Linear model essentially estimates the average marginal capital over the estimation period (analogous to the average marginal cost) and the tangent line at the average output level is a similar inference.

We would expect these two lines to be similar when the Log and Linear model specifications are both good fits. We have found that when the Log and Linear models could both be fitted robustly the estimates of the fixed common quantity of equipment are broadly similar. We have taken an average of these two estimates to be our final value.

Notes

¹ This paper was written with assistance from two colleagues; Matthew Corkery and Darren Waterman and was issued solely to Vodafone. PricewaterhouseCoopers LLP, its partners and staff neither owe nor accept any duty of care to any third party whether in contract or in tort (including without limitation negligence or statutory duty or howsoever otherwise arising) and shall not be liable in respect of any loss, damage or expense of whatsoever nature which is caused by any third party's reliance upon information derived from the report. If any third party wishes to rely upon the report or information derived therefrom, they do so entirely at their own risk.

This report includes the presentation of results from analysis that PricewaterhouseCoopers has carried out. All analysis is the responsibility of PricewaterhouseCoopers. Vodafone operating companies have been responsible for providing the underlying data for the analysis. The data and detailed results are commercially sensitive and therefore are not presented in this report.

² See Appendix 1

³ In addition, by looking at equipment quantities the bias of equipment cost inflation is removed. A potential issue with using equipment quantities instead of costs is the issue of changes in the productive capacity of different equipment types. If equipment became more productive we would observe (even in the absence of fixed equipment quantities) a reduction in the marginal equipment required to service an increase in output through time. If this was not taken into account this could mean that the analysis could overstate the level of fixed common costs. However, we believe that for Sites, Transceivers and Cells, which yield the most robust testable relationship with usage that this is not an issue. We understand that these equipment types have not materially changed in respect of productivity capacity through time and that these represent approximately 60% of total costs.

⁴ The data set began in 1996 Q1. The BH Erlang was 379 but the number of Sites, Cells and Transceivers was 0 so we have started from 1997 Q3.

⁵ With the exception of the regression of MSC against BH Erlang in the Netherlands where the intercept a, was significant at the 7% level.

⁶ Where no statistically significant relationship for a particular equipment type could be found we have assumed that the level of fixed common costs for this particular equipment type in that country is zero.

⁷ Since the information displayed is commercially sensitive the left hand scale has been removed.

Dan Elliott



Director, Frontier Economics

Dan Elliott is a Director of Frontier Economics, with over 15 years experience as a consulting economist. He is a leading expert in the field of utility regulation, in particular focussing on the telecommunications, water and transport sectors. Dan advised Vodafone in relation to the UK Competition Commission's call termination charges price review and has extensive telecommunications experience throughout Europe.

The importance of price elasticities in the regulation of mobile call termination

Ramsey pricing was an issue that was considered at some length during the UK Competition Commission (CC) enquiry, as a method of determining the proportion of fixed and common costs that should be allocated to call termination. The CC expressed concern over the wide range of elasticity estimates with which it was presented and ultimately rejected Ramsey pricing in favour of equal proportion mark-ups on the LRIC (EPMU).

Subsequently the issue of Ramsey pricing was discussed again by OfTel in its review of mobile call termination charges under the new regulatory framework. OfTel endorsed the view that the estimates of price elasticities were too unreliable to rely on for regulatory price setting.¹ The European Regulator's Group (ERG) has also taken the view that Ramsey pricing can be regarded as "practically unfeasible", because of the "detailed information about total costs, marginal costs and demand elasticities" that is needed.²

In addition the CC also argued that Ramsey was not even the right benchmark to use in setting termination charges for the following reasons:

"One way of thinking about the market in competitive (or contestable) terms might be to consider MNOs bidding for the right to terminate particular types of call. In this situation, MNOs will be willing to bid prices down to at least the level of their average costs – towards their LRIC plus a reasonable contribution to common and fixed costs; which is very close to EPMU (see paragraph 8.86). If one MNO were to bid a set of Ramsey prices for the right to terminate particular calls in these circumstances, the others would have very strong incentives to undercut it, meaning that a competitive market for termination would not sustain Ramsey prices."³

It would appear, therefore, that if Ramsey prices are difficult to identify *and* are not the appropriate benchmark anyway that there is little need for further debate over Ramsey pricing for call termination.

However, in my view, this would be quite wrong. While recognising that problems existed around the robustness of data for the UK investigation, there is every reason to expect that our ability to obtain robust estimates of price elasticities will improve significantly. As the issue of call termination regulation is reviewed in more countries, and as the length of the available time series of data increases.

Furthermore, the CC's argument that EPMU is intrinsically the more appropriate benchmark for call termination charges is incorrect. The general argument that one would expect competing firms to try to set Ramsey prices is in fact well known and well accepted; it pays firms in competitive markets to maximise the consumer surplus of their customers and this is best achieved by Ramsey pricing. The CC's argument that "competition" for termination would drive down mark-ups is wrong, because it fails to take into account the knock-on effect on other tariffs. As the CC itself recognised, the "waterbed" effect between tariffs offered by mobile operators means that if a firm were to undercut the Ramsey price in the way described, it would have to charge higher prices for other services, which would be less attractive to customers and therefore intrinsically less competitive.

Not only is the idea of EPMU as competitive pricing wrong in a single period, it is also wrong when thinking about how firms price products over time. Firms in competitive markets vary mark-ups across a product's lifecycle (as well as between

products) as the price-sensitivity of demand changes. In the early stage of a product's life demand may be uncertain and the price elasticity relatively high. At this point mark-ups will be low. As the product becomes accepted and demand strengthens elasticities will fall and mark-ups will tend to increase. Then, as newer alternative products are introduced, the price elasticity of demand will rise again, leading to a reduction in mark-ups. EPMU, by fixing mark-ups over time as well as between services, obstructs this form of inter-temporal Ramsey pricing, which is commonplace in competitive markets and facilitates the uptake of new products and technologies.

Given the importance of this issue, this paper assesses whether using EPMU was the right approach, given the intuitive and quantitative evidence that different price elasticities exist for the range of mobile services offered in the UK market.

The Intuitive story

In the UK investigation the CC concluded that there was insufficient evidence to conclude that the elasticities of different mobile services were different, and hence it applied EPMU to allocating fixed and common costs between services. Oftel has subsequently endorsed this view in its review under the new EU framework.

However, for it to be true that the (super-)elasticities of mobile services are all equal it really needs to be the case that the own price elasticities of the services are equal *and* that the cross-price effects between the services are negligible.

Own price elasticities

As regards own-price elasticities of mobile services, far from expecting these to be equal, there is every reason to expect that they will not be. There is, for instance, no logical reason why the own price elasticity of mobile subscription should be equal to the own-price elasticity of mobile originated or fixed-to-mobile (F2M) calls. Likewise there is no reason to expect different categories of call (like mobile-to-mobile (M2M) and F2M) to have equal price elasticities, because the services themselves differ from each other, the different circumstances under which such calls are made and the different costs of making these calls. It may also be the case that the profiles of customers making M2M and F2M are also different. There is no logical reason for arguing that the own-price elasticity of any one type of call will be greater or smaller than any other. Resolving that issue should be an empirical matter and the evidence presented to the CC (see Table 1) supports the view that there are substantial differences in own price elasticities.

There are many instances where similar products are priced with very different mark-ups, reflecting the different price elasticity of the customer groups who consume them. For example, the mark-up on branded and own-brand goods in supermarkets will frequently be very different, reflecting the different demand characteristics of the customer groups that buy them.

Cross price effects and super-elasticities

Even if it were the case that the own price elasticities of different mobile services were found to be equal, it remains the case that the "super-elasticity" of mobile services should be expected to be greater than for F2M calls, because of the existence of cross-price effects between subscription and mobile originated calls.

Put simply, an increase in the price of mobile subscription can be expected to reduce the number of mobile subscribers. This has a knock-on effect on the number of mobile originated calls made (because there are fewer callers and fewer people to be called). In addition, a reduction in mobile subscribers can be expected to reduce the volume of F2M calls (for the same reason that there are fewer mobiles subscribers to be called). These cross-price effects increase the super-elasticity of mobile subscription and consequently reduce the size of the mark-up for fixed and common costs that it is efficient to recover from subscription charges.

The same argument applies to the price of mobile originated calls. An increase in price directly reduces the volumes of calls made. In addition, by reducing the consumer surplus from calls, some marginal subscribers will react to the increase by ceasing to be subscribers. Hence there is a cross-price effect from the price of mobile originated calls to the number of mobile subscribers, which increases the super-elasticity of the price of mobile originated calls. In addition, the reduction in mobile subscribers resulting from an increase in mobile call charges will, by the process described above, have a knock on effect on the volume of F2M calls, which also increases the super-elasticity of the price of mobile originated calls.

Despite the fact that the CC assumed equal own-price elasticities, these cross-price effects mean that the super-elasticities are not equal. Specifically these effects increase the super-elasticities of mobile services relative to that of F2M calls. This is demonstrated in Annex 1 to this paper. Thus, while the specific numbers assumed by the CC are not the issue here, the existence of price interactions between mobile subscription and the volumes of calls (both mobile originated and F2M) mean that we would expect, other things being equal, that the super-elasticity of mobile services will be higher than for F2M calls. This, in turn suggests that optimal mark-ups for fixed and common costs should be weighted more towards F2M calls than is suggested by pricing according to the EPMU rule.

The empirical evidence

Chapter 8 of the CC's report reviews the elasticity assumptions presented by the various parties to the investigation, including both original research prepared for the investigation and some evidence already produced on the price elasticity of the relevant services in other countries.

These results are presented in Table 8.7 of the final report and are summarised in Table 1.

Table 1: Comparison of the various elasticity estimates presented to the CC

	DotEcon	Frontier Economics	Holden Pearmain	Access Economics *	Dr J Hausman †
Own-price elasticity of:					
Mobile subscriptions	-0.37	-0.54	-0.08	–	–
Mobile-originated calls	-0.62	–‡	-0.48	-0.8	-0.5 to -0.6
Fixed-to-mobile calls	-0.43	-0.18	-0.11	-0.08	–
Cross-price elasticity of:					
Mobile-originated calls with respect to the price of subscription	-0.25	-0.50	-0.13	–	–
Mobile subscription with respect to the price of mobile-originated calls	-0.48	–‡	-0.13	–	–
Fixed-to-mobile calls with respect to the price of subscription	-0.21	-0.23	–‡	–	–
Fixed-to-mobile calls with respect to the price of mobile-originated calls	-0.27	–‡	–‡	–	–

Source: UK Competition Commission, Table 8.7

* As noted in the text, the Access Economics estimates relate to Australia.

‡ Parties were unable to find statistically significant results for these elasticities.

† As noted in the text, the Dr Hausman estimates relate to the USA.

Note: The elasticities for DotEcon and Frontier Economics here are the implied elasticities that result from their econometric estimates and demand systems. This enables a direct comparison to be made with the Holden Pearmain study.

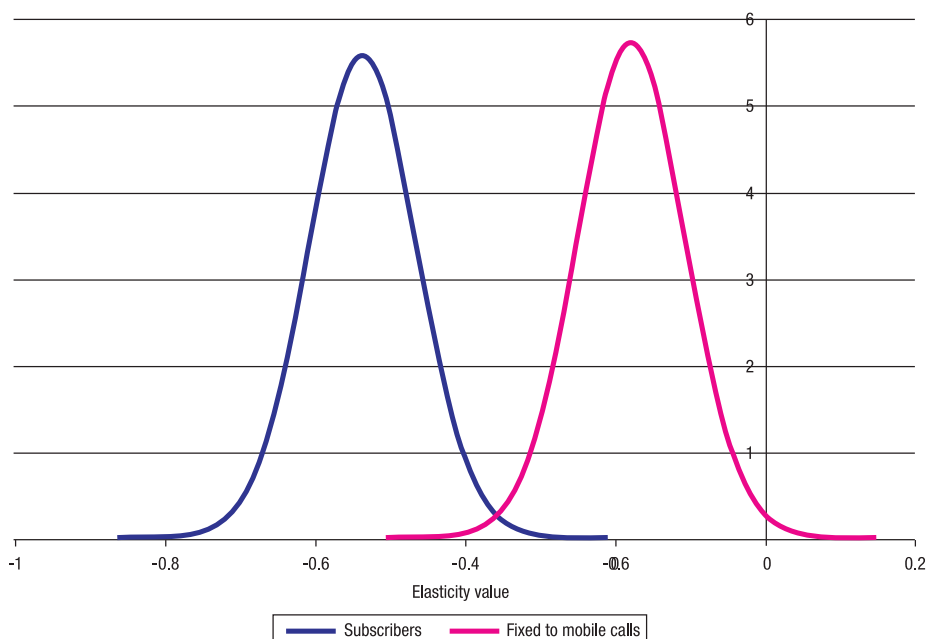


Figure 1: Probability distribution of parameter values

Source: Frontier calculations

The CC made much of the differences between the elasticities derived by the different parties.

There are well recognised difficulties in determining the precise value of all of the relevant own-price and cross-price elasticities. However, in my experience the variation in estimates of the relevant parameters is well within the bounds of uncertainty over which the CC and regulators are routinely expected to make a judgement in the course of such an investigation.

Furthermore, the process of measuring elasticities by econometric methods allows us to place “confidence intervals” around the estimates. This means that although we cannot be sure of the exact values of each parameter, we can get a sense of how likely it is that the true value differs from the estimate made by the model. Figure 1 illustrates the confidence intervals from the Frontier Economics modelling of mobile subscription and F2M calls.

The peak of each distribution represents the “expected” or most likely value of the parameter, given the data available. The value of curve at any point reflects the relative likelihood that the true elasticity is the value on the horizontal axis.

Figure 1 illustrates that the expected elasticity of mobile subscription is significantly higher than the expected elasticity of F2M calls, based on the data analysed by Frontier. Furthermore, the fact that the two distributions barely overlap at all indicates that the likelihood that the true values of these elasticities are in fact equal is very low indeed. The clear indication of the graph is that mobile subscription is significantly more price sensitive than F2M calls.

Given the empirical evidence it is difficult to see how it could justify the view that the elasticities are not different.

Conclusions

Despite its rejection by the CC and, subsequently by Ofcom, Ramsey pricing remains relevant as the issue of call termination regulation is reviewed in more countries.

Although there remain issues surrounding the robustness of the available data, this is a situation that is likely to improve as more data is collected in more countries and the length of the available time series increases.

Furthermore, examination of the data that has been presented in the UK shows that, although there is some spread in the results presented, there is a consistent pattern indicating that the own-price elasticities of mobile services are not equal. When the existence of cross-price effects is taken into account it seems quite clear that the super-elasticity of mobile services is likely to be significantly higher than that of F2M calls. This suggests that the EPMU approach taken by CC and Ofcom is not justified.

Not only does EPMU not reflect the way prices are set in a competitive market, but, given the empirical evidence it would seem that EPMU significantly under-estimates the correct level of mark-up on F2M calls for fixed and common costs.

Annex 1

This annex demonstrates that the CC’s assumptions regarding elasticities implied higher super-elasticities for mobile services than for F2M calls, despite the assumption of equal own-price elasticities.

This can be easily demonstrated with the use of algebra, but it can also be explained simply, by reference back to the columns of Table 2. This can be illustrated by reference to the assumptions, which the CC used for their welfare analysis, presented in Table 2.

Table 2 shows that the CC assumed that the price elasticity of subscription, mobile originated calls and fixed to mobile calls are all assumed to be -0.3 , which means that a 1% rise in any one of these prices is expected to reduce demand for that service by 0.3%.

Despite the fact that the CC assumed equal own-price elasticities, these assumptions do not support equal super-elasticities. This can be easily demonstrated with the use of algebra, but it can also be explained simply, by reference back to the columns of Table 2.

Examining the first column of figures, the own-price elasticity of mobile subscription is -0.3 . However, the next row shows that the cross-price elasticity of mobile outbound calls with respect to the price of subscription is -0.108 , while the last row shows that the cross-price elasticity of fixed to mobile calls with respect to the price of subscription is -0.0405 .

In total therefore, the CC is assuming that 1% rise in subscription prices does all of the following:

- it reduce the number of subscribers by 0.3%;
- it reduces the volume of mobile outbound calls by 0.108%; and
- it reduces the volume of fixed to mobile calls by 0.0405%.

The same logic applies to the second column: a 1% rise in mobile outbound prices does *all* of the following:

- it reduce the number of subscribers by 0.198%;
- it reduces the volume of mobile outbound calls by 0.3%; and
- it reduces the volume of fixed to mobile calls by 0.03%.

Table 2: Elasticities assumed by the CC in analysing the effect of a price cap

		Effect of price movement		
		Subscription price	Mobile call price	Fixed-to-mobile price
On quantity	Subscriptions	-0.300	-0.198	0
	Mobile calls	-0.108	-0.300	0
	Fixed-to-mobile calls	-0.045	-0.030	-0.300

Source: UK Competition Commission, Table 9.11

In contrast, the third column states that a 1% rise in fixed to mobile prices reduces the volume of fixed to mobile calls by 0.3%, but has no effect on the other two services.

Now, a super-elasticity measures the total impact of the movement in any one price on the demand for all services. It is a weighted sum⁵ of all own-price and cross-price effects in each column.

It follows therefore that, provided the cross price effects in the first two columns (from the prices of mobile services) are non-zero but the cross price effects in the third column (from the price of fixed to mobile calls) is zero, then the super-elasticity of mobile subscription and mobile outbound calls *must* be higher than for fixed to mobile calls.

Notes

- ¹ See, for instance, Ofcom's Explanatory Statement on Wholesale Mobile Voice Call Termination, published on 19.12.03: http://www.ofcom.org.uk/legacy_regulators/ofcom/mobile_call_termination/.
- ² "Consultation Document on a draft joint ERG/EC approach on appropriate remedies in the new regulatory framework", 21.11.03: http://www.erg.eu.int/doc/publications/erg0330_draft_joint_approach_on_remedies.pdf.
- ³ CC para 8.71.
- ⁴ CC para 8.86.
- ⁵ Weighted by the share of each item in total revenue.



Professor David Newbery

Professor of Applied Economics, University of Cambridge

David Newbery is Professor of Applied Economics at the University of Cambridge, England. He has been an economic advisor to Ofgem, Ofwat, and to the Office of Rail Regulation, a member of the Competition Commission and is a member of the academic panel of environmental economists, formerly at the Department of Transport and the Environment, now with DEFRA. He is chairman of the Dutch Electricity Market Surveillance Committee. He has managed a series of research projects on electricity and telecoms privatisation and regulation, and on road congestion, funded mostly by the UK Research Council since 1988. He pursues research in public policy, regulation, energy, environment, road pricing and finance, taxation, the economic transformation of Eastern Europe, and industrial organisation. His specific expertise relevant to the present paper is expertise in social cost-benefit analysis (his first book was on this subject, published in 1976), and extensive work on optimal pricing in distorted markets, and on optimal taxation. His most recent book, *Privatization, Restructuring and Regulation of Network Utilities*, MIT Press, 2000, deals extensively with telecoms regulation, as well as electricity and gas.

Application of Ramsey pricing for regulating mobile call termination charges

Executive summary

The UK Telecommunications Regulatory Authority (Ofcom, now Ofcom¹) has determined that Mobile Network Operators have Significant Market Power in the market for mobile call termination on 2G networks, and that these termination charges should be subject to a price cap. All parties to the dispute agree that a Ramsey mark-up on Long Run Incremental Cost (LRIC) would produce the same price as competition in the market for call termination. Ofcom accepts that Ramsey pricing would maximise social welfare if the network externality were addressed and other mobile markets were competitive, but argues that these other markets are not sufficiently competitive to ensure that Ramsey pricing of termination alone would maximise social welfare. Ofcom also argues that estimates of the price elasticities needed to set the Ramsey mark-up to recover the fixed and common costs are too unreliable for setting regulated charges. Instead Ofcom proposes an equi-proportional mark-up on termination LRIC, lower than any estimate of the Ramsey mark-up.

The argument that at present data and modelling limitations reduce the reliability of estimated demand elasticities is not a compelling argument for equi-proportional mark-ups, as that requires the implausible assumption that all relevant elasticities are equal. Instead telecoms regulators should follow best

practice in other regulated industries of accepting the correct principle for setting charges and then attempting to refine the implied estimates needed. Ofcom has done this in the past for cost modelling, recognising that Long Run Incremental Cost is a more suitable basis for setting price controls than the Fully Allocated Cost model favoured until recently by US Public Utility Commissions. All UK regulators have chosen the Capital Asset Pricing Model for estimating the weighted average cost of capital for setting price controls, and have employed increasingly sophisticated econometric techniques to estimate the relevant parameters.

Electricity and gas regulators in the UK and on the Continent increasingly use Data Envelopment Analysis and other econometric techniques to estimate efficiency frontiers in comparative benchmarking exercises for setting the X-factor in price controls. EU regulators collaborate to pool data and experience to refine their methods and estimates for setting price controls. In the same spirit, Ofcom should accept that the current UK mobile termination price control is at best a temporary solution forced by a lack of confidence in the robustness of the elasticity estimates needed for determining the theoretically correct Ramsey mark-up, and should work to improve the required estimates for setting future Ramsey mark-ups.

Common ground

There is considerable agreement between many of the parties to the mobile termination dispute about a number of propositions:

1. In the presence of fixed and common costs, the efficient way of recovering these costs would be Ramsey pricing, in which the mark-up over marginal costs (in equilibrium, long-run incremental costs, LRIC) would be inversely proportional to market demand elasticities.²
2. In a market where subscribers must choose between different suppliers to purchase the whole range of bundled services, effective competition in all service markets would force the suppliers to charge Ramsey prices for each service they offer to subscribers (i.e. subscription and calling).³
3. Termination is a bottleneck service and thus not subject to the same competitive pressures as subscriptions and calling. Mobile network operators (MNOs) can raise termination charges to fixed operators who are subject to regulation and unable to exercise countervailing power, and will have incentives to do so if the surpluses above cost thereby generated can be applied to fund competitive activities against other mobile operators. As a result, competitive forces will not ensure that MNOs set termination mark-ups at Ramsey levels without regulation.
4. New subscribers confer a network externality on existing mobile and fixed line subscribers, which they will undervalue in making the subscription decision. There is therefore a case for subsidising subscription relative to the incremental cost, and this adjustment forms part of the identification of appropriate Ramsey mark-ups.

Issues

The main issues under dispute are how to apply these agreed principles, given that it is now expected that *ex ante* regulation of mobile call termination will be justified by the finding of Significant Market Power (SMP). The first question is conceptual: what form of regulation best meets the objectives of the European Communications Directives, given that it must be justified and proportionate? The answer is that regulation should ideally deliver the same results as a competitive market, and should be limited to addressing the lack of competition in those markets where firms are found to have SMP. Thus the IRG (2003) lists as one of their principles for applying remedies for problems in mobile termination (MT):

IRG believes that charge control that in the long term sets MT charges at a competitive level is proportionate and justified because this obligation can effectively prevent the adverse effects of possible lack of competition on the wholesale call termination market.⁴

The only relevant market where firms have been found to have SMP is mobile call termination. It therefore follows that regulation should address the issue of SMP in that market, assuming that other markets are workably competitive.

Article 13 (2) of the Access Directive (2002/19/EC) requires that "National regulatory authorities shall ensure that any cost recovery mechanism or pricing methodology that is mandated serves to promote efficiency and sustainable competition and maximise consumer benefits." The Independent Regulators Group interprets this as follows:

IRG considers that high MT charges are a competition problem as long as they are set above the charge level of an effectively competitive wholesale MT market. At such a level the cost of call termination are efficiently set and the level of MT charges can be considered the outcome of a sustainable competition process.

It's IRG's view that MT charges should be set at a competitive level, i.e. the charge level that would be set if the wholesale MT market would be effectively competitive. In an effectively competitive market, excessive margins are competed away and prices driven down to the efficient level of cost plus a sustainable margin.⁵

The best form of regulation to promote efficiency is a price cap, as that simulates the effect of a competitive market, in which the MNO has no ability to raise prices without losing his market and profits. Again, to quote from the IRG:

In an effectively competitive market, excessive margins are competed away and prices driven down to the efficient level of cost plus a sustainable margin. However, in markets where competitive forces are restricted, it can be necessary to apply *ex-ante* regulation in the interest of the end-user, to make sure that margins are not excessive to the detriment of consumers. A price control regime, if appropriately designed, can be a proportionate response where competitive forces and other regulation are not sufficient to ensure that operator margins are not excessive (that prices are cost reflective with a reasonable return on capital for operators).

Therefore it is IRG's view that, if the market analysis of an NRA shows for its national market that a MNO has a dominant position on the wholesale MT market and finds that the level of MT charges of a MNO are significantly above a competitive level, a form of price control that forces MT charges down to a competitive price level may be appropriate and proportionate in the light of the regulatory framework considering the nature of competition in the relevant wholesale MT market. Such an obligation can achieve the objectives of promoting competition and protecting the consumers interests by setting MT charges at a competitive level.⁶

In the presence of fixed and common costs, commercial companies must mark-up prices above marginal costs if they are to remain financially viable. It then follows that the appropriate

regulation of mobile call termination charges is a price cap set as a Ramsey mark-up on termination LRIC.

Oftel⁷ accepts the logic of Ramsey pricing from the perspective of maximising social welfare:

Ramsey prices are a set of prices for a group of services that maximize social welfare when the presence of common and fixed costs across these services does not allow the adoption of marginal costs pricing (as the firm would not then breakeven).⁸

Oftel considered whether it should set Ramsey termination charges but reached the conclusion that this was not appropriate. Oftel's reasons are set out in detail in Oftel (2003, Annex K), and, briefly, fall under a number of headings:

1. The retail market is not competitive
2. Tariffs are multi-part and permit price discrimination
3. The estimation of elasticities poses big technical hurdles
4. The models for estimating the optimal mark-up are over-simplified
5. The size of the common and fixed costs is small

The first point raises interesting issues, for Oftel accepts that he does not intend to regulate the mobile retail market, where he considers that no MNO holds SMP (Oftel, 2003, K10). "The Director does not believe that this implies that there is no possibility of excess profits accruing to the MNOs as a result of high termination charges." (Oftel, 2003, 4.43.) As a result, "he believes that there is a strong risk that setting Ramsey termination charges would not maximise social welfare and thus, Ramsey would not be the efficient pricing approach for regulating termination charges." (Oftel, 2003, K10.)

This argument based on social welfare is conceptually quite different from the guiding principle of competition motivating the Communications Directives, and clearly set out in the Principles of the Independent Regulators Group. Perfect competition, provided all externalities are internalised, would deliver efficiency, which is not the same as maximising social welfare. Efficiency is, however, a sensible criterion for sector-specific policy makers such as telecoms regulators, as distributive justice (i.e. maximising social welfare) is better left to tax and expenditure policy, decided by central government. The IRG therefore argues for setting termination price controls "at a competitive level". If all mobile markets were competitive, then mark-ups in each market would be set at Ramsey levels. If a price cap is to be imposed on call termination services due to firms having SMP, then this price cap should logically be set at the Ramsey level. The fact that this may not maximise social welfare is not directly relevant – the regulator is not being given a hunting licence to leverage his justification for intervening in one market (mobile termination) to making further adjustments to try and improve outcomes in other markets where he has no direct justification to intervene. The requirement to promote competitive outcomes rather than the grander ambitions of

maximising social welfare is intended as a deliberate restraint on the power of the regulator. Regulation has costs, and should be limited to interventions where the benefits outweigh these costs, and then only when firms are found to have SMP in the relevant market.

The second point is that the models used to compute the Ramsey mark-up assume linear price schedules whereas in fact mobile service pricing is characterised by complex non-linear pricing. This may be a valid criticism of the specific calculations based on linear models, but is not a valid criticism of the principle of setting a Ramsey mark-up for termination. It is instead an argument for employing more sophisticated models. Conceptually, a variety of non-linear tariffs allow MNOs to segment the market into sub-markets, in each of which the tariff is effectively an access charge and a constant marginal calling cost. If different groups of consumers have different demands for access and calling, and can be allocated to different sub-markets, then mark-ups in each of these competitive sub-markets would be set at Ramsey levels, achieving overall greater efficiency than if these sub-markets were combined into a single undifferentiated market with a single access and marginal calling charge. Non-linear pricing can indeed be taken as evidence that competition is working more effectively (in achieving a higher overall level of efficiency) than simple linear tariffs more typical of monopoly. As Oftel has not argued for differentiating termination charges by similar sub-markets, it is not clear that there is any need to adjust the determination of the Ramsey mark-up for termination.

The third and fourth points are also criticisms of the difficulty of correctly estimating the relevant elasticities to determine the Ramsey mark-up. The last point concerns the materiality of the difference between the hard-to-estimate Ramsey mark-up and the simple equi-proportional mark-up proposed by Oftel, and again does not affect the principle of the desirability of setting a Ramsey mark-up. Just because it is difficult and has not been done satisfactorily to date does not mean that it should not be done in future in the UK, or elsewhere if call termination is to be regulated. The history of utility price regulation, and more generally of competition policy, is one of increasing sophistication in the economic and quantitative methods that have been applied. If regulators are to meet the test set by the Communications Directives, and specifically the agreed best practice of the Independent Regulators Group that the price control should simulate the effect of competition, they will have to rise to the standards set by best practice met by other aspects of price regulation.

Periodic reviews and the evolution of methodology

It seems unlikely that technical innovations will make the call charge termination market contestable in the foreseeable future, and so it is likely firms in this market will continue to have SMP, and to require regulation. If so, regulators and MNOs will be

faced with a periodic re-setting of the price-cap, as has become familiar for other regulated natural monopolies such as electricity transmission and distribution. Revisiting the question of how to set call termination charges creates opportunities for improving the method for setting these charges. Regulators in other countries may also be required to set call termination price controls, and will need to draw on experience elsewhere.

If we look back at the history of price-cap regulation since its introduction in the UK in 1984, we observe a period of learning, gradual refinement of the principals to be applied, and gradual improvements in the methods of estimating the appropriate levels for these price-caps. The same is true if we look at regulators in other countries, who both learn from experiences elsewhere and often suggest improvements. We should expect the same process of learning and refinement for setting mobile termination price-caps. The UK Competition Commission (CC) decision should be seen as contingent, subject to revision in the light of more careful reasoning by regulators and better empirical evidence.

Oftel set out its initial views in *Ramsey Prices and Network Externalities: Dr. Rohlfs' Analysis*, 23 May 2002:

Ramsey prices, including the implications of externalities are relevant in theory. But for practical reasons they are unlikely to provide a reliable basis for setting regulated charges.⁹

The CC (in its letter to Oftel of 4 July 2002) invited Oftel's reaction the proposition that the elasticity estimates needed to compute the Ramsey mark-up were too uncertain. Specifically the CC suggested that "Ramsey prices cannot be set because the estimates of elasticities are too uncertain. Econometric estimates of elasticities may be unreliable, as past behaviour, which was based on rapid growth, may not represent future behaviour. Furthermore, current estimates might not be relevant for the next few years, as future price changes may be greater than those experienced in the past."¹⁰

Oftel said in its response to this letter that it

agrees that econometric estimates are likely to be unreliable. This is not only because past behaviour may not represent future behaviour. Robust econometric estimates are usually extremely difficult to derive, because of a variety of factors including data deficiencies, complexity of the underlying relationships, etc. For an example of some of the problems that can arise see *Assessment of Demand Elasticity Estimates of .econ*, A paper by Dr John Hunter and Professor Christos Ioannidis, Brunel University, published by Oftel in September 2001.¹¹

Oftel sets out detailed reasons why it believes that "any pricing methodology based on the elasticities of the services is unlikely to provide a reliable basis for setting regulated charges."¹² These include the difficulty of identifying the relevant price, the sensitivity of the result to the specification and estimation of the demand function, deficiencies in data (short time-series),

the importance of network and non-price effects, only some of which are captured by subscriber numbers, the exclusion of some mobile services, and the failure to distinguish among the various differently priced mobile-originated calls.

These objections are contingent and some are exaggerated. For example, the claim (in K.36) that a 9 x 9 matrix of elasticities would be needed overstates (by a factor of 27) the requirements of an acceptable demand system. The assumption that equi-proportional mark-ups are preferable to Ramsey mark-ups assumes, incorrectly, that the evidence that demand elasticities are equal is stronger than that they are unequal. The difficulties and objections could (and should) be overcome by further empirical investigation, in much the same way that the price-cap methodology has been refined in the other network industries. The fact that developing and applying the correct methodology may be difficult and imprecise is not an acceptable (nor accepted) reason for the regulator to substitute an inferior if easier approach. The fact that the relevant demand elasticities have been estimated in various countries suggests that it can be attempted and that with the passage of time, better data will allow these estimates to be improved.¹³ As such estimates are replicated and the quality and quantity of data improves, so the reliability of the elasticity estimates should improve.

Evidence of increased econometric sophistication from competition authorities

The practice of competition law has a longer history in Europe than the practice of utility regulation, and thus gives more evidence of the role of quantitative economic analysis. The British Office of Fair Trading published *Quantitative techniques in competition analysis* (Research paper 17) in October 1999. The paper opens by observing that:

1.3 Over recent years, the use of quantitative analysis in antitrust has increased for a variety of reasons. These reasons include the development of modern and fairly reliable quantitative techniques, advancements in user-friendly software and cheap hardware, availability of more and better data and, not least, an increasing use of economists and economic evidence, by antitrust authorities and the companies concerned.

2.4 The use of quantitative techniques has differed from country to country and, in some instances, between different authorities within the same country. In the US, antitrust authorities and the courts have a longer tradition of relying on economic analysis and empirical verification. This is partly due to the increased influence of economists in the Department of Justice, which became noticeable during the 1970s, but is also due to the more litigious nature of US antitrust policy which is very demanding in terms of supporting economic and factual evidence. Expert testimony is more often required in a litigation setting where the adversarial process pitches expert against expert and where each party tries to expose the weakness of the other parties' arguments and evidence. An investigative

procedure poses different demands on the parties involved and does not allow them to influence the investigative process as much. On the contrary it is the investigating authority which drives the process and this is typically the case in Europe (MMC, the European Commission's DGIV, etc).¹⁴

The Report then gives a number of examples where demand elasticities (of the kind required to determine the correct Ramsey mark-up) were estimated and used in various competition cases. For example, estimates of demand elasticities have been used to define the relevant market using the now-standard SSNIP (Small but Significant, Non-transitory Increase in Price) test.¹⁵ Chapter 9 of the OFT Report discusses the estimation of residual demand elasticities to inform the competition authorities about the potential for exercising market power. Chapter 15 cites examples of more complex estimates of demand systems (of the kind needed for estimating relative demand elasticities for the various mobile markets). These are typically needed to examine the likely impacts of mergers between firms that produce a variety of products, and are well illustrated by the Kimberly-Clark/Scott merger case presented in the report.¹⁶

The significance of this increasing economic and quantitative sophistication is reinforced by noting that the European Commission now specifically advocates the use of quantitative techniques to provide evidence of demand substitution:¹⁷

There are a number of quantitative tests that have specifically been designed for the purpose of delineating markets. These tests consist of various econometric and statistical approaches: estimates of elasticities and cross-price elasticities for the demand of a product, tests based on similarity of price movements over time, the analysis of causality between price series and similarity of price levels and/or their convergence. The Commission takes into account the available quantitative evidence capable of withstanding rigorous scrutiny for the purposes of establishing patterns of substitution in the past.

The fact that the European Commission expects that quantitative evidence should be used in competition cases strongly suggests that it should also be taken into account in designing regulation that has the same purpose of delivering efficiently and competitively priced services.

Setting charges for public utilities: evidence of regulatory learning

The history of price-cap regulation since its introduction by Professor Littlechild for Oftel in 1984 has been one of applying economic principles to encourage efficiency, often in conscious distinction to principles of fairness that had evolved under cost-of-service regulation in the United States. As such, it provides many examples of the process of learning, improving empirical estimation, and refining the methodology.

Refinements in cost modelling

Whilst American regulators have been content to apply a Fully Allocated Cost (FAC) methodology in determining rates (tariffs or charges), Oftel has invested substantial effort and resources defending the merits of, and then estimating, Long Run Incremental Cost (LRIC) as the basis for its cost-recovery mechanism. While it is easy to compute costs using the FAC approach, it is intrinsically difficult to estimate LRIC for particular services (such as call termination). Indeed, Oftel has taken more than two years of intense joint work with the MNOs to identify the component costs needed to measure LRIC, and has not yet reached agreement on some of the more important details with the MNOs. That said, the methodology is accepted as intrinsically superior to FAC, precisely because it provides better incentives for efficiency. It is also accepted that the estimates are likely to continue to be improved for future price controls. The fact that the estimates are imperfect does not make them invalid for setting price controls.

Successive periodic reviews of all British network utilities indicate a steady increase in the degree of sophistication of the economic analysis of costs and of the design of suitable incentives to improve efficiency. For example, the UK gas regulatory authority (Ofgas) required Transco (the British gas network company) to develop (at very substantial cost) a computer model that could be used to determine the Long Run Incremental Cost of meeting an increase in injections at any node matched by a corresponding off-take at any other node. This model, Transcost, could then be used by Transco, Ofgas and industry participants to determine entry and exit prices for the National Transmission System. The fact that the model was made publicly available aided the transparency and credibility of the regulatory process (in sharp contrast to the earlier history of disputes over gas transport charges) but was achieved only after much research, programming and validation.¹⁸

Determining the cost of capital

The next example raises similar methodological and econometric difficulties as setting Ramsey mark-ups. A key element in setting price-caps for capital intensive networks is the determination of an adequate reward to shareholders' investment in the network. Investors must be assured that they will receive a fair rate of return on their investment. Every price control review must address the estimate of this fair rate of return, or the weighted average cost of capital (WACC). The CC reference on mobile call termination is no exception, and contains extensive discussions of the setting of the WACC.

In the early stages, there was considerable disagreement about the appropriate theoretical model to use, and despite continuing academic reservations, most parties now accept the capital asset pricing model (CAPM) as an appropriate workhorse, although this issue was again raised in the mobile termination reference. That requires estimating the risk-free rate of return, the debt and equity risk premium, the equity beta, and the allowance for tax treatment. The first four components are

contentious and subject to a degree of uncertainty. Some require econometric estimates. Views on both the relevant risk-free rate and the equity risk premium have changed considerably over the past fifteen years. Nevertheless, there is considerable agreement about the approach to take, and at any one time, reasonable agreement about the ranges in which the values lie. That is not to dispute that the final range can be quite wide (from 7.7% to 14.4% in the mobile reference – see CC, 2003, 2.242). Fortunately, there is a subsequent empirical test, as the purpose for setting the WACC is to ensure that investors will continue to be willing to invest. The share price of narrowly focussed utilities gives information about the market perception of the profitability of investment in the industry and provides a further test on the requirement that the utility can continue to raise capital to finance needed investment.

Setting the X-factor in the price control

A further example is the setting of the predicted efficiency growth factor, X, in the RPI – X price-cap formula. In the UK, this was initially introduced to ‘hold the fort’ until competition arrived, but at and after the first Periodic Review, it has been set to move prices closer to estimated efficient costs, with some modest allowance for underlying productivity growth. Over time, the methodology has been refined to use yardstick comparisons based on increasingly sophisticated econometric techniques such as Corrected Ordinary Least Squares (COLS), Data Envelopment Analysis and Stochastic Frontier Analysis.

One of the pioneers in this field was the Electricity Supply Association of Australia (ESAA) and, later, the Independent Pricing and Regulatory Tribunal (IPART) of New South Wales (also in Australia). In 1991 ESAA established a project to develop international benchmarks of ‘best practice’ as a guide for further reform. This was to estimate national performance indicators at the state level, measure total factor productivity growth by sector, and benchmark Australian performance against international comparators. London Economics published its report *Measuring the efficiency of the Australian Electricity Supply Industry* in August 1993 and completed its international benchmarking study in 1994. In 1999 London Economics published its *Efficiency and benchmarking study of the NSW distribution businesses* for IPART as an input into the determination of the X-factor in the price control. These studies set out and applied the econometric methodology to estimate productivity growth, and to make cross-company (and cross-country) comparisons of performance using Data Envelopment Analysis and Stochastic Frontier Analysis. The consultants tested the robustness of the estimates, while recognising that inevitably uncertainties would remain about the potential for future productivity improvements.¹⁹

The Dutch electricity regulator, DTe, similarly commissioned consultants to estimate the potential for cost reductions over the first price control period for electricity network companies, and published his findings in 2000 (DTe, 2000). The projected rate of cost reduction (i.e. the X-factor in the price control) was a combination of the extent to which each network company was

currently below the efficiency frontier (i.e. the lowest achievable level of costs), and the rate at which the frontier would shift over time (i.e. the rate of productivity improvement of an efficient network company). The first element is based on benchmarking, a complex econometric and linear programming analysis of the data of the companies to be regulated, and a set of comparator companies within and outside the country. The productivity growth estimates are based on past performance, and some judgement of the likelihood that these past rates will continue in the future. Both aspects are contentious, and require high quality quantitative analysis, but neither difficulty has prevented an increasing use of this approach.

The British electricity regulator, Offer (subsequently Ofgem) has been somewhat slower in following this approach. The distribution price control for the period 1999/2000-2004/5 was informed by a rather casual benchmarking exercise that was criticised by Cambridge Economic Policy Associates (Ofgem, 2003). CEPA proposed improvements and observed that over time, the quality of analysis was improving as regulators collected data more systematically, exchanged consistent data with each other, allowing the set of comparator companies to increase, and tested the robustness of the different approaches to measuring efficiency.

The accuracy of the proposed X factor can be judged at the end of the Price Control period by testing to see whether the anticipated efficiency gains have been achieved. Indeed, the market typically gives its verdict rather sooner by either increasing or lowering share prices based on their prediction of the tightness of the control. Over time, the accuracy of the regulators’ forecasts of predicted performance improvement can be measured and the quality of successive regulation improved.

The relevance of learning for Ramsey pricing of call termination charges

Oftel has been willing to estimate the size of the network externality and to reach conclusions about the appropriate size of the subscription subsidy to internalise that externality, even though both of these require estimates of demand elasticities. It would be illogical to argue on the one hand that demand elasticities cannot be relied upon for determining the Ramsey mark-up, while at the same time relying on similar demand characteristics to determine the size of the subsidy to address the network externality.

Regulators in other network industries accept that there will never be complete precision or agreement in determining either the WACC or the X factor, but they are nevertheless willing to present the evidence on the basis of which they reach a conclusion. Clearly, Oftel did not argue that because it was complicated to calculate LRIC, some simpler accounting alternative should be used instead. As the mark-up is as important in determining the final price as LRIC, logically Oftel should be prepared to co-operate with the industry in making the best estimate of the appropriate mark-up. Telecoms regulators,

like other regulators, will have to accept that there will never be complete precision in measuring demand elasticities (and the fixed and common costs), but they will become increasingly confident that they are not equal (as would be required from equi-proportional mark-ups, or EPMU) and will be able to place increasingly tighter lower bounds on the share of these fixed and common costs to recover from termination. Over time, one would hope that these estimates would converge on a more widely agreed level for that mark-up.

Ramsey pricing should lead to increased efficiency gains, and this can be tested in much the same way as setting the WACC or X factor can be tested by examining the subsequent productivity performance of the utility. Changes in termination charges should lead to a re-balancing of mark-ups on other services, and monitoring the mark-ups should allow better judgements about the size of the 'waterbed effect' and the extent to which other services will be Ramsey priced.²⁰ Just as it is difficult to judge the value of the information loss upon water company merger, so it may be difficult to agree on the changes in social net benefit of moving termination charges towards their efficient level, and for the same reason. The value of information and the impact of changes in call termination charges upon social welfare both rely on econometric estimates. Consumer surplus can only be estimated from the underlying utility functions representing preferences, and these can only be identified from demand responses. While it may be difficult to estimate the shapes of the entire demand curves, it should be sufficient to estimate their local shapes (around the current levels of demand) if all that is needed is an estimate of changes in consumer surplus. Electricity regulators across Europe increasingly exchange data and discuss methodological issues when benchmarking transmission and distribution companies, and one should therefore expect communications regulators to exchange data and refine their methodology in setting Ramsey termination mark-ups. To summarise, once it is agreed that the Ramsey mark-up (adjusted for network externalities) is conceptually the correct method under certain conditions (no evidence of significant market power in the remaining mobile market), it logically follows that the aim should be to produce the best estimate of that mark-up, and to commission further studies where these can usefully improve that estimate.

Improving the estimates of elasticities

Ofcom (and regulators in other countries) could encourage closer industry collaboration (or pooling of individual datasets) to create a more accurate and robust data set of prices and volumes for econometric analysis. This might usefully be supplemented by bespoke consumer research (such as conjoint analysis) and sharing of similar studies with other national regulatory authorities (NRAs). It would be reasonable to assume (until demonstrated otherwise) that consumer behaviour and hence consumer demands are similar in other countries at a similar stage of development. Comparisons with Called Party Pays (as in the USA) might further improve the accuracy of measuring demand elasticities, as callers there face different prices for making and receiving calls.

As the mobile market matures, each MNO is likely to concentrate its marketing attention on understanding consumer demands better, and is likely to increase the sophistication with which it analyses the rich data on calling patterns and their response to changes in tariff packages. These should in turn provide increasingly reliable estimates of demand elasticities (although care will be needed to distinguish firm and market level elasticities). NRAs could commission further work to combine this information, in much the same way that it has commissioned studies to estimate indices of mobile tariffs.

Conclusions

Ofcom and other communications NRAs should agree that if call termination is to be regulated, then the requirements of the Communications Directives imply that the form of regulation should be a price-cap based on a Ramsey mark-up on LRIC. In order to set this price control, NRAs will need to estimate call termination LRIC, estimate the size of the fixed and common costs that must be recovered by mark-ups on each service, and finally determine the share of the common and fixed costs to be recovered from call termination charges. This last step requires estimating the market price elasticities for the various services from data on prices and volumes. Pooling data from a large number of countries and companies is the most efficient way of improving the accuracy of these estimates. Direct econometric estimates of the demand elasticities may be supplemented by some of the useful indirect methods that have been adopted in other industries.²¹

References

- Armstrong, M. and J. Vickers (2001) "Competitive price discrimination", *RAND Journal of Economics*, Vol 32 No 3, Winter
- Australian Competition and Consumer Council (2000) *Incentive regulation, benchmarking and utility performance*, Utility Regulators Forum Discussion Paper, ACT Dickson
- Competition Commission, (2003) *Vodafone, O2, Orange and T Mobile*, London: Competition Commission (report available at http://www.competition-commission.org.uk/rep_pub/reports/2003/475mobilephones.htm#full)
- Doern, G.B., and S. Wilks (eds.) (1996) *Comparative Competition Policy: National Institutions in a Global Market*, Oxford: Clarendon Press
- DTe (2000) *DTe makes known the planned efficiency discounts for the electricity grids*, press release of 3 August 2000, at http://www.dte.nl/en/nieuws_en_publicaties/persberichten/00_02.asp
- ESAA/London Economics (1993) *Measuring the efficiency of the Australian Electricity Supply Industry*
- Grajek, M (2003) "Estimating Network Effects and Compatibility in Mobile Communications", Berlin: Wissenschaftszentrum fuer Sozialforschung
- Hausman J.J. and G.K. Leonard, 1997 Economic analysis of differentiated products mergers using real world data, *George Mason Review*, 5, 321-46
- Independent Regulators Group (2003) *Principles of Implementation of Best Practices on the application of remedies in the wholesale mobile voice call termination market* available at <http://irgis.anacom.pt/admin/attachs/359.pdf>
- IPART/London Economics (1999) *Efficiency and benchmarking study of the NSW distribution businesses*
- Okada, Y. and K. Hatta, (1999) 'The Interdependent Telecommunications Demand and Efficient Price Structure', *Journal of the Japanese and International Economies*, December; 13(4): 311-35
- Newbery, D. (1999), *Setting Transportation Tariffs using Transcost*, London: Ofgas
- Ofcom (2003) *Proposals for the identification and analysis of markets, determination of market power and setting of SMP conditions* Published on the Ofcom web-site on 19 December 2003 at http://www.ofcom.org.uk/legacy_regulators/oftel/mobile_call_termination/
- Ofgem (2003) *Productivity improvements in distribution network operators: final report*, submitted by Cambridge Economic Policy Associates Ltd, London: Ofgem, November
- OFT (1999) *Quantitative techniques in competition analysis*, Research paper 17, October, available from the OFT web-site <http://www.of.gov.uk>
- Oftel (2001) *Assessment of Demand Elasticity Estimates of .econ*, A paper by Dr John Hunter and Professor Christos Ioannidis, Brunel University, London: Oftel, September 2001, and available at <http://www.ofcom.org.uk/static/archive/oftel/publications/mobile/hunt0901.htm>
- Oftel (2002a) *Ramsey Prices and Network Externalities: Dr. Rohlfs' Analysis*, London: Oftel, 23 May
- Oftel (2002b) *Ramsey prices and the incentives of mobile operators*, London: Oftel, 12 February
- Oftel (2002c) *Ramsey pricing – Oftel's response to letter of 4 July*, London: Oftel, 12 July
- Shapiro, M.D. (1987) 'Measuring market power in US industry', National Bureau of Economic Research Working Paper: 2212 April.
- Sheffman, D.T. and P.T. Spiller, (1987), 'Geographic market definition under the US Department of Justice Merger Guidelines', *Journal of Law and Economics*, 30: 123-47
- Stigler, G.J. and R.A. Sherwin, (1985), 'The Extent of the Market', *Journal of Law and Economics*, 28: 555-85
- Vodafone (2003) *Review of price elasticities of demand for fixed line and mobile telecommunications services*, August, submitted to the New Zealand Commerce Department and available at <http://www.comcom.govt.nz/telecommunications/obligations/Vodafone%20paper%20on%20price%20elasticities%20for%20weighted%20revenues%20approach.PDF>

Notes

- ¹ The regulatory functions of Oftel were transferred to a new regulatory agency – Ofcom – at the end of 2003.
- ² Strictly, super-elasticities, that take into account the impacts of price changes on demand for all services. See Ofcom (2003, K1)
- ³ Armstrong and Vickers (2001) cited in Competition Commission (2003, 8.67). The subsequent discussion is somewhat confused, but can be reconciled with this claim by noting that MNOs would not choose to set Ramsey mark-ups for termination as they have SMP, but if they were forced to set Ramsey mark-ups for termination and all other services were competitive, and they could price discriminate between these other services, then overall competition would force them to set Ramsey mark-ups for all other services, delivering efficiency overall. Given that termination cannot realistically be made competitive, setting a price cap at Ramsey levels simulates the effect of competition. The claim by the Competition Commission at 8.71 that if termination were competitive, MNOs would have an incentive to undercut Ramsey mark-ups is met at 8.73 by pointing out that failing to recover the efficient (Ramsey) level of charges from one service (termination) would lead to a greater loss of profits from pricing above Ramsey levels for some other service.
- ⁴ IRG (2003), p. 3.
- ⁵ IRG (2003) p. 17.
- ⁶ IRG (2003), p. 23.
- ⁷ Ofcom took over Oftel's functions on 1 January 2004. All the documents cited here were produced by Oftel, but are now only available on the Ofcom web site.
- ⁸ Ofcom, 2003, Annex K, K1
- ⁹ Oftel (2002a) and Oftel (2002b). See also CC (2003, Appendix 1.1, chapter 9, Annex 5).
- ¹⁰ Oftel (2002c)
- ¹¹ Oftel (2001)
- ¹² Ofcom (2003, K21 et seq)
- ¹³ See for example Okada and Hatta (1999), who estimate both price elasticities and network effects for Japanese mobile telecommunications to determine a Ramsey optimal price structure. Vodafone (2003) provides a survey of estimates of price elasticities of demand, and notes that Okada and Hatta's estimates are implausibly high. This may be because of a rather short time-series and imperfect control of subscriber numbers by region. Grajek (2003) argues that properly controlling for network effects at the firm and industry level is critical for reducing biases that otherwise tend to increase the estimated elasticity, and discusses the problem of non-linear pricing.
- ¹⁴ For an overview see Doern, G.B., and S. Wilks (eds.) 1996. *Comparative Competition Policy: National Institutions in a Global Market*, Oxford: Clarendon Press
- ¹⁵ See, for example, Sheffman, D.T. and P.T. Spiller, 1987, Geographic market definition under the US Department of Justice Merger Guidelines, *Journal of Law and Economics*, 30: 123-47, and Stigler, G.J. and R.A. Sherwin, 1985, The Extent of the Market, *Journal of Law and Economics*, 28: 555-85
- ¹⁶ Hausman J.J. and G.K. Leonard, 1997 Economic analysis of differentiated products mergers using real world data, *George Mason Review*, 5, 321-46
- ¹⁷ Commission Notice on the definition of the relevant market for the purposes of Community competition law, OJ 1997 C372/5.
- ¹⁸ Newbery (1999), *Setting Transportation Tariffs using Transcost*, London: Ofgas gives an account, while the model can be accessed from Transco's website <http://www.transco.uk.com/>.
- ¹⁹ The Australian Consumer and Competition Commission (ACCC) prepared Incentive regulation, benchmarking and utility performance in 2000 for the Utility Regulators Forum, summarising the state of debate of various state level regulatory agencies and noting the importance of incentive-based regulatory methodology based on 'licensee-specific benchmarks'.
- ²⁰ The 'waterbed effect' (Oftel, 2003 describes it as "swings and roundabouts") is the extent to which changes in mark-ups on one service (like termination) are offset by changes in other mark-ups. In the extreme case overall contributions to the fixed and common costs will not vary as one mark-up is changed by regulation, just as pushing down on one part of a waterbed will cause it to bulge out somewhere else. Of course, the relative prices of different services will change and with them the overall efficiency of pricing, and hence the level of social welfare.
- ²¹ Thus Shapiro (1987) measures the market power of US industry, by measuring the ratio of the firm's to the market elasticity as a measure of non-competitive conduct. To implement this measure, both the firm's and the market elasticities of demand must be estimated. The elasticity faced by the firm can be estimated from the cyclical behaviour of productivity, while the market elasticity was estimated using an instrumental variables procedure exploiting a covariance restriction between productivity shocks and demand shocks.



Professor Patrick Rey

Professor of Economics, University of Toulouse
Research Director, Institut d'Economie Industrielle (IDEI)

Patrick Rey holds a PhD and a « Habilitation à diriger des Recherches » (University of Toulouse, 1992). He is currently professor of economics at the University of Toulouse and is also a professor at Ecole Polytechnique. His work has been published widely in renowned economic journals and he has served on the editorial board of several of them (*Review of Economic Studies* 1992-1999, *Journal of Public Economics* 1998-2002, *Journal of Industrial Economics* since 2001, *Journal of European Economic Association* since 2003). He has served as expert for OECD, the World Bank, the Commissariat au Plan and the European Commission, where he has been a member of the *Academic Advisory Group of the Merger Task Force* and is now a member of the *Economic Advisory Group on Competition Policy* of DG Competition. He also served on the *Commission Consultative sur les Réseaux et Services de Télécommunications* (hosted by the French telecom regulator).



Dr. Bruno Jullien

Director of Research, CNRS
Research Director, Institut d'Economie Industrielle (IDEI)
Director of GREMAQ (University of Toulouse)

Bruno Jullien graduated from Ecole Polytechnique (1982) and ENSAE (1984), and holds a PhD in economics from Harvard University (1988). He is currently at the University of Toulouse where he is heads the GREMAQ, the main research unit of the University. Bruno Jullien is also Professor at Ecole Polytechnique. His work has been published widely, covering a wide range of topics, with a particular focus on Industrial Organisation, contract theory, network economics, competition policy and regulation, and the economics of risk.

Bruno Jullien has served as expert for the Commissariat au Plan, the Agence de Régulation des Télécommunications (ART) and the European Commission where he acted as an economic expert on a case for the DG Competition, a contributor to reports and a member of an economic expert panel for the DG Information Society commissioner.

Mobile to mobile call termination

It is generally accepted that mobile operators have a joint interest in increasing fixed to mobile (F2M) termination rates, so as to extract revenue from the fixed line operator(s). However, there is still an ongoing debate on the impact on broader efficiency and social welfare outcomes of deregulating mobile to mobile (M2M) termination rates.

A growing literature¹ suggests that where competing mobile operators set termination charges by bilateral interconnection agreements, removing the option for operators to act

independently, then, in contrast with the situation for F2M termination, operators no longer have the incentive to increase termination charges.

Despite this, authorities have, so far, been unwilling to contemplate treating F2M and M2M termination differently in regulatory terms. A number of concerns are typically cited. First, that bilateral agreement of M2M termination rates could be used as a collusive instrument to dampen retail competition and raise mobile operators' profits. Secondly, that asymmetries

between networks could result in M2M termination rates being set inefficiently (in particular, that incumbent operators could choose to set use high M2M termination charges so as to restrict the entry/expansion of smaller competitors). Thirdly, if the retail markets in which operators are competing are not fully competitive, that this might affect how mobile operators set bilaterally agreed M2M termination rates.

These concerns have been expressed recently by a number of different groups, including Oftel/Ofcom, the European Regulators Group (ERG) of National Regulatory Authorities, the European Commission and the Independent Regulators Group (IRG).²

In the following discussion we attempt to address these issues. In general we find that the results of the literature are not as ambiguous as may have been suggested. Moreover, where uncertainties exist, it is by no means clear that they imply that mobile operators would choose to set excessively high M2M termination rates.

It is worth commenting at the outset that it is possible to view M2M termination in an entirely different light to F2M termination. F2M termination, if unregulated, presents an opportunity for mobile operators to exercise market power derived from the termination bottleneck. In contrast, M2M termination agreements are horizontal agreements between networks which are producers of highly complementary products (because of the network externality of interconnection). While the welfare effects of such horizontal agreements are not always unambiguous they are by no means always deemed to be bad.

A simple intuitive explanation of why F2M and M2M termination are different is that M2M termination clearly cannot be a source of profit for the mobile industry as a whole. By definition, the sum of all M2M termination charges within the mobile industry is zero. Hence the mobile operators collectively cannot be, *per se*, generating profits from M2M call termination.

The risk of “collusion” in setting M2M rates

While setting M2M call termination charges above cost cannot generate profits for the mobile industry directly, it remains true that M2M termination charges can have an impact on mobile operators’ pricing policies and, therefore, on their profits.

This issue lies at the heart of the main criticism levelled at allowing M2M call termination rates to be determined by bilateral negotiation: that the result is likely to be a “collusive” one, where mobile operators use M2M termination rates to push up call charges and thus weaken retail competition. However, what often fails to come across clearly in the debate is that this result that has been shown to be specific to only one form of competition between operators: when operators compete with each other on usage (or linear) tariffs alone.³ This result does not apply when operators compete with more sophisticated forms of

pricing policies (including two-part tariffs, quotas, subsidizing handsets, and so forth).

In general, operators are indeed expected to base their usage prices on their *perceived* marginal costs, which include the termination mark-up for the share of traffic that goes off-net; as a result, reciprocal termination mark-ups will generate higher profits from usage, as they induce the operators to set retail prices above the true cost.⁴ The literature⁵ however points out that when operators set separate charges for subscription and usage (two-part tariffs), then they will compete away profits from calls in customer acquisition and retention (e.g., through lower subscription fees or higher handset subsidies). Hence mobile operators would be misleading themselves to suppose they are making profits from call termination, without taking into account the indirect impact of these higher usage revenues on their other sources of revenue (e.g., subscription fees) or expenses (e.g., handset subsidies).

It has been suggested that the first case may be relevant to the mobile industry, because the majority of subscribers have pre-pay phones, and therefore only pay usage charges, while most contract customers pay for bundled minutes, consume only within their bundle and thus effectively only pay a monthly fixed charge.

However, this view overlooks the point made above, which is that with two-part tariffs operators cannot raise profits by raising retail call charges, because these profits are competed away in lower subscription charges. Therefore, the key question is the extent to which profits from usage can be dissipated through other dimensions of the pricing policies. Even in the case of pre-pay customers, mobile operators “charge” periodically for “subscription”, in the sense that operators compete in handset subsidies to attract and then retain customers; this feature is all the more important as customers periodically want to replace their handsets, as the technology and the offered services evolve over time. All that is required for the two-part tariff result to apply to pre-pay customers is that there be a relationship between the level of call charges and the price that mobile operators “charge” these customers for handsets. If this is the case, then setting call charges above cost could not be used to sustain excess profits in the mobile sector. In a competitive market there is every reason to expect to see such a relationship (between call charges and handset prices) because higher access charges induce higher usage prices and thus make customers more profitable to the operators, who would be expected to compete these profits away in attempting to attract subscribers by subsidizing handsets (setting prices below cost).

We also believe that these arguments apply to contract customers. Although this specific type of bundling contract has not been explicitly studied in the literature⁶, the same intuition applies: raising call termination charges may raise retail call charges but any profits can be competed away through all those pricing components that are not sensitive to traffic, including

fixed fees and handset prices. Hence there is no reason to believe that the use of bundles undermines the general conclusion about two-part tariffs.

In conclusion, once reciprocity is granted, we should think about mobile operators setting M2M call termination charges at the level that supports the profit maximising (or welfare maximising) level of retail tariff equilibrium rather than to generate profits from access (or to contribute to fixed and common costs). It is therefore wrong to apply to M2M termination charges the kind of test that could be inspired by standard, “one-way” bottleneck analyses. It would be misleading to think that each mobile operator has a systematic incentive to raise its termination rates to extract more revenue from the other mobile operators. This view would overlook the fact that mobile operators are actually “competitive bottlenecks”⁷. While they have control *ex post* over the termination of calls to their subscribers, they must also compete in order to attract and keep these subscribers. Any market power that the operator may exercise in the termination segment is bound to intensify retail competition.

The effect of asymmetries between networks in setting M2M rates

There are a number of ways in which actual mobile networks differ from those in the theoretical literature. In addition to the concerns already addressed, these additional factors have been cited as reasons why the existing theoretical literature cannot be relied upon as indicating what would happen if M2M termination charges were not directly regulated. We would divide these factors into two groups. The first group:

- traffic imbalances between operators;
- asymmetries in network size and the use of M2M termination charges to create barriers to entry or expansion; and
- cost asymmetries between networks,

we do not perceive to have a major bearing on the likely outcome of a bilateral process for setting termination rates. The second group:

- the effect of mobile network externalities and
- the implications of the use of differential on-net/off-net pricing;

raises more questions, because the effect of these factors is uncertain or the predicted outcome appears at variance with what we actually observe. This latter group represents areas where we feel more work can fruitfully be carried out so as better to understand the dynamic behaviour of mobile markets.

In the following sections we deal with each of these issues in turn.

Traffic imbalances

Most theoretical models of telecoms pricing assume that interconnection traffic is in balance. As a result no operator is a net buyer or seller of termination. However, in practice imbalances do exist between operators. In the short term it may be perceived that in these circumstances the net seller of termination will want a high termination charge and the net buyer will want a low charge. At best this could make the process of negotiating a termination charge sticky and at worst it could make the final outcome uncertain. OfTel (2003) cite this as being a reason why movements in M2M termination rates would be sticky in a downwards direction.

Dessein (2003) studies networks with a range of different types of consumer and allows for traffic imbalance at the customer level (although traffic is balanced in equilibrium at the level of the operator) and finds a neutrality result. Laffont-Rey-Tirole’s work on networks with asymmetric regional coverage results in traffic imbalances in equilibrium as a result of different (average) outgoing prices being offered by the two networks. This does not change however the general analysis of the preferred level of the termination charge between the two networks.

It does not seem, therefore, that there is a strong basis to suggest that traffic imbalances would lead operators as a whole to deviate from cost-based termination rates in a way that would reduce welfare.

Entry barriers

It has also increasingly a received wisdom that incumbent operators may have a strategic incentive to agree on “high” M2M call termination charges, in the knowledge that these would also apply to an entrant. This has been proposed as another reason why regulation of M2M termination charges is still required.⁸

We note that high termination charges only disadvantage an entrant if the entrant were expected to run an interconnection traffic deficit, which in turn depends on traffic imbalances. If interconnection traffic with the entrant is balanced then the entrant may actually prefer high termination charges⁹. More generally, there is no automatic link between network size and traffic balance; in particular, a small operator can for example “target” those subscribers that tend to receive calls more than they place ones, in order to generate a traffic surplus.¹⁰

However, a common entry strategy is to undercut the incumbents’ prices. If an entrant adopts this approach on call charges then there would be reason to expect the entrant to run an interconnection deficit (since its customers would be encouraged to call more often than the others). A high termination charge could then penalise the entrant, as well as restricting its ability to undercut the level of the incumbents’ call charges. But this

should be viewed in light of the above-mentioned possibility that the entrant could target customers that generate access revenue, i.e. those who receive more calls than they place. In addition, even in those situations where the entrant would nonetheless run an access deficit, it is still not clear that incumbents would prefer a high termination charge. Incumbents would have to trade off the damage to their own interests of accommodating an entrant with the damage they do to the market as a whole, including themselves, by setting an excessive termination charge.¹¹ The outcome of such a trade-off is uncertain.

Indeed, Carter and Wright¹² have pointed out that, with operators of different sizes, it may be socially optimal in certain circumstances to let the larger operator choose the reciprocal termination rate. In their set-up, the larger operator always prefers to set the termination rate at cost (in contrast, the other operator does so only if it is small enough). This results from a particular trade-off between the termination profit and the impact of termination rates on retail competition. While this result may be particular to the way in which the authors have modelled the asymmetry of the operators, and other ways of modelling network asymmetries may yield different results, their analysis does establish that it would be misleading to assert simply that larger or incumbent networks have an incentive to drive up M2M termination charges for strategic reasons.

Cost asymmetries between networks

Intuition suggests that minor differences in cost between networks would not lead to a major divergence in equilibrium, but the exact outcome could be hard to predict. The most efficient network will be more attractive; however, with non-discriminatory pricing call prices should differ and generate a traffic imbalance in favour of the least efficient operator, since it would charge higher prices. Thus the effects discussed above should apply. While there is no reason to believe a priori that cost asymmetries would dramatically change the conclusions, this may be worth investigating further as the implications of cost asymmetries remain so far poorly understood.

Mobile network externalities and differential on-net/off-net pricing

The theoretical result from the existing literature¹³ on competition with differentiated prices for off-net calls is that externalities tend to imply that the profit maximising termination charge is below cost while the welfare maximising level may be above cost. The main result where networks can differentiate on-net from off-net pricing concludes that networks would price off-net calls below cost (even in the absence of externalities).

Clearly, these results must be missing some aspect of the dynamics that actually drives off-net charges to exceed on-net charges.

This reveals the limitation of the existing work in the area of M2M termination charges. Although it is far from clear whether these observations have any bearing on how operators would set bilateral M2M call termination charges in the future, it is impossible to be entirely confident of the outcome when existing theoretical models do not fully explain existing pricing behaviour.

Imperfect competition

In the context of setting F2M termination rates, both Oftel and the Competition Commission have questioned the extent to which the UK mobile market is fully competitive. It has also been suggested that if the retail market were not fully competitive this would strengthen the case for regulatory intervention in setting termination rates.

Given these observations it is important to ask whether the case for deregulating M2M call termination relies on the retail mobile market being perfectly competitive?

In our view, the answer is that it does not. First we note that, due to spectrum scarcity and large sunk costs, there is only room for a limited number of operators in the mobile market, since these operators must recover fixed and sunk costs. Hence, what might appear as imperfect competition in terms of concentration can be the outcome of (possible very tough) competition in terms of investment, infrastructure, and so forth. Second, recognising this feature of the mobile telephony industry, the literature on M2M call termination usually relies on models of “imperfect” *ex post* competition, based on some form of differentiation among the operators¹⁴. We should actually point out that, in a hypothetical context of “pure and perfect” *ex post* competition, profits would be competed away – whatever the access charges the operators would first agree to; hence, by nature, operators could not have any incentive to manipulate the access charge and depart from the welfare-maximizing level: the whole issue would thus be void. The literature on this issue thus always assumes away this hypothetical scenario of perfect *ex post* competition, usually by accounting for some form of differentiation among the operators. Some of the literature has for example focused on a particular “Hotelling-type” model¹⁵ of Bertrand price competition among horizontally differentiated operators. These models produce a “profit neutrality result”: operators compete in two-part tariffs and while aggregate profits in the market are above the perfectly competitive level, due to the differentiation among the operators, they are completely independent of the level of the termination charge. In these models, therefore, imperfect competition has no bearing on the level of the call termination charge.

This profit neutrality is a model specific property – it relates to the fact that, in Hotelling-type models with a fixed number of subscribers, aggregate profits are mainly triggered by the degree of differentiation among competitors and independent of the general cost level in the industry and/or the surplus brought to consumers. However, if operators were to retain even a small

part of any increase in usage surplus (the profit neutrality result could be viewed as a special case, where mobile operators cannot do so), then they would have an incentive to raise usage surplus. In the absence of externalities, maximising total usage surplus can be achieved by setting retail call charges at cost, which, in turn means mobile operators would wish to set M2M termination charges equal to the (true) marginal termination cost. Hence, far from giving mobile operators a reason to want to set M2M termination charges high above cost, imperfect competition in the retail market may actually encourage mobile operators to set M2M call termination at low levels.

When subscription responds to pricing conditions, adding one subscriber benefits all others, which would call from a welfare standpoint for subsidizing access. But the same logic can apply as well to firms. And indeed, in a model where demand for subscription is elastic and the profit neutrality thus does not hold, Dessein (2003) concludes that the profit maximizing termination rate is below cost. This conclusion derives precisely from the network externality generated by the participation of a new customer, which benefits all mobile users.¹⁶

The fact remains that the level of the termination charge generally does affect the retail equilibrium. Depending on the exact nature of retail competition, as well as on the precise characteristics of demand and supply conditions, it may also affect equilibrium profits. If so, there is then a level of termination charge that maximizes the industry profits. Allowing the operators to adopt this termination charge may well benefit consumers, however. It remains possible that the characteristics of competition could lead to a divergence between private interest (setting termination charges so as to maximise profits) and the social interest (setting termination charges so as to maximise consumer surplus and/or total welfare), but quantifying this is a complex issue that relies on a fine analysis of the nature of retail competition and of the characteristics of the various operators¹⁷. It may therefore be difficult to determine the socially desirable level for the access charge. In particular, insisting on termination charges that reflect exactly the cost of termination may well reduce consumer surplus and total welfare, compared with the situation where the operators would freely set termination rates.

Conclusions

The first observation is that the issues raised by M2M termination differ drastically from F2M termination. The second main conclusion is that, except under very specific conditions, mobile operators would not achieve a collusive outcome by agreeing on high termination charges.

This conclusion has been shown to apply even where the mobile retail market is not perfectly competitive, on the basis on the analyses of competition with two-part tariffs.

It is also applicable to pre-pay and contract customers, notwithstanding the fact that pre-pay customers do not pay regular fixed (as opposed to usage) charges while many contract customers pay fixed monthly bills that cover the bundle of calls that they have purchased. The point is that changes in retail call charges should feed through to the retail price of handsets, or in the case of contract customers, through any other part of the bundle. Moreover, multi-part tariffs of all sorts should be expected to reduce the incentive for mobile operators to drive up retail call charges.

A number of factors have been discussed, such as asymmetries or traffic imbalance, but, based on current knowledge, these factors do not appear to support an urge for regulatory intervention, notwithstanding the fact that their implications for regulation are ambiguous.

It remains that some aspects such as cost asymmetries, traffic imbalance, or consumers' heterogeneity would require some further investigation before any definite conclusion could be drawn. What is clear, however, is that it is not possible to claim either that large networks have an unambiguous incentive to drive up the termination rates paid by small networks or that high termination charges act clearly as a barrier to entry or expansion.

Notes

- ¹ Laffont J.J., Rey P. and J Tirole (1998), "Network Competition I: Overview and Nondiscriminatory Pricing", *RAND Journal of Economics* 29 (1-37); Hahn JJ (2000), "Network Competition and Interconnection with Heterogeneous Subscribers, mimeo, Oxford University; Dessein W. (2003) "Network Competition in Nonlinear Pricing", *RAND Journal of Economics* 34 (1-19); Gans J. and S. King (2001), "Using Bill and Keep Interconnect Agreement to Soften Network Competition, *Economic Letters* 71 (413-420); Carter M. and W. Wright (2003), "Asymmetric Network Interconnection, *Review of Industrial Organization* 22 (27-46).
- ² European Regulators Group (ERG), "Draft joint ERG/EC approach on appropriate remedies in the new regulatory framework", November 2003. www.erg.eu.int/doc/publications/erg0330_draft_joint_approach_on_remedies.pdf.
- Valletti, T., "Obligations that can be imposed on operators with significant market power under the new regulatory framework for electronic communications – Access services to public mobile networks", September 2003
- http://europa.eu.int/information_society/topics/ecom/useful_information/library/studies_ext_consult/index_en.htm
- Independent Regulators Group (IRG), "Principles of Implementation and Best practice on the application of remedies in the mobile voice call termination market", November 2003. <http://irgis.anacom.pt/admin/attachs/359.pdf>
- Oftel, "Wholesale Mobile Voice Call Termination: Proposals for the identification and analysis of markets, determination of market power and setting of SMP conditions", December 2003. http://www.ofcom.org.uk/legacy_regulators/oftel/mobile_call_termination/
- ³ Armstrong M. (1998), 'Network Interconnection', *The Economic Journal* 108 (545-564), Laffont J.J., Rey P. and J Tirole (1998), "Network Competition I: Overview and Nondiscriminatory Pricing", *RAND Journal of Economics* 29 (1-37).
- ⁴ That is, even if termination mark-ups have no direct impact on profits when traffic is balanced, they have however an indirect impact through their effect on the operators' usage prices.
- ⁵ Laffont J.J., Rey P. and J Tirole (1998), "Network Competition I: Overview and Nondiscriminatory Pricing", *RAND Journal of Economics* 29 (1-37); Hahn JJ (1999), "Network Competition and Interconnection with Heterogeneous Subscribers, mimeo, Oxford University; Dessein W. (2003) "Network competition in Nonlinear Pricing", *RAND Journal of Economics* 34 (1-19).
- ⁶ In the studies by Hahn JJ (1999) and Dessein W. (2003), competing networks offer menus of fixed quantity/price contracts as a selection device among heterogeneous consumers. Actual bundles can be interpreted as menus of "three-part tariffs" that include a fixed fee, a quota of "free minutes" covered by the fee, and a marginal price for usage beyond the quota.
- ⁷ The term « competitive bottlenecks » has been used routinely in the literature to refer to these situations, as in the mobile telecommunications industry, where the operators need each other to access subscribers, but compete as well for these subscribers – see Laffont J.J. and J Tirole (1999), *Competition in telecommunications*, MIT Press, and Armstrong M. (2002), "The Theory of Access Pricing and Interconnection," in *Handbook of Telecommunication Economics*, eds M. Cave, S. Majumdar and I. Volgelang, North-Holland. This term has for example been used specifically to distinguish M2M from F2M termination by Valletti, T. (2003), "Termination charges: An Economist's view", presentation to the Global Communications Consortium Seminar *The Competition Commission's Inquiry on Termination Rates for Mobile Network Operators: An assessment*, held on Feb. 25 by the London Business School.
- ⁸ See references from Oftel, ERG and IRG above.
- ⁹ See Carter M. and W. Wright (2003), "Asymmetric Network Interconnection, *Review of Industrial Organization* 22 (27-46).
- ¹⁰ If customers have a "balanced calling pattern" (that is, their calls are evenly distributed among the population), then every user is equally likely to be called by any other one. Then the amounts of inbound and outbound calls between two networks are both proportional to the number of subscribers of each of the two networks, so that the traffic is balanced irrespective of network size (that is, denoting by N_A and N_B the number of subscribers in networks A and B, and assuming for instance that each customer calls every other customer for M minutes, the volume of traffic from A to B and the volume of traffic from B to A are both equal to $N_A \times N_B \times M$). Traffic imbalance arises rather from different calling pattern within two networks due to biased calling pattern, consumer heterogeneity and different compositions of subscriber bases (e.g., one network focussing more on "callers" than on "receivers").
- ¹¹ In particular, Laffont-Rey-Tirole (1998) point out that reciprocal access charges cannot be used to deter entry, since the entrant can always avoid incurring an access deficit by "match" the marketing and pricing policy of the incumbent.
- ¹² See footnote 10.
- ¹³ Laffont J.J., Rey P. and J Tirole (1998), "Network Competition II: Price Discrimination", *RAND Journal of Economics* 29 (38-56).; J. Gans and S. King (2001), "Using 'Bill and Keep' Interconnect Arrangements to Soften Network Competition," *Economic Letters*, Vol. 71, No.3, June 2001, pp.413-420.
- ¹⁴ Although fixed costs are not always modelled explicitly, they can be accommodated in the models with no change at all, because fixed costs can simply be recovered from excess profits. Note however, that this implies that the benchmark for the optimum M2M charge in the absence of externalities is traffic-sensitive cost and not a marked-up value to reflect fixed costs.
- The literature moreover points out that, under the Caller Pays Principle, there is no "perfectly competitive" equilibrium (ex post) whenever the access charge departs from cost. That is, too intense competitive pressures would destabilize any candidate equilibrium in pure strategies. It could remain equilibria in mixed strategies, where operators would choose their pricing policies "randomly" – a formal way to capture competitive instability.
- ¹⁵ The standard Hotelling model is as follows: firms have a constant unit cost of production (constant returns to scale) and are horizontally differentiated (which can be interpreted as being located at two ends of a "road"); consumers are distributed along this road and have a preference for the "closest" competitor (e.g., they must incur a "transportation cost" to go shopping); consumers are willing to buy one only unit, so that the price can be interpreted as a "subscription fee", and all consumers buy in equilibrium (that is, all consumers have a sufficiently high reservation price for the good). A particular (and unrealistic) feature of this Hotelling model is that the equilibrium profits are insensitive to a uniform increase in the cost of the firms. The literature on M2M termination has partly focused on a generalised Hotelling model where in addition consumers have an elastic demand for usage; however, when operators compete in two-part tariffs, usage prices then simply reflect perceived marginal costs (including termination charges) while subscription fees play the same role as the "prices" in the standard Hotelling model. As a result, equilibrium profits are again insensitive to the level of the (perceived) marginal costs); inflating the termination charge thus does not affect the profits but only results in higher usage fees, thereby introducing an undesirable distortion that reduces both consumer surplus and total welfare.
- ¹⁶ Similar conclusion are obtained by Schiff A. (2001), Two-way Interconnection with Partial Consumer Participation, mimeo.
- ¹⁷ For instance, W. Dessein (2003) argues that in a Hotelling model where the intensity of competition varies across categories of users along with the usage intensity, profits can be enhanced by setting the call termination charge above or below cost, depending on the correlation between usage intensity and "transportation cost".

Professor Stephen C Littlechild



Honorary Professor, University of Birmingham Business School and Principal Research Fellow, Judge Institute of Management Studies, University of Cambridge

Professor Littlechild has been working in the area of network economics and regulation for over 30 years and is an acknowledged world authority on the use of price controls in network industries. He was Director General of Electricity Supply (DGES), in charge of the Office of Electricity Regulation (OFFER), from its foundation in September 1989 to December 1998. In 1983 he proposed the RPI-X approach to price controls, which has since been widely adopted for regulating utilities in the UK and overseas. He was a member of the Monopolies and Mergers Commission for six years, prior to his role at OFFER, during which time he participated in reports on Manchester Airport (the first use of RPI-X control after BT) and British Gas (steps to secure gas competition), several efficiency studies of nationalised industries, and numerous monopoly and merger references.

Price controls on mobile termination charges

Executive summary

Regulation of mobile termination charges varies considerably across Europe. There is increasing intention to present a unified approach. But some approaches could have an adverse effect on competition and efficiency.

The UK now has considerable experience in setting price controls on utilities. Some controls are intended to stimulate productive efficiency in monopoly networks; others aim to facilitate a transition to competition. Controls on mobile termination charges have a yet different stated aim – to relate prices more closely to costs so as to increase allocative efficiency, and to reduce cross-subsidisation between users. Yet considerations of productive efficiency and competition are important here too.

In order to achieve these latter aims, price controls need to be related to the costs achievable by efficient companies, not to the costs actually incurred by a company regardless of efficiency. Controls on mobile termination charges in the UK are consistent with this principle: the allowed charges differ only to the extent that operators have different access to radio spectrum, which is largely outside their control.

Elsewhere in Europe the approach varies. There is a much greater variation between the termination charges of different operators, including between operators with access to the same

type of spectrum. This suggests that in some countries there is less concern to limit differences in charges to differences in costs outside the control of the operator. Such approaches provide less incentive to efficiency, and greater likelihood of distorting competition between operators. A policy of setting asymmetric price controls in this way is likely to be to the detriment of customers in the longer term.

1. Introduction

The mobile telephone sector in Europe presents an interesting paradox. On the one hand, the sector has become steadily more competitive and has expanded remarkably quickly and successfully. Competition has brought lower connection charges and call prices, cheap handsets, improved quality of service and a rapid rate of innovation. Not surprisingly, subscriber coverage has expanded significantly and the number of calls per subscriber has increased. It would be difficult to find other sectors of the economy that have performed so successfully

On the other hand, regulation of the mobile sector is still present, and in some respects is increasing in severity. For example, although some countries such as the UK have removed regulation of outbound calls, the UK Competition Commission, consistent with the recommendation of Oftel, has recently imposed price reductions on mobile termination charges of the order of 15% per year for each of the next four years, on all four

main mobile operators.¹ Other European regulators have also been actively controlling mobile termination charges, and there seems little prospect at present of the trend reversing.

Nonetheless, there are important differences between the regulatory approaches in Europe. For example, Oftel in the UK has taken a symmetric, forward-looking, multi-year approach to setting price controls. In contrast, some other regulators have tended to set prices asymmetrically, based on past costs and on an annual basis. The latter approach is liable to impact adversely on competition and to reduce the incentive to efficiency.²

There are now attempts to unify the framework of regulation across Europe. For example, the European Regulators Group has recently published for consultation a draft paper that seeks to promote a consistent regulatory approach on the choice of appropriate regulatory obligations (or remedies) to be imposed by National Regulatory Authorities (NRAs) on operators with significant market power.³

Which approach to adopt as a uniform one is a matter of some importance. The relationships between price controls, incentives and competition could have adverse effects if overlooked.

I have been asked to look at the basis on which price controls on mobile termination charges have been or should be set, with particular reference to the issue of symmetry or asymmetry as between different operators within each national market.

Section 2 notes the statutory criteria that telecommunications regulators need to consider. Sections 3 and 4 review the approach and experience in setting price controls on utilities since privatisation, contrasting the controls used in monopoly networks and those in potentially competitive industries. Sections 5 to 7 explain the relatively symmetric approach of the Competition Commission on UK mobile termination charges and the importance of such symmetry. Sections 8 and 9 describe and contrast the differentials in charges that characterise most other European countries. Section 10 points out the implicit disadvantages of setting charges on an asymmetrical basis. Section 11 summarises and concludes.

2. Regulatory criteria

The statutory duties of European telecommunications regulators differ in detail, but share many common features. EC law also provides an important element of commonality, with a view to providing a level of consistency across Europe. For example, Article 9, headed 'General responsibilities of the national regulatory authorities', says that they must

- encourage and secure adequate interconnection in the interests of all users
- act in a way that provides maximum economic efficiency and gives the maximum benefit to end-users, and

- take into account the need to stimulate a competitive market.

The questions then arise: how are these duties best discharged? How best to balance the interests of consumers with respect to price and the promotion of competition and efficiency?

In parallel with the privatisation of other utilities and the introduction of competition, many national Governments have introduced price controls designed to fulfil similar aims in these other sectors. It will therefore be helpful to look at the principles behind these controls, distinguishing between permanent controls aimed at protecting users of monopoly networks and transitional controls aimed at facilitating the introduction of competition. This then raises the question whether should there be a stronger attempt to secure consistency of approach in the setting of controls on mobile termination charges.

3. The origin of modern price caps

The RPI-X price control, that has become the standard regulatory approach in the UK and many other countries, was originally proposed at the time of the privatisation of British Telecommunications in 1983.⁴ It had several objectives, including:

- to assure customers that BT's prices (in real terms) would go down and not up after privatisation
- to provide better incentives to efficiency and innovation than the traditional US cost of service approach, and
- to facilitate the development of competition in the telecommunications sector.

Although it was hoped and expected that this approach would encourage lower prices and increased output, the aim was not primarily what economists would call "allocative efficiency"⁵. A satisfactory "distribution of income" was an important aim, in the sense that customers would get an acceptable "share of the cake". But the main aims were greater "productive efficiency"⁵ (lower costs) so as to "increase the size of the cake" itself, and greater competition so as to increase efficiency and to ensure that the industry discovered and met the needs of the customers and shared the benefits with them.

The purpose of the price control was therefore summarised as "to hold the fort until competition arrives". It was envisaged that it might be possible to remove the control once there was sufficient competition.

In the event, the Government allowed less scope for competition than the report recommended. Despite promises that the price control would be removed, it has remained in place, and has repeatedly been tightened. It has continued to stimulate greater efficiency by BT, and to pass a significant share of the benefits to customers. But the successively tighter controls, and sub-

controls on the prices of local calls and connections, may have tended to discourage new entry and competition in the provision of fixed line telecommunications services. In other words, in the pursuit of some statutory objectives (prices and efficiency), other objectives (competition and innovation) may have been downplayed, outweighed or overlooked.

4. The development of price caps in utility regulation

In subsequent applications of price caps – in the UK and elsewhere, and in different sectors – the two different aims of the price cap have been more clearly distinguished. That is, in some cases price caps have been used to promote efficiency in a monopoly business, in other cases to facilitate the transition in a potentially competitive business.

(a) Network monopolies

For network monopolies such as electricity transmission and distribution, gas transportation and distribution, and to some extent water and airports, it has been accepted that effective competition is unlikely to be feasible in the foreseeable future. Accordingly, the main aim of the network price controls is not to promote competition or even to replicate its effects. Nor is the main aim one of allocative efficiency, since the demand for these products is relatively inelastic (so that modifications to the pricing structure would not have a great effect on the pattern of output). Rather, the aims are to regulate an effective monopoly so as to promote greater efficiency, and to secure that the benefits of this are shared with customers.

To this end, RPI-X price controls are set at periodic intervals, typically five years. Within any industry, the levels of these controls may vary from one company to another, but only to the extent that the costs of the companies necessarily differ and are outside their control – for example, by virtue of the different geographical areas they serve. The controls do not differ according to the efficiency or inefficiency of the companies, or according to the costs that ought to be under their control. Steps are therefore taken to calculate, and to base the price controls on, the costs of an averagely efficient company, or of companies at the “efficiency frontier”. This is true for price controls on utility networks not only in the UK, but also in countries such as Norway, Sweden and Australia.

The reason for this approach is simple. If actual costs were simply passed through, it would remove or reduce the incentives of companies to become more efficient, and to innovate. This in turn would be against the interests of customers in future.

(b) Potentially competitive sectors

Other sectors like electricity generation and gas production, and the retail supply of electricity and gas, are actually or potentially competitive markets. Accordingly, where price controls have

been used, the aim has increasingly been to promote such competition, not to replace it. To that end, any controls have been envisaged as transitional.

The wholesale electricity generation and gas markets have never been price controlled because they were seen as potentially competitive from the beginning. The retail supply markets for electricity and gas were subject to “cost-pass-through” controls in the early years, before the markets were opened and retail competition allowed. As with all such controls, this limited profit margins but did not necessarily ensure efficient purchasing, so customers were still vulnerable to monopoly.

With the opening of the retail market for smaller customers (including residential), the pass-through price controls were at first replaced by fixed price caps. These provided better incentives on the companies to purchase efficiently. They were also set with a view to promoting competition.

For example, the electricity retail cap left sufficient margin to make it worthwhile for existing suppliers to compete for customers, and for new suppliers to enter the market. Although the price cap was subsequently tightened, room was still left for competition to grow. It was another case of holding the fort until competition arrived. In this case, after four years sufficient competition had developed that the retail price controls were removed altogether. The same applied to the gas price caps.

Other countries have tended to set rather tighter price controls on their competitive markets. The result is that although retail competition has often developed, it has not yet done so sufficiently vigorously that the regulator has felt able to remove the price cap.

Considerations of allocative and productive efficiency are not overlooked in regulating these competitive businesses. However, these aims are seen as best promoted by greater competition.

5. The Competition Commission approach to mobile termination charges

How are price controls set in the case of mobile telephone termination charges? They do not fit easily into either of the two categories described above.

Is call termination a potentially competitive market? Certainly the UK mobile sector as a whole is potentially and increasingly competitive, as the Competition Commission affirms. However, within the sector, the Commission held that each mobile operator has a monopoly of termination to its own subscribers, so that mobile termination services are not competitive. As a result of the Calling Party Pays system, each operator can and does exercise market power. Furthermore, since the Commission found that there was no realistic prospect of competition to provide termination services in the foreseeable future, there seemed no point in setting a price control to promote competition.

Are termination services therefore to be regarded as a monopoly, where the aim should be to increase efficiency and pass the benefits to customers? The Commission did find that the provision of termination services was a monopoly, but it did not hold that the sector was inefficient, either in its production costs generally or in termination services in particular. It did not see the aim of a price control as being to increase productive efficiency.

Nor did the Commission see the main aim as being to ensure a better distribution of income as between customers generally and the mobile operators. It did not find that the operators were making or were liable to make excess profits. Although it held that termination charges were significantly above the marginal costs of providing termination services, it also found that the operators tended to compete away these profits in trying to attract subscribers. In particular, it found that the operators cross-subsidised handsets.

The Competition Commission explained (para 2.509) that it thought it right to set termination charges that attributed costs on the basis of who caused them, with a view to securing that

- consumers do not pay too much for fixed-to-mobile or off-net calls
- consumers who make more of those calls do not unfairly subsidise other consumers
- cost-reflective call charges should minimise distortions in the volume and pattern of calling
- less resource-intensive technologies should not displace more resource-intensive ones, and
- there should be less incentive to subsidise handsets.

The main aims of the Competition Commission were thus two-fold: to increase allocative efficiency by ensuring that the price of each service more closely reflected the cost of that service, and to improve the distribution of income as between the consumers of different mobile services by reducing the extent of cross-subsidy between the different services.

6. The continuing importance of productive efficiency and competition

Whether allocative efficiency is numerically very important compared to productive efficiency, and whether it is the role of a regulator to make fine distinctions between different classes of customers, are matters on which different views might be held. It is not clear that the net welfare gains from lower termination charges are very significant in this particular case⁶. It is thus debateable how far it is worthwhile for a regulator to take significant interventionary measures to attain these goals.

Whatever view one takes on that, however, it is important for a regulator to take account of possible implications for all the statutory duties. In particular, even when taking measures to deal with allocative efficiency, internal cross-subsidies, distribution of income, and so on, a regulatory authority needs to look also at the effects of such measures on productive efficiency and on competition in the sector as a whole.

There are two particular concerns with respect to controls on mobile termination charges. The first reflects a lesson learned from network price controls. If the price controls on the termination charges are set to reflect the actual costs of each company, this will not provide an incentive to increase efficiency. On the contrary, it will be an incentive not to bother to increase efficiency. The less efficient companies will see no need to catch up with the better ones, the better ones will see no need to innovate and become even more efficient. Over time, prices to customers may, in one way or another, be expected to reflect costs in the industry. So the losers will be customers as a whole.

Second, such an approach would have adverse implications for competition between the mobile operators. The less efficient operators would not only get their higher costs covered, they would pass on these higher costs to the other mobile operators who needed to terminate messages with their subscribers. Consequently, the costs of the more efficient operators would be unduly increased. In parallel, the costs of the less efficient operators would be unduly reduced by the lower termination charges of the more efficient operators. These modified differentials in costs would in turn impact on prices to subscribers. This would further distort the pattern of competition between mobile operators: not only would the less efficient operators be protected against their lower efficiency, they would actually gain relative to the other operators.

A regulatory authority that wishes to promote greater allocative efficiency and reduced cross-subsidisation thus also needs to take care to protect productive efficiency and competition. As in setting network price controls, particular attention must be given to the determinants of an operator's costs, and to distinguishing those costs that reflect efficiency as opposed to those costs that reflect conditions beyond the control of the operators themselves.

7. The UK Competition Commission calculation of price controls

How did the UK Competition Commission deal with these issues? It concluded that a price control was necessary, in the form of a price cap calculated to bring termination charges down to what it called "the fair charge". This comprised "a reasonable estimate of LRIC [Long Run Incremental Cost], plus an allowance for network fixed and common costs and relevant non-network costs, plus a markup for the network externality". (para 2.505)

Other regulators have taken different approaches in some respects. For example, electricity, gas, water and airport

regulators have typically started from the total costs of actual networks, rather than the incremental costs of hypothetical networks. These other regulators have also had to allocate fixed and common costs, but they have generally not had to deal with arguments for mark-ups for network externalities.

Telecommunications regulators in the US and elsewhere have frequently used a LRIC approach. This may reflect the more rapid rates of technological progress and cost reduction in that the telecommunications sector, and the view that users of these networks should not be denied the lower prices that should result under competition. It also implies that price controls should reflect efficient costs, not the actual costs incurred by any particular company. And in the absence of good reason to the contrary, price controls reflecting efficient costs should be uniform rather than asymmetric.

How did the Competition Commission approach the issue of symmetry between mobile operators with respect to reflecting costs in price caps? It argued that “the only differences we should allow are those that are absolutely outside the companies’ control: that is, differences due to the allocation of spectrum” (para 2.538). It rejected arguments for differences due to allegedly different costs of capital and market shares. It took the view that all companies could achieve the efficient scale of output and costs of operation, at least within a short time. On this basis, it set just two caps: one for the two operators with combined 900/1800 MHz systems and another for the two operators with 1800 MHz systems.⁷ No cap was applied to the new entrant Hutchison 3G.

Oftel subsequently endorsed the Commission’s approach and proposed a set of price controls for the period to 2005/2006. For 2004/05, average termination charges should not exceed 5.14 ppm (pence per minute) for combined 900/1800 MHz operators and 5.81 ppm for 1800 MHz operators. These charges would fall to 4.61 ppm and 5.19 ppm, respectively, in 2005/06. (All these figures are in 2000/01 prices.) Thus, the higher charge (applying to 1800 MHz operators) would be some 13.0% above the lower charge in 2004/5, falling to 12.6% above it in 2005/06. These caps refer to the average termination charge over the course of a year: they do not constrain the structure of these charges, hence operators are free to vary them by time of day or year, subject to meeting the overall average level.⁸

This approach seems consistent with the concerns above. It provides an incentive for the smaller or less efficient operators to become efficient and it does not place an undue discriminatory burden on more efficient competitors. As and when relevant spectrum is no longer scarce, or is freely tradable, there would be scope to remove even the remaining distinction between operators in the regulated price cap.

8. Policies in Europe

Policy in European countries varies significantly. In some countries there is regulation of the termination charges of all

mobile operators, in other countries regulation of only the largest operators, in yet other countries regulation of none. There may or may not be a distinction between the termination charges of operators with 900/1800 MHz networks and those with 1800 MHz networks. Partly in consequence, there is a great variety of termination charges both across Europe and within most countries.

The first column of Table 1 shows the percentage difference between the highest and lowest average termination charge in 12 of the largest mobile (cellular) markets in Western Europe.⁹ In three countries (Ireland, the Netherlands and Switzerland) the highest average termination charge is some 10% to 13% greater than the lowest. In other countries, like France, Germany, Italy and Spain, the highest is about 25% above the lowest. In the UK, Belgium, Greece and Portugal the highest ranges from 37% to 46% greater than the lowest. And in Austria it is about 75% higher. Is this extensive range sensible, and consistent with a uniform regulatory approach in future? If not, how should the disparities be dealt with?

It seems helpful to analyse the situation in terms of three distinct categories of operators. Category A is the set of mobile operators that have 900 MHz networks or combined 900/1800 MHz networks; these are typically (although not always) the earlier operators with larger market shares that may have access to the lowest costs. Category B is the set of operators with 1800 MHz networks; these are typically (though again not always) the later entrants with smaller market shares that might be expected to have somewhat higher costs by virtue of their less favourable access to spectrum. Category C is the set of new entrants, defined as those established since 2000; they have various different technologies but very small market shares as yet (none exceeding 6%).

Peak and off-peak termination charges vary considerably from one operator to another, presumably dependent upon their customer base and strategy. For present purposes the average of these peak and off-peak charges, weighted by usage, gives a more accurate picture of the level of charges of each operator, and of the differentials between them, than using peak or off-peak charges alone. The average charge better indicates the potential effects, on operator revenues and competition, of the differentials between operators.¹⁰

For each country, Table 1 analyses termination charges in terms of three components:

- the ratio of the highest to lowest termination charge among category A operators (see column 3 in Table 1)
- the ratio of the highest category B termination charge to the highest category A termination charge (see column 5 in Table 1)
- the ratio of the category C termination charge to the highest category B charge (see column 7 in Table 1)

Table 1 Mobile Termination Charge Ratios in European Countries¹¹

Country	Overall Ratio ^a (Max to min)	900/1800 operators (Category A)		1800 operators (Category B)		Entrants (Category C)		Overall Ratio excluding new entrants ^e (Max B to min A)
		Ratio ^b (Max A to min A)	Market share ^f	Ratio ^c (Max B to max A)	Market share ^f	Ratio ^d (Max C to max B)	Market share ^f	
	%	%	%	%	%	%	%	%
Austria	74.4	22.7	(29.3)	0		42.2	(6.2)	22.7
Belgium	46.2	23.1	(32.2)	18.75	(12.0)	na		46.2
France	23.4	0		23.4	(15.7)	na		23.4
Germany	25.2	0		25.2	(12.2)	na		25.2
Greece	40.5	0		0		40.5	(1.8)	0
Ireland	12.0	0.8		na		11.1	(4.0)	0.8
Italy	25.4	0		24.4	(17.1)	0.8	(0.6)	24.4
Netherlands	12.9	0		12.9	(14.4)	na		12.9
Portugal	32.7	32.7	(21.1)	na		na		32.7
Spain	26.4	2.7	(53.9)	23.1	(20.2)	na		26.4
Switzerland	10.3	10.0	(20.5)	0.3	(17.2)	na		10.3
UK ¹²	37.2	0		17.8	(24.7)	16.5	(1.0)	17.8

na – not applicable since no operators in that category. a – Represents the percentage difference between the operator with the highest termination charge and the lowest termination charge in each country. b – Represents the ratio of highest to lowest termination charge among category A operators. c – Represents the ratio of the highest category B termination charge to the highest category A termination charge. d – Represents the ratio of the highest category C termination charge to the highest Category B termination charge. e – Represents the ratio of the highest to lowest termination charge, excluding the charges made by entrants. f – Market share is the market share of the operator with the highest average termination charge in that category.

- finally, the last column (column 9 in Table 1) is the ratio of the highest to lowest termination charge, excluding the charges made by small new entrants.
- new entrants have been identified in five countries, with markups over 1800 MHz termination charges ranging up to 42%, but their market shares are very small;

These ratios are all expressed as percentage markups. The Table thus shows how far there is differentiation between the charges of 900/1800 MHz operators, how far 1800 MHz operators charge above 900/1800 MHz operators, how far new entrants charge higher still, and what the total difference is between highest and lowest termination charge. In parentheses after each of the entries with positive mark-up is the market share of the operator with the highest average termination charge in that category.

The picture emerging from Table 1 is more refined than the brief indication of differences cited earlier:

- within most countries, there is no difference, or only a small one, between the average termination charges of the 900/1800 MHz operators. However, for one country (Switzerland) the highest charge is about 10% above the lowest charge in this category, for two countries (Austria and Belgium) the differential is about 23%, and for one country (Portugal) the differential is nearly 33%;
- several countries have no 1800 MHz systems, and in three others (Austria, Greece and Switzerland) the highest charge by such operators is no different from that of the 900/1800 MHz operators. But in six countries (Belgium, France, Germany, Italy, Spain and UK) the differential is in the range 18% to 25%;

A clearer picture is obtained by excluding these very small new entrants. The overall picture as given in the last column of Table 1 may then be summarised as follows. In two countries there is no or negligible difference in termination charges across operators; in three countries the highest termination charges is (or will be) some 10 to 13% above the lowest (the UK inadvertently being at 18% at present but scheduled for 13%); in five countries that ratio is about 24%; and in two countries it is 37% and 46%.

9. Comments on European policies

It is not surprising that operators with smaller market shares tend to have higher termination charges, at least in an unregulated market. Economists have argued that customers will base their demand for calls on the average price of mobile calls if they are unaware which mobile network they are calling. Operators will take account of the effect of increasing their termination charges only insofar as this impacts on the weighted average price, and smaller operators will have a smaller impact on that average. It has elsewhere been shown that peak termination charges in Europe do indeed tend to be higher for smaller operators. There is also evidence that “asymmetric regulation of the larger operators will, ceteris paribus, induce the smaller operators to increase their termination rates even further.”¹³

Non-regulation of new entrants is understandable. It is unwise to deter innovation and potential sources of future competition. The market shares of new entrants are very small – not exceeding about 6% in the present classification. The high termination charges set by new entrants are unlikely to be a significant burden on competitors or on average price, and customers potentially benefit from lower connection charges or handsets that the new entrants may be able to offer.

However, as the one-time new entrants grow larger, and secure higher market shares, the situation changes. Their higher termination charges become increasingly burdensome on the operators and customers of other networks, who have to pay these charges. The asymmetric regulation of the larger networks, requiring them alone to reduce termination charges – and, according to some, even inducing the smaller networks to set higher termination charges than they otherwise would do – increasingly tends to distort the process of competition. This seems likely to be the case in the five countries that allow 1800 MHz operators to make significantly higher termination charges than 900/1800 MHz operators. In these five countries, 1800 MHz operators with differentials of 19% to 25% above 900/1800 MHz networks have far from negligible market shares ranging from 12% to 23%.

Typically, the higher termination charges of these 1800 MHz operators are not regulated, so the distortion derives from the asymmetric (more severe) regulation of the 900/1800 MHz operators. However, a regulated distortion can also apply within the 900/1800 MHz networks themselves. In Switzerland, for example, such an operator charges about 10% above other operators in that category, and has a market share of about 20%. More extreme are the 900/1800 MHz operators in Austria and Belgium with regulated termination charges about 23% above that of the lowest operator in the same category, despite having about 30% market shares. Where the termination charge is regulated on an asymmetric basis with reference to differential costs, there is a clear danger that this will reduce incentives to efficiency as well as distort the competitive process.

The overall pattern of these differentials suggests that, even where the National Regulatory Authorities (NRAs) are setting price controls on mobile termination charges, most other European countries are doing so on a different basis from the UK. It is not clear why the costs of efficient GSM 900/1800 MHz operators, and therefore their regulated prices, should be significantly different from each other. It would be surprising if charge differentials of the order of 19% to 25% can be explained by differences in access to radio spectrum as between 900/1800 MHz and 1800 MHz operators, when the modelled cost differential was put at 13% in the UK. Nor is it clear what other cost differences that lie beyond the control of the companies could explain these differences in charges.

10. Implications for future policy

It thus seems that many countries are regulating mobile termination charges on an asymmetric rather than symmetric basis. Some operators are being required to reduce their charges while others are allowed – either by explicit regulation or by the absence of regulation – to set higher charges, often very much higher. It seems unlikely that the extent of the differentials can be fully explained by cost differences between efficient operators, having taken account of differential access to radio spectrum. In consequence, such asymmetric regulation may be expected to impair the competitive process and reduce the incentives to efficiency.

If price controls are to be applied to termination charges, a policy of setting price caps symmetrically is therefore to be preferred, with exceptions only for new entrants, and for cost differences reasonably and necessarily attributable to differences in access to scarce and untradeable spectrum. Such a policy would be more conducive to competition and efficiency, and would better promote the interests of customers generally.

European regulators have recently proposed various different approaches for the future. They differ in robustness on this particular issue. Ofel has recently restated clearly the case for a symmetric efficiency-oriented approach for the UK.¹⁴ In contrast, the draft approach by the European Regulators Group (ERG) seems a little complacent. It claims that “With a cost-oriented access price, excessive pricing is made impossible and allocative efficiencies are reduced”.¹⁵ However, it fails to recognise the potential for productive inefficiencies and the threat that “cost-oriented” price controls pose to competition if they are related to actual and not efficient costs. The ERG suggests that NRAs might allow higher charges for smaller and newer players in order to promote competition. Fortunately, it recognises that “the problem is ... when the grace period should end ... as otherwise more efficient operators in the market might be put at a competitive disadvantage”.¹⁶ The report of the Independent Regulators Group is similarly lacking in precision. It advocates termination charges set at competitive levels in the long term. However, it suggests that a variety of costing methods could be used to calculate these, without clearly explaining the significance of symmetry and efficient benchmarks. It also suggests that “national or market specific circumstances can justify different remedies in order to achieve a competitive level”,¹⁷ which leaves a disconcertingly large degree of regulatory discretion.

Whether price controls are the only or best solution to this issue is another matter. The UK Competition Commission found no basis for assuming that competition to provide termination charges would be effective in the foreseeable future.¹⁸ It may therefore seem unlikely that competition can be relied upon to rescue customers from poorly set price controls on termination charges, or to reduce or remove the need for such price controls. However, the Competition Commission did acknowledge that the Calling Party Pays (CPP) principle was the cause of the

market power over termination charges, and it also found that the alternative approach of Receiving (or Mobile) Party Pays (RPP or MPP) would develop competition in this market.¹⁹ Rather surprisingly, the Commission rejected this policy after minimal consideration.²⁰ The RPP policy deserves further exploration. In the meantime, it is all the more important that, if price controls are to be used, they should be set with a view to promoting efficiency and competition.

11. Conclusions

Despite the rapid growth and superior performance of the mobile sector across Europe, price controls are ubiquitous, especially price caps on termination charges. Yet the precise basis for these controls evidently varies considerably across Europe. So too does the size of the differential between the highest and lowest termination charge within any country. There is now a proposal to promote a consistent regulatory approach across Europe.

The stated rationale for these price caps on mobile charges has perhaps been set out most clearly in the UK. This rationale is different from the rationale for price caps in the rest of the utility sector. It is not primarily to promote efficiency in a monopoly business and to ensure that savings are shared between investors and customers. Nor is it to “hold the fort” in a potentially competitive business until competition actually arrives. Instead, it is to modify the structure of charges so as to improve allocative efficiency and to redistribute income between customers themselves. Other European regulatory groups have similarly emphasised allocative efficiency. However, whatever the stated rationale, considerations of productive efficiency and competition should not be ignored when setting these price controls.

In the UK the regulated differential between the termination charges of mobile operators reflects only costs assumed to be beyond the control of these operators. In contrast, in many other countries the differential is so great that it seems likely that those operators with higher costs get higher price caps (where they are capped at all), regardless of the reason for these higher costs. Such an asymmetric approach is not conducive to the competitive process nor to improved efficiency in the mobile sector as a whole. It also invites operators and others to argue for yet further justifications for price differentials. Such policies are likely to be to the detriment of mobile customers in future.

Notes

¹ Vodafone, O2, Orange and T-mobile, Competition Commission, OfTel, London, December 2002.

² There is also considerable variation and asymmetry in the regulation of fixed as well as mobile services - for example in the reciprocity or otherwise of call termination charges between incumbent operators and alternative access network operators.

³ European Regulators Group (ERG), Consultation Document on a Draft Joint ERG/European Commission approach on appropriate remedies in the new regulatory framework, 21 November 2003. Available at www.erg.eu.int/doc/publications/erg0330_draft_joint_approach_on_remedies.pdf

⁴ Stephen C Littlechild, *The Regulation of British Telecommunications Profitability*, Department of Industry, February 1983. Reprinted in Ian Bartle (ed.), *The UK Model of Utility Regulation*, CRI, The University of Bath, 2003.

⁵ Allocative efficiency means an efficient allocation of resources between different goods and services. It is generally taken to imply price equal to average or marginal cost so as to give accurate cost-reflective signals to market participants.

⁶ See, for example, James Mirrlees “Witness Statement”, High Court of Justice Queen’s Bench Division Administrative Court in the Appeal case between Vodafone Limited against the Competition Commission and the Director-General of Telecommunications”, March 5, 2003.

⁷ The two earlier entrants Vodafone and O2 (formerly BT) had 900 MHz systems and subsequently added some 1800MHz capacity. The two later entrants Orange and T-Mobile have only 1800 MHz systems but have more usable spectrum. I am advised that 900 MHz systems generally require a smaller network than 1800 MHz systems at low levels of output, but in the UK require a larger network at high levels of output. There is a difference of view as to whether underlying costs are materially different at present levels of output, and further empirical work would be required to resolve this.

⁸ The price cap reductions are complicated by the fact that OFTEL missed the deadline for imposing price cap reductions in 2003/04, and has proposed two reductions in 2004/05 in order to meet the original price cap numbers for 2005/06. The present allowed average differential between 1800 MHz and 900/1800 MHz operators is 17.8% but is scheduled to reduce to 12.6% in 2005/06. See Table 1 and Table 3 in Appendix H, OfTel (2003), *Wholesale Mobile Voice Call Termination, Proposals for the identification and analysis of markets, determination of market power and setting of SMP conditions*, Explanatory Statement and Notification, 19 December 2003, available at http://www.ofcom.org.uk/legacy_regulators/ofTel/mobile_call_termination/

⁹ The data are for the largest 12 mobile markets, excluding Sweden (9th largest) but including Ireland (16th largest). The Swedish NRA is currently in the process of reviewing termination rates for all MNOs, and the current rates in that market are not reflective of formal rate setting activity by the NRA.

¹⁰ In most countries, including in Europe, the termination charges for fixed to mobile calls are the same as those for mobile-to-mobile calls. The Portuguese data in Table 1 refer to termination charges for fixed to mobile calls (the charges for mobile-to-mobile calls are symmetrical there).

¹¹ Sources: Mobile Communications Guide to western European cellular subscribers, 1 July 2003, for operators, technologies, number of subscribers (hence market shares) and date of launch. Vodafone staff for estimates of weighted average termination rates (hence ratios thereof). Note that Cullen data used by Dewenter and Haucap (see below) refers to mobile termination rates only at peak.

¹² The intended regulated differential in UK was lower, as explained in footnote 8 and accompanying text.

¹³ Ralf Dewenter and Justus Haucap, “Mobile Termination with Asymmetric Networks”, Social Science Research Network, FAF Economics Working Paper No. 23, October 2003, available at <http://ssrn.com/abstract=458222>

¹⁴ E.g. “... firstly, it is desirable for Orange and T-Mobile to have the same TAC (target average charge), reflecting the identical efficient costs that they incur as 1800 MHz operators. Secondly, although the fair charge differs in each year for combined 900/1800 MHz operators (Vodafone and O2) and 1800 MHz operators (Orange and T-Mobile), it is desirable to align the target charges of the two types of operators in this first period so that the TAC for each is the same amount above the fair charge. ... Thus one type of operator would not have an advantage over the other which potentially might result in a distortion in retail competition.” (para H.3, p. 276) “... permitting 1800 MHz operators to continue to set termination charges even further above cost that [than] combined 900/1800 MHz operators would risk a distortion of competition in the retail mobile market.” (para H.31, p. 284), OfTel, *Wholesale Mobile Voice Call Termination, Proposals for the identification and analysis of markets, determination of market power and setting of SMP conditions*, Explanatory Statement and Notification, 19 December 2003, available at http://www.ofcom.org.uk/legacy_regulators/ofTel/mobile_call_termination/

¹⁵ European Regulators Group (ERG), Consultation Document on a Draft Joint ERG/European Commission approach on appropriate remedies in the new regulatory framework, 21 November 2003, p. 115. Available at www.erg.eu.int/doc/publications/erg0330_draft_joint_approach_on_remedies.pdf

¹⁶ *Ibid.* p. 116.

¹⁷ Independent Regulators Group (IRG), Principles of implementation and best practice (PIBs) on the application of remedies in the mobile voice call termination market, 20 November 2003, p. 3, available at <http://irgis.anacom.pt/admin/attachs/359.pdf>

¹⁸ Surprisingly, despite OfTel’s statutory duty to promote competition, the Competition Commission did not explicitly and systematically consider the effect of its policy or of alternative policies on the development of competition, nor did it mention taking any steps to promote competition.

¹⁹ Competition Commission report paras 2.488 – 2.492.

²⁰ The report adduced no substantive evidence of the claimed inadequacies of RPP and overlooked (arguably even misrepresented) the successful experience of RPP in the USA (there called MPP). E.g. “Mobile party pays (MPP) creates better incentives than calling party pays (CPP) for mobile network operators to place downward pressure on termination rates. Cellular telephone use in the United States and Canada has continued to increase at a significant pace despite the MPP regime and now far exceeds mobile telephone use in countries with CPP regimes.” Sidak, J. Gregory and Crandall, Robert, “Should Regulators Set Rates to Terminate Calls on Mobile Networks?” *Yale Journal on Regulation*, Vol. 21, 2004. available at <http://ssrn.com/abstract=462041>



To keep the environmental impact of this document to a minimum, we have given careful consideration to the production process. The paper used was manufactured in the UK at mills with ISO14001 accreditation. It is 75% recycled from de-inked post consumer waste. The document was printed in accordance with the ISO14001 environmental management system, using inks which are vegetable based. All the steps we have taken demonstrate our commitment to making sustainable choices.

Designed and produced by Barrett Howe Plc

