The Policy Paper Series • **Number 7** • **November 2007**

The Economics of Mobile Prices

Foreword

I hope you enjoy our seventh 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 Group

Contents

This paper can be seen online at www.vodafone.com/publicpolicyseries

Published by Vodafone Group Plc

ISBN 978-0-9552578-3-4

Dr. Jonathan Sandbach

Head of Regulatory Economics,
Vodafone Group Public Policy

Jonathan Sandbach has been a professional economist working in the telecommunications industry for over 20 years. This has included positions within British Telecom, Cable and Wireless and currently Vodafone, where he is Head of Regulatory Economics. He has also been a senior consultant at National Economic Research Associates, where he advised a number of telecommunications operators and national regulatory agencies worldwide. Jonathan Sandbach holds degrees in Mathematics, Statistics and Economics from the University of York, the London School of Economics, and a PhD from City University London.

Mobile price structures: two sides of the market

Telecommunications networks are examples of two-sided markets: providing communication services to their own subscribers and, over the same platform, providing connectivity to their own subscriber base for users of other networks. The two markets are linked: more subscribers on a particular network means more opportunities for users of other networks to make calls and, looking ahead, more opportunities for providers of mobile content to supply their services. This linkage provides the essence of a two-sided market.

Whenever we look at two-sided markets, one striking observation is immediately apparent: the structure of prices (i.e. who pays for what) is fundamentally important for the development of the market. If credit card companies charge annual fees, this will affect both the number of customers they acquire, and appeal of the network to merchants. Similarly, if media channels increase subscriber or viewer charges, they risk losing advertising revenue. In both these simple cases, the two-sided network provider will aim to balance prices in order to encourage take-up on both sides.

The first two papers in this pamphlet are written by economists who have researched two-sided markets in telecommunications and other industries. The views expressed are those of the authors and not Vodafone, but they demonstrate the importance of the two-sided analysis of telecommunications markets that Vodafone has always advocated.

The first paper, by Dr David Evans, places mobile telecommunications networks in this context. The conclusion is simple, but important for any price regulation of the industry: trying to equate price structures to cost structures will move the industry away from what would otherwise be its competitive equilibrium development path. Regulation in the context of two-sided markets is much more difficult and can easily do more harm than good.

One particular aspect of mobile telecommunications is the "waterbed", a term first coined in meetings between Vodafone and Professor Paul Geroski of the Competition Commission in the UK in 2001. Historically, interconnection charges for voice termination have been set by mobile operators above cost, with the surplus used to subsidise acquisition and retention of subscribers (a typical characteristic of two-sided markets). The waterbed refers to the impact that lowering termination rates would have on subscriber prices. Lower termination rates means network operators have less incentive to compete for subscribers who bring with them termination revenues (i.e. push down on one side of the bed, and the other rises). In theory, the waterbed should exist if either of two conditions is satisfied: either there must be competition for subscribers between mobile networks or, even if there is no competition, the mobile subscriber market must be less than fully saturated (defined in the sense that the elasticity of subscription equals zero, rather than 100% penetration). Both these situations will mean that network operators will increase mobile prices in response to a fall in termination rates. The only situation in which the waterbed should not exist is where there is a single monopoly supplier of mobile subscriptions and the market is fully saturated.

The waterbed has long been hypothesized as a feature of the mobile network industry. Observing the waterbed empirically,

however, is statistically challenging. Other trends (i.e. a general reduction in mobile prices due to increased competition, and economies of scale as traffic volumes have grown) mask the impact of the waterbed on subscriber prices. Furthermore, since mobile termination rates are influenced by factors other than regulation (e.g. network costs), and these same factors also influence subscriber prices, any simple correlation or ordinary least squares regression of mobile termination rates on subscribers prices will be statistically biased.[1]

The second paper in this collection, by Professor Tommaso Valletti and Dr Christos Genakos, applies a rigorous econometric analysis to test for the first time the existence of a waterbed. In doing this, the analysis addresses the statistical problems. The authors use a model that takes account of time and country effects, and so allows the true waterbed effect to be observed. The model also uses the standard econometric technique of instrumental variables to eliminate the statistical bias that is inherent in any simple correlation or regression model between incoming and outgoing prices.

Genakos and Valletti conclude that the waterbed does "exist" and is "high but not full". In fact, the magnitude of the waterbed is itself a function of the degree of competition for mobile services (measured by the number of competitors) and the degree of market saturation. Their modelling results also appear to confirm the theoretical applicability of the waterbed effect: that it is stronger whenever (1) there is a higher degree of competition for mobile subscribers (measured by the number of competing networks); or (2) the market is less than fully saturated. If either of these two conditions is satisfied (but not necessarily both), the waterbed is high.

There are two further interesting results from the Genakos and Valletti analysis. Firstly, some light is shed on the dynamic path of the waterbed. Whilst there is some evidence that mobile network operators are able to anticipate reductions in termination rates, most of the waterbed effect feeds through progressively over a period of 18 months after the rate reduction has been implemented. This emphasizes that we should not expect immediate adjustments to mobile prices following a reduction in termination rates. Secondly, the magnitude of the waterbed can be larger than we would have expected from simply the relative proportions of terminating and subscriber revenues. This is explained by price elasticity effects. To illustrate this consider the impact of a 10% reduction in the mobile termination rate. Due to the small price elasticity the reduction in terminating revenues may be 8% (an elasticity of -0.2). Terminating revenues in the mobile industry are generally around a quarter of subscriber revenues, and so we might expect a full waterbed to be seen in an increase in subscriber revenues of 2%. However, subscriber numbers have their own price elasticity, and so to achieve an increase in revenues of 2% will require an increase in subscriber prices of 4% (assuming an elasticity of -0.5). Genakos and Valletti do not estimate a subscriber elasticity, but they do observe a magnitude for the waterbed that exceeds what we would expect given the relative size of terminating and subscriber revenues. This suggests that the relative price elasticities are important.

The third paper, written by me, looks at the issue of pass-through of changes in mobile termination rates (typically reductions) to fixed network tariffs for end customers. This is important because consumer benefits from lower mobile termination rates can only be realized if there is a full pass-through of the termination rate to fixed-to-mobile and mobile-to-mobile (off-net) retail prices or, at least, to the overall bundle of prices supplied to callers to the mobile network. In the absence of strong competition between fixed network operators (which appears not to be an effective constraint in many countries), this is one area that may require regulatory action.

To the extent that price regulation is contemplated it must take account of the two-sided market, and price structures must reflect more than costs, particularly the linkages between demands on both sides of the market. The final paper in this collection takes the work of previous papers and uses it to present a model illustrating the impact of the two-sided market on "optimal" mobile termination rates. This model shows that the two-sided model of mobile networks, and the existence of a waterbed, should have significant impact on the levels that regulators choose to set regulated mobile termination rates. The only situation in which this is not the case is when the price elasticities on the subscriber side of the market are zero (for both increases and decreases in prices). This could be the case in a fully saturated market, but even this is unlikely given that many existing subscribers may only remain on the network at a price that reflects an existing subsidy from termination revenues. Price increases would lead to a reduction in the number of subscribers. Ofcom estimates, based on its market research, that 34% of the existing mobile subscriber base in the UK can be considered "marginal", in the sense that they would not renew their mobile handset and service if it were lost and they received no subsidy.[2] More recent market research by Vodafone in other developed markets suggests that this result is not unique to the UK.[3] The estimate of marginal subscribers from the UK 2006 research is also exactly the same percentage estimated four years previously for the UK market by Competition Commission. Despite the proportion of the adult population owning a mobile telephone growing from 68% to 81% between 2002 and 2006, it seems that the proportion of these subscribers that can be considered "marginal" remains high and constant. It appears that marginal customers are a feature of mobile markets even when they are mature.

Using existing estimates of elasticities for call termination and mobile subscriptions, and assuming complete pass-through of any termination rate reduction to fixed network end user tariffs, the paper illustrates the extent to which acknowledgement of a waterbed effect will impact regulated termination rates. This is in sharp contrast to current regulatory trends where termination rates are being regulated down to levels of long run incremental costs plus mark-ups for common costs. Even where an allowance has been made for an "externality mark-up" (as in the UK, giving a regulated price of 5.1p[4]), the efficient (welfare maximizing) price is at least twice this level, even if the waterbed is only 50% effective, and around three or four times this level if the waterbed is fully effective. Our modelling further shows that the optimal mark-up over cost varies

significantly between countries, depending on the quantity of fixed to mobile calls (relative to other call types) and the level of mobile prices, suggesting that simple cross-country benchmarks should be applied with caution.

If we accept that pass-through of mobile termination rates reductions to fixed-to-mobile call prices (or more generally prices for any other fixed network operator service) is incomplete, as the evidence appears to suggest in some countries, the case for regulating termination rates above cost becomes even stronger. Any economic rents arising from fixed-to-mobile calls can be channeled to either fixed or mobile operators according to how the termination rate is set. A high termination rate will channel rents to mobile operators, who will then pass this on to lower consumer prices through the waterbed effect, reflecting the intensity of competition in mobile markets. A lower termination rate, on the other hand, will channel rents to fixed network operators. If competition is not fully effective in the fixed network sector, the rents will not be fully passed through to consumer prices. The asymmetric competitive situation between fixed and mobile sectors should lead regulators to exercise caution before assuming that lower mobile termination rates are good for consumers.

Notes

1 In technical econometric terms, the explanatory variable (mobile termination rates) will be correlated with the model residual, resulting in biased OLS estimates.

2 See "Mobile call termination", Ofcom, 27 March 2007, See Table A16.3. Ofcom assume a handset cost averaging £70 per subscriber. The precise definition of a marginal subscriber, therefore, is one that would not be prepared to pay this amount to subscribe or re-new a subscription to the mobile network.

3 For example, see "Quantifying the Number of Marginal Subscribers in the Australian Mobile Market", a report submitted to the ACCC, available at www.accc.gov.au/content/item.phtml?itemId=794815&nodeId=762ce931df658dd7e0de9dda2f2030b9&fn=Vodafone%20Submission%20-%20Annex%20C%20Marginal%20Subscribers.pdf

4 See "Mobile call termination", Ofcom, 27 March 2007, See page 2.

Dr. David S. Evans

Jevons Institute on Competition Law and Economics, University College London

David S. Evans leads LECG's global antitrust practice. An authority on industrial organization, he is the co-author or editor of six books, including Invisible Engines: How Software Platforms Drive Innovation and Transform Industries (MIT Press 2006), which won the best business book of 2006 award from Association of American Publishers, and Catalyst Code: The Strategies of the World's Most Dynamic Companies (Harvard Business School Press, 2007). He has also written more than 70 articles in professional journals in economics and law. His most recent work has concerned the economics of multi-sided platform businesses, product design, and tying and bundling.

Dr. Evans holds several academic positions. He is the Executive Director of the Jevons Institute on Competition Law and Economics at the University College London where he is also a Visiting Professor. He is also a Lecturer at University of Chicago Law School. He is the Chairman of the editorial board of Competition Policy International. Dr. Evans has BA, MA and Ph.D. degrees from the University of Chicago.

Do mobile operators have a dominant position in a market for the wholesale termination of calls from fixed to mobile?

1. Context

This paper was written in 2004, originally for a case in Sweden, and has since been submitted in various other regulatory proceedings.

2. Introduction

The European Commission[1] and the PTS in Sweden, amongst other regulators, have claimed that mobile operators are dominant in the provision of wholesale termination for voice calls, including calls from fixed networks. I consider the implications of this claim against the backdrop of the nature of competition between mobile networks.

Mobile calls are services provided by platforms that operate in what economists refer to as a 'two-sided' market. In two-sided markets competition will act to keep the overall level of prices in line with costs, as it does in other types of market; the individual prices of services supplied to the two sides, however, can differ from the individual costs, no matter how strong competition among platforms. What is more, the prices that maximize consumer welfare will often differ from cost-reflective ones. These characteristics have particular relevance to the way in which regulators and competition authorities should consider two-sided markets.[2]

When considering call termination on mobile networks, the European Commission, the PTS, and other regulators have adopted an approach that leads to claims that a company is dominant when it can set prices that differ from costs on only one side of a two-sided market. Yet this relies upon a competitive benchmark that is unrelated to the competition actually faced by mobile operators. And it relies upon a benchmark that is unrelated to socially optimal prices. If this approach were generally adopted by regulators and competition authorities in other markets, it would lead to an excessive number of restrictions and interventions in cases where there is quite clearly no problem to fix. The fact that we do not generally see such interventions only serves to confirm that the approach adopted by the European Commission and the PTS for mobile call termination represents a significant departure in the application of

competition law and economics and confirms my view that they have erred in intervening in this case.

The rest of this paper is structured as follows. Section 3 summarises some of the important features of competition between mobile platforms. Section 4 introduces two-sided markets and show that mobile networks fit within the standard framework of two-sided markets. Section 5 discusses the optimality of prices in two-sided markets, and shows that there is no reason in general to think that cost-based prices are better for consumers than the prices set by competing platforms. Section 6 draws out the implications for assessing dominance in two-sided markets, highlighting the problems with approaches that seek to identify dominance on an individual side of a two-sided market. Section 7 concludes.

In preparing this paper, I have read the Commission's Draft Recommendation of June 2002,[3] and its final Recommendation of 11 February 2003,[4] Ofcom's Statement on Wholesale Mobile Voice Call Termination dated 1 June 2004 ("The Statement"), as well as prior consultations leading to the statement.[5] I have also read the decision handed down by the telecom regulator in Sweden concerning similar issues.[6] References to regulators throughout the paper should be understood as applying particularly to these three cases, although I understand that other European regulators have adopted a similar approach.

3. Competition between mobile phone networks

A mobile network provides services to two distinct customer groups. We call the mobile network in question the "home network" and call other networks "other networks."

i Subscribers receive a mobile telephone number and handset when they sign up for service. Subscribers can originate and receive calls from other subscribers on the home network and from individuals on other mobile and fixed networks that have interconnection agreements with their home network.

ii Subscribers on other networks can make calls to subscribers on the home network as long as those other networks have interconnection agreements with the home network.

Calls to mobile from fixed networks are the result of "joint demand" by a fixed subscriber and a mobile subscriber. That is, the product can only exist if an individual puts herself in a position to receive calls and if another individual places a call. This feature means that mobile networks compete in a two-sided market.

The mobile network is remunerated for providing these services in a number of ways. Subscribers typically pay a fee to join the home network, a monthly rental charge, and per minute fees for originating calls to other members of the home network or to members of other networks. In some cases subscribers get a handset when they join, and the cost of this handset is higher than the signing fee, leading to claims that subscribers receive subsidised handsets.

Mobile networks are also paid when their subscribers receive calls. When a subscriber of another network makes a call to a subscriber of the home network, that network pays a fee to the home network for the delivery of the call to its subscriber. These are called termination charges. The other networks pass these fees on in whole or in part to their subscribers, since throughout Europe, the party who initiates the call bears the costs of the call. This is known as the calling party pays (CPP) principle. Accordingly, I will focus in this paper on the response of fixed-line subscribers to the final price of a call to a mobile network, rather than on the negotiation between the fixed network and the mobile network.

In many countries competition between mobile operators is robust. In Sweden, five main mobile network operators – TeliaSonera AB, Tele2 Sverige AB, Vodafone Sverige AB, Hi3G Access AB and Telenor Mobile Sverige – all compete for subscribers,[7] with significant numbers of customers changing their home network every year.[8]

4. Two-sided markets

As noted above, competition between mobile phone networks occurs within what economists have termed a two-sided market.[9] A market is two-sided when there are two distinct groups of customers, those customers need each other in some way, and a "platform" can bring the two sides together and harvest the externalities between them in ways they cannot do for themselves.[10] It should be clear from the previous section that mobile networks have these characteristics.

Many other industries are based on platforms that compete in two-sided markets. These include:[11]

i Exchanges. These are platforms that facilitate trading between buyers and sellers. They include financial exchanges such as the London Stock Exchange (LSE), shopping malls such as Covent Garden, and auction sites such as eBay.

ii Software platforms. These are platforms that *inter alia* enable software developers to write applications that people can use and that enable people to run those applications on their computers. They include video game consoles such as Sony Playstation, computer operating systems such as Windows, and PDA and mobile phone operating systems such as the PalmOS and Symbian.

iii Advertising-supported media. These are platforms that enable advertisers to meet viewers.[12] Newspapers such as the *Financial Times*, magazines, free-television stations such as TV-5, and Internet portals such as Google are platforms where advertisers can meet viewers.

iv Payment-media. These are platforms that provide a medium of exchange for buyers and sellers. Payment card systems such as Visa, MasterCard, EuroCheck Eufiserv and Maestro help participating cardholders and card-accepting merchants to transact with each other.

v Communications networks. These are platforms that enable people to communicate with one another. They include telephone networks, whether mobile like Vodafone or fixed such as France Telecom, and instant messaging such as AOL Instant Messenger. For any particular communication there is a "caller" (or "sender") and a "receiver" who have different preferences and could be charged different sums.

Mobile networks

A mobile phone network is a two-sided platform in the business of helping people who want to receive calls "get together" with, amongst others, people who want to make calls to them from their fixed line.[13] That they fit within the standard two-sided market framework is clear from the comparison with two classic two-sided markets – shopping malls and payment cards.

Shopping malls bring together consumers and merchants. Consumers benefit from shopping at retailers in the mall, as well as from related amenities such as parking and restrooms, while retailers benefit from access to customers. Retailers pay rent to the shopping mall, while customers pay nothing to enter and often receive parking and other services for free.[14]

The shopping mall customers are analogous to mobile subscribers, but instead of subsidised handsets and low outgoing call charges, they receive free amenities such as parking and restrooms to encourage them to join. The retailers are similar to fixed line callers. If a retailer wants to reach those shopping mall customers, it must pay the access fee set by the mall (i.e. the rent).[15] Retailers are, of course, free to reject the rents set by shopping malls and operate instead (or in addition) on the high street or the Internet. By doing so, however, they lose access to those consumers who have chosen to shop in a given mall at a given time.[16]

Payment cards also operate in two-sided markets. One side of the market consists of consumers who may choose to join one or more payment card networks. On the other side, merchants decide whether to accept cards from each network for payment. In many countries, merchant fees account for the substantial majority of total revenues for a card system.[17] As with shopping malls, payment card systems facilitate transactions between retailers and consumers.

Retailers can choose to decline any given card brand. By doing so, they cannot offer customers the option of using that brand for payment. Some customers may be willing and able to pay with cash, checks, or other card brands, but others will not be. Again, as with mobile termination, retailers that want to allow payment via a card brand must pay the fees set by the system.[18]

5. General economic principles of two-sided markets

Two-sided markets are more complicated to analyse than the more familiar one-sided markets. Nevertheless, economists have subjected them to intense scrutiny over the past few years, and some clear results have emerged concerning the pattern and optimality of prices.

The pattern of prices

In considering platform pricing policies the economics literature distinguishes between:

i The overall pricing levels charged by platforms. This is the weighted sum of prices to the two customer groups.

ii The pricing structure. That is, the relative prices charged to the two sides.

I also note that platform businesses may charge two different kinds of prices: an access charge to the platform and a usage fee based on interacting with the other side.

The characteristic that distinguishes two-sided markets from traditional one-sided markets is that the pricing structure matters.[19] Consider lowering the price on one side and raising it on the other by the same amount so that the price structure has changed while the price level has not. The market is said to be two-sided when this shift affects the level of output on a platform. For instance, if a shopping mall decided to lower rents and introduce an equivalent admission charge for shoppers to enter the mall, it is clear that the overall attractiveness of the mall to both shoppers and retailers will be affected. In the same way, a mobile network is two-sided because the allocation of prices between the mobile subscriber and the people who call them affects both the number of subscribers who join and the volume of calls on the network.[20]

Economics has shown that two fundamental principles apply to the prices charged to the two sides.[21] These principles hold irrespective of whether the prices are set by a monopoly platform, through competition among platforms, or by a welfare-maximizing social planner:

i Interdependent prices. The prices to the two sides are determined interdependently. For example, if a newspaper raises the subscription price it will get fewer readers and will have to lower its advertising rates. Any change to the price on one side will change the price on the other side. We will see below that several telecom regulators accept this interdependence in the case of mobile telephony.

ii No cost causation. There is no direct relationship between the price on one side and the incremental cost of serving that side. For many two-sided platform services, customers consume the service "jointly," so that it is not meaningful to even talk about one side causing the cost of their consumption. For example, an eBay transaction only takes place when a buyer makes a purchase from a seller.

These principles help explain what might appear to be pricing anomalies in many real-world markets served by two-sided platforms. There are a number of platform industries in which one side pays nothing for the service or even receives inducements to take the service: application developers for computer operating systems, shoppers at shopping malls, American Express charge-card customers, and people who use Adobe Acrobat to read PDF files. There are other examples, where it is clear that one side does not make a direct contribution to profit or to the coverage of common costs: buyers of video game consoles, such as Sony Playstation, pay the manufacturing cost of the console (or sometimes less) and revenues come primarily from game developers who pay licensing fees; most of the revenues for newspapers and magazines commonly come from advertisers. Table 1 lists platform businesses and their pricing policies.

The social optimality of prices

The pricing levels in two-sided markets do not raise any novel issues.[22] The pricing structures, on the other hand do. Prices that appear to favour one side are common in practice in platform-based businesses, as we have seen. Sometimes it seems as if one side is subsidizing the other side. It is, therefore, reasonable to ask whether pricing in two-sided markets is prone to significant market failure or, to put the question another way, whether such pricing structures can themselves be taken as evidence of market failure.

Table 1: Two-sided platforms in Europe

Two-sided Platform	European examples	Side one	Side two	Side that gets charged little
Residential property brokerage	www.home-for-sale-uk.com	Buyer	Seller	Side one
Apartment brokerage	www.vitrine.be (Belgium) www.agent-immobilier-france.com (France)	Renter	Owner/ Landlord	Typically side one
Newspapers and magazines	Le Soir (Belgium), Le Monde (France)	Reader	Advertiser	Side one
Network television	RTBF (Belgium), Arte (France, Germany)	Viewer	Advertiser	Side one
Portals and web pages	Flix.de (Germany), www.oasi.net (Italy)	Web "surfer"	Advertiser	Side one
Operating system	Windows	Application user	Application developer	Side two
Video game console	Xbox (Microsoft), Playstation (Sony)	Game player	Game developer	Neither– both sides are a significant source of platform revenue.
Credit card	Visa, EurocardMasterCard	Cardholder	Merchant	Side one

Source: Draws from David. S. Evans, "The Antitrust Economics of Multi-Sided Markets", *Yale Journal on Regulation*, Vol. 20, No. 2, (Summer 2003), p. 337.

Economic theorists have examined this question and have generally found that regardless of the degree of market power there is no basis for believing that pricing in two-sided markets is subject to systematic bias. There are several related findings:

i The socially efficient prices are subject to the same seeming anomalies as the privately efficient prices discussed above. They, too, may be less than zero or below incremental cost for certain customer groups. Therefore, there is no basis for inferring from a particular pricing pattern that there is a market failure.

ii The theories do not find that a social planner would systematically have a more skewed or less skewed pricing structure than a profit-maximizing firm. That is, socially efficient prices are not systematically more or less tilted to one side than privately profit-maximizing prices.[23]

Three examples highlight the prevalence of two-sided markets. The important findings are that socially efficient prices will not, in general, reflect costs, that the 'competitive benchmark' for 'two-sided markets cannot reflect costs; and that prices that depart from some measurement of cost cannot be adduced as evidence of market failure in two-sided markets.

Adobe Acrobat

Adobe Acrobat is, by far, the leading program for creating and reading documents, where the format is preserved.[24] For example, Adobe Acrobat software can be used to create a version of a Microsoft Word document that looks the way it looks in Word. This version can be viewed by anyone with Adobe Acrobat Reader, without relying on any Microsoft Word software. This is the de facto standard for disseminating formatted documents today. Adobe has chosen to give the product away to the reader side of the market and to collect revenues from the creator side.

It could be argued that the readers are paying too little, while creators are "forced" to pay if they want access to consumers with Adobe Acrobat Reader.[25] Yet consumers would probably be worse off if prices on the two sides reflected costs, so that readers were charged for the software. This would have the effect of lowering the willingness of readers to join the Adobe Acrobat platform, thereby lowering the value of Adobe Acrobat to document creators. Creators would then pay less for such a product, but the value of the platform would have been significantly diminished for both sides.

The Economist

Advertising-supported media provides a more common example. A magazine, such as *The Economist*, is a two-sided platform that connects subscribers and advertisers.[26] Consumers subscribe to magazines for the content and may or may not find the advertising desirable.[27] Companies pay the magazine for access to its subscriber base. Advertisers have other means of reaching consumers, but if they want to advertise to these consumers when they read this magazine, their only choice is to pay the magazine's rates. Many advertising supported publications, such as magazines and newspapers, rely on the advertiser side for the bulk of their revenues.[28] Raising prices to consumers would lower the number of subscribers, with advertisers less willing to pay as much for a smaller audience.[29]

Yellow Pages

My final example is a particular type of advertising-supported media: yellow pages directories. Companies pay suppliers to take out advertisement in the directories, which are delivered for free to households. Many households undoubtedly benefit from having a yellow pages directory they can refer to, and presumably, the companies that pay to advertise in them also believe they are getting a valuable service. However, it is also likely that many households would be unwilling to pay a price for the directory that reflected the costs involved in printing and delivering their copy. If prices on the two sides of the market reflected costs then one outcome could well be that too few households would be willing to pay at all, making the directory unprofitable and leading suppliers to exit the market, to the detriment of householders and advertisers alike.

6. Identifying dominance in two-sided markets

With this background in mind, I address a practical issue that concerns the application of competition law in two-sided markets, where entry on a single side is not possible (i.e., companies have to serve both groups of customers if they are to provide any services at all).[30] Under what circumstances should a company be found to be dominant?

This question is important because competition law imposes restrictions and obligations on dominant companies. The answer turns on the nature of the competitive benchmark that competition law protects, since regulators often assume that dominant companies are the ones that have the freedom of action to depart from that competitive benchmark.

Lack of independence

Dominance has been defined by the Court in *United Brands* as a position of independence:

> *"65. The dominant position thus referred to by Article [82] relates to a position of economic strength enjoyed by an undertaking which enables it to prevent effective competition being maintained on the relevant market by affording it the power to behave to an appreciable extent independently of its competitors, customers and ultimately of its consumers."* (ECJ in *United Brands v Commission*, Case 27/76 [1978] ECR 207.)

According to this definition a firm is dominant under competition policy if it has independence of action – that is, if it is largely unconstrained by the decisions of its customers to switch to substitutes and by actual or prospective rivals who may undercut its prices and thereby vie for its customers.

In two-sided markets, one must be careful in applying the notion of dominance to a particular side, since, as noted above, prices on one side are inextricably linked to prices on the other side. An increase in prices on one side may be profitable when looked at only on that side. However, that increase in price will reduce the number of customers on that side and, therefore, reduce the value of the platform to and demand elasticity of customers on the other side. As a result, the price increase to one side may not be profitable when changes in demand on both sides are taken into account. These effects will be accentuated when there is competition from other platforms. In fact, *so long as there is competition on one side of the platform*, it is not possible to have independence of action even if a firm does not face competition on the other side (as has been claimed, for example, by the regulators with respect of mobile call termination).

This lack of independence can mean that individual mobile operators cannot set termination charges on one side of the market substantially below those of their rivals without affecting their ability to compete profitably for subscribers on the other side. Competition between mobile operators acting on both sides of the market will act to keep the overall return from recruiting a subscriber low. An individual operator that departed from market norms by setting low termination rates would earn even less from recruiting each subscriber, and so would be at a significant competitive disadvantage. If an operator tried to set cost-reflective prices while its competitors did not, it would have difficulty covering costs, if competition for subscribers were sufficiently robust.

This lack of independence also means that mobile operators would jeopardise efforts to recruit subscribers if they set termination rates that were too high. Subscribers buy mobile phones to receive calls, as well as to make them. If an operator sets a high price, which results in fewer calls, it would be less attractive to subscribers, other things equal. This constraint, which is generally recognised by regulators, will supplement the more standard, one-sided market, constraint: the prospect that customers would switch to competing suppliers in response to changes in prices.[31]

Regulators have often accepted the existence of some or all of these effects but they have claimed that these constraints are not sufficient to prevent dominance.[32] I show below that these claims are flawed, as they are based on a fundamental misconception about the nature of competition in two-sided markets.

Competitive benchmarks

Regulators claim that mobile operators are dominant in the provision of call termination services, because the way they define relevant markets effectively identifies a dominant company as one that is able to set prices that depart from the individual costs of terminating calls on their networks. The argument is in two steps. First regulators define markets using the SSNIP test, which asks whether a hypothetical monopolist could set prices above a competitive benchmark. Ofcom describes its definition of a separate market in the December consultation:

> *"A product is considered to constitute a separate market if a hypothetical monopoly supplier could impose a small but significant, non-transitory price increase ("SSNIP") above the competitive level without losing sales to such a degree as to make this unprofitable."* [33]

The Commission makes similar statements in its guidelines on market analysis, as does PTS.[34] Since dominant companies are ones with strong positions on relevant markets, this definition effectively ties the notion of dominance to the ability to set prices above the competitive level. This definition makes it very important that regulators have in mind the right competitive benchmark against which to assess prices.

Second, regulators use a competitive benchmark that has very little to do with the actual competition faced by mobile operators; that is, inter-network competition. Neither the Commission nor PTS specifies the competitive benchmark they use. However, their conclusions imply that they believe competitive prices would be very different from those set in the absence of regulation.[35] Ofcom, on the other hand, is very clear. In its December consultation, it claims that competitive prices reflect the costs of each service, when considered individually and as measured by long-run incremental cost (LRIC):[36]

> *3.19 The Director's view is that the most appropriate basis for assessing whether charges are cost reflective is forward looking long run incremental costs (LRIC) plus a mark-up for common costs. LRIC-based charges most accurately reflect the resources consumed by the provision of services and correspond most closely to the level that would occur in a fully competitive market. Hence, the Director has carried out a detailed modelling*

of the LRIC of UK 2G mobile networks and has estimated the LRIC of voice call termination for a 2G operator, also taking into account cost data from the MNOs. The Director has then added a mark-up for common costs.

But if this is Ofcom's position, then it has a very unrealistic view about how competition works in two-sided markets. Prices might be related to the individual costs of call termination in a hypothetical world in which technology were different and subscribers could use their handsets to make outgoing calls on one network and receive incoming calls on another one. But in the real world, competition among mobile networks, no matter how fierce, will never lead to prices that match the costs of the individual services (nor would such prices be socially desirable), even though competition may mean that the overall price level set by a platform tracks overall costs.

The implications of a cost-based competitive benchmark

Not only do regulators use a very artificial benchmark, but the benchmark is also undesirable. When there is competition on at least one side of a two-sided platform market, there are good public policy reasons why the competitive benchmark used in competition law assessments of dominance and market power should be the pricing levels and structures that are established by competing platforms.

There are a number of practical problems should competition law instead identify a company as dominant whenever it is able to set prices that depart from costs.

First, courts and authorities could not cope with the competition cases arising from such a rule. Two-sided platform businesses are numerous, and prices almost never reflect the costs of the individual sides of the market. Examples include all advertising-supported media; matchmaking services, including exchanges, estate agencies, and nightclubs; transaction services ranging from payment cards, mobile phone payment devices, and personal check clearance; software platforms, including those used on mobile phones, personal digital assistants, and computers; video games; and numerous internet-based businesses.

Second, the rule would impose costly restrictions on the prices that companies set, without providing any clear benefits for consumers. Competition law restricts the ability of dominant companies to set high prices (to avoid claims of excessive pricing), and low prices (to avoid claims of predation). Perversely, these further restrictions may lead prices to depart from those that would result from unrestricted competition. As in traditional markets, one cannot be certain that the prices set by competing platforms would necessarily be the ones that an omniscient social planner, cognizant of all possible market failures and holding all relevant data, might set. However, there is no basis in economics for believing that prices that are more cost-reflective would generally be better, and they would not, by definition, be "competitive."

Third, the rule would impose restrictions that are intended to prevent dominant companies using their position to exclude rivals, even though it is very hard to see how the mobile operators can use their position in call termination to prevent competition either from the fixed network, pagers, faxes or email as alternative means of reaching subscribers. Similarly, it is hard to believe that a mobile network can undermine competition from rival mobile networks through its choice of call termination price for calls from fixed to mobile.

Fourth, the rule may impose obligations that companies cannot comply with. It may be very difficult, if not impossible, for a platform facing competition to comply with restrictions that prevent it from setting prices much higher than costs unless its rivals all followed the same rule. For example, suppose a firm lowered the price on one side of the market and increased it on the other side (perhaps in an effort to make prices align with costs on both sides). If its rivals do not follow, the firm will lose sales and profits, unless it has made its platform more desirable. For example, suppose *The Times* lowered its price to advertisers but raised its price to subscribers to align costs and prices better. It would lose subscribers and become less valuable to advertisers. If newspapers and other advertising-supported media with which it competes did not adopt the same change in pricing structure, *The Times* could lose advertising revenues despite the lower advertising prices, which in a one-sided market would lead to higher advertising volume, as well as subscription revenues, since consumer demand will decrease given the higher subscription prices.

7. Conclusion

In this paper, I have shown that mobile operators compete in a two-sided market, and that in such markets neither competitive prices, nor socially optimal ones, reflect only the costs on the individual sides of the market. Against this backdrop I have considered some of the implications of the finding by various regulators that mobile operators are dominant providers of call termination from fixed networks. This finding is based on an approach to identifying dominance and a competitive benchmark that would find a company dominant whenever prices departed from costs. In doing so, I argue that telecoms regulators are misapplying a single-sided market benchmark to a two-sided market and ignoring the interdependency of pricing, which is fundamental both to a proper understanding of two-sided markets and to the proper consideration of dominance in European competition law and economics. This error would have very profound and, in my view, adverse implications for the level of regulation in Europe, if it were applied generally to two-sided markets. The very fact that we do not see such interventions in other markets suggests that the regulators have erred in the application of economics and of competition law principles to mobile telephony.

Notes

1 Commission Recommendation of 11 February 2003 on relevant product and service markets within the electronic communications sector susceptible to ex ante regulation in accordance with Directive 2002/21/EC of the European Parliament and of the Council on a common regulatory framework for electronic communication networks and services, OJ L 114/45 8.5.2003.

2 See Jean-Charles Rochet and Jean Tirole, "Two-Sided Markets: An Overview," mimeo, Institut D'Economie Industrielle, (2004); David S. Evans, "The Antitrust Economics of Two-Sided Markets," Yale Journal on Regulation, Vol. 20, No. 2, (Summer 2003); and Julian Wright, "Access Pricing Under Competition: An Application to Cellular Networks," Journal of Industrial Economics, Vol., No. 3, (September 2002)

3 Draft Recommendation on Relevant Product and Service Markets within the electronic communications sector susceptible to ex ante regulation in accordance

with Directive 2002/21/EC of the European Parliament and of the Council on a common regulatory framework for electronic communication networks and services, 17 June 2002 ("Draft Recommendation").

4 Supra note 2.

5 Competition Commission, "Cellnet and Vodafone: Reports on references under section 13 of the Telecommunications Act 1984 on the charges made by Cellnet and Vodafone for terminating calls from fixed-line networks," Presented to the Director General of Telecommunications," December 1998 ("CC Report 1998"); Competition Commission, "Vodafone, O2, Orange and T-Mobile: Reports on references under section 13 of the Telecommunications Act 1984 on the charges made by Vodafone, O2, Orange and T-Mobile for terminating calls from fixed and mobile networks. Presented to the Director General of Telecommunications," December 2002 ("CC Report 2002"); Oftel, "Review of Mobile wholesale call termination markets," 15 May 2003 (the "May Consultation"); and Oftel, "Wholesale Mobile Voice Call Termination, Proposal for the Identification and analysis of markets, determination of market power and setting of SMP conditions. Explanatory Statement and Notification," 19 December 2003 (the "December Consultation").

6 PTS (The Swedish National Post and Telecom Agency), Decision on matters 04-7287/23, 04-7228/23, and 04-6952/23, 6 July 2004 ("PTS Decision").

7 PTS Decision, p. 14.

8 PTS, "Swedish telecommunications market first half-year 2003," 18 December 2003, p. 36.

9 Jean-Charles Rochet and Jean Tirole, "Platform Competition in Two-Sided Markets," Journal of the European Economic Association, Vol. 1, No. 4, (June 2003); Jean-Charles Rochet and Jean Tirole, "Two-Sided Markets: An Overview," mimeo, Institut D'Economie Industrielle, (2004); Mark Armstrong, "The Theory of Access Pricing and Interconnection," In: Handbook of Telecommunications Economics, (2002), Ch. 8; Julian Wright, "Access Pricing Under Competition: An Application to Cellular Networks," Journal of Industrial Economics, Vol., No. 3, (September 2002).

10 'Two-sided markets' differ from, and should not be confused with, what are sometimes called markets and aftermarkets, an example of which would be the market for photocopiers and the aftermarket for replacement toners. As with two-sided markets, the prices of photocopiers and toners are 'interdependent'. But in the copier case, both the photocopier and the toner are consumed by the same customer. Two-sided markets bring different groups of customers together. Furthermore, in markets and aftermarkets, it is generally possible to supply only one of the markets (e.g., selling toner without also selling photocopiers). In two-sided markets, the platform cannot exist unless it meets the needs of both sides of the market.

11 See generally, David S. Evans, "The Antitrust Economics of Two-Sided Markets," Yale Journal of Regulation ,Vol. 20, No. 2, (Summer 2003).

12 The viewers may not particularly want to meet some of these advertisers but are coaxed to do so with valuable content.

13 Market participants and industry developers are increasingly recognising the platform features and capabilities of phone networks. Regulators also increasingly recognise the platform features of mobile networks (see Ofcom's review of progress towards a digital UK, available at http://www.ofcom.org.uk/research/industry_market_research/m_i_index/cm/overview/rmd/2_2/?a=87101; and Ofcom's research on Driving Digital Switchover available at http://www.ofcom.org.uk/research/dso_report/section2. See also Steven Hope, "SDR Technology: What does it offer the Mobile Operator?" available at http://www.ofcom.org.uk/static/archive/ra/topics/research/topics/converge-new-emerging/sdr/4-hope.pdf).

14 Peter Pashigian, and Eric D. Gould, "Internalizing Externalities: The Pricing of Space in Shopping Malls," Journal of Law and Economics, 41, (1998), pp. 115–142.

15 The timing differs in that fixed line callers make calls after mobile customers have signed into a mobile network, whereas retailers sign leases before customers decide where to shop. Nevertheless, the reason retailers sign leases is to gain access to customers they know the mall will attract.

16 Retailers may be able to sell to those same customers in other locations at other times (in the same way as a fixed line caller may reach a mobile receiver other than via a fixed to mobile call), but they lose the opportunity to make the sale to those customers at that time (when the customers are presumably looking seriously to make purchases).

17 Jean-Charles Rochet & Jean Tirole, "Cooperation Among Competitors: Some Economics of Payment Card Associations," Rand Journal of Economics, 33, (2002); Richard Schmalensee, "Payment Systems and Interchange Fees," Journal of Industrial Economics, 50, (2002); Rochet, Jean-Charles and Jean Tirole, "An Economic Analysis of the Determination of Interchange Fees in Payment Card Systems," Review of Network Economics, Vol. 2(2), (2003b) pp. 69–79.

18 Proprietary systems, such as American Express or Diners Club, generally set fees directly to both cardholders and merchants. Card associations, such as MasterCard/Europay and Visa, set interchange fees that merchant banks pay to cardholder banks, which help determine the relative fees on the two sides of the system.

19 Jean-Charles Rochet and Jean Tirole, "Two-Sided Markets: An Overview," mimeo, Institut D'Economie Industrielle, (2004), pp. 10–13.

20 CC Report 2002, supra note 6, ¶ 1.7. The report notes the concern that the pricing structure in place encouraged overuse of mobile calls.

21 The optimal overall price level is determined by an analogue of the standard Lerner formula. The individual prices depend on the incremental costs of supplying each side and characteristics of the demand for each side, including the interaction between these demands. The optimal pricing structure depends on the relative demand externalities of each group of consumers. The optimal pricing conditions are shown in Rochet and Tirole (2003), supra note 10; Rochet and Tirole (2004), supra note 10. Other papers that investigate optimal pricing decisions in platform markets are Bernard Caillaud, and Bruno Jullien, "Chicken & Egg: Competition Among Intermediation Service Providers," RAND Journal of Economics, Vol. 34(2), (2003), pp.309–328; and Geoffrey Parker and Marshall Van Alstyne, "Information Complements, Substitutes, and Strategic Product Design" Working Paper No. 299, Tulane University and University of Michigan, (2000).

22 Platform businesses may earn risk-adjusted competitive rates of return. Or a platform business may be able to charge prices that result in profits that, measured ex post, exceed competitive levels. These may reflect compensation for making risky investments or rewards for risky innovation and therefore may reflect competitive returns when measured ex ante. Or they may reflect supra-competitive rates of return measured ex ante, perhaps as a result of a government grant of monopoly, luck, or anticompetitive practices. But these are the same issues that regulators and competition authorities face for many situations.

23 In the special case of linear demand, Rochet and Tirole find that the socially and privately optimal pricing structures are the same holding the overall price level constant (Jean-Charles Rochet and Jean Tirole, "Platform Competition in Two-Sided Markets," Journal of the European Economic Association, Vol. 1 Issue 4, (June 2003)).

24 "Acrobat 6 Professional is a totally stunning bit of software. It took a great amount of foresight to come up with the idea of Portable Document Format (PDF) ten years ago and the fact that it has become such a de facto standard is proof of what a great original idea it was. Acrobat 6 takes this success and really embeds it into both key applications as well as the operating system. Creating a PDF is now easier than doing a screen-grab, it's even easier than doing a 'SaveAs some other file format' in any application. I think the new Acrobat turns an industry-standard into an industry-foundation, with PDF looking to be the secure 'wrapper' of choice, in which single or amalgamated project documents will be exchanged." (Martyn Day, "CADserver," July 2003)

25 Many consumers today have the software installed, while others can download it for free from Adobe's web site.

26 Jean-Charles Rochet and Jean Tirole "Platform Competition in Two-Sided Markets," Journal of the European Economics Association, Vol. 1, (2003) pp. 990–1029; Julian Wright and Ulrich Kaiser, "Price Structure in Two-sided Markets: Evidence from the magazine industry," CEPR Discussion paper, (2004–08).

27 Even when consumers do not want advertising, they value the content that is paid for by the advertising more than any disutility from the advertising.

28 Approximately 80 percent of newspaper revenue comes from advertisers (Lisa George & Joel Waldfogel, "Whom benefits whom in daily newspaper markets?" National Bureau of Economic Research, Working Paper No. 7944, (2000)). In fact, one of the reasons some publications charge at all is because advertisers want some assurance of the quality of the subscriber base (that subscribers actually read the publication).

29 Of course, as in all of these examples, further lowering prices to the low price side would further increase consumers on this side. However, two-sided platforms have already chosen prices that maximize the value of their platforms. Artificially changing the pricing structure is, therefore, likely to diminish the value of the platform.

30 As is clearly the case with mobile phone telephony, where no operator can realistically provide an 'incoming calls only' service or an 'outgoing calls only' service.

31 Commission guidelines on market analysis and the assessment of significant market power under the Community regulatory framework for electronic communications networks and services (2002/C 165/03).

32 Both the Commission in its Draft Recommendation, and Ofcom in the consultations leading up to its Decision accept that some callers to mobile may switch to alternative means of reaching a subscriber if prices of mobile termination went up, and that some potential mobile subscribers may take into account the level of call termination. But, for example, Ofcom notes that the numbers involved are very small. "Ofcom's February 2004 survey of mobile users (forthcoming) found that, on an unprompted basis, only 2 per cent of users mentioned that one of the reasons for choosing a particular network was that it would be cheaper for others to call them. Only 9 per cent claimed that the cost of other people calling them was a significant factor in choosing their network, and that only 11 per cent had actually found out how much it would cost people to call them." (The Statement, supra 5, p. 13).

33 December Consultation, ¶. 2.4.

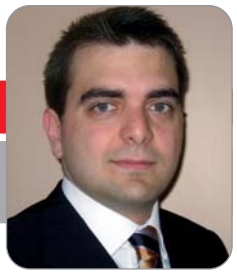
34 OJ C 165/6 11.7.2002, Commission guidelines on market analysis and the assessment of significant market power under the Community regulatory framework for electronic communications networks and services, ¶ 40–42, and PTS Decision section 2.2

35 In its Recommendation, the Commission claims that the competitive constraints on call termination are insufficient, by which it presumably means insufficient to constrain prices to competitive levels.

36 December Consultation, ¶ 3.19.

Dr. Christos Genakos[2]

Lecturer, Economics Department, Cambridge University



Prof. Tommaso Valletti[3]

Professor of Economics, Imperial College London

Christos Genakos is a Lecturer at the Economics Department of Cambridge University and a Fellow and Director of Studies in Economics at Selwyn College. He is also a senior research economist at the Centre for Economic Performance (CEP) at the London School of Economics. He received his PhD from the London Business School, after completing his MSc at the University College London. His research focuses on Industrial Organization and Applied Microeconomics. His recent work includes the welfare analysis of mergers in the PC industry, examining foreclosure incentives in the PC and server operating systems market, estimation of the "waterbed" effect in mobile telephony and evaluation of international management practices.

Tommaso Valletti has a magna cum laude degree in Engineering from Turin and holds a MSc and a PhD in Economics from the London School of Economics, where he also taught until 2001. He is Professor of Economics at Imperial College London, and also Professor of Economics at the University of Rome. He is a Fellow of CEPR.

Tommaso is a member of the panel of academic advisors to Ofcom, the UK communications regulator. He is also a member of the panel of academic advisors of the UK Competition Commission. He was a board director of Consip, the Italian Public Procurement Agency, in 2002-2005. He has advised numerous bodies, including the European Commission, OECD, and the World Bank on topics such as network interconnection, mobile telephony markets, and spectrum auctions.

Testing the "waterbed" effect in mobile telephony[1]

Abstract

This paper examines the impact of regulatory intervention to cut termination rates of calls from fixed lines to mobile phones. Under quite general conditions of competition, theory suggests that lower termination charges will result in higher prices for mobile subscribers, a phenomenon known as the "waterbed" effect. The waterbed effect has long been hypothesized as a feature of many two-sided markets and especially the mobile network industry.

Using a uniquely constructed panel of mobile operators' prices and profit margins across more than twenty countries over six years, we document empirically the existence and magnitude of this effect. Our results suggest that the waterbed effect is strong, but not full. We also provide evidence that both competition and market saturation, but most importantly their interaction, affect the overall impact of the waterbed effect on prices.

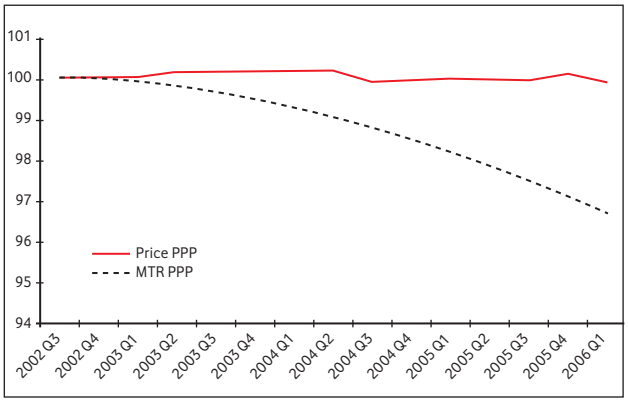
1. Introduction

Mobile termination charges[5] have become the regulators' focus of concern worldwide in recent years. Regarding the fixed-to-mobile termination rates especially, a large theoretical literature has demonstrated that independently of the intensity of competition for mobile customers, mobile operators have an incentive to set charges that will extract the largest possible surplus from fixed users.[6] This competitive bottleneck problem provided justification for regulatory intervention to cut these rates. However, reducing the level of termination charges can potentially increase the level of prices for mobile subscribers, causing what is known as the "waterbed" effect. The main purpose of this paper is to examine the existence and magnitude of the waterbed effect in the mobile telephony industry.

Both regulators and academics have recognized the possibility that this effect might be at work and be strong in practice. The first such debate started in 1997 in the UK with the original investigation by the Monopolies and Mergers Commission (now Competition Commission).[7] Another example is the New Zealand Commerce Commission which, in its 2005 investigation, initially took the position that mobile subscription prices would rise in response to a cut in termination rates only if mobile firms operated in a perfectly competitive environment. The Commission was subsequently convinced that the waterbed effect is a more general phenomenon, but there remained doubts about the importance of such an effect. The most recent termination rate proposals by the UK regulator Ofcom considered the issue of the waterbed in order to analyse the impact of regulation of call termination. Ofcom acknowledged the importance of the waterbed effect, but questioned whether the effect was "complete", arguing that this can only be the case if the retail market is sufficiently competitive.[8]

Yet, despite the importance of the waterbed effect for welfare calculations, there is no systematic evidence on its existence or magnitude. Casual empiricism suggests that mobile subscription prices have been decreasing quite steadily over time in virtually every country, despite the regulation of mobile termination rates. At the same time, though, the industry has become more competitive, with additional entry, tougher competition, etc., exerting a countervailing force. As an example, Figure 1 plots the evolution of subscription prices and termination rates in France. While termination rates have been cut over the years, prices to medium user customers have remained more or less constant. Does this imply there is no waterbed effect? Not necessarily as competition in the industry might also have intensified and other trends, such as economies of scale due to growth in traffic volumes, may also mask the impact of the waterbed on subscription prices.

Figure 1 Average price and MTR decline (France, Medium User)



Notes: Figure 1 presents normalized (at the beginning of the period) PPP-adjusted average prices (total bill paid by medium usage consumers) and MTR rates for France based on the Teligen and Cullen International dataset.

In this paper we analyze the impact of fixed-to-mobile termination rate regulation on prices and profit margins on a newly constructed dataset of mobile operators across more than twenty countries during the last decade. The timing of the introduction of regulated termination rates, but also the severity with which they were imposed across mobile firms, varied widely and has been driven by legal and institutional aspects of each country. Using quarterly frequency data and employing panel data techniques that control for unobserved time-invariant country-operator characteristics and general time trends, we are able to identify and quantify for the first time this waterbed effect. Our estimates suggest that although regulation reduced termination rates, this also led to an increase in mobile outgoing prices on average. This waterbed effect is shown to be robust to different variable definitions and datasets.

However, although the waterbed is shown to be high, our analysis also provides evidence that it is not full: accounting measures of profits are positively related to MTR, thus mobile firms suffer from cuts in termination rates. Finally, our empirical analysis also reveals that both competition and market saturation, but most importantly their interaction, affect the overall impact of the waterbed effect on prices: the waterbed effect is stronger the more intense competition is in markets with high levels of market penetration and high termination rates.

Our paper is related to an emerging literature on "two-sided" markets that studies how platforms set the structure of prices across the two sides of the business (see Armstrong, 2006, and Rochet and Tirole, 2006). Telecommunications networks are examples of two-sided markets: providing communication services to their own customers over the same platform and providing connectivity to their customer base to other networks. The two markets are linked: more subscribers on the network means more opportunities for users of other networks to make calls. Whenever we look at two-sided markets, the structure of prices (i.e., who pays for what) is fundamentally important for the development of the market. In mobile telephony, typically it is only senders that pay (under the Calling Party Pays – CPP – system), while receivers do not. This is why termination rates are not the locus of competition and, if left unregulated, they will be set at the monopoly level.[9]

This is also a case where the mobile firms sell two goods with interdependent demand: at any given termination rate, the volume of fixed-to-mobile calls that an operator receives depend on the number of mobile subscribers on its network. In a sense, mobile subscribers and fixed-to-mobile calls are complements, as an increase in the number of subscribers will cause an increase in the volume of fixed-to-mobile calls.[10] Our work therefore also contributes to the more general understanding of two-sided markets.

2. Two simple models of the waterbed effect

In this section we discuss two simple but related models that give rise to the waterbed effect. The first one is a perfect competition model, where the waterbed effect arises from the zero-profit condition. The second model analyzes a monopoly situation, where the waterbed effect arises via an increase in the 'perceived' marginal cost of each customer. The aim of this section is to show how the waterbed effect can emerge under a rather wide range of circumstances.[11]

Perfect competition embodies the assumption of a "full waterbed" since any termination rent is simply passed on to the customer. Hence, if there is a full waterbed, profits should not be affected by the level of the MTR. Still, a full waterbed effect does not imply a straightforward magnitude of the elasticity (defined as the percentage change in mobile subscriber pirces in response to a 1% change in the MTR). The waterbed elasticity could be above or below 1, in absolute value, depending on the relative sizes of (a) termination revenues relative to costs and (b) price elasticities for subscriptions and incoming calls.

A similar argument can be made in the case of pure monopoly. Each time a customer is attracted, it comes with a termination rent: the higher the rent, the lower the perceived marginal cost. If regulation cuts termination rents below the profit maximising level, this is 'as if' marginal costs increase, and as a consequence retail prices will increase as well. Hence, the waterbed phenomenon is also expected under monopoly.

Notice that our analysis has focused on an "uncovered" market, in the sense that there is always some customer who does not buy any mobile service. This assumption may be called into question as in many countries penetration rates now exceed 100%. While this does not alter our analysis in the case of perfect competition, the monopoly example requires this further qualification. In this limiting situation, a waterbed effect will not exist where there is a monopoly supplier in a fully penetrated market as the price set will be related only to the willingness to pay of the last subscribing customer and not to the termination rate.

In summary, the waterbed effect would arise under the two extreme cases of perfect competition and monopoly. These simplified models are admittedly unrealistic to describe the complex world of mobile telephony, but appealing as they generate the waterbed effect under very different assumptions. Mobile markets worldwide are dominated by a small number of firms. Competition among them is expected to be somewhere between the two extreme scenarios of perfect competition and monopoly. Under these more general (oligopolistic) market conditions, the same economic logic applies. We therefore expect the waterbed effect to be a robust phenomenon even after introducing complexities into the theoretical model that would make it a better and more realistic description of the industry.

Hence, our main predictions that we bring to an empirical test are:

1. A waterbed effect exists under quite general market conditions. Lower termination rates induced by regulation should be associated with higher retail prices to mobile customers. We also warn against a too simplistic interpretation of the waterbed price elasticities, since in general one should not expect a 1:1 effect even in a model with perfect competition, since demand elasticities and cost shares will have an impact too.
2. For low levels of market penetration, the impact on retail prices, via the waterbed effect, exists independently from the level of competition. As far as profits are concerned, when the industry is perfectly competitive, exogenous changes in termination rates have no impact on profits. On the other hand, when the industry is not competitive, profits are negatively affected by regulatory cuts in termination rates.
3. For high levels of market penetration, we expect an increase in competition to make the waterbed effect stronger. The waterbed effect is always expected to be in operation under competition for any level of market penetration. However, in the limiting case when the market is fully covered, a monopolist sets its prices just to ensure that the last customer subscribes to the services, in which case termination rates have no impact on mobile retail prices. Therefore, when relating the magnitude of the waterbed effect to the intensity of competition, we will want to control for the market penetration in a given market, since this is a good proxy for subscription demand elasticity at different stages of the product life cycle of mobile telephony.

3. Econometric specification and data

3.1 Estimation strategy

As an initial investigation we looked at the elasticity of calls from fixed to mobile phones. If demand for these calls were very elastic, for instance because fixed users calling a mobile customer could find alternative ways of contacting that customer, there would be a doubt that the waterbed effect could exist at all. The equation we estimate takes the form:

$$(1) \quad \ln Q_{ct} = \alpha_0 + \alpha_1 \ln MTR_{ct} + \beta \ln Penetration_{ct} + \gamma \ln Subscribers_{ct} + d_c + d_t + \varepsilon_{ct}$$

where c = 1, 2, ..., denotes the different countries and t = 1, 2, ..., denotes time. Q_{ct} denotes the total quantity of fixed to mobile calls (to Vodafone customers). α_1, β and γ can be directly

interpreted as elasticities. It is important to notice that the price in equation (1), the MTR, is regulated, therefore it can be taken as exogenous. We also do not directly observe the price of fixed-to-mobile calls charged to the fixed users (which would include also the so called "retention rate" set by the fixed network operator). However, this is not a major problem for two reasons. First, the retention rate is typically a small portion compared to the MTR. Second, what matters for the mobile firm is precisely the MTR and how demand reacts to changes in the MTR.

Turning now to the waterbed, our empirical strategy is in two steps. In the first step, the analysis is based on the following regression equations:

(2) $\ln P_{ujct} = \alpha_{ujc} + \alpha_t + \beta_1 \text{Regulation}_{jct} + \varepsilon_{ujct}$

(2a) $\ln \text{EBITDA}_{jct} = \alpha_{jc} + \alpha_t + \beta_1 \text{Regulation}_{jct} + \varepsilon_{jct}$

The dependent variable in (2) is the logarithm of outgoing prices ($\ln P_{ujct}$) for the usage profile u = {low, medium, high} of operator j in country c in quarter t. The dependent variable in (2a) is the logarithm of earnings before interest, taxes, depreciation and amortization ($\ln \text{EBITDA}_{jct}$) of operator j in country c in quarter t. EBITDA is defined as the sum of operating income and depreciation and we use it as a proxy for profits. The main variable of interest, Regulation_{jct}, is for the moment a binary indicator variable that takes the value one in the quarters when mobile termination rates are regulated.

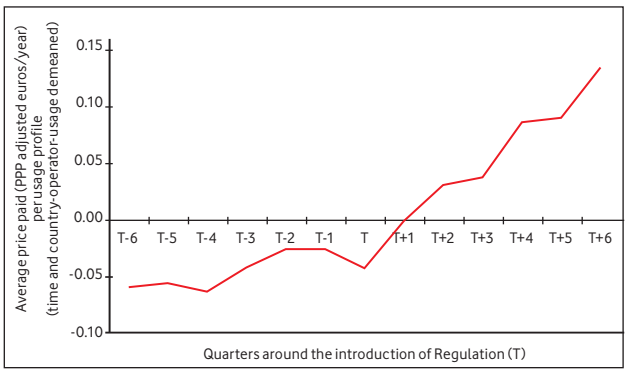
Both regressions (2) and (2a) constitute a difference-in-difference model, where countries that introduced the regulation are the "treated" group, while non-reforming countries (always regulated or always unregulated) are the "control" group. Due to the inclusion of (usage-)country-operator and time fixed effects, the impact of regulation on prices (or profits) is identified from countries that introduced this regulation and measures the effect of regulation in reforming countries compared to the general evolution of prices or profits in non-reforming countries. The "waterbed" prediction is that, *ceteris paribus*, the coefficient on regulation should have a positive sign in (2), and a zero or negative effect in (2a) depending on whether the market is competitive or not.

This difference-in-difference specification allows us to control for time-invariant country-operator characteristics that may influence both regulation and prices or profits. Furthermore, the specification also accounts for common global trends.

One important concern regarding this difference-in-difference specification is that the unbiasedness of the estimator requires strict exogeneity of the regulation variable. For example, our results would be biased if countries and operators, which have witnessed slower decrease in prices (including fixed-to-mobile prices) than comparable countries, were more likely candidates for regulation. The direction of causation here would be reversed: *because* of high retail prices, then fixed-to-mobile termination rates are regulated.

There are two ways we can address this concern. Firstly, according to theory, the intensity of competition should *not* matter as to whether or not to regulate MTRs. Unregulated MTRs are always "too high", independently from the level of competition (though the level of competition might affect the optimal level of regulated MTR). In principle, therefore, we should expect every country to regulate MTRs sooner or later, which is indeed what we observe in the data. Secondly, what we observe empirically is the exact opposite of the above prediction. Figure 2 plots the average (time and usage-country-operator demeaned) prices in countries that have experienced a change in regulation, six quarters before and after the introduction of regulation. As we can see, compared to prices in the rest of the world, average prices in countries that experienced a change in regulation were actually lower before the introduction of regulation. Moreover, in line with our predictions, the introduction of regulation has a clear positive impact on prices (the waterbed effect) that becomes stronger as regulation becomes progressively more binding over time. Hence, classical reverse causality seems to be less of a concern in our context.

Figure 2: Average price around the introduction of regulation



Notes: Figure 2 plots the evolution of time and country-operator-usage demeaned average logarithm of the PPP adjusted price paid per usage profile six quarters before and after the introduction of regulation of fixed-to-mobile termination charges based on the Teligen data corresponding to the best deals available at every period.

Most importantly for establishing causality, the regulation variable should be "random". This (non-selectivity) assumption is quite restrictive because regulatory intervention does not occur randomly, but is the outcome of a long regulatory and political process. However, this process regarding mobile termination rates has been driven in practice by legal and institutional aspects. The UK has been at the forefront of this debate and started regulating MTRs already back in 1997. Other countries followed suit. Importantly, the European Commission introduced a New Regulatory Framework for electronic communications in 2002. The Commission defined mobile termination as a relevant market. Procedurally, every Member State (EU 15 at the time) was obliged to conduct a market analysis of that market and, to the extent that market failures were found, remedies would have to be introduced. Indeed, all the countries that completed the analysis did find problems with no single exception, and imposed (differential) cuts to MTRs (typically, substantial cuts to incumbents and either no cut or only mild cuts on entrants). Hence, the timing of the introduction of regulated termination rates, but also the severity with which they were imposed across mobile operators has been driven by this regulatory process and varied widely across countries with no systematic pattern.

Moreover, conditional on (usage-) country-operator and time fixed effects, the regulation variable should be uncorrelated with other time-varying factors. In other words, the main criticism of our framework is that we do not allow for joint country-time fixed effects. A spurious correlation pointing towards a high waterbed would arise if, for example, a country is not regulated but is competitive and has low prices, while another country is regulated with low MTR but is also quite concentrated, so it has high prices: we attribute econometrically higher prices to the waterbed (via regulation), even if - in principle - the waterbed effect did not exist at all. While this may not be very plausible (typically, countries with low MTRs are also competitive, at least anecdotally, which should give rise to the opposite bias), it is important to bear in mind this caveat when interpreting our results. In addition, we tried to alleviate this data limitation problem as much as possible by splitting our sample of countries into three macro regions (Western Europe, Eastern Europe, and Rest of the World) and introducing regional-time control variables. Despite this not being the ideal solution, our results become stronger, as we will demonstrate in the next section.

Before we discuss the various data sources, it should be stressed that using only a binary indicator for regulation is quite restrictive. It does not allow us to distinguish between countries that have introduced substantial price cuts in MTRs and countries that have regulated MTRs too but only mildly. For this reason, we also experiment with two other measures of the impact of regulation.

In the spirit of Card and Kruger (1994), we construct two additional indexes. The first one is:

$$\text{MaxMTR index}_{jct} = \begin{cases} 0 \text{ if } MTR_{jct} \text{ is unregulated} \\ \dfrac{MaxMTR_{ct} - MTR_{jct}}{MTR_{jct}} \text{ if } MTR_{jct} \text{ is regulated} \end{cases}$$

In other words, when the country is unregulated, the index takes a value of zero. If instead the country is regulated, we construct an index that takes larger values the more regulated a mobile operator is, compared to the operator that is regulated the least in the same country and period.

This index takes advantage not only of the different timing of the introduction of regulation across countries, but also of the widespread variation on the rates imposed across operators within countries. This variation in regulated MTRs was particularly evident in countries where there was a large asymmetry between the "large" incumbents and the "small" entrants. While from a theoretical point of view the "monopoly bottleneck" problem exists independently from the size of an operator, in practice, regulators have been more reluctant in cutting the MTRs of the new entrants. They did this most likely with the idea of helping entrants secure a stronger position in the market. Thus new entrants have been either unregulated for many periods (while the incumbents were regulated at the same time), or they have been regulated nominally but only very mildly, while much more substantial price cuts were imposed on the incumbents. Hence, in this index, the highest MTR within a country at every period becomes the benchmark for comparing how tough regulation has been on the rest of the firms.

Our second regulation index is based on the same principle, but restricts the sample to only those countries for which we know with certainty that there is at least one fully unregulated operator. For example, the UK was one of the first countries to introduce termination rates regulation, but throughout this period mobile operator 3 (Hutchison) was left completely unregulated. Thus, for the purposes of this index we use the termination rates that this firm was charging as a benchmark for all the other firms. This exercise severely restricts our sample size, but makes the identification even more transparent and exogenous. Hence, the second index is:

$$\text{UnregulatedMTR index}_{jct} = \begin{cases} 0 \text{ if } MTR_{jct} \text{ is unregulated} \\ \dfrac{UnregulatedMTR_{ct} - MTR_{jct}}{MTR_{jct}} \text{ if } MTR_{jct} \text{ is regulated} \end{cases}$$

In other words, the index takes the value of zero when the country is unregulated. If instead the country is regulated, we construct an index comparing the rate each operator is regulated to the one charged by the unregulated firm in the same country and period. Both these indexes, allow us to get different measures of the severity of regulation in each country and period.

Finally, in the second step, our analysis is based on the following instrumental variable (IV) regression models:

$$(3) \quad \ln P_{ujct} = \alpha_{ujc} + \alpha_t + \beta_1 \ln(MTR)_{jct} + \varepsilon_{ujct}$$

$$(3a) \quad \ln EBITDA_{jct} = \alpha_{jc} + \alpha_t + \beta_1 \ln(MTR)_{jct} + \varepsilon_{jct}$$

The idea here is to estimate the waterbed effect on prices directly through the MTRs using regulation as an instrumental variable. Regulation is a valid instrument as it is not expected to influence prices other than the impact it induces via MTRs. This is because regulation acts on prices only indirectly via reducing MTRs, while regulators do not intervene in any other direct manner on customer prices.

3.2 Data

For the purpose of our analysis we matched three different data sources. Firstly, we use Cullen International to get information on mobile termination rates. Cullen International is considered the most reliable source for MTRs and collects all termination rates for official use of the European Commission. Using this source and various other industry and regulatory publications, we were also in a position to identify the dates in which regulation was introduced across countries and operators.

Secondly, quarterly information on the total bills paid by consumers across operators and countries is obtained from Teligen. Teligen collects and compares all available tariffs of the two largest mobile operators for thirty OECD countries. It constructs three different consumer usage profiles (large, medium and low) based on the number of calls and messages, the average call length and the time and type of call. A distinction between pre-paid (pay-as-you-go) and post-paid (contract) is also accounted for. These consumer profiles are then held fixed when looking across countries and time.

Thirdly, we use quarterly information taken from the Global Wireless Matrix of the investment bank Merrill Lynch (henceforth, ML). ML compiles basic operating metrics for mobile operators in forty-six countries. For our purposes, we use the reported average monthly revenue per user (ARPU) and the earnings margin before interest, taxes, depreciation and amortization (EBITDA). Through this source we also obtain information on market penetration and number of mobile operators in each country, together with the number of subscribers and their market shares for each operator.

All consumer prices, termination rates and revenue data were converted to euros using the Purchasing Power Parities (PPP) currency conversions published by the Organization for Economic Cooperation and Development (OECD) to ease comparability. None of our results depends on this transformation. More detailed data description, together with the dates of the introduction of regulation and summary statistics, can be found in the Appendix of the fuller version of the paper.

The various data sources have different strengths and weaknesses regarding our empirical question. The Teligen dataset has two main advantages. First, by fixing *a priori* the calling profiles of customers, it provides us with information on the best choices of these customers across countries and time. Second, the prices reported in this dataset include much of the relevant information for this industry, such as inclusive minutes, quantity discounts etc. (although it does not include handset subsidies). However, this richness of information comes at the cost of having data for only the two biggest operators of every country at each point in time. For instance, if a country, such as the UK, had five mobile operators, possibly regulated differentially over time, only two observations per customer profile would be available. This reduces the variability and makes identification of our variables of interest harder, especially given that the biggest mobile operators are often regulated at the same rate.

On the contrary, the ML dataset provides us with information on actual revenues rather than prices. The dependent variables that we use are primarily EBITDA (a measure of profit and cash flow) and ARPU (which consists of all revenues, including revenues from MTR). These are aggregate measures encompassing all revenues associated with mobile voice services. Therefore, they have to be interpreted as measures of an operator's revenues and profitability rather than the total customer bill. Both these measures suffer from endogeneity problems which could introduce bias and inconsistency in our results. However, this dataset contains information on almost all mobile operators in each country and hence it allows us to exploit more within-country variation.

4. Benchmark results

Starting with the demand for fixed to mobile calls from specification (1) we use the detailed cross-country dataset from Vodafone containing information on the quantity of fixed to mobile minutes, and on the revenues accrued from these phone calls, which is our measure of mobile termination rates, together with the number of Vodafone subscribers at each period. We also matched the Vodafone dataset with the dataset from Merrill Lynch, which contains information on the overall market penetration in each country. We do not fully report results here[12], but all elasticities estimated in specification (1) were statistically significant at the 1% level, and the results were quite realistic: there is an inelastic demand for fixed-to-mobile calls (elasticity of -0.216)[13], and as the market becomes more mature (high market penetration), the total number of fixed-to-mobile calls goes down (elasticity of -0.264) - most likely reflecting more mobile-to-mobile calls instead of fixed-to-mobile calls. Finally, fixing the penetration level, more Vodafone subscribers increase the total number of calls to Vodafone's subscribers (elasticity of 0.381).

Table 1 reports our benchmark results from the waterbed specification (2) using the price information from Teligen as the dependent variable. The data for this table consists of the best possible deals for each user profile among all possible contracts available, both pre-paid and post-paid. For that reason, we also add a binary variable ($Pre\text{-}paid_{jct}$) indicating whether the best deal was on a pre-paid contract or not.[14] The estimated waterbed is 0.133 and strongly significant in column 1, where we utilize the simplest specification with a binary indicator for regulation. That means that the introduction of regulation of MTRs increased bills to customers by 13% on average. Notice that the coefficient on pre-paid is negative but insignificant, indicating that prices on the best pre-paid deals were no different than those on monthly contracts.

In column 2, using the *MaxMTR* index we obtain again strong evidence of the waterbed effect. Similarly, in column 3 when we severely restrict our sample to only those countries we know with certainty had at least one unregulated mobile operator, we still get a positive and significant effect.[15] Notice also that the coefficient on pre-paid becomes now negative and significant, indicating that pre-paid customers were getting significantly better deals from the two main mobile operators when they were faced with an unregulated competitor. It seems likely that incumbents were offering significantly better deals to (the more elastic) pre-paid customers as a way of attracting consumers and putting pressure on the prices charged by their unregulated competitors.

In the last two columns, for reasons already discussed in the previous section, we estimate an even more restrictive version of our model by allowing for regional-time fixed effects. Essentially, our sample of countries can be naturally divided into three macro regions: Western Europe, Eastern Europe and Rest of the World (Australia, New Zealand and Japan). Western European countries have been all subject to the New Regulatory Framework adopted by the European Commission, while other Eastern European countries have only recently been subject to regulation with the accession of new member States. Controlling for these regional effects in columns 4 and 5, results in an even stronger waterbed effect, without reducing its statistical significance.[16]

Table 1: Estimating the "waterbed" effect (Teligen)

	(1)	(2)	(3)	(4)	(5)
Estimation method	OLS	OLS	OLS	OLS	OLS
Dependent variable	lnP_{ujct}	lnP_{ujct}	lnP_{ujct}	lnP_{ujct}	lnP_{ujct}
$Regulation_{jct}$	0.133***			0.152***	
	(0.033)			(0.033)	
$MaxMTR\ index_{jct}$		0.290***			0.316***
		(0.068)			(0.066)
$UnregulatedMTR\ index_{jct}$			0.127**		
			(0.051)		
$Pre\text{-}paid_{jct}$	-0.045	-0.051	-0.127***	-0.052	-0.056
	(0.040)	(0.041)	(0.044)	(0.039)	(0.040)
Observations	1734	1734	450	1734	1734
Country-Operator-Usage	150	150	36	150	150
Within-R^2	0.220	0.234	0.367	0.252	0.267

Notes: The dependent variable is the logarithm of the PPP adjusted total bill paid by consumers with different usage based on the Teligen data corresponding to the best deals available at every period. All equations include country-operator-usage and a full set of time dummies (first three columns) or a full set of region-time dummies (last two columns). All countries in the sample were divided into three macro regions: Western Europe, Eastern Europe and Rest of the World (RoW); see text for more details. Standard errors adjusted for heteroskedasticity and autocorrelation of unknown form and clustered by country-operator-usage are reported in parenthesis below coefficients: *significant at 10%; **significant at 5%; ***significant at 1%.

Next, we look at the impact of regulation on profitability measures using specification (2a). Table 2 reports the effect on EBITDA.[17] Column 1 shows that regulation had a negative effect on profit margins, although the data is considerably noisier. Using our two indexes, instead of the binary regulation variable (columns 2 and 3), reveal again a negative relationship, though the effect is not statistically significant.

In columns 4 and 5, the inclusion of the regional-time fixed effects increases the magnitude of the coefficients without affecting much their statistical significance. If markets were fully competitive there should be no impact on profits. Thus, these results suggest that competitors seem to have some degree of market power.

Table 2: Estimating the "waterbed" effect (Merrill Lynch)

	(1)	(2)	(3)	(4)	(5)
Estimation method	OLS	OLS	OLS	OLS	OLS
Dependent variable	$lnEBITDA_{jct}$	$lnEBITDA_{jct}$	$lnEBITDA_{jct}$	$lnEBITDA_{jct}$	$lnEBITDA_{jct}$
$Regulation_{jct}$	-0.125*			-0.138*	
	(0.070)			(0.076)	
$MaxMTR\ index_{jct}$		-0.024			-0.054
		(0.133)			(0.139)
$UnregulatedMTR\ index_{jct}$			-0.148		
			(0.236)		
Observations	1135	1135	319	1135	1135
Country-Operator	67	67	16	67	67
Within-R^2	0.209	0.203	0.281	0.215	0.209

Notes: The dependent variable is the logarithm of the EBITDA from the Merrill Lynch dataset. All equations include country-operator and a full set of time dummies (first three columns) or a full set of region-time dummies (last two columns). All countries in the sample were divided into three macro regions: Western Europe, Eastern Europe and Rest of the World (RoW); see text for more details. Standard errors adjusted for heteroskedasticity and autocorrelation of unknown form and clustered by country-operator are reported in parenthesis below coefficients: *significant at 10%; **significant at 5%; ***significant at 1%.

In our second step, using specifications (3) and (3a) we report the results from the IV regressions in Table 3. The first three columns use the same Teligen data as before, whereas the last three columns examine the effect on EBITDA. First stage results across all columns confirm that regulation has a significantly negative effect on MTR as expected. In addition, regulation does not seem to suffer from any weak-instruments problems as indicated by the first stage F-tests. Column 1 shows that that regulation through MTR has indeed a negative and significant effect on prices. The magnitude of the elasticity of the waterbed effect is above 1. Over the period considered, regulation has cut MTR rates by 11% and, at the same time, has increased bills to mobile customers by 0.11 x 1.207 = 13.3%.

The elasticity of the waterbed effect is lower at 0.938 and 0.334, in columns 2 and 3 respectively, using the more sophisticated indexes of regulation, but still negative and highly significant. The effect on accounting profits is positive and significant in column 4, and positive but not significant with the more nuanced measures of regulation. Table 4 also provides evidence that the results remain unchanged and if anything become stronger when we estimate the more restrictive version of our model that includes region-time fixed effects.

Table 3: "Waterbed" effect through MTR

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation method	IV	IV	IV	IV	IV	IV
Dependent variable	$\ln P_{ujct}$	$\ln P_{ujct}$	$\ln P_{ujct}$	$\ln EBITDA_{jct}$	$\ln EBITDA_{jct}$	$\ln EBITDA_{jct}$
$\ln(MTR)_{jct}$	-1.207***			1.127*		
	(0.411)			(0.603)		
MaxMTR $index_{jct}$		-0.938***			0.070	
		(0.278)			(0.392)	
UnregulatedMTR $index_{jct}$			-0.334**			0.620
			(0.133)			(0.862)
1st Stage Coef.	-0.110***	-0.310***	-0.382***	-0.111***	-0.335***	-0.239**
	(0.024)	(0.035)	(0.028)	(0.037)	(0.051)	(0.098)
1st Stage F-test	21.83***	78.85***	188.24***	8.90***	43.88***	5.90**
	[0.000]	[0.000]	[0.000]	[0.004]	[0.000]	[0.028]
Observations	1734	1734	450	1135	1135	319
Clusters	150	150	36	67	67	16

Notes: Columns 1, 2 and 3 utilize the Teligen data as in Table 1. The dependent variable for these columns is the logarithm of the PPP adjusted total bill paid by consumers with different usage for the best deals available. Columns 4, 5 and 6 utilize the Merrill Lynch dataset as in Table 2. The dependent variable for these columns is the logarithm of the EBITDA. All regressions use the "Regulation" dummy as the instrumental variable. All equations include either country-operator-usage (Teligen) or country-operator (Merrill Lynch) and a full set of time dummies. P-values for diagnostic tests are in brackets and italics. Standard errors adjusted for heteroskedasticity and autocorrelation of unknown form and clustered by either country-operator-usage (Teligen) or country-operator (Merrill Lynch) are reported in parenthesis below coefficients: *significant at 10%; **significant at 5%; ***significant at 1%.

Table 4: "Waterbed" effect through MTR (Regional-Time Controls)

	(1)	(2)	(3)	(4)
Estimation method	IV	IV	IV	IV
Dependent variable	$\ln P_{ujct}$	$\ln P_{ujct}$	$\ln EBITDA_{jct}$	$\ln EBITDA_{jct}$
$\ln(MTR)_{jct}$	-1.529***		1.415*	
	(0.496)		(0.757)	
MaxMTR $index_{jct}$		-1.076***		0.187
		(0.283)		(0.473)
1st Stage Coef.	-0.100***	-0.294***	-0.098**	-0.288***
	(0.023)	(0.032)	(0.038)	(0.052)
1st Stage F-test	18.15***	85.18***	6.47**	30.43***
	[0.000]	[0.000]	[0.013]	[0.000]
Observations	1734	1734	1135	1135
Clusters	150	150	67	67

Notes: Columns 1 and 2 utilize the Teligen data as in Table 1. The dependent variable for these columns is the logarithm of the PPP adjusted total bill paid by consumers with different usage for the best deals available. Columns 3 and 4 utilize the Merrill Lynch dataset as in Table 2. The dependent variable for these columns is the logarithm of the EBITDA. All regressions use the "Regulation" dummy as the instrumental variable. All equations include either country-operator-usage (Teligen) or country-operator (Merrill Lynch) and a full set of region-time dummies. All countries in the sample were divided into three macro regions: Western Europe, Eastern Europe and Rest of the World (RoW); see text for more details. P-values for diagnostic tests are in brackets and italics. Standard errors adjusted for heteroskedasticity and autocorrelation of unknown form and clustered by either country-operator-usage (Teligen) or country-operator (Merrill Lynch) are reported in parenthesis below coefficients: *significant at 10%; **significant at 5%; ***significant at 1%

We must remark that the ML dataset is probably less reliable than the Teligen dataset, so we take our conclusion on accounting profits more cautiously. In addition, all these results have to be qualified as termination rents could be also exhausted with non-price strategies, i.e., increasing advertising, or giving handset subsidies that we cannot control for. However, we do not expect handset subsidies effects to be too relevant, for instance, for pre-paid customers, and the test on EBITDA should take these additional factors into account. If handset subsidies were linked to inter-temporal subsidies (short-run losses are incurred to get long-run profits from captive customers), our results on profitability are, if anything, biased downwards. This is because a cut in MTR would look more profitable as fewer losses are made in the short run. Therefore our result on profitability would probably look stronger if we could account for handset subsidies.[18]

Taken together these benchmark estimates confirm our theoretical intuition that there exists a strong and significant waterbed effect in mobile telephony. However, this effect is not full as competing firms seem to enjoy some degree of market power.

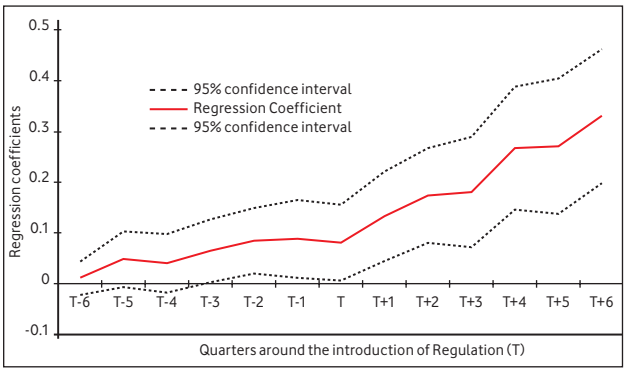
5. Dynamic regulation effects

The effect of regulation on prices might not be just instantaneous. On the one hand, termination rates are typically regulated over some periods using "glide paths", in which charges are allowed to fall gradually towards a target over that period. The temporal adjustment path is known and anticipated by operators, at least before a new revision is conducted. On the other hand, there could also be some inertia. For instance, customers may be locked in with an operator for a certain period, therefore there would be no immediate need for mobile operators to adjust their prices as these customers would not be lost right away. Alternatively, when termination rates change, it may take some time for operators to adjust retail prices because of various "menu" costs. Hence, we would like to investigate whether firms anticipated regulation (possibly by trying to affect the outcomes of the regulatory process) and indeed whether the effect of regulation was short-lived or had any persistent long term effects. To quantify these dynamic effects of the waterbed phenomenon, we define binary indicators for twelve, non-overlapping, quarters around the introduction of regulation and a final binary variable isolating the long-run effect of regulation.

Figure 3 plots the regression coefficients on these binary indicators together with their 95% confidence interval. As expected, regulation has no effect on prices six to four quarters before the actual implementation. However, there is some small but statistically significant anticipation of the regulatory intervention three to one quarters before. As discussed before, for the large majority of countries regulation was preceded by a long consultation period between the regulator and the various mobile operators. Our results reveal that operators started adjusting their price schedules slightly upwards even before the actual implementation of the new termination rates.

However, it is the actual implementation of the regulation that has the biggest impact on prices as revealed by the immediate increase on the coefficients after regulation. In other words, regulation is binding from the beginning and as it tightens up over time, the waterbed effect increases. As we can see in figure 3, regulation also seems to have a large and very significant long-run waterbed effect. The coefficient which quantifies the effect of regulation on prices post the sixth quarter after its introduction is strongly significant and implies a long run elasticity of the waterbed effect of 33%. Note that this coefficient is not directly comparable to the previous estimates of the waterbed effect, as it incorporates the effect not only of the introduction of regulation, but also of the progressive tightening of termination rates. What is crucial is that prices seem to respond continuously with every tightening of the rules giving rise to a waterbed phenomenon that is not a one-off event.

Figure 3: The evolution of the waterbed effect



Notes: Figure 3 plots the regression coefficients on binary variables six quarters before and after the introduction of regulation. The dependent variable is the logarithm of the PPP adjusted total bill paid by consumers with different usage based on the Teligen data corresponding to the best deals available at every period. All equations include country-operator-usage and a full set of time dummies. Standard errors are adjusted for heteroskedasticity and autocorrelation of unknown form and clustered by country-operator-usage.

6. Interaction with competition and further evidence

6.1 Competition and market penetration

Having established that the waterbed effect exists and has a strong long run effect, we now want to investigate in greater detail how competition affects this phenomenon. Competition is obviously expected to have a direct impact on prices: the more competitive the market, the lower the prices to customers. Besides this effect, however, if termination rates are "high" (e.g., unregulated) or a substantial mark-up is allowed, competition is expected to have an additional impact via the waterbed effect: the more competitive the industry, the lower the prices will be, on top of the direct effect, as any termination rent will be passed on to the customers. As discussed in Section 2, a waterbed effect is expected to exist also under monopoly, though the effect is milder as some rents will be kept by the monopolist. However, the waterbed effect is not expected to be very relevant under monopoly when the market is very saturated and the monopolist still has an interest in covering it. Hence, in our empirical specification it is crucial to control for subscription penetration levels. Our specification reads:

$$(4)\quad \ln P_{ujct} = \alpha_{ujc} + \alpha_t + \beta_1 \ln(MTR)_{jct} + \beta_2 \ln(Competitors)_{ct} + \beta_3 \ln(Penetration)_{ct} + \gamma_1[\ln(MTR)_{jct} \times \ln(Competitors)_{ct}] + \gamma_2[\ln(MTR)_{jct} \times \ln(Penetration)_{ct}] + \gamma_3[\ln(Penetration)_{ct} \times \ln(Competitors)_{ct}] + \delta[\ln(MTR)_{jct} \times \ln(Competitors)_{ct} \times \ln(Penetration)_{ct}] + \varepsilon_{ujct}$$

Equation (4) is an extension of our previous specification (3) with the aim to specify a particular channel that might affect the intensity of the waterbed effect. Our proxy for the intensity of competition is simply the number of rival firms ($Competitors_{ct}$) in each country and period. The number of mobile operators in a country can be taken as exogenous as the number of licences is determined by spectrum availability. Over the period considered, several countries have witnessed the release of additional licences. The degree of market saturation/maturity is measured as the percentage of the population with a mobile

phone ($Penetration_{ct}$). Our main coefficient of interest is δ, where MTR is interacted both with the intensity of competition and with the degree of market saturation.

Results are reported in Table 5. Column 1 is the baseline waterbed effect, comparable to that of column 1 in Table 3, restricted to the sample of firms and countries for which we have information on all these variables. Column 2 shows that a larger number of competing firms exerts the expected negative impact on prices. In column 3, the coefficient on the interaction between the competition variable and MTR is positive but insignificant, whereas in column 4 when we introduce all interaction terms, this coefficient becomes positive but barely significant.[19]

As we discuss in our theoretical section, the effect of competition on termination rates would differ depending on the level of market saturation and for that reason in column 5 we introduce our preferred specification which includes this triple interaction term. Our coefficient of interest, δ, is negative and strongly significant indicating that the waterbed effect is stronger the more intense competition is in markets with high levels of market penetration and high termination rates. This result is in line with our theoretical predictions where we pointed out the need to control for penetration levels when comparing competitive markets with concentrated ones.

Notice that the direct waterbed effect still exists in all markets, as β_1 is negative and very significant. The rest of the coefficients are also reassuring. We find that competition has a strong, negative direct impact on prices, besides any waterbed effect (β_2 = -0.344) and that prices are also systematically lower in more mature markets (β_3 = -3.228). When MTR is simply interacted with competition, not controlling for penetration levels, there is no statistically significant relationship.

We also find a positive and significant coefficient on the simple interaction between MTR and saturation (γ_2 = 1.422) and on the interaction between the number of competitors and market saturation (γ_3 = 2.346). Although these coefficients are not our main focus, a couple of comments are in place. A positive coefficient on γ_2 indicates that the waterbed effect is lower in higher penetration markets. Intuitively, low penetration markets usually consist of heavy users for whom the waterbed effect is expected to be strong. But as the market becomes more saturated, this typically involves attracting marginal users who make and receive very few calls. Hence, we expect the waterbed effect to decrease as the market becomes more saturated because of the different types of consumers that are drawn into the mobile customer pool. On the contrary, we have no prior expectations on the coefficient γ_3 as there is no strong reason to believe that, controlling for the number of competitors, the impact of competition should be more or less intense as the market saturates. On the one hand, a negative coefficient would arise if operators become less capacity constrained and compete more fiercely. On the other hand, if operators in mature markets tend to collude more easily over time, the result would be a positive coefficient.

Finally, in column 6, where we use as an instrument the *MaxMTR* index instead of the binary variable *Regulation*, we confirm the conclusions previously drawn. Results are virtually unaffected for the majority of the coefficients, with the direct waterbed effect (β_1) and the coefficient on the triple interaction (δ) becoming even stronger.

Therefore, in line with our theoretical predictions, our empirical analysis reveals that both competition and market saturation, but most importantly their interaction, affects the overall impact of the waterbed effect on prices.

Table 5: Competition and waterbed effect

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation method	IV	IV	GMM	GMM	GMM	GMM
Dependent variable	$\ln P_{ujct}$	$\ln P_{ujct}$	$\ln P_{ujct}$	$\ln P_{ujct}$	$\ln P_{ujct}$	$\ln P_{ujct}$
$\ln(MTR)_{jct}$	-1.580**	-1.282**	-0.733**	-0.775***	-0.585***	-1.026***
	(0.587)	(0.525)	(0.285)	(0.235)	(0.223)	(0.220)
$\ln(competitors)_{ct}$		-0.289*	-0.473***	-0.522***	-0.344**	-0.339*
		(0.173)	(0.180)	(0.178)	(0.173)	(0.188)
$\ln(mkt\ penetration)_{ct}$		-0.768	-0.533	-1.785***	-3.228***	-3.707***
		(0.483)	(0.371)	(0.563)	(0.840)	(0.882)
$\ln(MTR)_{jct} \times \ln(competitors)_{ct}$			0.093	0.168*	0.098	0.117
			(0.097)	(0.087)	(0.083)	(0.086)
$\ln(MTR)_{jct} \times \ln(mkt\ penetration)_{ct}$				0.168	1.422***	1.792***
				(0.141)	(0.364)	(0.413)
$\ln(competitors)_{jct} \times \ln(mkt\ penetration)_{ct}$				0.962**	2.346***	2.527***
				(0.441)	(0.557)	(0.587)
$\ln(MTR)_{jct} \times \ln(competitors)_{jct} \times \ln(mkt\ penetration)_{ct}$					-0.895***	-1.191***
					(0.248)	(0.293)
Observations	1371	1371	1371	1371	1371	1371
Clusters	141	141	141	141	141	141
Sargen-Hansen test of overidentifying restrictions	-	-	4.244	4.418	6.071	3.654
			[0.374]	[0.220]	[0.108]	[0.301]

Notes: The dependent variable is the logarithm of the PPP adjusted total bill paid by consumers with different usage for the best deals available from the Teligen data. All equations include country-operator-usage and a full set of time dummies. P-values for diagnostic tests are in brackets and italics. Standard errors adjusted for heteroskedasticity and autocorrelation of unknown form and clustered by country-operator-usage are reported in parenthesis below coefficients: *significant at 10%; **significant at 5%; ***significant at 1%.

7. Conclusions

Regulation of fixed-to-mobile termination charges has become increasingly prevalent around the world during the last decade. A large theoretical literature has demonstrated that independently of the intensity of competition for mobile customers, mobile operators have an incentive to set charges that will extract the largest possible surplus from fixed users. This competitive bottleneck problem provided scope for the (possibly) welfare-improving regulatory intervention. However, reducing the level of termination charges can potentially increase the level of prices for mobile subscribers, the so-called "waterbed" effect.

In this paper we provide the first econometric evidence that the introduction of regulation resulted in a waterbed effect. However, although the waterbed effect is high, our analysis also provides evidence that it is not full: accounting measures of profits are positively related to MTR, thus mobile firms suffer from cuts in termination rates. Finally, our empirical analysis also reveals that the waterbed effect is stronger the more intense competition is in markets with high levels of market penetration and high termination rates.

Our findings have three important implications. First, mobile telephony exhibits features typical of two-sided markets. The market for subscription and outgoing services is closely interlinked to the market for termination of incoming calls. Therefore, any antitrust or regulatory analysis must take these linkages into account either at the stage of market definition or market analysis.

Second, any welfare analysis of regulation of termination rates cannot ignore the presence of the waterbed effect. Clearly, if the demand for mobile subscription was very inelastic, the socially optimal MTR would be the cost of termination (though the regulation of MTR would impact on the distribution of consumer surplus among fixed and mobile subscribers). If, instead, the mobile market was not saturated and still growing there would be a great need to calibrate carefully the optimal MTR. We acknowledge that this calibration exercise is very difficult and must be done with great caution. It is therefore all the more important that further analysis is undertaken to understand the behaviour of marginal users that might give up their handsets when the waterbed effect is fully at work.

Third, our analysis on the existence and magnitude of the waterbed effect is also relevant in the current debate of regulation of international roaming charges. The European Commission has voted in 2007 to cap "roaming charges"[20] of making and receiving phone calls within the EU. The aim is to reduce the cost of making a mobile phone call while abroad and hence encourage more overseas (but within EU) phone use. Hence, a reduction in roaming charges may cause a similar waterbed phenomenon, whereby prices of domestic calls may increase as operators seek to compensate for their lost revenue elsewhere. While the magnitude of the waterbed effect caused by this new legislation is debatable, our results demonstrate that regulators have to acknowledge its existence and carefully account for it in their welfare calculations.

Future research should concentrate on two aspects that we consider to be the limitations of this paper. On the one hand, more detailed information would allow researchers to overcome our data limitations. Having price data on a larger number of mobile operators within countries, would allow for joint country-time fixed effects to be properly controlled for in the empirical specification. Furthermore, to investigate the marginal consumer's behaviour before and after the introduction of regulation and their elasticity regarding the waterbed effect, more detailed consumer-level information is required. On the other hand, given the non-linear retail price schedules and the complex incentives schemes (handsets, personal vs. business buyers' contracts, etc.) provided by mobile operators, more detailed customer information at a country level would allow us to model more satisfactorily the effect of competition and market penetration on the waterbed effect.

References

Armstrong, M., 1998, "Network Interconnection in Telecommunications," Economic Journal, 108: 545-564.

Armstrong, M., 2002, "The Theory of Access Pricing and Interconnection", in M. Cave, S. Majumdar and I. Vogelsang (eds.) Handbook of Telecommunications Economics, North-Holland, Amsterdam.

Armstrong, M., 2006, "Competition in Two-Sided Markets", RAND Journal of Economics, forthcoming.

Card, D. and A.D. Krueger, 1994, "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania", American Economic Review, 84: 772-793.

Hausman, J. and J. Wright, 2006, "Two-sided markets with substitution: mobile termination revisited", mimeo, MIT.

Laffont, J.-J., Rey, P. and Tirole, J., 1998, "Network competition I: Overview and non-discriminatory pricing", RAND Journal of Economics, 29: 1-37.

Rochet, J.-C. and J. Tirole, 2006, "Two-Sided Markets: A Progress Report", RAND Journal of Economics, forthcoming.

Valletti, T. and G. Houpis, 2005, "Mobile termination: What is the 'right' charge?" Journal of Regulatory Economics, 28: 235-258.

Wright, J., 2002, "Access Pricing under Competition: an Application to Cellular Networks", Journal of Industrial Economics, 50: 289-316.

Notes

1 We would like to thank Steffen Hoernig, Tobias Kretschmer, Marco Manacorda, Elias Papaioannou, Jonathan Sandbach, Jean Tirole, John Van Reenen, Julian Wright, and seminar audiences in Barcelona, Rome, Paris and the 8th CEPR Empirical IO meeting for helpful comments and discussions. We are also grateful to Bruno Basalisco for research assistance. We acknowledge research funding from Vodafone. The opinions expressed in this paper and all remaining errors are those of the authors alone.

2 Selwyn College, University of Cambridge & Centre for Economic Performance, London School of Economics, e-mail: cg370@cam.ac.uk, Web: http://www.sel.cam.ac.uk/personal/genakos/

3 Tanaka Business School, Imperial College London, University of Rome "Tor Vergata" & CEPR, e-mail: t.valletti@imperial.ac.uk, Web: http://www3.imperial.ac.uk/people/t.valletti

4 A fuller version of this paper is available from the authors' web pages.

5 These are the charges mobile operators levy on either fixed network operators or other mobile operators for terminating calls on their networks.

6 See, for example, Armstrong (2002), Wright (2002), Valletti and Houpis (2005) and Hausman and Wright (2006). Armstrong and Wright (2007) also provide an excellent overview of the mobile call termination theoretical literature and policy in the UK.

7 The term "waterbed" was first coined by the late Prof. Paul Geroski, chairman of the Competition Commission in the UK, at the time of the first investigation on interconnection charges in the mobile industry.

8 See "Mobile call termination, Proposals for consultation", Ofcom, September 2006.

9 The U.S. is a noticeable exception in that there is both a RPP (receiving party pays) system in place and, in addition, termination rates on cellular networks are regulated at the same level as termination rates on fixed networks. The U.S. also has a system of geographic numbers that does not allow to distinguish between calls terminated on fixed or mobile networks. For these reasons, the U.S. is not included in our sample. Most of the mobile world is under a CPP system.

10 It is important to be very careful with the use of standard definitions taken from normal "one-sided" markets. In this example, the notion of complementarity between mobile subscribers and fixed-to-mobile calls is more controversial if one starts instead with a price increase of mobile termination.

11 For an analytical treatment, see the fuller version on the authors' web pages.

12 Results are available on request.

13 This is the elasticity with respect to MTR (T). To get the elasticity of fixed-to-mobile calls with respect to the end user price (P), one has to multiply it by the "dilution factor" P/T and the "pass-through rate" dT/dP. In the case of the UK, Ofcom have assessed a dilution factor of approximately 1.5. Ofcom also believes that pass-through of the termination may be less than complete (i.e., $dP/dT < 1$, or $dT/dP > 1$), since BT's price regulation applies to a whole basket of services. However, in other European countries the fixed network retention (P - T) is itself directly regulated.

14 It is important to mention that the MTR is applied uniformly and does not distinguish, say, between calls to heavy users on contracts and calls to low users on prepaid. However, the waterbed price reaction of the mobile firm to changes in MTR can in principle differ by type of user or call, since their profile of received calls can differ, or the intensity of competition can differ by type of user too.

15 The elasticities are not directly comparable as the regulatory variables have different mean values.

16 We do not report the results of column 3 with the regional-country fixed effects because the Western Europe region binary indicator includes all the countries that had one operator being not regulated.

17 Similar results on the impact on ARPU are shown in the longer version of this paper.

18 All our analysis is related to the regulation fixed-to-mobile termination rates and not to mobile-to-mobile termination rates. This should not raise particular concerns in our analysis for two reasons. First, in many jurisdictions mobile-to-mobile rates are not regulated, apart from imposing reciprocity, and therefore cuts in fixed-to-mobile rates do not apply to other types of calls. Second, if for some reasons termination of both types of calls are regulated at the same level, theory says that a change in reciprocal mobile-to-mobile rates should have no obvious impact on profits and tariffs (just a re-balancing in the various components of the customer's bill). If firms compete in two-part tariffs, the impact of reciprocal access charges on profits and bills is neutral (see Armstrong, 1998, and Laffont et al., 1998). Thus we really interpret our empirical results as the impact of the regulation of fixed-to-mobile termination rates on prices and profits.

19 The instruments used are discussed in the fuller version available from the authors' websites.

20 These are the charges made to customers when using their phones outside their home country, i.e., an Italian customer making/receiving a phone call in Greece.

Dr. Jonathan Sandbach

Head of Regulatory Economics,
Vodafone Group Public Policy

Pass-through

Abstract

This paper explores the theory and evidence on the nature of "pass-through" of mobile termination rates in fixed network operator tariffs. Fixed network operators receive a cost saving from reduced mobile termination rates, which they may or may not pass onto their own subscribers, through reduction in fixed-to-mobile call prices, or other prices in their service portfolio. In theory any company that is freely setting prices to maximize profits in a competitive or monopoly market will adjust those prices in response to changes in marginal cost. However, the empirical evidence suggests that without regulation, pass-through of mobile termination rate reductions into consumer prices for fixed-to-mobile calls may not occur, or may be incomplete. This suggests that some fixed network operators may be using the opportunity of mobile termination rate reductions to move their fixed-to-mobile retail prices nearer to profit maximizing or monopoly levels, and calls into question the idea that fixed network operators are constrained by competition in their pricing of fixed-to-mobile calls.

Why pass-through?

Whenever Mobile Termination Rates (MTRs) change, Fixed Network Operators (FNOs) experience a change in their costs. This cost saving can be either: (1) passed directly onto consumers through a reduction in fixed-to-mobile (F2M) prices; (2) passed onto consumers through price reductions spread across the whole portfolio of FNO services; or (3) retained as profit with no alteration to end prices.

In some countries FNO regulation prescribes case (1) through regulated fixed network mark-ups or retentions. Other countries include the fixed network retention within an overall basket of regulated FNO services, thus providing the FNO with a choice of case (1) or case (2). In other countries the fixed network retention is unregulated, leaving the FNO to choose between all cases. Table 1 summarises the position in a selection of OECD countries with CPP (Calling Party Pays).

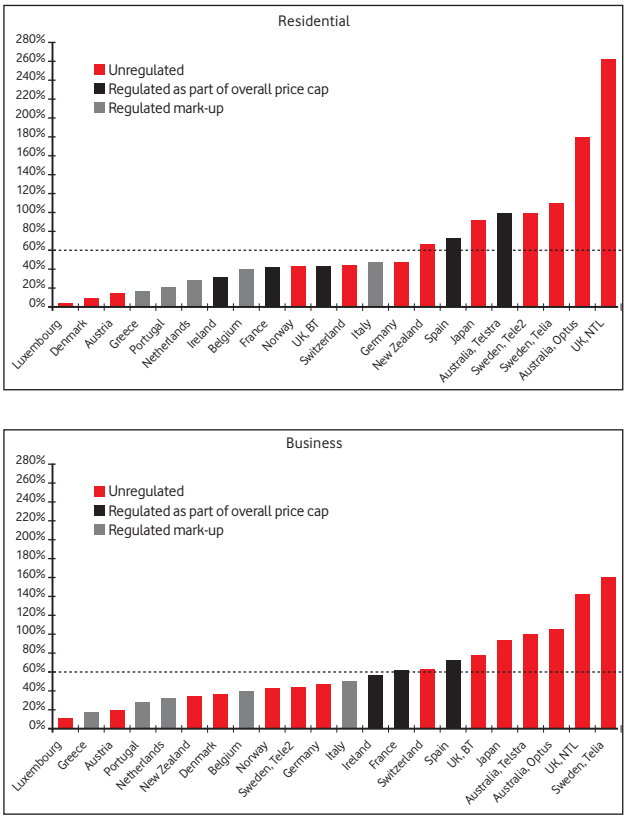
Table 1: Retail price regulation of F2M calls

FNO retention directly regulated	FNO retention regulated as part of general price cap	FNO retention unregulated
Belgium	Australia	Austria
Greece	France	Denmark
Italy	Ireland	Germany
Netherlands	Spain	Hungary[1]
Portugal		Luxembourg
		New Zealand
		Norway
		Sweden
		Switzerland
		UK

Source: Vodafone

Chart 1 shows the magnitude of the FNO F2N call price retention (in percentage terms) over the MTR. Countries where the retention is specifically regulated are shown in blue, countries where it is regulated as part of a general retail price basket are shown in green, and countries where it is unregulated are shown in red. Chart 1 shows a tendency for retentions to be significantly higher in countries where it is unregulated.

Chart 1: Magnitude of FNO mark-up (retention) over MTR



Source: Vodafone analysis using data for 2005

Economic theory shows that whenever an FNO has freedom in setting prices in a competitive or monopoly market, those prices will always be changed in response to a change in input costs (in this case the MTR). The precise relationship between changes in prices and changes in costs will depend on the price elasticity. If demand is very elastic (perhaps due to intense competition between FNOs), then margins will be kept low, and changes in the MTR will be directly reflected in equal changes in prices (full pass through). If demand is inelastic, a profit maximizing MNO will seek to charge a high mark-up over the MTR (as shown in Chart 1 for the "unregulated" FNOs) but, as before, when the MTR changes, the profit maximizing price will also change.

The situation is much the same if the FNO is regulated only on an overall basket of its retail prices. The FNO has freedom to set whatever balance of prices it chooses, but can be expected to alter this balance of prices in response to changes in the costs of one individual input (the MTR).

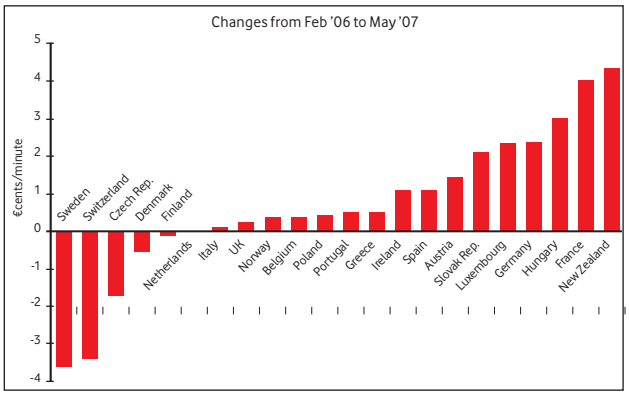
The only situation in which we would not expect pass-through is when the FNO is not formerly regulated on its retention, but has been "obliged" to set a price below the profit maximising level. In this case, an FNO may use a reduction in the MTR as an opportunity to move closer to a profit maximizing price structure by not reducing the F2M price. If this were the case, the implication would be that the existing F2M price was constrained more by "political acceptability" rather than competition.

Empirical measurement of pass-through

It is not possible to measure pass-through of changes in the MTR to the broad level of all FNO prices. This is because F2M calls contribute only a small part of the FNOs' overall revenue base, and MTR's form only a small proportion of the overall FNO cost base. As we have seen, however, if pass-through does occur we would expect to see it concentrated on F2M call prices. Even a profit maximizing monopolist would adjust prices in response to a change in costs.

It is possible to directly observe pass-through of changes in the MTR to changes in F2M prices, by tracking changes in the FNO mark-up (or retention) over time. This is done by Chart 2 for the period from February 2006 to May 2007. In all these countries there have been significant MTR reductions over this period.

Chart 2: Changes in the FNO mark-up (retention) over MTR



Source: F2M prices have been measured from the Teligen T-Basket analysis, by taking the difference in the Residential Composite Basket with and without calls to mobiles, and dividing this by the number of calls to mobiles in the basket. It thus represents the incremental revenue per minute earned by FNOs from F2M calls, which is the measure we seek for our purposes. The MTR is a subscriber weighted average of the rates of all MNOs in each country.

FNOs in some countries have actually reduced their retentions on F2M calls. These notably include Sweden and Switzerland where, although not regulated[2], the mark-ups were originally set at high levels (in excess of 6c), and so it is understandable that there has been pressure for reductions. Similarly the mark-up in the Czech Republic has been reduced from a high base of over 6c.

There are then a group of countries where the mark-up has stayed roughly constant. This includes countries where the mark-up is directly regulated (Belgium, Greece, Italy, Italy and Portugal).

In all other countries, however, there have been significant increases in the FNO retention on F2M calls, resulting from failure to pass-through MTR reductions. This group includes both countries where the mark-up is unregulated (Germany, New Zealand, Hungary and Luxembourg) and at least one country where it is only regulated as part of a larger basket of FNO tariffs (France).

In fact, in the cases of Germany, France, New Zealand and Luxembourg pass-through has been well below 50% (i.e. the FNOs have passed on less than 50% of MTR reductions).

There must be some caveats before drawing conclusions from Chart 2. F2M prices will be influenced by variables other than the MTR (for example, the changes in the degree of competition between FNOs, network and non-network productivity improvements, etc.). However, these effects would, if anything, have put pressure on FNOs to reduce further F2M prices. In contrast, however, between February 2006 and May 2007, most European FNOs have increased the mark-up they charge on F2M calls.

Conclusion

The empirical evidence suggests that without regulation, pass-through of any MTR reductions into consumer prices for F2M calls may not occur, or may be incomplete.

We have seen that in theory any company that is freely setting prices to maximize profits will adjust those prices in response to changes in marginal cost. Therefore, the only explanation for the behavior in Germany, Luxembourg and New Zealand is that F2M prices are set below profit maximizing levels, and the FNOs used the opportunity of an MTR reduction to move prices closer to a profit maximising level.

This calls into question (1) the likelihood of consumer benefit from MTR reductions in countries where the FNO retention is not specifically regulated, and (2) the idea that FNOs are constrained by competition in their pricing of F2M calls.

Notes

1 Other than regulation of designated USO tariffs.

2 It is possible that Swisscom and Telia, although not regulated on F2M prices, felt it necessary to be seen to act in consumer interests to avoid explicit regulation.

Welfare implications

Abstract

This paper explores the consequences for regulation of the "waterbed" effect in the tariffs of the mobile network operators, and shows how this should impact the quantum of any regulated mobile termination rate. Even if the waterbed is only partially effective (e.g. 50% effective), on plausible values for other market parameters (e.g. price elasticities), the optimal level of the mobile termination rate will typically be at least double the underlying cost.

Introduction

Existing regulatory practice holds that mobile call termination must be priced at cost in order to ensure an economically efficient pricing structure.[1] This conclusion ignores the impact of the waterbed effect, long recognized in the mobile industry. In simple terms, the waterbed applies when, faced with changes to mobile termination rates (MTRs), mobile network operators (MNOs) respond by increasing other prices. The theoretical explanation is that any change in the economic rents being earned on call termination will affect the intensity of competition for subscribers (who bring with them termination revenues) or, even without competition for mobile subscribers, will alter the revenue that the monopoly network operator would require to profitably connect a subscriber to the network. In either case, a lower MTR will result in higher mobile prices (than would otherwise be the case).

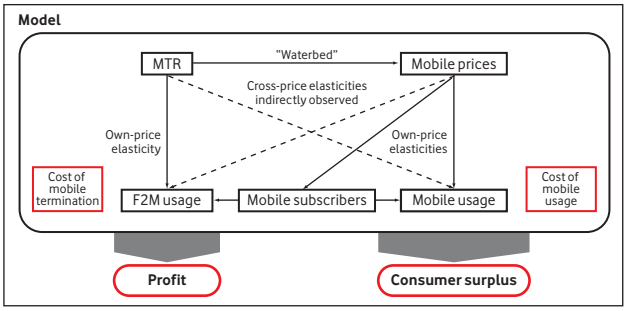
The waterbed has already been discussed and quantified in the paper of Genakos and Valletti earlier in this pamphlet. The purpose of this present paper is to quantify the impact the waterbed should have on the "optimal" level of any regulated MTR. We show that if the waterbed is recognized, it should result in a material mark-up to MTRs. Although the magnitude of this mark-up will be dependent on price elasticities which are difficult to measure with precision, the mark-up is nevertheless material for plausible elasticity assumptions.

Modelling the MNO waterbed

We have constructed a model of the competitive equilibrium between competing mobile operators.

Chart 1 summarises the linkages in the model.

Chart 1: Model overview



There are three own-price elasticities that we include in the model. The first is the own-price elasticity of F2M calls with respect to the MTR.[2] Genakos and Valletti estimate this elasticity to be -0.216, capturing the pass through of MTR to F2M prices, although other research by NERA using a fuller dataset of all major operators in the UK, Netherlands and Belgium estimates a lower value of -0.05. The second and third are the own price elasticities of mobile subscription and usage per subscriber with respect to mobile prices. Previous econometric studies have estimated a subscription elasticity of -0.55[3] and up to -0.9[4] for usage per subscriber respectively (using aggregate data for the whole industry level – firm specific elasticities will be higher). We also assume that a 10% increase in subscribers will cause a 10% increase in usage. This is a reasonable assumption, combining the fact that new subscribers will have a lower usage level, but will provide more people for existing subscribers to call. Analysis undertaken by Vodafone shows that whereas in developing countries the latter effect dominates, in developed countries these two effects off-set each other.

The number of mobile subscribers also has an effect on F2M usage. More mobile subscribers will mean more calls to those mobile subscribers. The mobile subscriber elasticity on F2M calls was estimated by Genakos and Valletti to be 0.381.

Substitution between mobile and F2M calls also needs to be captured. Ideally we would want to include these effects through direct cross-price elasticity estimates within the usage demand functions for F2M and mobile originated calls. In practice, however, these cross price elasticities are difficult to estimate as MTR and mobile call prices tend to fall in parallel, since both are affected by the same underlying cost trends (enforced in the former through cost based regulation, and in the latter through competition). Genakos and Valletti estimate a negative elasticity of -0.264 on F2M calls from increases in the overall subscriber base, capturing the cross-price elasticity effect.

Finally, the MTR and mobile prices are also linked through the so-called "waterbed". For example, an increase (decrease) in the MTR will (given inelastic demand) provide additional (less) economic rent, to be competed away by MNOs in their mobile prices. We look at the impact of different waterbed assumptions, ranging from zero to 100%.

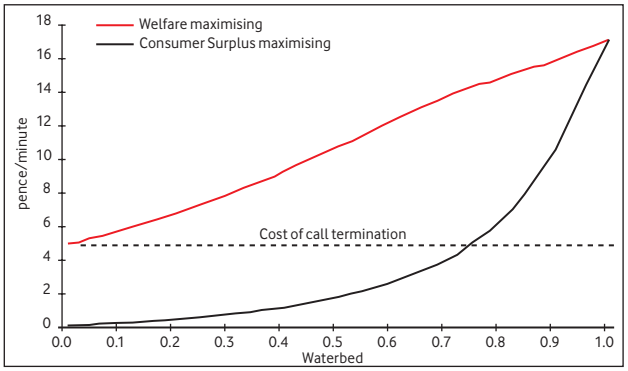
We have initially calibrated the model using the same data as Ofcom for the purposes of the UK mobile call termination rate review, covering quantities, prices and costs.[5] We then re-calibrated the model for a selection of other countries, taking account of the differences in quantities of subscribers, F2M call volumes, MTRs and average revenue per user (as a proxy for mobile prices). Most of this calibration data was sourced from Wireless Intelligence, with the exception of F2M call volumes and MTRs, for which we used Vodafone's own internal sources.[6]

Before we look in detail at the country results it is worth reviewing the consequences on the waterbed effect for one country in detail. We choose here to look at the UK. We look at a continuum of waterbed effects from (a) none at all, i.e. mobile prices are invariant of the MTR; to (b) complete waterbed effect, i.e. MNO profits are invariant to the MTR.

Chart 2 shows the welfare maximizing and consumer surplus maximizing level of the MTR as the effectiveness of the waterbed increases. Consumer surplus has been calculated for both mobile services and F2M calls, assuming that MTR reductions are fully passed through to F2M price by the fixed network operators. The previous paper show that this may not be the case where the FNO is not subject to explicit regulation on the mark-up (or retention) it charges on F2M calls. If this is the case, we would expect consumer surplus to be maximized by a higher MTR, especially if the mobile waterbed is high (since part of the benefit of lower MTRs will be passed to fixed network operator profits).

If there is no effective waterbed, the optimal MTR is consistent with textbook economics. Consumer surplus is maximized by a zero MTR (which would, of course, result in the MNO making a loss on each call terminated), whilst welfare (the sum of consumer and producer surplus) is maximized when the MTR is set at cost (4.9p in this case). However, as the waterbed becomes effective, the optimal level of the MTR increases. Increases in the MTR allow reductions in mobile prices, creating additional consumer surplus. Therefore, the consumer surplus maximising MTR and the welfare maximizing MTR will both increase, the former more than the latter due to the loss of producer surplus (profit) at high MTRs when there is an effective waterbed. When the waterbed is fully effective, profits are unaffected by the MTR, and so consumer surplus and welfare maximizing MTR converge at 17.1p.

Chart 2: Optimal MTR and the waterbed



We also looked at the sensitivity of the optimal MTR to changes in the assumption of the subscriber elasticity (assumed to be -0.55 in Chart 2). Chart 3 shows the welfare maximizing and consumer surplus maximizing level of the MTR as the elasticity varies between 0.0 and -1.0, whilst fixing the waterbed at 50% effective. Note that even when the elasticity on the number of subscribers is zero, subscriber usage still responds to changes in mobile prices in our model. If the number of mobile subscribers is inelastic consumer surplus is maximized by an MTR below cost (but a little above zero because of the usage elasticity), whilst the welfare maximizing MTR is a little above cost for similar reasons (if we were to assume zero elasticities for both subscriber and usage the welfare maximising MTR would equal cost). However, as the magnitude of the subscriber elasticity increases in absolute terms, the level of the optimal MTR rises.

Chart 3: Optimal MTR and the subscriber price elasticity

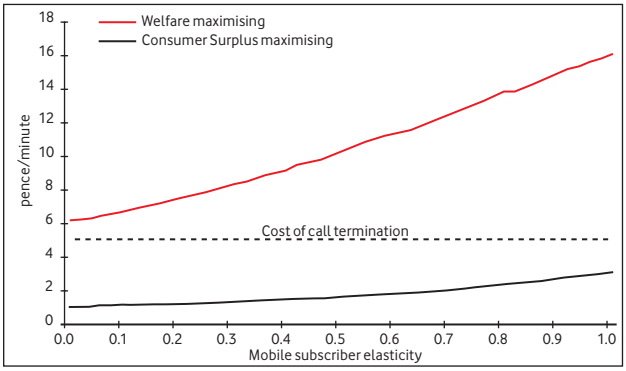


Table 1 provides results for the model calibrated to a number of countries. In all cases a termination cost of 4.9p has been assumed,[7] and the same price elasticities. The cases differ by the demand-side calibrations in each country – particularly the volumes of incoming and outgoing traffic and revenues.

Table 1: Welfare maximising MTRs

Pence/minute	Waterbed effectiveness		
	None (0%)	Partial (50%)	Full (100%)
Albania	4.9	13.0	38.8
Belgium	4.9	12.6	26.5
Czech Republic	4.9	11.7	18.8
France	4.9	11.2	17.4
Germany	4.9	10.8	17.6
Greece	4.9	11.8	22.0
Hungary	4.9	11.1	19.3
Ireland	4.9	10.5	15.5
Italy	4.9	10.7	18.0
Malta	4.9	11.8	20.6
Netherlands	4.9	9.3	11.7
Poland	4.9	10.9	25.1
Portugal	4.9	12.3	21.2
Romania	4.9	11.6	12.1
Spain	4.9	11.9	20.9
Switzerland	4.9	11.4	16.5
UK	4.9	10.6	17.1

Source: Vodafone analysis

The illustrated results assume subscriber and usage elasticities of -0.55 and -0.3 respectively – if these elasticities were zero, then the welfare maximizing MTR would equal cost irrespective of the waterbed assumption. However, under our price elasticity assumptions we see that the optimal MTR is significantly above cost in all cases.

Another feature of the results in Table 1 is the wide spread between countries, despite the same assumed underlying cost level. This is a consequence of a range of different factors in each country:

- different levels of F2M calls (per subscriber or as a proportion of total traffic), ranging from low levels in Albania where the fixed network is weak, to much higher levels in Western European countries. A lower level of F2M calls effectively increases the above cost mark-up required on each call to achieve the same level of externality benefit through the waterbed effect;
- different levels of mobile prices, ranging from low levels in Poland and Romania, to up to three times higher in Switzerland and some other western European countries. Higher mobile prices indicate a greater consumer value for mobile services, thus increasing the externality benefit from subsidizing subscriptions through a higher MTR.

This shows the importance of individual market characteristics in determining the correct level of the MTR in individual countries.

Conclusion

NRAs typically ignore the two-sided market that exists in mobile telecommunications when setting a regulated rate for mobile call termination. The principal effect that is omitted is the so-called waterbed, whereby MNOs use the surplus revenue on call termination to reduce outgoing prices and thus increase their subscriber base. Even in markets that appear fully saturated, it is important to maintain this price structure in order to achieve optimal levels of economic welfare.

Ignoring the waterbed will lead NRAs to set regulated MTRs, based solely on Long-run Incremental Cost (LRIC), at less than 50% of their optimal level (i.e. optimal rates will be more than twice LRIC levels). Furthermore the optimal level of the MTR varies significantly by country, even when costs are assumed to be the same. This shows the importance of individual country analysis, and very cautious use of cross-country benchmarks.

Our analysis has ignored the issue that, for some countries, pass-through of MTR reductions to F2M calls (or prices of any other fixed network operator service) may not be complete. If this is so, as it may be in a number of countries, welfare will be maximized by setting a still higher MTR. By doing this the economic rents arising from F2M calls will be channelled back to consumers through the competitive mobile waterbed (lower mobile prices), rather than being retained by fixed network operators in higher profits.

Notes

1 There are a few exceptions in the recognition by some regulatory authorities of the existence of a network externality effect, justifying a small mark-up over cost. For example, in the UK Ofcom include a mark-up in regulated mobile termination rates. See "Mobile call termination", Ofcom, 27 March 2007, See Annex 16.

2 In fact, the price of an F2M call will be the sum of the MTR and the fixed network's retention. For the purposes of this analysis we assume that the fixed network retention will be a constant mark-up, and so need not concern us. In many countries the fixed network retention is regulated. In other countries, however, this assumption is optimistic. As discussed in the previous paper, reductions in the MTR may be passed through to other fixed network operator services, or used to increase fixed network operator profits. Where the latter is the case, welfare benefits of lower MTRs will be lost to consumers, and overall welfare will also be lower since F2M prices will remain high.

3 Hausman, J. (1997). Other estimates include Madden and Coble-Neal (-0.25) and DotEcon (-0.37).

4 Estimates of price elasticities on mobile usage per subscriber are not readily available (due to the lack of published usage data). However, Vodafone's own econometric analysis typically yields estimates of -0.9.

5 Ofcom "Mobile call termination", Ofcom, 27 March 2007, Annex 19.

6 The reliance on Vodafone data dictated the choice of countries.

7 In practice termination costs will differ significantly between countries.

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. All the steps we have taken demonstrate our commitment to making sustainable choices.

Designed and produced by Barrett Howe Plc