No. 2133

THE THEORY OF ACCESS PRICING: AN OVERVIEW FOR INFRASTRUCTURE REGULATORS

Tommaso M Valletti and Antonio Estache and

INDUSTRIAL ORGANIZATION

Centre for Economic Policy Research

ISSN 0265-8003

THE THEORY OF ACCESS PRICING: AN OVERVIEW FOR INFRASTRUCTURE REGULATORS

Tommaso M Valletti and Antonio Estache

Discussion Paper No.2133
April 1999

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: http://www.cepr.org

CEPR Discussion Paper No. 2133

April 1999

ABSTRACT

The Theory of Access Pricing: an Overview for Infrastructure Regulators*

The paper provides policymakers and regulators with an overview of the more relevant theoretical issues related to the pricing of access to ensure that the political debate around practical concerns is solidly grounded. The paper discusses in detail the importance of access pricing in the context of: 1) a liberalized and vertically separated industry, 2) liberalized but vertically integrated industries, 3) unregulated access (private negotiations).

JEL Classification: D4, L1, L9
Keywords:access pricing, regulation

Tommaso M Valletti
Department of Economics
London School of Economics
Houghton Street
London
WC2A 2AE
UK
Tel: (44 171) 955 7551
Fax (44 171) 831 1840
Email: t.valletti@lse.ac.uk

Antonio Estache
World Bank
1818 H Street NW
Washington
DC 20433
USA
Tel: (1 202) 473 6370
Fax (1 202) 334 8350
Email: aestache@worldbank.org

*This paper is produced as part of a CEPR research network on The Evolution of Market Structure in Network Industries, funded by the European Commission under the Training and Mobility of Researchers Programme (Contract No ERBFMRX-CT98-0203) This paper was prepared while the first author was a visiting scholar in the Infrastructure Regulation unit of the Economic Development Institute of the World Bank. We are indebted to Ian Alexander, Phil Burns, Claude Crampes, Richard Green, Jean-Jacques Laffont, Martin Rodriguez-Pardina and Xinzhu Zhang for very useful discussions and/or comments. The paper represents however our views only and does not engage in any way the institutions we are affiliated with.

Submitted 29 March 1999

NON-TECHNICAL SUMMARY

An important component of policies to promote effective competition in all segments of network industries such as electricity, telecommunications or railways, is a regulatory environment guaranteeing that competitors have access to the services of 'potential bottleneck facilities' too costly to duplicate. Fair access rules to these facilities and in particular fair access prices will generally improve economic efficiency by easing competition in markets both upstream and downstream of the bottleneck. This is true whether the industries are vertically unbundled/separated or not. More specifically, appropriate access pricing rules are needed whenever a dominant firm controls the supply of one or more inputs that are vital for its competitors.

A failure to design these access rules properly is one of the key reasons why the potential gains from the restructuring and deregulation of network utilities are not maximized and/or shared fairly between the users and the owners of these essential facilities. This is also why the access problem is at the top of the agenda of many regulators in developed countries. Regulators in Canada, the UK, the US and EU countries, for instance, are all spending significant amounts of resources to address this issue properly in the key infrastructure sectors. The access problem is currently less well understood in developing and transition countries, although significant efforts are being undertaken in various Latin American countries to assess it as part of their efforts to address second generation issues in regulation. It is also becoming a key issue in Eastern Europe as an increasing number of countries are trying to get ready to take the necessary formal steps to smooth their admission to the EU.

The main purpose of this paper is to provide policy-makers and regulators with an overview of the more relevant theoretical issues related to the pricing of access to ensure that the political debate around practical concerns is solidly grounded. The paper first clarifies the linkages between industry structure and access pricing, then it discusses in details the importance of access pricing in the context of a liberalized and vertically separated industry (i.e. an industry with a vertical unbundling of its key activities into its main components such as generation, transmission and distribution in the case of electricity). The paper also covers the case of liberalized but vertically integrated industries (i.e. when at least one of the firms in the industry provides both competitive and bottleneck services as in the case of companies controlling transmission as well as some generation or some distribution of electricity). The situation with unregulated access (private negotiations) is also considered. Finally, the issue of the implementation of the theoretical principles is discussed.

After discussing a number of theoretical findings, Ramsey prices stand up as a good benchmark because they say something that should not be neglected*:*

markets are related so that demand and supply cannot be considered in isolation. When financial constraints have to be met, Ramsey prices provide the second best guidelines that take into account both allocative and productive efficiency. This is however not easy to implement since they require a good deal of information. Difficulty, however, does not imply unfeasibility. While it is true that elasticity of demand is difficult to forecast for new innovative services (especially in the telecommunications sector), in some utilities (gas, electricity) patterns of demand are rather standard and predictable so that the regulator could and should try to produce such estimates. In particular, a price cap mechanism may be the regulatory tool that brings Ramsey prices without requiring the regulator to explicitly calculate them.

The paper delivers one main message: *the access charge is often performing too many tasks*. Different goals and policy objectives lead to alternative ways of calculating optimal charges. While it is true that theory is extremely useful to understand the mediating function of access prices, we stress that one first fundamental step should precede any access distortion: *whenever possible, the use of access pricing as an instrument for the promotion of too many goals should be resisted and other instruments should be used*. Regulators should be aware that there is a sequencing of events that can reduce the complexity of the access problem. For instance, if the regulator believes there are barriers to entry, the tax/subsidy issue of the entry barrier should be addressed directly and be made explicit, rather than burying it into the access pricing problem. The latter could indeed be the only option available, but only after having realised that other options are not feasible. A similar argument can be made for universal service obligations. In other words, by understanding the links between different problems, new instruments become available that allow us to fine tune the regulatory process.

Introduction

An important component of policies to promote effective competition in all segments of network industries such as electricity, telecoms or railways, is a regulatory environment guaranteeing that competitors have access to the services of "potential bottleneck facilities" too costly to duplicate. Fair access rules to these facilities, including fair access prices, will generally improve economic efficiency by easing competition in markets both upstream and downstream of the bottleneck. This is true whether the industries are vertically unbundled/separated or not. More specifically, appropriate access pricing rules are needed whenever a dominant firm controls the supply of one or more inputs that are vital for its competitors. Examples include: gas transportation, electricity transmission, local telecommunications access or railway track. At a more general level, access pricing is part of the anti-trust concern of market foreclosure which is central in the so-called "essential facilities doctrine" covered by the US legislation.[2] In a broader context, access pricing is related to a variety of competition policy issues that include quantity discounts, cross-subsidies, tie-ins, refuse to deal or to unbundle, exclusive dealing and predatory pricing.

A failure to design these access rules properly is one of the key reasons why the potential gains from restructuring network utilities are not maximized and/or shared fairly between the users and the owners of these essential facilities. This is also why the access problem is at the top of the agenda of many regulators in developed countries. Regulators in Canada, the US or some EU countries for instance are all spending significant amounts of resources to address this issue properly in the key infrastructure sectors. The access problem is currently less well understood in developing and transition countries, although significant efforts are being undertaken in various Latin American countries to assess it as part of their efforts to address second generation issues in regulation.[3] It is also becoming a key issue in Eastern Europe as an increasing number of countries are trying to get ready to take the necessary formal steps to smooth their admission to the EU.[4]

Our main purpose is to provide policymakers and regulators with an overview of the more relevant theoretical issues related to the pricing of access to ensure that the political debate around practical concerns is solidly grounded.[5] The paper deals mainly with infrastructure services and it will not venture into the vast literature on anti-trust economics.[6] The rest of the paper is organized as follows. After a brief introduction to clarify the linkages between industry structure and access pricing, section 2 discusses the importance of access pricing in the context of a liberalized and vertically separated industry (i.e an industry with a vertical unbundling of its key activities into its main components such as generation, transmission and distribution in electricity). Section 3 covers

---

[2] The essential facilities doctrine has developed in the US through the application of the Sherman Act 1890. The Act declares illegal every contract in restraint of trade or commerce and prohibits every attempt to monopolise any part of the trade or commerce among several States or with foreign nations. The doctrine first surfaced in the 1912 case *United States v Terminal Railroad Association* (TRA). TRA was a joint venture that acquired terminal railroads on either side of the two bridges crossing the Mississippi River into and out of St Luis. That junction was of extreme commercial importance: 24 railroads converged there and the cost of building a further bridge was too high. By acquiring control of the bottleneck, TRA effectively controlled all railroad traffic converging in St Louis. This was made particularly explicit by a clause that allowed TRA members to exclude non-members. The Supreme Court found that the joint venture agreement was in breach of the Sherman Act. The Supreme Court did not require the breaking up of TRA since there were some benefits from joint ownership, but that access to the terminal should be just and reasonable for every firm needing access.

[3] See for instance, Estache and Rodriguez-Pardina (1998).

[4] For a recent overview: see Bruce *et al.* (1998).

[5] This paper updates and complements an extremely useful and user-friendly first survey by Cave and Doyle (1994).

[6] This does not imply that anti-trust problems are not relevant to network utilities, rather that the nature of access within such industries is so peculiar to require in many occasions a specific regulatory authority, otherwise the general guidelines of competition law could suffice.

the case of liberalized but vertically integrated industries (i.e. when at least one firm in the industry provides both competitive and bottleneck services such as a company controlling transmission as well as some generation or some distribution of electricity). Section 4 discusses the situation with unregulated access. Finally, section 5 covers some useful theoretical contributions to the implementation of the principles discussed in the previous sections. A more technical annex illustrates the degree of technicality that regulators will have to get into when applying the principles discussed here. The paper concludes with some very brief comments in section 6.

1. Industry structure and access pricing

Many possible industry structures are studied by academics. They generally vary according to the degree of regulatory intervention and to the discretion permitted to the bottleneck owner. Despite differences, they all share two common features that are relevant to most practitioners:

- the existence of vertically related markets;
- the existence of an essential facility provided only by one firm.

Markets are said to be *vertically related* when the production or supply of final goods or services involves different activities from "upstream" to "downstream" that are linked to each other in a clear sequence. For instance, in the electricity industry the chain of activities that includes generation, transmission, distribution, metering and billing jointly delivers electricity power to a final user. Some of the activities are potentially competitive (e.g. generation, metering and billing), some others exhibit increasing returns to scale, making it likely that a single firm owns them because their duplication would be too wasteful (e.g. transmission and distribution).

The control of an essential facility in the production chain is quite common. The joint ownership of electricity or gas transmission and generation is common even in restructured countries and can damage competition in generation for instance. In telecoms, for the call of the customer of any telephone company (fixed, mobile, etc.) making a call to a customer of some other operator to be completed requires that the first company has "access" to the other operator's subscriber. Even if the two end users belong to the same operator, access may be required. This is for instance the case of a call between two mobile users when the delivery of the call is made over the interconnected fixed network. To deliver the final service (a completed call), various links and switches will be used. If alternative networks exist, the call can be routed in different ways. It is likely, however, that some parts of the network are not duplicated. The local loop involves high fixed costs and it is a candidate to be a bottleneck in the hands of a single firm.

With vertically-related activities and bottlenecks, there are two main types of regulation:

- regulation of structure;
- regulation of conduct.

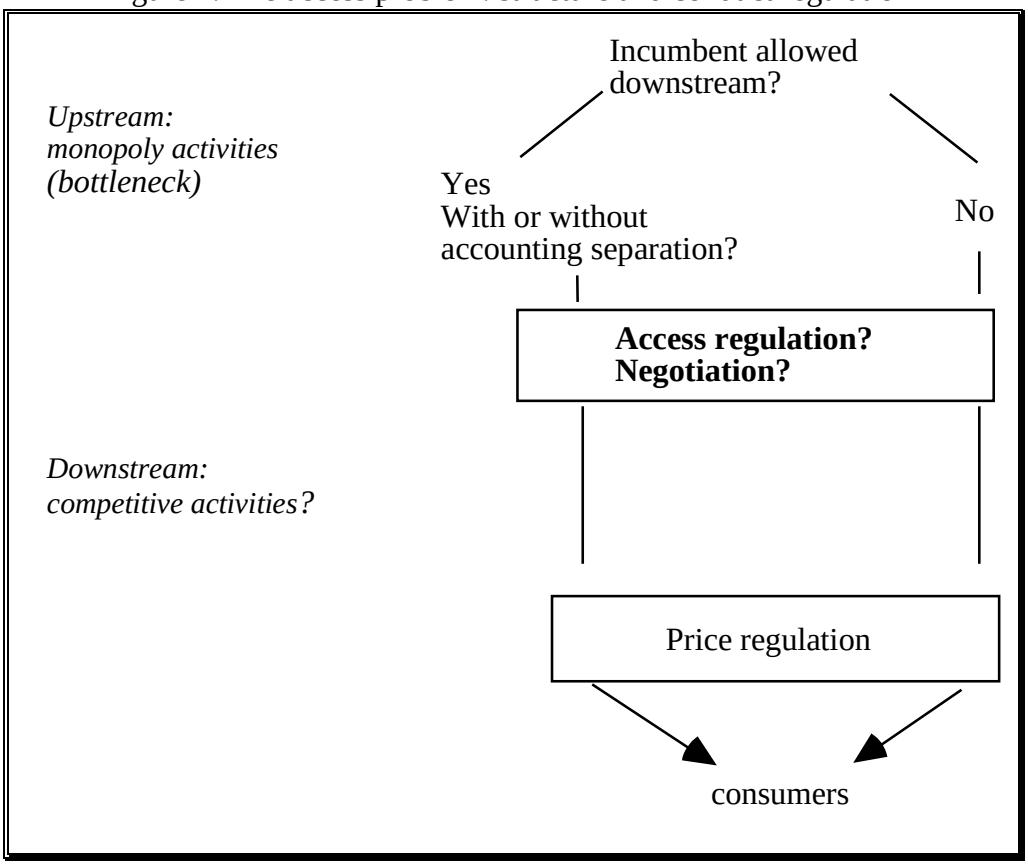
The *regulation of structure* includes merger controls, the removal of entry barriers, restrictions on the line of business or the break up of an integrated incumbent. For our purposes, it addresses two fundamental questions at the downstream level, namely whether the upstream monopolist is allowed into the downstream market, and if and how entry is regulated downstream. The *regulation of conduct* concerns the pricing behaviour of firms both in terms of their level and their structure. The constraint on prices can be both at the final and at the intermediate level.[7]

In practice, a mix of structure and conduct regulation is likely to be adopted, together with other instruments such as accounting separation. In fact, often, a regulator may leave access charges to be

[7] A broader view on regulatory issues can be found in Armstrong *at al.*, 1994; Laffont and Tirole, 1993.

set by private negotiation and intervene only if parties fail to reach an agreement. In sum, the analysis can quickly become complicated. The broad issues are summarised in Figure 1 which defines the organization of the rest of the paper.

Figure 1. The access problem: structure and conduct regulation[8]



To address these possibilities, the paper distinguishes between two particular vertical structures:

- vertical separation/unbundling with liberalisation,
- vertical integration with liberalisation.

These two cases are not exhaustive by any means, but they are particularly important in practice and the corresponding stylised models are useful to capture the relevant problems. In principle, there should be a third broad structure, a vertically integrated monopoly without liberalisation. This is not considered here because access pricing would only become a problem of internal transfers between the firm's divisions.[9]

---

[8] We do not get into quality regulation but it is clear that technical or service quality can both be adjustment variables for the service providers and hence need to be monitored quite seriously.

[9] The discussion on the choice between structures (separation vs. integration) is beyond the scope of this paper. We take a certain structure for given, and address the access problem within that structure. The desirability of integration versus separation can become easily a controversial issue. We just recall that integration of an incumbent firm is to be preferred when there are important economies of scope between the provision of bottleneck services and potentially competitive ones. Other standard benefits of integration include many forms of co-ordination, both in terms of pricing and other activities. For instance, an integrated firm internalises many external effects, so that its decisions are more efficient in terms of network co-ordination and compatibility between parts, network integrity, service reliability, timetabling. An integrated firm may charge lower final prices if it avoids a chain of successive mark-ups over costs imposed by separate different firms. From the legal point of view, allowing an incumbent firm to participate in the downstream market, does not necessitate to draw potentially arbitrary boundaries between activities. When boundaries are ill-defined, there could be costly litigation battles or lengthy hearings, or the firm may be induced to bypass constraints using inefficient investments (think of the money

2. Access Rules for vertically unbundled industries

Most restructuring of infrastructure activities results in some degree of vertical unbundling of the sector aiming at freeing the potential competitive forces wherever possible. Previously vertically integrated electricity and gas public utilities have had their various activities dis-integrated vertically to allow competition in generation and in distribution for instance. Vertical unbundling of infrastructure industries often results in the creation of some type of potential bottleneck activity. This could be an electricity transmission company structurally separated from downstream competitive activities that require the essential service, i.e. the distribution companies who purchase electricity from the generators upstream from the transmission company. It could also be the owner of rail tracks who sells access to various companies owning rolling stock. Because of the separation, regulation is somehow simpler than the case of an integrated incumbent. Most of the concerns arising from anti-competitive behaviour are now absent since, under many circumstances, the upstream firm has no reason for favouring one particular downstream firm at the expenses of the others. This section is however useful to introduce problems that would *also* arise with an integrated incumbent. Box 1 may also be useful to show how these questions arise in practice as illustrated by a discussion of access pricing issues in the UK telecoms context.

Box 1: The importance of market structure in practice: lessons from UK telecoms*

In the UK telecoms industry, the basket of services included in the price-cap regime is now very narrow (the coverage is less than 20% of total revenues of the incumbent operator, BT—British Telecom) because the regulator wants to restrain from regulating prices in competitive activities. The regulator has even announced that any retail price control will end from the year 2001, thus abandoning the price-cap mechanism in place since the privatisation BT. This shift in the regime is due to the fact that BT is now facing competition from a number of well-established operators. However, the industry will not be subject only to ordinary competition law guidelines, but the regulator will still regulate the interconnection charges. This is because BT still owns 90% of subscriber lines, with the implication that the majority of calls are either initiated or terminated on BT's network. The regulatory problem is complicated by the fact that competition, despite being more intense than a few years ago, is far from perfect. The number of license holders in the UK is very high, but the majority are niche players and only a few operators are big enough to contrast BT in an effective way. As a consequence, in the final markets conditions do not resemble those of a textbook example of pure competition, rather the market is better described by an oligopoly with a limited number of firms that can exercise some degree of market power.

If there was competition in the final market, the situation would be simpler. Under proper accounting separation, the downstream division of the incumbent could not be put at any advantage compared to the rivals. Price competition would eliminate extra-profits downstream and the access charge should only deal with the recovery of fixed costs in the bottleneck. However, in practice it is likely that firms can enjoy some degree of market power because of the investments involved. Market power implies that the final allocation is going to be reduced compared to the desirable one. Hence access charges have to perform additional roles, taking into account the strategic interaction of firms in the final market.

The example shows that the regulation of access in the presence of unregulated and imperfectly competitive final markets, is not a mere academic exercise. It matters in practice. Only when demand is linear, products are homogeneous and there are constant returns to scale both upstream and downstream is there is no difference between vertical integration and vertical separation: price competition brings prices down to costs, and there are no further fixed costs to recover. This is seldom the case in practice…suggesting that most telecoms regulators will have to address the issue

* For details see Valletti (1998).

spent by public telephone operators on Video-on-Demand when they are not allowed to provide broadcasting services). The boundary itself is likely to change over time, and technological advances can make a restriction obsolete in a very short period.

2.1 The basic framework

Consider an upstream firm, providing access to a downstream sector that in turn sells to the final users. The access price has to be regulated to induce firms to take efficient decisions and to attain levels of production in the interest of consumers. The main difference between the regulation of a standard monopolist (i.e. when final prices are regulated) and the regulation of access considered here, is that downstream firms have interrelated demand for access. Imagine, for instance, two rail companies offering passenger services only. They both need access to tracks and stations. What "company 1" can get depends on the services it sells but it also depends on what the other company sells. The access charge enters both directly into the firm's profit and indirectly since it influences the rival's choices. The effect becomes particularly important if the downstream market is not perfectly competitive, so that strategic considerations affect the picture and add a twist to the analysis. We address these points below in a slightly more rigorous way.

a. What happens in the best theoretical situation (the first best)?

In order to provide one unit of final good, downstream firms need one unit of the upstream input. This input is produced by the bottleneck owner at a unit cost c_0. The bottleneck owner also incurs a fixed cost F. This F can be interpreted as the set up cost of the network, or some other costs deriving from social obligations that cause losses to the bottleneck provider. . The users are charged a unit access charge denoted by a..

If all firms in the competitive sector are similar (in terms of technology and thus of costs) and their products are identical, firms undercut each other until price competition drives to zero all extra profits in the final market. The price charged to final users ends up equal to the marginal cost of each firm. This amounts to the sum of the access charge and any other cost incurred in order to transform the intermediate good. If we denote by c all the other unit costs (besides the unit access charge), the final price would be $p = a + c$. The lower the access charge, the lower the final prices and the higher the total quantity consumed by the end-users.

Without any other source of distortion in the functioning of the market, the best that could be done is to follow a *marginal rule*: the price to the final user should be set equal to the marginal cost of production. The resulting allocation is said to achieve what is known as the *first best*.[10] The access price should be set equal to the marginal cost of production ($a = c_0$) and since the downstream sector is perfectly competitive, the consumer price would be $p = c_0 + c$.

b. What happens when the theoretical first best cannot be achieved?

By setting the access charge equal to the marginal cost of producing the input, the incumbent would recover only the variable costs and would make a loss equal to the fixed cost F If this loss cannot be paid for by direct subsidies by the government, theory generally moves on to the next reference situation (*second best*). It seeks to set the efficient access price subject to a budget constraint (i.e. the need to avoid losses) faced by the firms. This means that all the firms must find the economic resources to sustain their activities in the market under consideration. If competitive firms are present in the downstream sector, the downstream firms do not make losses. The only firm that

[10] As the terminology suggests, there is no other way to improve the situation for society as a whole. The optimal situation is achieved when the "marginal" benefit to society (consumers' valuation) of one more unit of production is exactly identical to the marginal cost to society (firms' production costs for that extra unit).

remains to be considered by a regulator is the upstream monopolist. Since the total quantity supplied in equilibrium by the downstream firms sector depends on the access charge ($Q^S(a)$ denotes this quantity), the incumbent monopoly can break even only if *the access price recovers fixed costs on average*: $a = c_0 + F/Q^S(a)$. [11] This is an average pricing rule. Since the quantity supplied has to be equal by the quantity demanded by the users, denoted as $Q^D(p)$, and that $p = a + c$, we finally get the following implicit formula for the access charge:

$$\mathbf{(1)} \qquad \boldsymbol{a = c_0 + F/Q^D(a + c).}$$

c. What happens when different downstream users value the final goods differently?

In practice, the downstream operators faced by regulators often supply different kinds of final service. An electricity distribution company has residential, commercial or industrial clients for instance. The same Watts of energy can be demanded by any of these users but the way they value energy (their reservation values) can be very different. Similarly, access to the local loop in telecommunications networks is required both to complete a call originated by a mobile operator and to have access to the Internet using a dial-up connection. Even if the physical characteristics of the final good are the same, the economic valuation of final users can be very different.

When users' valuations differ, economic theory suggests that different prices should be charged to final users in order to recover fixed costs. This (second-best) strategy would follow what are known as *Ramsey prices*. The problem is still to achieve the best allocation, but in the absence of government transfers, the firm must be able to break even using only revenues from users. The idea is to miminize the distortions from marginal cost pricing while allowing the incumbent firm just to break even. . The rules shows that the percentage mark-up over the marginal cost (i.e. the distortion) is inversely related to the elasticity of demand. This means that if one has to set the price higher than marginal costs in order to cover fixed costs, theory suggests that it is better to do so with those consumers not very responsive to price changes (those with a low price elasticity of demand) . In fact, even if they end up paying more than in the first best, they will not change their consumption levels in a dramatic way . On the other hand, if prices are increased for users that are very responsive to price changes (with high price demand elasticity), even a mild change in price would distort consumption away from the ideal consumption.

This notion can be applied to access prices as well. When the downstream sector is competitive, downstream firms are simply like middlemen between the bottleneck monopolist and the final users. The access charge to the firm selling to users with rigid demand should therefore be higher than the access charge paid by another firm selling to consumers that are more price-sensitive.

While Ramsey charges are attractive conceptually, they are not always easily politically or legally easy to implement. First, there could be distributional concerns. Users with inelastic demand should pay more only if this is acceptable also after taking into account equity considerations. Many of the infrastructure sectors deliver services that are vital: it is sufficient to think of water to see that demand could be extremely rigid. According to the Ramsey rule, prices should be set very high, which is clearly unacceptable among poor people. However, this is not a substantive objection to Ramsey charges. They can be accommodated to take into account the distributional content of a good,

[11] It may be worth to point out here for practitioners that they must recognise that the cost of capital matters and it must be included in the allowed cost items since they are not abnormal profit.

and in that case they would simply take a more complicated expression.[12] Second, Ramsey prices are not always legally feasible. Indeed, discriminating among different downstream firms according to the elasticities of their final customers can, in some countries, raise problems according with anti-trust legislation. Even more fundamentally, in order to implement Ramsey prices, a great deal of information is required. The regulator should know the cost of the regulated firm and also the different elasticities of demand of final users. This kind of information is more likely to be in the regulated firm's hands rather than the regulator's. Hence there are additional constraints arising from asymmetric information between the regulators and its regulated industries (more on this later).

2.2 Identifying and dealing with trade-offs

Even with a clear structural separation of the various activities of an industry, there are trade-offs between at least two types of efficiency: allocative efficiency (the best product mix for society for a given level of resources scarcity) and productive efficiency (the cheapest cost, for a given output mix). Since access prices are an integral part of the cost structure of downstream firms, wrong regulatory decisions can allow regulated firms to reduce the direct linkages between final prices and cost structures, influencing the two types of efficiency differently . Concerns arise mainly in two contexts that need to be tested by regulators:

- when the degree of effective competition in the downstream market is restricted by one players,
- when the level of entry costs in the downstream market is high in comparison to the potential benefits of additional entry.

…and the optimal regulatory decision to adress these two types of concern will often drive the access charge in opposite directions.

First, a firm that has some market power in the downstream market is able to add a mark up to its own marginal cost. This is a common concern in the telecom industry for instance where incumbents have tended to have some effective degree of market power in particular among residential users. This type of imperfect competition leads to the so-called *double marginalisation problem*. If the access charge already includes a mark up on the marginal cost of access provision, the final effect on retail prices will be further magnified.[13] If it is true that high mark ups are added to own costs, the regulator may consider that a way to deliver final prices that are not excessively inflated is to *decrease the cost of inputs, including access prices*: low access prices are beneficial to end-users since they reduce the underlying structure of costs. The problem of recovering the bottleneck fixed costs remains, however, so that network access can be subsidised only if other means are available to repay the incumbent. This is where trade-offs become important for the incumbent as well as for the regulator as discussed next (since how much to cross-subsidise and from which sources to cross-subsidise become important for any regulated industry concerned with market power). This may mean that allocative efficiency is not necessarily the ideal one.

Second, imperfect competition and market power in the downstream sector can arise from technology constraints. Downstream firms may have to incur fixed costs that are non-recoverable (set up costs). Since each downstream firm has to pay an entry cost that is sunk, the concern arises that

---

[12] The tendency in developed countries is now to separate the questions of equity and efficiency that should be handled respectively by a fiscal authority and by an industry regulator. While in this paper we abstract from equity considerations, it is clear that one cannot forget about the impact that restructuring tariffs would have on those belonging to the lowest tail of income distribution. But this is a wider debate that needs to consider a wider range of instruments.

[13] Consider a good that costs $1 per unit and it is sold to a wholesale seller that sells it to a retail seller. If every firm adds a mark up of 20% on its own costs, then the producer would sell it for $1.20 to the wholesale seller. The latter would ask $1.44 to the retailer that finally would charge $1.73 to the final users, i.e. a total 73% mark up on the cost of production.

too many firms may enter. The gains that are brought by entry have to be related to the increase in competition (lower final prices) but they have to be traded against the cost of entry, i.e. the unrecoverable cost. The latter expenses could well be used in a more useful way elsewhere in the economy. As a result, when downstream firms also incur set up costs, the regulator may decide that a *higher access price* than suggested by average cost pricing can be instrumental to reduce the waste of duplicated resources that produce homogeneous goods. On the other hand, if entry brings about the benefit of product variety (heterogeneous final goods) the previous argument on wasteful duplication is diluted.

2.3 Summing up the main lessons for regulators of unbundled sectors

The main lessons from theory for regulators of network industries that have benefited from unbundling and are separated into bottleneck and competitive activities are summarised in Table 1. The main point for regulators is that if government subsidies are not available and the regulator still wants to follow a marginal cost pricing rule, the likelihood of underinvestment in the upstream sector is high since fixed cost would not be recovered by private operators and these would go bankrupt. If subsidies to the incumbent are not possible, an alternative is to ensure that unit access charges are at least equal to the average cost of the bottleneck owner. Another alternative available when different services are produced with the essential input (and when distributional considerations are ignored), is to allow access charges to follow an inverse elasticity rule in which the more a good is needed by a downstream user, the higher the access charges that the bottleneck owner should be allowed to levy from that specific user.

The previous rules change when more imperfections are brought into the picture. With imperfect downstream competition or downstream entry costs, a compromise has to be found between allocative and productive efficiency. In general this means that the regulator will have to decide whether to undercut or overshoot the optimal access price to address a specific source of market imperfection. Market power brings two kinds of distortions: mark ups above marginal cost, and, in the presence of entry costs, cost duplication. The latter happens because too many firms, attracted by potentially extra-normal profit, end up entering a market. The principles are the same as before. Access should be subsidised to offset mark ups (but then alternative instruments have to be used to repay eventual bottleneck fixed costs), the opposite should happen if entry is not socially desirable.

Table 1. Access charges with structural separation

	Access charge	Eventual problems
First best	**Marginal cost**	***F* not recovered in the absence of transfers**
Second best (one final good)	**Average cost**	
Second best (many final goods)	**Ramsey charges**	
Imperfect downstream competition	**Lower**	***F* not recovered in the absence of transfers**
Downstream entry costs	**Increase**	**Not true if entry brings product variety**

3. Access rules for vertically integrated industries

While voluntary provision of access can be a very realistic scenario in the presence of a vertically integrated incumbent, (see Box 2), more often than not, the access pricing issue is quite serious in the

presence of vertically integrated industries. This section considers the common case in which after competition is allowed in the competitive segments of an industry, such as electricity distribution, the vertically integrated incumbent is allowed to compete in the final goods market against entrants that need access to a vital bottleneck owned by the incumbent. Another frequent case occurs in the railway industry when shipper who own a railroad are often tempted to use this ownership to place their competitotrs at a disadvantage.

Box 2: Voluntary access provision can work if the rules are right

Consider a transition case in which an incumbent who, before deregulation, was the sole supplier of the final service. After deregulation, it remains the bottleneck owner but can also supply the final service through an independent subsidiary that pays the same access charge as entrants do. Theory as well as intuition suggest that deregulation will be opposed by incumbent firms when it reduces the pre-deregulation "quasi-rents" such as monopoly extra-profits, the recovery of fixed costs, compensations for fulfilling universal service obligations, etc... Once the incumbent is required to provide access to its essential facilities, the process of assessing access charges can become good news for the incumbent. If the access charges are not compensating the incumbent enough, an uncertain transitory phase is likely, where the incumbent would adopt various strategies (legal battles, lobbying, bribes, etc.) to avoid or delay the changes. In this situation, the regulator needs to assess the private decision of the incumbent of giving access to his own bottleneck.

Spulber and Sidak (1997) give some useful advise to practitioners and explain how to set access charges, taking into account that the incumbent should at least preserve its previous quasi-rents. Their analysis thus involves the imposition of an additional constraint, the voluntary access provision. The exact level of the regulated access charges depends on the details of the industry under consideration but the most interesting result is that it is possible to find access charges such that the incumbent would not oppose entry, and such entry would happen only if firms are more efficient than the incumbent. Access charges can both provide incentives to entry and be such that efficiency gains are shared with the incumbent that, as a consequence, is not reluctant to entry. The result does not depend on the intensity of price competition or on the degree of differentiation among products.

The fact that the bottleneck owner is allowed to compete against other firms means that there is a danger that the incumbent will set access charges which make further entry difficult. In a limiting situation, the integrated incumbent may even deny access on reasonable terms. This reasoning may suggest that the access price should be set low, in order to contrast the anti-competitive attitude of the incumbent. However, if access is set too low, inefficient entry may occur. Moreover, if fixed costs are involved in the bottleneck, the regulator should ask how much the entrants should contribute to repay the fixed cost of a service that they use in order to supply their customers. Since vertical integration with liberalisation is the quintessential structure for the analysis of the access pricing problem, we will go into the details of several relevant cases.

3.1 Regulated final products

Box 3 contains the notation used in a stylised model that will be developed in different steps in the remaining parts of this section. We discuss here the benchmark situation with a "benevolent" regulator quite familiar with the cost structure of service providers as well as their efforts levels to minimize costs. This regulator fixes *all the prices* to maximise an unweighted sum of consumer well being and total industry profits, subject to a break-even constraint for the incumbent ("a second best situation"). The fact that the regulator is "benevolent" means that he is trying to balance the interest of the various parties involved, firms and consumers, without favouring anyone in particular. In a nutshell, the optimal theoretical access charge in that situation is given by equation (2):

$$\textbf{(2)} \qquad \boldsymbol{a = \text{direct cost} + \text{modified Ramsey term}}$$

Box 3: Vertical integration with liberalisation – A stylized model*

Suppose there are two related industries, one upstream (the bottleneck) and one downstream (the potentially competitive sector). The bottleneck is fully controlled by an incumbent firm denoted by *I*. In the downstream market, the incumbent faces competition from an entrant denoted by *E* which needs access to the bottleneck input in order to produce the final good. In principle, the bottleneck input can be sold directly to the public (this is the case, for instance, for local calls in telecommunications), this means that there are three final markets:

- the bottleneck, denoted by 0;
- the downstream good produced by the incumbent, denoted by 1;
- the downstream good produced by the entrant, denoted by 2.

All activities exhibit constant returns to scale, except from a fixed cost *F* which is incurred in the bottleneck. Denote by c_i the marginal cost of production in the final market *i*, q_i the total quantity supplied to end users, and p_i the corresponding price. Finally, one unit of the bottleneck is needed to produce one unit of a final good, and *a* is the access charge paid by the competitor to the monopolist. Given this notation, the total profit of the incumbent and of the entrant are respectively:

$$\pi_I = p_0 q_0 + p_1 q_1 + a q_2 - [F + c_0(q_0 + q_1 + q_2) + c_1 q_1],$$
$$\pi_E = p_2 q_2 - (c_2 + a) q_2 .$$

* For a more technical analysis refer to Armstrong *et al*. (1996) and Laffont and Tirole (1994 and 1996).

The first thing to understand is that the optimal access charge is derived together with the prices of final goods. Every customer participates to the recovery of fixed costs. In order to reduce distortions, customers that are not very price sensitive are required to contribute more to such recovery. As a result, mark ups over marginal costs are higher in "inelastic" segments. In the case of telecoms for example, the bottleneck (local calls) can be sold directly to final users. Hence the charge for local calls also ends up with a mark up. This means that the recovery of fixed costs comes from three sources: two sources directly accrue to the incumbent through his subscribers (imagine local and long-distance users), the third source comes from the entrant's users and it is passed to the incumbent via the access charge.

Since markets are related, it is quite intuitive to find similar relationships among optimal prices. Optimal charges derive both from demand and supply considerations, and this is what the formula is saying. Ramsey prices put a strong emphasis on *interdependencies* between markets and on the *simultaneity* of optimal price setting. The intuition for the result is as follows. Imagine that the customers of the entrant firm are not particularly price responsive. Then they can be charged a high price. Part of such high price is passed on to the incumbent via a high access charge. This is in the interest of society as a whole because it allows other prices to be reduced without violating the incumbent's budget constraint. If consumers of the product supplied by the incumbent are price sensitive, they are charged a low final price, so that also their consumption is not particularly distorted. Box 4 provides a more formal presentation of the derivation of the optimal Ramsey prices.

Ramsey charges are based both on cost and demand conditions, however in practice many regulators do not allow usage-dependent pricing for the fear that the incumbent could engage in anti-competitive practices. Such a fear is not always reasonable. It is not entirely clear, for instance, why an incumbent fixed telecom operator should have to offer mobile operators the same conditions it offers to other operators when they terminate their calls on its network. If demand to and from mobile users is not very price sensitive, the incumbent could cover a higher portion of its fixed costs on that segment via higher access charges, thus allowing for price reductions in more price responsive segments. As long as the regulator can ensure that market power is not abused (e.g. by setting excessive access charges that exclude entrants), economic theory calls for different charges whenever demand elasticities are different.

Box 4: Derivation of Ramsey charges-How complex are they?

Suppose a benevolent regulator fixes *all the prices* in order to maximise an unweighted sum of consumer well being and total industry profits, subject to a break-even constraint for the incumbent. Optimal *Ramsey prices* result:

$$\frac{p_0 - c_0}{p_0} = \frac{\lambda}{1+\lambda}\frac{1}{\hat{\eta}_0}$$

(1)
$$\frac{p_1 - c_0 - c_1}{p_1} = \frac{\lambda}{1+\lambda}\frac{1}{\hat{\eta}_1}$$

$$\frac{p_2 - c_0 - c_2}{p_2} = \frac{\lambda}{1+\lambda}\frac{1}{\hat{\eta}_2}$$

where λ is the implicit value to the economy (the shadow price) of the budget constraint (which only matters when there is a deficit) and $\hat{\eta}_i$ is the price elasticity of demand in final market *i* modified to account for substitution possibilities among goods. $\hat{\eta}_i$ are usually called "superelasticities" in the literature and they reflect the global impact of a change in the price of a good on the incumbent's profit. In practice, if demands are independent, they are equivalent to normal elasticities (this is the likely case for the bottleneck). If goods are substitutes, as it is likely between final products 1 and 2, they are smaller than ordinary elasticities ($\hat{\eta}_i \leq \eta_i = |\frac{\partial q_i / q_i}{\partial p_i / p_i}|$).

The interpretation is very close to ordinary Ramsey prices: given that the first best is not attainable because fixed costs have to be recovered, it is necessary to introduce some distortions. Allocation is distorted the least if mark ups are high in those markets where users are not particularly price sensitive. A price higher than marginal costs allows cost recovery without affecting too much the consumer behaviour. The entrant is a price-taker so he will produce according to a marginal rule (the price p_2 received on the additional unit produced, has to equal the cost of the additional unit amounting to $a + c_2$). This allows to determine the optimal access price:

(2)
$$a = p_2 - c_2 = c_0 + \frac{\lambda}{1+\lambda}\frac{p_2}{\hat{\eta}_2}$$

In order for Ramsey charges to be calculated, the regulator needs to know:

- the marginal costs;
- the elasticity of demand and the substitutability among products;
- the value of λ that can also be interpreted as the cost of public funds associated with the distortionary nature of taxation. In developed countries it is estimated around 0.3, and it is higher the less efficient the fiscal system is.

This formula also gives, as a particular case, the solution when charges are chosen by an unregulated upstream provider ($\lambda \to \infty$, so that the whole coefficient $\lambda / (1+\lambda)$ tends to 1). This means that an unregulated firm would restrict output too much and charge excessive high prices both at the intermediate and final level. This simple observation is the main reason for access control.

To sum up, it is important to keep in mind that in the kind of industry structure discussed here, access is priced above marginal cost *not* because the incumbent exerts monopoly power (in particular if the regulators understands its cost structure quite well) but because deficits are socially costly and the charge performs as a tax used to raise money that repays the deficit. The charge is particularly high when it does not distort too much the allocation in the final market (the elasticity of the entrant's customers is low) or when the budget balance is particularly severe (equivalently, the social cost of public funds is high). By increasing *a* beyond its marginal cost of production, the retail prices can be reduced, which is in the interest of final users.

Note, however, that there are a couple of common instances in which increasing access charges to pay for deficits has alternatives and before starting to "cheat" on the access charge, it is worth to investigate the nature of the fixed costs to be recovered. If they are not pure bottleneck costs and they stem from other types of restraints placed on the incumbent, a removal of the constraint itself could

ease the regulatory task. One such case in point is the need for the incumbent to recover costs associated to social obligations. Uniform nation-wide pricing requirements, low connection fees to residential users, or universal service obligations, impose a real burden that has to be recovered elsewhere. If this is the origin of the access deficit, the first question to ask is whether one should remove the cause. Sometimes the answer is simple: universal service funds are feasible, without having to distort the access charge.

A second more complex common situation arises with tariff deaveraging. Because of the history inherited from the past, in many cases tariffs are not in line with costs and that the business sector is subsidising the residential one. This is a source of trouble during a liberalisation process, because the entrants try to attract only the most attractive users and "cream skimming" exacerbates the access deficit of the incumbent. Imagine an incumbent that breaks even over two segments: some extra-normal profit is made in one segment and it exactly offsets the losses in another segment. Imagine that the incumbent is regulated and cannot change his prices because of political constraints. Then the wedge between price and cost in the most profitable market can be exploited by an entrant that skims between the two segments. In the limit, entry can also be inefficient and it is just caused by regulatory constraints. If entry occurs and the incumbent is not allowed to change the prices in the subsidised segment, the incumbent could possibly go bankrupt. On the other hand, if the incumbent was exploiting some excessive market power that was not sufficiently regulated, entry would be beneficial to bring prices down. If it is acceptable to rebalance tariffs, the deficit may disappear. Only when this is not feasible, for instance due to very good reasons such as equity considerations, or abrupt changes that are unacceptable for political reasons, we go back to feasible second best as discussed next.

3.2 Narrowing the focus of regulation: the Efficient Component Pricing Rule

One creative second best solution commonly considered in regulatory circles arises if the regulation of access price is separated from users' prices. Supposing that the final product prices are already fixed, then access price has no effect on allocative efficiency. The regulator may still be concerned with cost recovery and productive efficiency, that is to say with efficient entry and cost minimization. The pricing policy that concentrates *only* on productive efficiency is the popular and controversial *ECPR* (Efficient Component Pricing Rule) also known as the Baumol-Willig rule, the imputation rule, the margin rule, or the parity-pricing formula.[14] The rule is very simple and states that when final products are homogeneous ($p_1 = p_2 = p$) and the market is contestable[15], the access charge should be equal to the difference between the final price and the marginal cost on the competitive segment:

$$\mathbf{(3)} \qquad a = p - c_1 = c_1 + (p - c_0 - c_1)$$

ECPR can be read in many equivalent ways:

- As a margin rule, it says that the margin of the incumbent in the final market (p - a) should be equal to its marginal cost in the downstream activity (c_1).

14 The rule was originally introduced by Willig (1979) and Baumol (1983). More recently, it has been popularised by Baumol and Sidak (1994). For a critical view, see Kahn and Taylor (1994), Economides and White (1995), and the response by Baumol *et al.* (1997).

15 In a contestable market there is free entry of firms and there are no unrecoverable sunk costs, so that firms are allowed to "hit-and-run".

- As a parity principle, the bottleneck owner imputes itself for the bottleneck input the same price at which entrants buy the input.[16]
- From the point of view of a potential entrant, he would find it profitable to enter only if it is viable, that is if total unit cost is less or equal than the final price: $p \geq a + c_E \Rightarrow c_E \leq c_1$. In this respect, the rule sends the right signal to new entrants. Entry is profitable only for those firms that are more efficient than the incumbent in the downstream activity.
- Alternatively, the rule says that the access charge should be equal to the direct cost of providing access (c_0) + the opportunity cost of providing access. The opportunity cost is (p - c_0 - c_1) since this is the reduction in the incumbent's profit caused by the provision of access. Entry does not alter the bottleneck cost recovery since the rule is neutral for the incumbent's revenue.

In words, ECPR gives this basic message:

(3') ***a = direct cost + opportunity cost***

The simplicity of the formula explains in part its popularity. Revenue neutrality for the incumbent, on the other hand, is also the criticism made by opponents: *if* the incumbent is earning supernormal profits, they will continue to be earned also in presence of potential entrants. In this respect, the rule guarantees monopoly rents! However, the observation is not completely appropriate because ECPR *assumes* that final prices are optimally set. More serious is the criticism that ECPR becomes irrelevant precisely in the proper context developed by its proponents. In fact, if the entrant is more efficient than the incumbent, all downstream production is delegated to the entrant: the incumbent disappears from the downstream sector, the industry becomes in practice vertically separated and the incumbent's regulated price is irrelevant. If the entrant is less efficient, he will never produce so that the industry is in practice fully integrated and the access price becomes irrelevant.

Recently, some authors have argued in favour of an adaptation of ECPR. In particular, Sidak and Spulber (1997) have supported the idea of a "market determined" ECPR (M-ECPR). The difference between ECPR and M-ECPR arises when the entrant can provide the final good at a lower price than the incumbent. M-ECPR sets the access charge equal to the difference between the entrant's final price and the marginal cost on the competitive segment:

(4) $\boldsymbol{a = p_2 - c_1}$

It is clear that M-ECPR delivers a different outcome than ECPR when prices do not remain constant after deregulation. If the original price of the incumbent was set at a very high level, M-ECPR would reduce part of the initial distortion. In this respect, M-ECPR corrects in part the fact that ECPR does not induce competition in the final market when the incumbents enjoys extra-rents. However, most of the criticisms made to ECPR do apply also to M-ECPR.

ECPR continues to be central in the debate on access pricing both in theory and in practice.[17] Its main contribution relates to the idea that the provision of access causes not only an incremental direct cost, but also an "opportunity" cost that is particularly realistic when entry occurs in the more

[16] Note that imputation, i.e. the bottleneck division of the incumbent charging the same charge to its downstream division and to the entrant, can be monitored only if there is proper accounting separation of the two incumbent's division.

[17] ECPR was central in a famous recent litigation between the Telecom Corporation of New Zealand and its rival Clear Communications.

profitable segments of a market. The following sections try to compare ECPR with the results obtained using the previous benchmark model. We will then allow for extensions in order to qualify the notion of "opportunity" cost under a wide range of circumstances.

3.3 Ramsey Access Pricing and ECPR compared

Both the Ramsey formula (2) and ECPR (3) imply that the access charge should be set higher than the direct marginal cost of providing access. How do the two formulations compare with each other? Under rather general conditions, the inverse elasticity rule can be manipulated to get $a > p - c_1$. The optimal access price is *higher* than that prescribed by ECPR, since this allows to keep the balance of the budget while reducing the final user price. Note that this is not due to losses in the local market ($p_0 > c_0$), rather to the correction of distortions in the final market caused by deviations from marginal cost.[18]

Box 5: A formal comparison of Ramsey Charges and ECPR

From eq. (2) and substituting p_2 from eq. (1) in Box 3, the Ramsey access charge can be rewritten as:

$$a = \frac{c_0 + k_2 c_2}{1 - k_2} \text{ where } k_2 = \frac{\lambda}{1+\lambda}\frac{1}{\hat{\eta}_2}$$

On the other hand, if we apply ECPR as given by $a = p_1 - c_1$, we get:

$$a = \frac{c_0 + k_1 c_1}{1 - k_1} \text{ where } k_1 = \frac{\lambda}{1+\lambda}\frac{1}{\hat{\eta}_1}$$

The two formulae coincide only if $c_1 = c_2$ and if $k_1 = k_2$, i.e. when entrant and incumbent have the same cost structure and the two final products are perfect substitutes for users (i.e. goods are homogeneous).

3.4 What happens when firms produce multiple products and have access alternatives?

The optimal Ramsey formulae can be generalised to allow for more complex cases. In particular, entrants may offer new product varieties, so that consumer choice increases. Entrants may also be able to supply the bottleneck themselves, though using less efficient technologies. For instance, in the telecommunications industry, the set of entry strategies of actual firms is quite rich, ranging from complete facilities bypass to complete service resale.[19] The optimal Ramsey formula is then modified to the following:

---

[18] Equations (2) and (3) mainly differ for the presence of the Ramsey term. ECPR seems to be more cost-based since it does not require knowledge of demand elasticities. This is only apparently so, because the contestability framework in which ECPR is embedded assumes that at the point of entry the entrant's elasticity of supply is infinite, so that the Ramsey term vanishes. Box 5 shows the conditions under which Ramsey and ECPR provide the same result are met in practice only by chance. In the case there is no deficit to recover ($\lambda = 0$), the two formulations are identical, and marginal pricing rules allow to reach the first best. When perfect competition is workable, extra rents are driven to 0, and the opportunity cost of entry is also zero. However, this situation cannot arise in the presence of increasing returns to scale, which is the standard case when bottlenecks exist.

[19] An example of complete service resale is a service provider that relies fully on other firms' infrastructure to supply value-added services to clients. When new facilities are built, it is common to distinguish between bypass and network duplication. The former relates to large users that decide to provide the bottleneck themselves, for instance in the telecoms industry large users may establish a radio link to connect with long-distance operators without requiring interconnection with the incumbent in the local loop. Network duplication occurs when a bottleneck is supplied by a competitor, rather than by customers.

$$(5) \qquad a = c_0 + \sigma(p_1 - c_0 - c_1) + \textit{Ramsey Term}$$

The expression says that the optimal access charge can still be thought in terms of ECPR. On top of the direct cost of providing access, there is a term corresponding to the opportunity cost to the incumbent. Such term is multiplied by a factor σ called the "displacement ratio". The displacement ratio determines how much sales the incumbent firm loses as a result of supplying access to its rivals.[20] Finally the formula also involves the usual Ramsey term that takes into account the differentiated demand response to price changes.

The displacement ratio is generally *less* than 1, according to the degree of product differentiation, bypass opportunities and technological substitution. In all cases, the opportunity cost of supplying access to rivals is reduced. This is because there is not a one-for-one displacement of the incumbent's sales. Under product differentiation, each unit supplied by the rival does not cause a unit reduction in demand for the incumbent. When there is no substitution between goods, there is no loss of demand for the incumbent ($\sigma = 0$).

The access charge is also set lower than the standard case to prevent wasteful bypass. If *a* is too high, entrants would be given incentives to build inefficient facilities in the bottleneck. Remember that one of the main roles of *a* is to discourage inefficient entry downstream. Hence a new trade-off arises between low charges that may attract inefficient entry in the competitive market and high charges that attract upstream entry and duplication of access facilities. This trade-off can be eliminated in part if *quantity discounts* are allowed in the pricing of access. Since large buyers of access are also those who are most likely to bypass the bottleneck, they would find bypass less convenient if the access tariff is decreasing at the margin. At the same time, the access price for small volumes would be sufficiently high to discourage the entrance of inefficient firms. However, a potential danger should also be pointed out. By favouring bigger firms, access discounts may introduce artificial increasing returns to scale in the downstream sector.

In summary, bypass considerations lead to a strong case for allowing quantity discounts (or menu of two-part tariffs) when they enable greater allocative efficiency, but it should be checked that they are not a cause for less competitive industry structures.[21] This is what electricity regulators often have to look at when large users are allowed to compete with distribution companies on the generation markets. An additional by-product of this competition is that small users (mostly residential) end up having to pay for a larger proportion of the access charge which would otherwise be divided among all users, including large users.

---

[20] In more formal terms, it measures the change in the incumbent's sales of final good divided by the change in the incumbent's sales of input to entrants as *a* changes. The displacement ratio can be decomposed in 3 multiplicative terms: $\sigma = \sigma_d \sigma_t \sigma_b$ where σ_d is the effect of substitutability between final goods for final consumers ($\sigma_d = 1$ only when the incumbent and the entrant offer identical, perfectly substitutable products); σ_t represents the technological possibility for the incumbent to change the input mix keeping output constant when the access charge is increased ($\sigma_t = 1$ when there is no possibility of such substitution in the input mix); finally σ_b represents the entrant's ability to bypass the bottleneck ($\sigma_b = 1$ when there are no opportunities for bypass).

[21] Quantity discounts (optional tariffs) can also occur in the pricing of final goods. On the one hand, quantity discounts are likely to promote allocative efficiency because customers can select from a richer menu of tariffs, on the other hand they can be used by the incumbent to target selected categories of users only, with a predatory intent that damages new entrants. A reasonable balance between these two opposite tensions has been adopted in the UK telecommunications industry where the regulator has decided that quantity discounts are a legitimate response of the incumbent in more competitive environments (subject to ordinary competition law rules), but they are not counted toward the calculation of the price-cap index.

3.5 What happens when the regulator cannot fully observe the cost structure?

The analysis has concentrated so far on the effects that access prices have on the final consumption of goods and on the incumbent's budget. This section considers how problems of moral hazard and adverse selection affect access charges.[22] The regulator is unable to monitor the firm's effort to reduce costs and has less information than the firm about its technological efficiency. The literature on regulation in the presence of asymmetric information points at a fundamental trade-off between incentive provision and rent extraction. In general, effort is induced by fixed-price contracts which make the firm residual claimant for its cost savings. On the other hand, extra-rents are better limited by cost-plus contracts. In our context, the optimal formula for the charge modifies to:

(6) ***a = direct cost + modified Ramsey term + incentive correction term***

The implementation of the latter term is done offering a menu of incentive contracts. The reasoning is that different firms with different cost characteristics, should be induced to freely choose different contracts, in particular high-powered schemes (i.e. schemes with strong incentives to maximise efforts to cut costs such as price caps) should be picked up by low-cost producers. The regulator does not want to distort the more efficient firms in their decisions, even if this decision may leave rents to them. This is the motive for not having the incentive correction term for the most efficient firm. The less efficient firms, on the other hand, should voluntarily select a contract that gives them an incentive not to produce too much and, at the same time, does not allow for excessive extra-profits (this is what rate of return regulation does).[23]

a. What happens when costs separation is possible but effort can't be observed?

Imagine first that the costs of the various activities, c_0 and c_1, are observed separately (for instance it is possible to distinguish between the transmission and generation costs in an electricity network), but the regulator cannot distinguish whether the level of each cost depends on cost-reducing effort or "luck". Incentive considerations ask for a general reduction of quantities of bottleneck inputs. Because of increasing returns to scale, this means that *the access charge is increased* for the competitor but simultaneously the incumbent is penalised with high prices for its own product. Because access costs are the same for the incumbent's own product and the entrant, the regulated incumbent cannot claim simultaneously an expensive production of the intermediate good (a high c_0 that damages rivals) and high efficiency for its own product that would leave a rent. If an issue is a potential anti-competitive behaviour of the incumbent trying to foreclose rivals, the

---

22 In actual transactions, one or more of the parties involved may have relevant private information that can be exploited. *Adverse selection* refers to problems of pre-contractual opportunism. The term was coined in the insurance industry to deliver the idea that the selection of people who purchase insurance is not a random sample of the entire population, but rather a group of people who know their personal situations. Even if there is no private information when an agreement is made, there can still be the possibility of self-interested misbehaviour. The latter is called *moral hazard* and it is a form of post-contractual opportunism that arises because it is not possible to observe perfectly actions that have consequences for all the parties involved. Also this term originated in the in the insurance industry where it referred to the tendency of insured people to change their behaviour in a way that could lead to claims against the insurance firm.

23 Under some conditions, pricing and incentives are decoupled. In such case, price formulae are the same as the previous sections with perfect information, while incentives are provided by the cost reimbursement rule. Without going into the details, we just remark that the incentive correction term depends on technology. More precisely it depends on the rate at which the firm must substitute effort and productivity to keep the same level of cost.

authority is able here to infer anti-competitive practices by comparing high access charges with low downstream prices.

b. What happens when neither costs, nor effort can be observed?

The case in which the regulator cannot separate the costs of the upstream and downstream activities, is also known the "sub-cost observability" problem. It is relevant when new facilities are built by the incumbent to provide access for entrants (and this is the case in many developing countries infrastructure sectors). Here, the incumbent can claim to have a high c_0 without having to punish itself in the final product market. There is a real danger that the incumbent firm succeeds in squeezing rivals out of the market, still enjoying extra-profits on its downstream activities.

This discussion leads to the issue of the incumbent's discretion over how fixed costs are allocated to the various services. When discretion is limited, there is good case in favour of using regulatory tools requiring the incumbent firm to report separate accounts for its bottleneck and competitive activities. This is the so-called *accounting separation*. Using accounting data, the incumbent has to show that the same access price is charged to rivals and to its own downstream division. In principle, it is a very good idea: the benefits of integration are preserved and the scope for anti-competitive behaviour is reduced. Separation provides further benefits since it makes costs more transparent to everybody and it improves the information flows to the regulator. When competitors use technologies similar to the incumbent's in the competitive segment, there may also be additional gains from "yardstick" competition (the regulator can use the entrants' performance as a benchmark to control the incumbent's efficiency).

Sub-cost observability, however, warns that accounting separation without controllability performs very badly and may not be an effective regulatory tool. The ownership structure affects the incentives to engage in accounting manipulations and to claim higher than true costs, in particular manipulations are more likely to happen under vertical integration. In other words, accounting separation and structural separation are not equivalent and the latter may still be preferred under many occasions. When manipulations are difficult to detect, the simple use of ECPR is also very dangerous: it gives the incumbent an incentive to exaggerate as much as possible the importance of opportunity costs, shifting costs from the downstream to the upstream section of the firm's accounts. In this context, there are also some relevant dynamic considerations discussed in section 3.7.

In summary, when regulators have imperfect information on costs and on efforts to cut these costs by the service providers, extra rents have to be left to the incumbent for incentive reasons and such rents depend positively on the incumbent's output. The access price has a direct impact on entrants' production (a high access charge reduces output) and, indirectly, on the incumbent's (the incumbent increases output to react to a reduction in the competitors' production). Hence incentives should depend on cost characteristics and output produced by *all* firms. When the same essential facilities are used by all firms, the incentive reward should decrease with the level of access price (increase with rivals' output). If cost-padding is a problem, low-powered incentive schemes should be preferred otherwise the access costs would be inflated too much. Accounting separation deserves its popularity when manipulations between divisions can be detected, if not a call for structural divestiture of the incumbent firm rather than separate accounts seems more appropriate.

3.6 What happens when investment in bottlenecks activities matters?

In the previous sections, we emphasised the role of access charges when the issue is to compensate the incumbent for the use of its infrastructures. But very often the access problem is mainly related to the need to design the access charges as a signalling instrument both for the incumbent and the entrant with respect to their investment decisions in situations in which one of the parties may have some degree of exclusivity over the implementation of the investment decision. Say for instance, electricity transmission is the exclusive responsibility of an integrated monopoly with competitors both upstream and downstream of the bottleneck facility. Who should be paying for the cost of expanding transmission capacity and how? A similar situation could arise in cases in which a single company has a monopoly over some rail infrastructure.

a. The incumbent's viewpoint

Incumbents may be faced with two types of situations. First, their investments made in the past (before deregulation) are not still fully amortised at the time of deregulation and access prices have to be designed to reflect the residual financing gap. The problem results from the fact that regulated incumbents have made investments approved by the regulator in the past under the assumption that all costs would be recovered and when competitive entry was not fully anticipated. This is the problem of *stranded costs*. How much compensation should the incumbent receive, if any? How should those costs be apportioned between the incumbent's customers and entrants who need access to the bottleneck?

The answer to the previous questions is very delicate. If regulators are poorly informed about the investments' costs, the incumbent can deliberately overstate them or he can engage in wasteful practices (gold plating, i.e. the purchase of more expensive options among inputs when less expensive ones would suffice). When the incumbent has such an information advantage, regulatory approval should not, by itself, guarantee cost recovery. However, a call for low access charges is not entirely convincing. The problem is a form of regulatory failure, not a competition issue. Because these investments were approved by regulators, one cannot apply the argument that firms in competitive markets should bear the full risk of their investments. The regulatory contract should guarantee compensation for all prudentially incurred costs. The size of the access deficit to be recovered can only be determined once the terms of the contract governing the relationship between regulators and regulated firms are clearly specified. Box 6 presents an example that applies to a pre-deregulation phase, in anticipation of a more competitive phase in the future and shows when it may make sense to have the incumbent receive less than compensatory prices.

The second situation an incumbent may have to face is an even less pleasant one. Once relevant investments are sunk, there may be post-contractual opportunism on the part of competitors. The entrants do not want to contribute to the financing of the new investment and the incumbent ends up stuck with a larger than fair share of the cost. If anticipated by the incumbent and if the bottleneck has no other potential use (asset specificity), this risk deters initial and future investment. This is referred to as *the "hold-up" problem*.[24] This classic problem can be avoided by the regulator by ensuring that the access charges are not too low and putting this commitment on paper in a regulatory contract. The need to anticipate the hold up risk is particularly important during transition periods to competition

[24] One possible solution to opportunism when asset-specific investments are involved, points towards the typical advantage of integration over separation. Vertical integration is beneficial when it overcomes externalities between the parties involved: hold up problems and double mark ups can be avoided, there is better risk-sharing between upstream and downstream divisions. However, there are also flaws as long as an integrated firm is tempted to get competitive advantages by raising its rivals' costs.

since investments of monopolist operators in the past were probably made with the conviction that they would be recovered under a different (more protected) market structure.

Box 6: Multiperiod Ramsey pricing and stranded costs

How should fixed costs be recovered over time in the final prices charged to end customers and in the access charges paid by new entrants? Imagine there are two periods, the pre-deregulation period and the post-deregulation period. The incumbent firm has to invest in durable facilities. Entry may materialise in the second period with some probability. If entry happens, new firms have lower production costs than the incumbent, for instance because they discover a new technology. If entry occurs, efficiency requires that the entrants use the incumbent's facilities and completely displace the incumbent's production. However, the incumbent has to recover its initial investment. This means that regulated access and final prices are subject to a multiperiod break-even constraint for the incumbent. The break-even condition involves the expectation over the appearance of competitive alternatives. This problem is studied by Wildman (1997). He shows the solution is an application of standard Ramsey principles to a two-period situation. More interestingly, as long as there is a positive probability of entry, the optimal regulated price falls. The intuition goes as follow. If there was no prospect of entry, the regulated price would be set just in order to break even. On the other hand, if there is entry, the incumbent makes some profits because he benefits in part from the increased efficiency of the entrant. This extra-rent is not more than compensatory because it has to be contrasted with a lower regulated price when competitors do not materialise. The results show that the incumbent has a strong incentive to encourage access to competitors when they can contribute to efficiency enhancement (cost reductions). This depends on the fact that more efficient entrants completely displace the incumbent. A part from the robustness of the result, the most relevant message to be derived from the analysis above is that past costs can only be determined with respect to the incentive system in effect at the time.

Overall, long-term enforceable contracts would improve the situation by clarifying the commitments to be made by all parties involved-although they may create scope for collusion when enforceability is problematic. Clearly, the best solution would be the presence of a workable wholesale market. In the presence of a truly competitive downstream sector, the chance of being engaged in relation-specific investments decreases. The investor can find more than one party to deal with, hence there is less room for opportunistic behaviour so that investment is undertaken. Moreover, for those utilities like gas and electricity where risk and price volatility are of primary importance, liquid *spot and forward markets* allow investors to adopt positions to suit their needs.

b. The entrant's viewpoint

Next, the regulator needs to recognise that the entry decision is often not cost-less for entrants. To assess these costs properly, the regulator needs to be able to distinguish between sunk costs (i.e. infrastructure investments) and other barriers to entry, such as customers' cost of changing supplier, additional advertising for unknown products, or even the incumbent manipulating access quality. Entry assistance may be called for where consumers value the product diversity provided by new entrants, or to foster learning-by-doing. Low access charges can be used to guarantee that the entrant breaks even. However, whenever possible, other tools are better. Direct entry subsidies can do the job, but they are unlikely to be legal. In that case, inefficient entry barriers could be directly tackled. This is happening, for instance, in the telecommunications sector, when regulators pursue *equal access* (mechanism designed to reduce customer inertia) as well as *number portability* (the user keeps the number if he decides to change operator so that at least end-user switching costs are reduced).

Entrants also make assessment of the degree and distribution of risk in the business they are getting into. All access rules discussed so far are *usage based*. The place the burden of risk on the incumbent. For instance, if demand is lower than expected, it is the incumbent who ends up providing unused costly capacity. In some situations, the incumbent can shift some of the risk by levying

capacity based charges. These can be thought of in terms of rental charges based on some anticipated share of capacity rented from the incumbent. When feasible, capacity interconnection agreements can be quite effective at balancing the risk. Since information is spread unevenly among parties, every time the better informed party bears some risk, we should expect efficiency gains. When the entrant are likely to be better informed about the potential market, the incumbent should not be required to produce forecasts about rivals' demand. In principle, entrants would also benefit from increased flexibility in their way of pricing. Usage charges often come with a structure that reflects the incumbent's underlying tariff structure (this is most evident with ECPR). Under capacity agreements, buyers of access are free to set their individual tariffs, without being anchored to the incumbent.

To conclude, it may be worth stressing that while the recognition of the notion of risk is crucial to the assessment of access prices, their design should not provide a strict guarantee either to incumbent to recover its fixed costs or to the entrant in order to successfully compete against the incumbent. Commercial risks always exist in the market place and they should of primary concern to the firms rather than the regulator.

c. The regulator's viewpoint

Consider now the implications on access charges deriving from repeated interactions between regulated firms and regulators. With asymmetric information, the regulator always tries to infer something about the firm's cost, e.g. today's performance will be used to set tomorrow's target. The issue now is that current high outputs will induce the regulator to insist on future high production pushing the firm to underproduce compared to a situation in which information is common to regulators and regulated firms. The result, known as **"ratchet effect"**.

The presence of rivals adds complications, but the message is similar: to drive rivals out, the incumbent has to decrease prices, but this potentially reveals valuable information to the regulator. In this respect, one could think that risks of **predation** are mitigated. Unfortunately, this is not true when we discuss the role played by access charges. The incumbent can claim high bottleneck costs and corresponding high access charges, without revealing too much information on its overall efficiency. The concerns anticipated above on sub-cost observability are confirmed: access prices can be used as a potentially dangerous predation tool.

3.7 Summary of the main lessons for regulators of vertically integrated firms

Table 2 summarises the key results obtained in this section. A quick glance suggests that the rather complex theoretical contributions deliver one message: *the access charge is often performing too many tasks*. Different goals and policy objectives lead to alternative ways of calculating optimal charges. While it is true that theory is extremely useful to understand the mediating function of access prices, one first fundamental step should precede any access distortion: *whenever possible, the use of access pricing as an instrument for the promotion of too many goals should be resisted and other instruments should be used*. Regulators should be aware that there is a sequencing of events that can reduce the complexity of the access problem. For instance, if the regulator believes there are barriers to entry, the tax/subsidy issue of the entry barrier should be addressed directly and be made explicit, rather than burying it into the access pricing problem. The latter could indeed be the only option available, but only after having realised that other options are not feasible. A similar argument can be made for universal service obligations. In other words, by understanding the links between different problems, new instruments become available that allow to fine tune the regulatory process.

Table 2. Access charges with vertical integration

Basic case	Access charge:	Potential problems:	Eventual remedies:
First best	marginal cost	require lump sums, otherwise fixed cost not covered	• tariff rebalancing • USO funds
Second best	Ramsey	• informational content • may not be sustainable	price cap
Productive efficiency	ECPR	partial rule	
Extensions:			
Entry promotion for: • Product variety • Entry barriers • Learning-by-doing	decrease	fixed cost may not be recovered	 • direct subsidies • equal access
• Bypass • Cost duplication	Increase	small entrants disadvantaged	quantity discounts
Risk and hold up (incumbent)	Increase		• long-term contracts • spot and forward market • capacity charges
Asymmetric information	Incentive regulation	Predatory behaviour	• accounting separation • floors and ceilings
Market power	decrease	fixed costs may not be recovered	price regulation

The other main messages of this section can be summarized as follows. When financial constraints have to met, Ramsey prices provide the second best guidelines that take into account both allocative and productive efficiency. These are however not easy to implement since they require a good deal of information. Difficulty, however, does not imply infeasibility. While it is true that elasticity of demand is difficult to forecast for new innovative services (especially in the telecommunications sector), in some utilities (gas, electricity) patterns of demand are rather standard and predictable so that the regulator could and should try to produce such estimates. In section 5 we will discuss how a price cap mechanism may be the regulatory tool that brings Ramsey prices without requiring the regulator to explicitly calculate them.

Ramsey prices stand up as a good theoretical benchmark because they say something that should not be neglected*: markets are related so that demand and supply cannot be considered in isolation.* This is something that ECPR does not do since it *only* focuses on productive efficiency. It should be clear that ECPR willingly narrows its considerations on the allocation of production between the bottleneck proprietor and its rivals. Final markets are not discussed, not because unimportant, but because ECPR's proponents say that access is the wrong instrument at the final level. This is why, in the end, ECPR ignores the fact that profits generated on access can be used to lower retail prices of the incumbent.

ECPR is very influential however and is likely to remain central in the debate on access. It has many merits and some potential flaws. On the theoretical side, it introduces the powerful concept of opportunity costs which differ as services have different degrees of substitutability or in the presence of bypass possibility. On the practical side, ECPR is a rather simple rule and it gives valid guidelines if there is no problem of recovery of fixed costs. If this is the case (but the question should be clearly stated), static productive efficiency remains the only goal, and ECPR is suitable for its promotion. The risk of ECPR is to be misunderstood and misapplied. In particular, it performs badly in situations the rule is not designed for and these are common. If not used jointly with complementary instruments (i.e. final price regulation, price floors/ceilings), it allows monopoly rents and anti-competitive conduct.

Table 2 also reports situations in which access should be decreased to promote entry. Again, the first question has to be why entry should be promoted. If this is because entry brings benefits from product variety but there are barriers to entry, a simultaneous effort should be made to remove these (e.g. by mandating equal access). A direct subsidy, if feasible, is more visible and potentially more accountable, without having to distort prices. If entry is needed to bring final prices down, one has to assess the alternative of regulating such prices in a more direct and effective way. Only when the most immediate and appealing options are not available, distorted access charges make sense.

4. What if access was left unregulated?

The difficulty of regulating access to a single bottleneck is pushing many governments to promote competition between networks, even if is at the cost of some degree of duplication of infrastructures. In the US railroads, track owners often sign voluntary agreements to use each other's tracks ir shipping needs and this is supported by a what is generally a two-part taiff with an annual fixed charge to cover non-variable capital costs and a variable charge depending on the number of trains or wagon to cover variable costs and wear and tear on the track.

Some governments view this also as a way of trying to reduce the need for government agencies to monitor or regulate the market. In the telecoms sector, as clients of competing operators try to get in touch with each other, mutual provision of access is the central issue. This is quite different from the situation in which an integrated monopolist controls the local network and is required to interconnect with entrants competing on complementary segments such as long distance. The new situation is such that there may no longer be a need for regulation in two way bottlenecks situations.

The debate is thus whether access charges should be, in this context, negotiated freely between operators as in New Zealand, or monitored closely as in the US of UK? In addition, are there high risks of collusion (a "raise each other's cost" effect for given market shares can used by the colluding firms to justify higher retail prices)? This section shows that the need for government intervention (through a sector regulator or a competition agency) and the possibility of unregulated access charges, depends on the answers to two separate questions:

- Are independent firms likely to find an agreement?
- If so, is the agreement efficient?

Although it is hard to provide exhaustive answers anticipating any type of situation (see Economides, 1996, for an extensive survey), we summarise here some of the most common situation in which government intervention should still be envisaged. The main guiding principles from theory can be structured as follows. In general, firms seek agreements or other forms of coordination when they can mutually benefit from the network. Access means that separate markets can communicate, but also that, at the same time, there can be competition for such markets. Hence, there is a tension between a market expansion effect and the resulting increase in competition driven by various factors.

4.1 Network externalities for market expansion for the prospects of agreement

In the presence of strong network externalities—i.e. the more people are connected the more valuable the service is to each user--, the willingness-to-pay of consumers rises with the size of the market, the expansion effect is likely to dominate but this does not mean that market expansion will always happen. Indeed, while this is true if each firm brings its own "captive" market of similar size, firms with different market sizes and types may have contrasting ideas about access. A big firm does not benefit much by providing access to a small firm, and the latter can often survive only by getting

such access. As a result, incumbents are reluctant to give access to small entrants supplying the same product. The prediction on a refusal to interconnect can be contrasted with many regulated industries, in which regulators intervene only if parties are unable to seek an agreement on a commercial basis.

In the telecommunications industry, for instance, it is quite easy to find (unregulated) agreements to terminate calls from fixed networks to mobile users. In this example, the operators are not competing fiercely against each other (if subscribers to fixed and mobile networks do not easily substitute between them). Both parties can benefit from successfully delivering a call, however this does not necessarily imply that the agreed terms are in the interest of society as a whole. This is most obvious in international telephony where operators in different countries are clearly not competing against each other. Access charges are always agreed upon in bilateral settlements, however the charges are usually set at a high level in order to induce the operators to charge high prices also to international users. On the other hand, when there is strong competition among incumbents and new entrants, interconnection is not very likely to be agreed upon because of diverging interests. In practice, few agreements have been signed so far with the consequence that access regulation has been quite active and intrusive.

4.2 Market power and non-price discrimination built in negotiated agreements

When prices (including access) are not regulated and entrants offer products very similar to the incumbent, a regulator or a competition agency can expect the latter to charge very high access charges in order to make entry impossible and can detect it by looking at the separate accounts of the upstream and downstream activities of the incumbent. These would show that the incumbent does not impute to its downstream division the same access charge required from the entrant. The main challenge that the regulator or the competition agency then faces is that a firm with market power can try to soften downstream competition even if access charges are transparent. One way is to affect the rivals' costs using non-price devices, such as poor quality, delays in processing orders, etc... Economides (1998) shows that the incentive for an incumbent monopolist to non-price discriminate against downstream entrants is so strong that in the limit the monopolist raises the cost (by reducing quality) until the rivals are driven out of business.

4.3 Information sharing and the likelyhood of negotiated agreements

If it is difficult for firms who do not have access to equivalent information to find agreements (calling for intervention), the reverse is also true: similarly informed and similar firms are quite keen on interconnecting with each other. Imagine two telecommunications operators, each one with its bottleneck (local customers). By interconnecting, captive customers can call more people and at the same time the operator will receive interconnection charges from calls originated on the rival's network. Does this double coincidence of wants also ensure that access charges are efficient? At first sight, it may appear that access charges are irrelevant. Each operator receives as much as it pays to its rival (charges from incoming calls exactly compensate charges paid for terminating calls). It is tempting to argue that on average there is no access revenue or deficit, hence there should be no worry.

This reasoning is wrong since access charge can be used as a *collusive device.* High interconnection charges sustain collusion by making price undercutting very costly. The reasoning goes as follows. If a firm deviates from collusion and lowers its price, it will attract many customers. The deviant's network then originates more calls than it receives, and this results in a net outflow of calls with corresponding expensive interconnection payments. In the end, deviation is not profitable

(see Armstrong, 1998; Laffont *et al.*, 1998). These results show that the coincidence of interests between parties does not always lead to efficient agreements. The policy implication is that, despite direct competition for end users, *the need for regulation of access remains*, even among firms that are symmetrically placed.[25] If access charges are privately and independently determined, complications still can arise from double marginalisation problems (each firm, in its pricing decisions, fails to take into account the rival's profits, so that double mark ups can bring prices above the monopoly level). However, it seems that the collusive power of access charges diminishes with the common suggestion made to consider non-linear pricing as common as two part tariffs in the reform process.

4.4 Regulatory instruments for negotiated contracts

When contracts are negotiated, the regulators or the competition agency do not have the control of prices since these are supposed to be driven by the market. So which regulatory instruments can they rely on? When networks are of different size, regulators can impose *reciprocity* in the setting of charges as an (imperfect) tool to limit interconnection refusal and entry deterrence. However, two problems may emerge. First, even if reciprocity on access charges can be monitored, the incumbent may still try to non-price discriminate against rivals without violating reciprocity clauses. Second, the imposition of the reciprocity principle commonly considered by regulators can lead to a market "monopolization" by the low quality network. In telecoms, this is the case if the costs for terminating calls on the high quality network are very important. Third, after a transition phase, when networks are likely to be of a similar size, reciprocal access pricing has the bad property of supporting collusive outcomes. Some practitioners have proposed to apply an ECPR concept: the access charge has a *ceiling* given by the network's final price minus the network's cost in the competitive segment. Generally, more research is needed on this application but an important point on the interpretation of ECPR in the presence of network competition could be useful to keep in mind: once consumers are connected to a network, each network has a bottleneck that is essential for completing calls. The opportunity cost concept is then related to the loss of a completion of a call. *However, firms compete (ex-ante) for customers, so that market share losses are in terms of customers rather than calls.* Ex ante and ex post perspectives lead to different ceilings, respectively based on marginal cost and average cost, where the latter also includes fixed costs of connecting a customer. The real danger for applying ECPR in a context it is not designed for, is that it can soften price competition by giving commitment to collusive devices. Given an agreed access price that sustains monopoly profits, ECPR would say that operators are prevented from lowering their prices below the monopoly level.

5. From theory to practice: implementation challenges

The paper has so far reviewed a number of theoretical access pricing rules which address various types of situations. The next step is to put them into practice through approximations. For the time being, even if research has not come out with "the" rule, some experiences with the existing partial rules provide valuable insights on implementation strategies and challenges. There are three key areas in which the success in meeting the implementation challenges are being reasonably promising:

---

[25] Our discussion is also relevant when the bottleneck is jointly owned by all firms in the industry (for instance, in international gas transportation networks). With a dominant firm, we go back to the unregulated vertical integration case (the only difference being that the dominant firm takes into account only a share of the bottleneck profits). If ownership is even, there is no reason why joint ownership will solve the access issue, and bottleneck members will set high charges to monopolise the final market.

progress made in measuring costs, in understanding the demand side and in monitoring anti-competitive behaviour.

5.1 Calculating and allocating costs

The first step in the implementation of any of the rules is to figure out what the costs are and how they are related to access services.[26] If this were an easy task, one could forget about demand considerations and apply some *accounting methods*. The access charge would then be equal to the cost allocated to each service. This is far from being an uncontroversial issue for several reasons. First, the cost in question includes both *incremental costs* (defined as costs directly related to the increase in production caused by the demand for access) and *joint and common costs* (costs incurred in the supply of a group of services that cannot be directly attributed to any one service; they typically derive from economies of scope). The latter have to be allocated in the right proportions to the various activities and this is a very delicate operation.

The easiest way is to adopt *fully distributed costs* (FDC). Under FDC, the common and joint parts are allocated according to various measures: output shares (uniform mark up per unit), directly attributable costs, revenues or price-proportional mark ups. All these rules are mechanical, therefore easy to implement but completely arbitrary. It is simple to understand why they are immediately dismissed by economists: cost minimisation is not encouraged and demand is not accounted for.[27] At the same time they are relatively simple, familiar and well understood. They still deserve some popularity when they are the only feasible practice, which may indeed be the case in developing countries. Moreover, under FDC, there is a commitment to allow the incumbent to recover its investments, which can be desirable in some circumstances. In any case, after an initial use justified on practical grounds, rules based on FDC should be replaced by other methods. In particular, the degree of arbitrariness is somehow reduced by better accounting systems, such as *activity-based costing*, that try to relate in a causal way common costs to the production of a good.

An improvement in the accounting cost method is already an achievement because more direct costing can be practised, still it is subject to criticism when it is based on accounting book values (historical cost accounting). In many occasions, the replacement cost of the bottleneck is different than its historic cost. Access charges based on historic costs can then send wrong signals to entrants, attracting too many inefficient firms or discouraging potentially efficient ones. To overcome these difficulties, current cost accounting methods can be employed. They are accounting methodologies where assets are valued and depreciated according to their current replacement cost. Typically, they involve the valuation of the firm's existing assets at the cost of replacing them with assets which serve the same function and are likely to incorporate the latest available technology. Such forward-looking approach is fundamental for the calculation of *long-run incremental costs* (LRIC). They are often advocated as the best base for access charges if one wants to promote entrants' competition. Proponents of "pure" LRIC also believe that networks' economies of scale are not so pronounced, so that fixed cost recovery is not a problem. Once the incumbent faces revenue requirements, LRIC plus (often uniform) mark ups may be used.

---

[26] It may be worth to highlight that the structure of access prices will depend on the allocation of costs that flows from the cost models or approximations discussed in this section. By underpricing peak demand for instance, a regulated incumbent may be manipulating and trying to increase the allowed asset base on which the rate of allowed rate of return is calculated.

[27] One effect of FDC is that more elastic segments often end up paying too much than they should, in practice they subsidise price-insensitive segments of the market.

Box 7: Accounting concerns in practice: Charging for access to Railtrack in the UK*

The discussion of the cost of capital cannot be separated from the asset valuation issue, a common topic of debate both among academic and practitioners. In the UK rail sector for instance, the hot issue during 1998 was the preparation for the review of Railtrack's access charges. The main concern in that context was not only the determination of the appropriate return to capital but also the determination of the appropriate regulatory asset base. The debate surged because the company wants the return to be calculated on the book value of the assets while the regulator and the competition agency think the appropriate base is the value of the company at the time of flotation. The regulator also had to be quite explicit about how access charges had to be set and this included an explicit discussion of the degrees of competition to which the various assets of the track owning company were subject to. It ended up identifying separately the revenue (and hence implicitly the access charge) allowed for station and track charges billed to franchised train operators.

* For details, see the Regulator's web side: www.rail-reg.gov.uk, see also Kennedy (1997).

Overall, LRIC still represents the dominant paradigm in the telecommunications industry (for instance it is adopted in the US, UK, and it is advocated by the EC since 1994). However, there are not many arguments to support it from an economic point of view and this is in stark contrast with the practical consensus it has received. The determination of LRIC still involves considerable discretion, and the uniform mark-up can be just justified on simplicity grounds. The LRIC is in fact a sort of FDC method, but taking a forward-looking perspective rather than a backward-looking accounting one. Moreover, LRIC severely limits firms from making money in the bottleneck, thus raising the question about who will actually invest in upgrading the bottleneck infrastructure in anticipation of constrained profits. At the moment, LRIC does promote competition on existing networks, and this may provide an answer for its popularity. However, we have already said that, even if there is a justifiable desire to promote competition, this does not imply that the access charges should be distorted to achieve such a goal. Despite these controversies, calculations of LRIC have until recently been done using engineering network models. This is very important because these models provide a common ground for confrontation between the regulator and industry players: causal relationships become clearer, allowing us to reduce the amount of common costs to be allocated in arbitrary ways. Better accounting methods, in pair with proper accounting separation, also deliver the desirable outcome of discovering existing cross-subsidies.

In the telecoms industry, additional research is also being done at FCC and at the World Bank to design cost models with endogenous switching points in the network which should allow a better assessment of all costs, including opportunity costs.[28] These new models for regulators are a major change since in the past they were relying on companies for cost estimates. One of the key novelties of these new models over previous models is that they allow the explicit optimization of switch locations (instead of taking them for granted as built in models proposed by the companies themselves). The result is a more effective estimate of the forward looking cost accounting and a more realistic estimate of the potential gains from competition to the incumbents.[29] The risks involved in getting the regulators to work with these models are quite hotly debated by both practitioners and academics. The main focus of the debate is again on the likely disagreements that are likely to emerge on the regulators' set of assumptions for some critical variables such as the cost of capital, amortization rates and inputs costs. These were some of the issues being dealt with in Argentina during the June '98 gas distribution price revision and the September '98 electricity

28 See FTC web site for more details; at the Bank, an applied research program co-managed by Estache and Guasch and based on theoretical research by Gasmi, Laffont and Sharkey is underway in the field with applications to Argentina and Peru.

29 One of the key punchlines of the recent developments in costs models is to show that for realistic estimates of the costs of public funds in developing countries, explicit cross-subsidies may have a very strong role to play under many realistic restructuring arrangements in telecoms.

transmission price review and that experience would make a strong case for a benchmark model in the hands of regulators for the right to deliver USO under various market organizations.

5.2 Towards a usage-based solution to the access pricing problem: the global price-cap

As we know from the theory, usage-based rather than purely cost-based methods should be used to set access charges. Among usage-based approaches, ECPR seems to be the easiest to implement, given that it requires much less information than Ramsey prices. In section 3 we already pointed out that the latter can be criticised on the ground that elasticities are difficult to estimate. Since it is reasonable to assume that the firm has better knowledge than the regulator about demand conditions, the difficulty can be overcome by delegating pricing decisions to the firm itself. Rather than imposing a particular set of prices, under many circumstances it is better to leave the firm free to charge whatever relative prices it chooses, subject to a constraint on a price index. This concept is rather familiar for the setting of price caps on final products and it can be extended also to access services.

The most innovative proposal is to have a *global price cap* on the entire incumbent's range of products (Laffont and Tirole, 1996 and 1998). The rationale is that access services, in the eyes of the incumbent, are just a particular type of service. The bottleneck input should then be treated as a final good and included in the computation of the price cap. The approach requires that a weighted average of all these prices not exceed a certain figure. For instance, using the notation introduced in section 3.1, and denoting as w_i the weight given to service i (i = 0, 1, access) as $\bar{p}$ the cap, the incumbent firm can set any price it wants as long as:

$$w_0 p_0 + w_1 p_1 + w_a a \leq \bar{p}$$

When a cap is properly set according to the previous equation, the regulated firm is *induced* to choose the optimal Ramsey prices, without the need for the regulator to know demand functions. This is because the regulated firm can use its own private information about costs and consumer preferences to set intermediate and final prices that satisfy the global cap and that such prices are exactly the Ramsey ones: because of the constraint imposed by the global cap, the firm internalises consumer surplus in proportion to the weights set in the cap. The global cap is only an average value over a basket of services, and the incumbent is left with the flexibility of increasing or decreasing all its prices as long as their weighted average satisfy the cap.

The proposal, while a major improvement over the ECPR, still has its own share of problems relating to the common ways of implementing price caps. For instance, the normal caps often used in practice just concentrate on final goods (the term $w_a a$ is not included), which distorts prices.[30] The properties of global caps remain encouraging even if they depend on strong assumptions. On top of the standard type of criticism that can be made to price cap mechanisms (weights in practice are based on realised outputs, rate-of-return flavour when the regulator cannot ignore the firm's past performance)[31], one additional concern is raised with global caps: the incumbent can engage in predatory practices. By raising the access price and lowering the final product price, the global cap can be satisfied while performing a price squeeze that damages new entrants.

---

[30] Prices are increased in competitive segments of the market and lowered in non-competitive ones.

[31] The weight of each good for computing the cap should be exogenously fixed at a level proportional to the forecasted quantity of the good, i.e. $w_i = q_i^e / q_{tot.}^e$ where q_i^e and $q_{tot.}^e$ are the estimates respectively of the quantity produced of good i and of total production. A precise estimate may not be available so that past observations are used in place, inducing distortions.

5.3 Monitoring anti-competitive behaviour: partial caps or adjusted global caps?

The concern that the bottleneck owner has the opportunity to distort prices and raise competitors' cost, has led some commentators to dismiss global caps in favour of *partial price caps*. Separate partial caps on final goods and access services can concentrate on two distinct markets, respectively final and intermediate goods. However, it is probably a better practice not to dismiss a global cap, rather to supplement it with other instruments. ECPR can be used as an upper bound that protects in part from predation (the access price is tied to final prices, so that the former cannot be increased without simultaneously increasing the latter).

Other standard devices that could be appended to caps are obtainable from forward-looking costs described before. They can perform two important functions, namely to obtain *price floors* and *price ceilings* used to fight cross-subsidies, predation and exercise of monopoly power on final users. Price floors are requested to retain some regulatory controls on the structure of charges when the incumbent is left free to set its access fees. In particular, *final* prices above LRIC can be required as a safeguard against predatory behaviour. Ceilings are intended to prevent incumbents from exploiting monopoly power in the final market. A well accepted idea is that final prices cannot be higher than their *stand-alone cost*, that is to say the cost of a single-product firm providing the good in question.

A certain number of criticisms could be made to the use of such instruments. For instance, a price below LRIC may not be a signal of predation but rather be a legitimate competitive response. Our interest being on access charges—rather than on antitrust concerns--, we just repeat that access prices should not serve too many purposes: auxiliary goals need supplementary instruments. Only when these other instruments should corrections in the access rules be made.

5.4 Looking for creative solutions

A final bridge between theory and practice is emerging from an impressive amount of work being undertaken to address a common problem with transmission congestion in electricity. Most of the recent literature makes it quite clear that when there is a seller and/or buyer market in an unregulated electricity market power-and when there is a scarcity rent from lack of investment in transmission--which is quite common in developing countries-- the solution to the common problems of congestion in transmission is quite important for the success of reforms aiming at increasing long run efficiency. New solutions to the problems posed by congestion pricing are being covered by the literature to reduce the need to worry about access pricing in a traditional way. A recent and stimulating contribution is provided by Joskow and Tirole (1998). They show that the allocation of some type of tradable transmission rights can work. They also show however that with this new approach, the traditional problems associated with the design of access prices are replaced by challenges due to the difficulty of picking the appropriate allocation rule when there is some degree of market power. How much it matters depends on the underlying configuration of the market power problems (voting rules, the nature of the rights--physical vs. financial--, the microstructure of the market for rights--essentially the extent to which there is free-riding in the financing of transmission capacity--). They demonstrate (in a simplified 2 nodes-model) that misallocations are more severe when the rights are given to a supplier with market power at the expensive importing region node or with a buyer with market power at the cheap node in an exporting region. In general, the specific allocation of physical rights seems to be much more damaging than the allocation of financial rights (because it adds factors with potential implications on the generators' market power). Although much more research is needed to be able to derive definitive conclusions from this type work, it shows that regulators do have alternatives they can pick from.

6. Concluding comments

The access pricing problem is clearly one of the most important but also one of the most controversial questions in regulated infrastructure services. Part of this complexity stems from the fact that access rules can only be discussed in practice with reference to the rest of the regulatory environment since regulators have multiple goals and constraints. This survey should have made it clear that access rules should not be assigned too many expectations. However, while much progress is still to be made to find practical solutions to the problem, we hope that this survey has also shown that there are a few things that access prices already do and these should be done rather than ignored while waiting for a full and encompassing solution. Finally, since there is so much learning that is taking place from practitioners, theory is likely to have to continue to expand its limits to keep up with the creativity of the sectoral specialists and three to four years from now, this survey will have to be updated significantly to reflect these improvements in our understanding of the problem.

References

Armstrong, M. (1997), "Competition in Telecommunications," *Oxford Review of Economic Policy*,13: 64-82

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

Armstrong, M.and J. Vickers (1998), "The Access Pricing Problem with Deregulation: A Note," *Journal of Industrial Economics* 46: 115-121

Armstrong, M., S. Cowan and J. Vickers (1994), *Regulatory Reforms: Economic Analysis and British Experience* MIT Press, Cambridge (MA)

Armstrong, M., C. Doyle and J. Vickers (1996), "The Access Pricing Problem: A Synthesis," *Journal of Industrial Economics* 44: 131-150

Baumol, W.J. (1983), "Some Subtle Issues in Railroad Regulation," *International Journal of Transport Regulation* 10: 341-355

Baumol, W.J., J.A. Ordover and R.D. Willig (1997), "Parity pricing and Its Critics: A Necessary Condition for Efficiency in the Provision of Bottleneck Services to Competitors," *Yale Journal on Regulation* 14

Baumol, W.J. and J.G. Sidak, (1994), "The Pricing of Inputs Sold to Competitors," *Yale Journal on Regulation* 11: 171-202

Bruce, R., I. Kessides and L Kneifel (1998), "Overcoming Obstacles to Liberalization of the Telecom Sector in Estonia, Poland, the Czech Republic, Slovania and Hungary", The World Bank, October, mimeo

Cave, M. and C. Doyle (1994), "Access pricing in network utilities in theory and practice", *Utilities Policy*, 4 (3): 181-189

Economides, N. (1996), "The Economics of Networks," *International Journal of Industrial Organization* 14: 673-699

Economides, N. (1998), "The Incentive for Non-price Discrimination by an Input Monopolist," *International Journal of Industrial Organization* 16: 271284

Economides, N. and L.J. White (1995), "Access and Interconnection Pricing: How Efficient is the 'Efficient Component Pricing Rule'?" *Antitrust Bulletin* 40: 557-579

Estache, A. and M. Rodriguez-Pardina (1998), "Light and Lightning at the End of the Public Tunnel: The Reform of Electricity in Mercosur Countries ", The World Bank, mimeo

Joskow, P. and J. Tirole (1998), "Transmission Rights and Market Power on Electric Power Networks", mimeo, IDEI, Toulouse

Kahn, A. and W. Taylor (1994), "The Pricing of Inputs Sold to Competitors: A Comment," *Yale Journal on Regulation* 11: 225-240

Kennedy, D. (1997), "Regulating Access of the Railways Network", *Utilities Policy*, 6 (1): 57-65

Laffont, J.J. and J. Tirole (1993), *A Theory of Incentives in Procurement and Regulation*, MIT Press, Cambridge (MA)

Laffont, J.J. and J. Tirole, (1994), "Access Pricing and Competition," *European Economic Review* 38: 1673-1710

Laffont, J.J. and J. Tirole (1996), "Creating Competition Through Interconnection: Theory and Practice," *Journal of Regulatory Economics* 10: 227-256

Laffont, J.J. and J. Tirole, (1998), *Competition in Telecommunications*, Munich Lectures (forthcoming)

Laffont, J.J., P. Rey and J. Tirole, (1998), "Network Competition: I. Overview and Nondiscriminatory Pricing; II. Discriminatory Pricing," *RAND Journal of Economics* 29: 1-56

Sidak, J.G. and D. Spulber, (1997), "The Tragedy of the Telecommons: Government Pricing of Unbundled Network Elements Under the Telecommunications Act of 1996," *Columbia Law Review* 71: 1081-1161

Spulber, D. and J.G. Sidak, (1997), "Network Access Pricing and Deregulation," *Industriaand Corporate Change* 6: 757-782

Valletti, T. (1998), "The Practice of Access Pricing: Telecoms in the UK", The World Bank, EDIRP mimeo

Wildman, S.S. (1997), "Interconnection Pricing, Stranded Costs, and the Optimal Regulatory Contract," *Industrial and Corporate Change* 6: 741-755

Willig, R.D. (1979), "The Theory of Network Access Pricing," in H.M. Trebing (ed.) *Issues in Public Utility Regulation* Michigan State University Public Utility Papers

Annex

ANNEX TO BOX 2. VOLUNTARY ACCESS PROVISION

This box contains an analysis for the more mathematically-inclined reader and it draws on Spulber and Sidak (1997). The notation that we will use is the following. Imagine the monopolist owner produces a bottleneck service at a unit variable cost equal to c_0. The incumbent can also produce the potentially competitive component at a unit cost c_I. When the price of the final service is p, the market demand is $Q = D(p)$. The monopolist also incurs a fixed cost F for the bottleneck. Let $R(p)$ represent the incumbent's net revenue from the service:

$$R(p) = (p - c_0 - c_I)D(p)$$

$R(p)$ can be seen as the quasi-rent accruing to the incumbent and it is sufficiently general to include different cases that go from the unregulated firm that behaves as a monopolist to the opposite situation of a regulated operator. The only restriction is that it should cover at least the fixed costs: $R(p) \geq F$.

There are two periods, before and after deregulation, denoted respectively as 0 and 1. Before deregulation, the price is p_0 and the associated quasi rent is $R_0 = R(p_0)$. After deregulation, entrants supply a quantity X and the final price equals p. Excluding revenues from access, the incumbent quasi-rent in period 1 is $R_1(p) = (p - c_0 - c_I)(D(p) - X)$.

As a result, when access is supplied, the opportunity cost that is foregone in the final market is:

$$\Delta = R_0 - R_1$$

We can now state an important condition. If a is the unit access charge, the incumbent is willing to supply access as long as the net access revenues are in excess of the opportunity cost of supplying access:

(A1) $$(a - c_0)X \geq \Delta$$

The previous condition is required for access to be voluntary. We now study the optimal access charge, taking into account the voluntary access condition (A1).

Entrants incur a unit cost of producing the final service equal to c_E which can differ from c_I. After deregulation, the post-entry price and the quantity supplied by the entrants depends on the access charge. The entrants' profits amount to:

$$\pi(a) = (p(a) - c_I - a)X(a).$$

It is reasonable to assume that $c_E \leq c_I$ and $p_1 \leq p_0$, i.e. the entrants are at least as efficient as the incumbent in the potentially competitive sector and the regulator does not allow the post-entry final price to rise above its initial level.

The regulator has to set the access charge that maximises social welfare subject to the voluntary access condition. When social welfare is given by the sum of consumer surplus (the area under the demand curve) and firms' profits, the problem takes the following structure:

$$\max_a W(a) = \int_{p(a)}^{\infty} D(x)dx + R_1(p(a)) + (a - c_0)X(a) + \pi(a)$$
$$\text{s.t. } (a - c_0)X(a) \geq \Delta$$

Different cases can arise according to the nature of competition among entrants in the final market.

Case 1. Price competition with homogenous final goods. All the products supplied in the final market are identical and firms compete *à la* Bertrand. That is to say that each firm sets its own price in order to maximise profits taking as given the price of the rivals. This situation describes very tough competition among identical firms and the equilibrium entails price cutting until firms make no extra-profits. This occurs when the unit price is equal to the unit total cost:

(A2) $$p^B = a + c_E$$

where the superscript B stands for Bertrand. The incumbents' sales in the final market are completely displaced by the more efficient rival ($X^B(a) = D(p^B)$), so that the voluntary access conditions simplifies to $\Delta \geq R_0$ and the welfare-maximising access charge takes the following form:

(A3) $$a^B = c_0 + R_0 / D(a^B + c_E)$$

From the previous rule, it can be seen that the first-best (access equal to the marginal cost) can be attained only when R_0 equals zero. For instance, it can never be achieved when there are fixed costs F to be recovered. As entrants become more efficient, the final price declines for two reasons: there is a direct effect from c_E and an indirect effect through the access charge (lower costs allow the incumbent's quasi-rent to be recovered over more units of output). Also note from eq. (A2) that the access charge can be rewritten as:

(A4) $$a^B = P^B - c_E$$

i.e. the unit charge is equal to the entrants' margin. This can be related to the ECPR that we discuss in Section 3.2. It should be noted that the margin rule (A4) is related to the entrants' cost and to the post-entry price, in contrast with ECPR in its original formulation. The Bertrand case analysed is in line with M-ECPR and the optimal access charges are less than the ECPR charge. The reason is that a lower access charge leads to bigger quantities of access demanded. The lower margin on access is then made on a higher volume, so that the incumbent's budget constraint is still satisfied.

We study now whether entry occurs and whether it is efficient. *Assume* that entry occurs so that the incumbent is displaced in the final market and $p^B < p_0$. Given the access charge and the expression for the initial quasi-rent, the last inequality can be rewritten as:

$$c_E + c_0 + R_0 / D(p^B) < c_I + c_0 + R_0 / D(p_0)$$

that can be rearranged to obtain:

$$\frac{R_0}{D(p^B)D(p_0)}[D(p_0) - D(p^B)] < c_I - c_E$$

The left hand-side of the inequality is negative given the initial assumption. Hence it has to be that $c_E < c_I$. The result is that if entrants are more efficient, the access rule (A3) ensures that entry occurs and that it is socially efficient.

Case 2. Quantity competition with homogenous final goods. This case is relevant when firms maximise their profits with respect to own quantity. One can think of capacity being chosen in this fashion. Quantity competition (also known as Cournot competition) is a good way of modelling imperfect competition when the final good is homogenous for consumers and it represents a kind of competition that is less tough than the Bertrand case analysed before. In order to tackle the problem of excessive entry, we introduce an additional feature: each entrant has to incur a fixed entry cost denoted as f (set-up cost). Entry is free, so that in equilibrium each downstream firm earns zero extra-profits (otherwise additional entry would occur).

Denoting by q^C the Cournot output per firm, and by n^C the number of downstream firms, the following conditions express respectively the Cournot first-order condition (obtained from the maximisation of the profit of a generic downstream firm taking as given the output of the rivals) and the no-entry condition (zero profit):

(A5) $$\begin{aligned} &p'(n^C q^C)q^C + p(n^C q^C) - c_E - a = 0 \\ &[p(n^C q^C) - c_E - a]q^C - f = 0 \end{aligned}$$

Multiplying the last expression by n^C and setting $X^C = n^C q^C$, we can derive the expression for the market price:

(A6) $$p^C = c_E + a + n^C f / X^C$$

By implicitly differentiating the last expression, one can easily show that $\partial X^C / \partial f < 0, \partial X^C / \partial c_E < 0$, i.e. total output decreases with two measures of downstream firms' efficiency. Since the zero-profit condition holds also for the incumbent's subsidiary, as in the Bertrand case, Δ simplifies to R_0.

The regulator's problem is now to maximise social welfare $W(a)$ subject to the Cournot conditions (A5) and the voluntary access condition (A1). One of two cases can arise:

Case (i). The voluntary access condition is not binding. The expression for the optimal access charge is:

(A7i) $$a^C = c_0 - \frac{p''(a^C)X^C(a^C)^2}{2n^C(a^C)}$$

Case (ii). The voluntary access condition is binding (for instance because F is high enough), then:

(A7ii) $a^C = c_0 + \frac{R_0}{X^C(a^C)}$

We are skipping some technical details, but it can be shown that, as long as $R_0 > 0$, the voluntary access condition is binding when demand is linear ($p'' = 0$) or when it is convex ($p'' > 0$). Hence case (i) only applies when demand is concave ($p'' < 0$), so that a^C is *always* set greater than c_0.

Recalling that $\partial X^C / \partial f < 0, \partial X^C / \partial c_E < 0$, it is evident from (A7ii) that the access charge increases when downstream firms are either not very efficient compared to the incumbent (high c_E) or they have to incur substantial fixed costs (high f).

We consider now the impact that the fixed costs f have on welfare. In the pre-deregulation period, output is Q_0. Let f_0 be the critical level of set-up costs such that after deregulation total output is unchanged: $X^C(f_0) = Q_0$. We consider only the more interesting case of actual entry costs that are below the critical level: $f < f_0$. This implies that $X^C(f) > Q_0$ and $p^C(f) < p_0$. The last inequality, simplifying the expression for p^C given by eq. (A6) and using eq. (A7ii), is:

$$c_E + c_0 + R_0 / X^C + n^C f / X^C < c_I + c_0 + R_0 / Q_0$$

The previous inequality holds also when, in the left-hand side, X^C is replaced by Q_0. This allows to rearrange the inequality to get:

$$(c_I - c_E)Q_0 > n^C f,$$

in other words entry is efficient because the efficiency gains obtained by producing at a lower cost downstream are greater than the fixed costs paid by the entrants. Note that efficient entry is associated to the access charge given by eq. (A6ii), not to *any* access rule.

In this box we have considered Bertrand and Cournot competition, the main results still hold for the case of differentiated products (see Spulber and Sidak, 1997).