TILBURG UNIVERSITY

TILEC

TILEC Discussion Paper

DP 2003-007

Access to Telecommunications Networks

By
Marcel Canoy, Paul de Bijl and Ron Kemp

October 2003

ISSN 1572-4042

Access to Telecommunications Networks

Marcel Canoy*, Paul de Bijl** and Ron Kemp***

‘There are things known, there are things unknown, in between are doors’

Jim Morrison

* CPB Netherlands Bureau for Economic Policy Analysis, The Hague

** CPB Netherlands Bureau for Economic Policy Analysis, The Hague, currently at Tilburg University, TILEC (Tilburg Law and Economics Center)

*** EIM, Zoetermeer

Paper prepared for the European Commission, DG Competition

March 2003

Contents

1 Introduction

Access has been a key issue for regulators right from the start of telecommunications market liberalisation. Despite some notable submarket successes, even by 2002 – more than ten years after the first steps towards liberalisation – incumbent operators still maintain a strong hold on many aspects of the sector, and infrastructure competition has not matured across the industry. Most notably, there has been little infrastructure competition within the local loop, while its unbundling progresses slowly (Buigues, 2002). As a result, incumbent operators with local loops are able to engage in anti-competitive behaviour, e.g. by providing access against unfavourable terms. Because ex post application of the Competition Law can have only a limited effect in such situations, regulators are tasked with setting the terms and conditions of access charges.

Although the legitimacy of involving regulators in access charges can be easily understood, things are a little trickier when it comes to determining just what that involvement should entail. Tasking them with setting access charges poses particular problems for a number of reasons. Firstly, the economic theory on determining optimal methods of setting access charges is not always clear-cut. Secondly, even in cases where the theory is unequivocal, the practice is not as easy as theory may suggest. The 'difficulties of the daily regulation business' is not just the usual cliché. Admittedly, the usual problems do exist here: required data is often simply not available and time is lacking. But on top of that, access regulation often serves too many goals. Access charges are used to allocate scarce capacity, to provide incentives for productive efficiency, to promote entry, to

promote investments by incumbents as well as by entrants – and all taking into consideration the fair allocation of common costs while also meeting equity goals such as a universal service obligation. This is quite a burden for an access charge, which is, in its simplest form, a one-dimensional price. Goals can so easily conflict. Low access charges can be good for stimulating service competition, but by the same token they might hamper infrastructural investment. The very least any regulator has to do is to determine just which goals it wants to reach with an access charge, and, in case of conflicts, which goals take priority. Sometimes 'smart regulation' is possible in attaining seemingly conflicting goals: an access charge which starts low but rises over time can promote services competition in the short term (enhancing static efficiency), while also providing incentives for entrants and incumbents to invest in infrastructure (enhancing dynamic efficiency). But more often than not, choices have to be made and trade-offs emerge.

This chapter provides an overview of access charge theory, mindful of the policy perspective. We draw a distinction between one- and two-way access.[1] One-way access refers to an entrant who needs access to an infrastructure owned by an incumbent, though the reverse is not true. Examples are Unbundled Local Loop and Carrier Select. Two-way access refers to network interconnection, where operators need mutual access in order to terminate calls on each other's networks. Policy implications can be quite different in one- and two-way access.

Determining access prices implies taking into consideration both the short and the long term. Particularly in the early days, both economic theory and regulation practices focused strongly on static efficiency (i.e. the short term).

However, dynamic considerations are important and should not be neglected for a number of reasons, one of which being that if infrastructures competition matures, less regulation is needed. For this reason this chapter devotes extra attention to dynamic considerations.

An interesting issue, briefly addressed here, is why the European Commission is involved in access policy. Typically, few worthwhile cross-border trade or arbitrage opportunities exist, and neither is there a 'policy-race-to-the bottom' argument for centralisation. Thus there would seem to be a strong case here for decentralisation. But there seems little point in having the national regulators constantly reinventing the wheel. Access problems in Finland appear very similar to those in Greece. The problems are highly complex, so that when it comes to knowledge, achieving economies of scale is an important reason for the European Commission's involvement.

The main sources used for this paper are surveys of the economic literature on telecommunications by Armstrong (2001) and Laffont and Tirole (2000), and policy-oriented research by Cave et al. (2001) and (to a lesser extent) Bennett et al. (2001), as well as a number of policy documents, most notably from the European Commission and OFTEL.

This paper is organised as follows. Section 2 discusses one-way access. Section 3 discusses two-way access. Section 4 analyses four special topics, namely fixed-mobile termination, margin squeeze, the allocation of common costs, and static versus dynamic efficiency. Section 5 contains the conclusion.

2 One-way access

2.1 Introduction

One-way access is characterised by a vertically-integrated incumbent, usually the former state-owned monopolist, with a local access network. One or more entrants will not have such a local network, nor will they be able to build one in a reasonably short term. Entrants may have their own long-distance network, but this is not necessary[2]. However, in order to compete with the incumbent in the retail voice telephony market and to match the incumbent's pattern of offers, they need access to the incumbent's local access network.

One-way access is common in competition's early phases, when entrants have not yet been able to roll out (some of) their own local connections; but is also relevant in the market's mature stages. An important example of one-way access is Unbundled Local Loop (ULL): an entrant gets access to the incumbent's local network copper-cable pairs, and has control over the use of (some of) the copper-cable pairs' total frequency spectrum. The main implication of this is that it enables new network operators to offer high bandwidth directly to consumers, thereby fostering competition for high-bandwidth services.[3]

Why is so much policy attention devoted to ULL? In the first place, ULL allows for a direct and comprehensive relationship with the end-consumer without requiring a full rolled-out network. It also puts pressure on operators to offer consumers a whole new range of services, using (some of) the total bandwidth, such as fast Internet, receiving richer content such as video on demand, etc.

Although there might be alternative technologies, which can provide broadband services, such as cable and UMTS, ULL seems to be deployed most rapidly.

At the moment, the roll-out of broadband services (e.g. ADSL) is relatively limited (Buigues, 2002). This might be related to the difficulties new entrants face: behaviour of the incumbents such as excessive pricing, delays in delivery, predatory pricing or price squeezes and refusals to supply the necessary information or space in the incumbent's locations. All these aspects deserve the regulators' attention .

Carrier Select is another example of one-way access: an entrant has originating and terminating access to the incumbent's local network. By dialling a prefix (usually consisting of four digits), consumers can indicate that they want their call to be carried by the entrant instead of by the incumbent. Such access generally works well after some initial problems and other matters needing attention (such as margin squeeze, or scarcity).

Because of the asymmetry between the bargaining positions of incumbents and entrants, as well as between their interests, it is highly unlikely they will agree on the access price level. Specifically, the incumbent has a strong incentive to raise the access price as high as possible, perhaps even to foreclose entry, whereas entrants prefer cheap access.[4] Their interests are thus strongly opposed, necessitating access price regulation. The most notable central issue is that of optimal access policy.

2.2 Access pricing

2.2.1 First-best solution

The ‘first-best’ solution, in other words the theoretical solution where there are no imperfections, is to make the access price equal the marginal access cost. Marginal cost here refers to traffic-dependent cost, which is practically zero if one applies a strict definition (roughly speaking it would equal the electricity cost per time unit of transmitting a signal). Given that the sales/marketing departments making pricing decisions typically attribute various fixed costs to telephony traffic, it may be advisable to apply a somewhat broader definition of marginal cost. Furthermore, where there is a substantial fixed cost in serving an additional consumer, a traffic-independent ‘access price’ may also be needed, such as the monthly lease price for a local line in the case of an unbundled local loop. This discussion covers traffic-dependent access prices, however. With an access price based on marginal costs, there are no retail price distortions, and entrants receive all the right signals: they can make positive profits only if they are more efficient than the incumbent. This access price is also fair and non-discriminatory, because it is the same for all entrants, and is not usage-based. In this first-best world, the incumbent’s fixed network cost is covered by the state making a lump-sum payment to the incumbent.

The first-best concept can rarely (if ever) be implemented in practice. An obvious problem with this solution is that lump-sum transfers from the government are usually not feasible, certainly not without creating large distortions elsewhere in the economy while seriously impeding the incumbent’s

incentive structure of the incumbent. Therefore, oneOne thus has to look for an access price above marginal cost, that which would helps to cover the incumbent's fixed network cost.

2.2.2 Ramsey pricing

The practical problems involved in setting access prices equal to marginal costs, bring about the necessity of introducing mark-ups to enable the incumbent to recover its network costs. On a more general level, one can try to set not only the incumbent's access prices, but also its retail prices to such an extent that welfare is maximised, subject to the constraint that the incumbent recoups its fixed cost. Note that welfare, defined as the sum of consumers' and producers' surplus, refers here to static efficiency. The solution to this problem is known as Ramsey pricing. Expressed in another way, the problem is to find the price structure for a multi-product (i.e. access and retail services) regulated incumbent, which will maximise welfare, given that the incumbent has to break even overall. A priori, Ramsey pricing implies that there should be mark-ups in the incumbent's access prices as well as in its retail prices. The central idea is that to maximise welfare, all the incumbent's prices (both wholesale and retail) should play a part in the recovery of fixed costs. In particular, the optimal access price will be equal to the marginal cost of access plus a 'Ramsey term' (i.e. a specific mark-up).

In their construction, Ramsey prices (wholesale and retail) reflect underlying costs as well as demand characteristics (see Laffont and Tirole, 2000, section 2.2). Hence, Ramsey prices are both cost-based and usage-based. This implies that they are, in principle, compatible with a firm's standard pricing

practices, since these also take into account costs and demand. More precisely, it can be shown that welfare-maximising prices result if the firm maximises its profits under the proviso that it offers a certain minimum level of surplus to its customers (the 'Ramsey-Boiteux' social welfare level). Note that the regulator is supposed to have full information covering cost and demand characteristics. Intuitively, if the price elasticity for a specific service is relatively low, then the mark-up in the price for this service can be relatively high. This is optimal for welfare, since the prices of services with higher elasticities can be reduced while still satisfying the incumbent's cost recovery constraints. For instance, it may be optimal to increase the access price above marginal cost in order to reduce retail prices. In general, mark-ups can be higher when they lead to less distortion in the optimal allocation.

It is important to note the potential problems attached to Ramsey pricing (see Laffont and Tirole, 2000, section 3.4). First, Ramsey pricing's informational requirements are very high; that is to say, they require the regulator to have knowledge on cost levels and demand elasticity – information the regulator generally doesn't have. However, one should not rule out the possibility of obtaining reasonable estimates of this data. Failing this the regulator can delegate pricing decisions to the better-informed firm, within the constraints imposed by Ramsey pricing. In observing this, it is usually apparent that unregulated firms have access to fine-tuned and sophisticated pricing tactics, suggesting that they have access to far more information than the regulators. In particular, pricing behaviour tends to reflect cost levels, elasticities, competitive pressure, and more. It can be shown that the structure (not the levels) of an unregulated monopolist's

prices is the same as the Ramsey price structure. Thus a regulator may determine the Ramsey pricing structure by imposing a global price cap so that (i) access is treated as a final item and included in the price cap, and (ii) the weighting of the price cap is determined externally based on forecast quantities. If the weighting is set appropriately, then the operator internalises net consumer surpluses in maximising its profits. Although the principles for setting the optimal weighting are fairly straightforward, the information required (on forecast quantities) depends on actual costs and demand elasticities, and may therefore be difficult to obtain (see Laffont and Tirole, 2000, section 4.7). Nevertheless, this is much less demanding than obtaining the information needed to set an optimal partial price cap.

Secondly, usage-based access prices are discriminatory access prices, and would therefore appear to contradict the non-discrimination policy principle. More precisely, Ramsey pricing prescribes that the charge paid by an entrant must depend on its use of the service. The access price typically depends on the incumbent's price-cost margin in the relevant retail market, on demand-side substitution possibilities, on supply-side bypass possibilities, and on the elasticity of access demand. For instance, customers of services which are not very price sensitive contribute more to cost recovery. From an economic perspective, this dependence simply maximises welfare: access prices should be higher when they are used for services for which demand is less elastic. However, applying Ramsey prices can imply a (very) skewed distribution of prices (see e.g. Jullien 2001) and hence it often has a dubious political feasibility.

Overall, from an economic perspective the problems mentioned above are not necessarily vexatious for policy makers. By construction, Ramsey pricing is the best way to set access and retail prices simultaneously; it is the 'least-bad' departure from first-best prices. At worst, if application turns out to be prohibitively difficult in practice or too unattractive politically, Ramsey pricing can (or should) serve as a useful benchmark for access regulation.

2.2.3 ECPR

Whereas Ramsey pricing aims at choosing optimal access and retail prices, it may happen that the access regulation problem is separated from retail pricing. In such an event, the Efficient Component Pricing Rule (ECPR) is a popular pricing rule, since it provides a link between the access and retail pricing.

ECPR, which is also known as the Baumol-Willig rule, has its background in the theory of contestable markets.[5] A retail market is said to be 'contestable' if there is potential hit-and-run entry, which constrains the incumbent's retail price – in other words where the threat of quick entry disciplines the incumbent's pricing behaviour. An underlying assumption is that potential entrants take the incumbent's price as given. Thus the theory is that for a fixed price charged by the incumbent, a more efficient firm can enter and capture the market; there is no 'in-market' competition.

Another ECPR underlying assumption is that access price regulation is separated from price setting in the retail market. In particular, retail prices are fixed in advance by the regulator. Thus ECPR's prescription is to choose the access price, which maximises welfare given the incumbent's retail prices. The

regulator is not concerned with overall welfare maximisation, as in the case of Ramsey pricing, but aims instead at cost recovery and productive efficiency.

Assuming a homogeneity in final products (which seems reasonable for voice telephony) and that the market is contestable (a 'heroic' assumption, as discussed above), ECPR prescribes that the access price should not be higher than the incumbent's opportunity cost, which is equal to its marginal cost of access plus its forfeited retail mark-up. One could equally state that the access price should not be higher than the incumbent's retail price minus its cost in the competitive retail activity.

An attractive feature of ECPR is that entrants receive the right signals, that is, they enter only if they have a cost advantage. Another positive feature, which is sometimes put forward, is that entry is revenue-neutral for the incumbent. But it is not clear why one should want to achieve revenue-neutrality. If the incumbent's profits are excessive, they will also remain so under ECPR if market entry is possible. ECPR and Ramsey pricing do not generally coincide, although this divergence can possibly be restored if one makes specific assumptions about symmetry and the absence of entrants' market power. An important difference is that ECPR ignores the relationship between the wholesale and retail markets. ECPR is a partial rule, in contrast to Ramsey pricing. If retail prices are regulated at Ramsey levels, then ECPR is optimal.

2.3 Dynamic pricing rules

The access pricing rules discussed in the previous subsection were based on the implicit assumption that firms do not invest or develop new technologies. Their

purpose was simply to compensate the incumbent for providing entrants with access to its network, such that its fixed investments could be recouped. Since telecommunications markets change very rapidly precisely because of investments and technological progress, dynamic considerations should also play a role in access regulation. In particular, it is important to note that any access price affects operators' (potential) profits, and hence also their incentives to enter the market, to invest in new technologies, to roll out networks, to maintain and upgrade existing networks, and so on. To deal with such dynamic issues, regulators have developed specific access pricing rules, discussed below. More generally, when assessing the effects of access regulation on firms' incentives to invest, one must distinguish between:

- the effects on entrants' incentives to roll out networks themselves (versus using an incumbent's existing network), and
- the effects on an incumbent's incentives to maintain and upgrade its existing network.

Among other things these effects depend on both current and anticipated access prices. For instance, if the access price is low and firms anticipate it remaining low, then the incumbent is not very eager to invest in its existing network, while entrants feel no urge to roll out their own networks. On the other hand, short-term competition will be intense, as entrants can easily compete by using the incumbent's network. In the short term this is good for consumers, but it is less certain if this also holds true in the longer term. If the access price is high

and firms expect it to remain thus, then it makes more sense for entrants to start rolling out their own networks, which in turn imposes a discipline on the incumbent to invest as well, in order to remain competitive. But in the short term it's expensive for entrants to start building up market share (e.g., by starting with Carrier Select services while their networks are not yet ready), thus negatively affecting consumers over the longer period. The following subsections discuss the dynamic consequences of access rules.

2.3.1 Backward-looking access pricing rules

Backward-looking cost-based access pricing rules are based on the incumbent's actual or historical costs (also called embedded costs). An example is Embedded Direct Costs (EDC). A backward-looking rule can be called fair in the sense that the incumbent is compensated for its actual network investments. But on an ongoing basis it may provide weak incentives for reducing costs, since the implicit message is that any cost will be reimbursed. One would also intuitively expect backward-looking access prices to be relatively high, compared with forward-looking prices (see below), making it harder for entrants without networks to compete with an incumbent. Recent experience with local-loop unbundling has shown that this is not always true however (see also the next section).

2.3.2 Forward-looking access pricing rules

Forward-looking cost-based access pricing rules are based on state-of-the-art, currently available technology. Hence, they explicitly take technological progress into account. If a new network can be rolled out at half the cost of the incumbent's network, then the cost of the new technology serves as the relevant benchmark. Accordingly, forward-looking rules incorporate cost efficiency; they can be used to correct for an incumbent's possible inefficiencies. Arguably, they can be used to mimic competition, which does not yet exist. It is even possible that an incumbent's access service, using an existing or obsolete technology, becomes a loss-making activity because of the downward pressure on access prices. Hence, in theory forward-looking rules give the incumbent an incentive to keep up with technological progress, and to keep investing in its network. A downside of forward-looking rules is that they tend to neglect so-called 'stranded assets'. If an investment has a fifty-fifty chance of success, then the revenue needed if the investment is successful should take into account the possibility that it might have failed. If access rules do not take this into consideration, the result may be under-investment and risk-averse behaviour.

The main example of a forward-looking rule is Long Run Incremental Cost (LRIC). LRIC prices can be substantially lower than the incumbent's actual cost levels. However, it should be noted that LRIC can also lead to substantially higher prices (e.g. because of increasing labour costs, or better but more expensive technology), especially when it comes to access to the local loop. LRIC does not appear to be well-suited to incorporating corrections for quality differences between old and new technologies. It may thus be more appropriate

for interconnection fees (i.e. two-way access prices; see the next section) than for pricing local loop access.[6]

Despite its intuitive appeal, LRIC has further drawbacks (see also Laffont and Tirole, 2000, Leo et al 2002). For instance, it should be noted that the cheapest technology available cannot be derived from standard accounting systems. Hence the regulator must determine just how efficient the incumbent should be (i.e. the regulator does not determine the desired efficiency level, but can put a lower boundary on it). In addition, LRIC may not allow the incumbent to make a profit margin on access. If this happens, the incumbent has strong incentives to deny access to entrants by using non-price anti-competitive practices, such as refusals to deal and delays in interconnection. Accordingly, a possible consequence of LRIC (and other forward-looking rules) is that the regulator has to continue to play a key role in managing entry. In other words, LRIC may imply that regulation remains heavy-handed, in contrast with the intention of gradually withdrawing regulation as competition matures.

LRIC's disadvantages originate in the implicit assumption that markets are contestable. Since telecommunication markets are typically not contestable, LRIC fails to take into account that assets can become stranded and that operators need risk premiums to cover that eventuality. It follows that applying LRIC without taking these dynamic considerations into account could be fairly disastrous.

To summarise: forward-looking rules such as LRIC have an appeal for interconnection. For access to the local loop there are serious dangers, in particular if used without taking stranded assets and common costs into consideration.

Overall, one cannot draw a clear-cut conclusion on the dynamic considerations. The problem is that access prices, especially in situations of one-way access as described above, often have to perform too many simultaneous tasks. Those tasks may conflict, as indicated above. The policy chapter 4 will consider possible solutions to this problem.

3 Two-way access

3.1 Introduction

Two-way access refers to a situation where there are two or more operators, each with their own infrastructure (of long-distance and local access networks), needing mutual links so that any consumer can call anyone else. These operators compete for network subscribers and need each other in order to offer their customers maximum network benefits (interoperability). The typical two-way access situation is network interconnection.

Another important two-way access example is fixed-mobile terminating access: a network operator can charge a consumer for making use of the network (e.g. a two-part tariff). The network operator can also charge other parties to reach their subscribers. There is competition for subscribers, but no competition for reaching the subscribers of a network. The characteristics of fixed-mobile call termination are that there is competition for mobile subscribers, but no competition for providing access to mobile customers once they have subscribed

to a particular mobile network. The main difference with network interconnection is that the services are provided to non-competing operators.[7] For example, in fixed-mobile termination, the mobile operator needs to access the fixed network and vice versa, though to a certain extent they are not exactly competing head-to-head. Much of the theory discussed in this chapter applies to fixed-mobile termination. For this reason the present authors have opted to deviate from the categorisation introduced by Armstrong (2001) who offers a separate discussion on 'competitive bottlenecks' (of which fixed-mobile termination is considered to be the main example). There are indeed a number of interesting and separate policy questions related to fixed-mobile termination. For example, access to the incumbent's fixed lines has had greater regulation than access to mobile. This creates some policy questions but does not necessarily warrant new theory. The policy aspects of fixed-mobile termination will be discussed in chapter 4.

Another two-way access situation concerns international network interconnection, which was relevant even before the start of liberalisation. The crucial difference with the examples above (and the reason the authors do not consider this to be an example of two-way access) is that there is no competition for subscribers. This creates its own problems, which will not be discussed here. The same applies to international roaming. If a subscriber roams on another network to enable his or her mobile to function abroad, the visited network operator bills the home network operator for this call and vice versa. Since international roaming is a competitive service in principle, and since possible related problems are beyond the access discussions of this chapter, it will not be pursued in this section.

As with one-way access, the central question is the extent of the access price, or in this case the terminating access price. A difference here is that by definition, there is always more than one access price – each operator charges its own terminating access price. Hence, an additional issue is whether terminating access prices should be reciprocal, that is, the same for all operators. A third issue concerns the so-called 'missing price'. Because of the 'calling party pays' principle (CPP), there is no fee for receiving a call. This fact will have great relevance in the discussion.

3.2 Access pricing

Two important policy questions govern terminating access prices. The first one is the optimal level of access prices: which levels of access prices maximise welfare (or, depending on the regulator's objective, consumers' surplus) in such a way that total industry profits are sufficient to cover fixed costs? The second issue concerns the discretion of firms in setting access prices. Should operators be allowed to choose access prices, for instance by negotiation? Or should these prices be regulated?

The first economic theory articles on access pricing in telecommunications focused mainly on symmetric, mature markets – in other words, operators are assumed to be identical in terms of networks, costs and market coverage, with the possible exception of an asymmetry introduced by assuming horizontal product differentiation (see Laffont et al., 1998, Armstrong, 1998, and Carter and Wright, 1999). Consider a duopoly, and suppose that access prices are set before the operators start to compete in the retail market. The operators compete by setting

linear prices (i.e. they charge customers a per-minute price only and not a subscription fee). Then, if access charges are chosen non-cooperatively by the operators, not surprisingly access prices turn out to be symmetric (and therefore also reciprocal) in equilibrium. Moreover, the operators have incentives to set the access price above the marginal access cost. This is because they benefit from raising their rival's cost, which generates higher profits without actually colluding in the retail market. By way of explanation, note that a high access price increases the cost of a unilateral retail price cut, because such a price cut increases the net outflow of calls to the rival operator, which is costly because of the access mark-up. In the absence of a joint and common cost, welfare is not maximised by these above-marginal cost access charges. Quite the opposite is in fact the case: welfare is maximised by setting the access price below marginal cost in order to compensate for the mark-up generated by the imperfect retail market competition (see Laffont and Tirole, 2000, ch. 5). In the presence of a joint and common cost, the optimal access price can be below or above marginal cost, because a mark-up on access contributes to cost recovery.

If operators compete in two-part tariffs (consisting of a subscription fee and a per-minute price), they have no incentive to charge an access mark-up. The reason is that, although a higher per-minute price (induced by an access mark-up) increases the per-customer profit, competition in subscription fees becomes more intense as a result of an access mark-up. In equilibrium, operators optimally set per-minute prices equal to marginal call costs, and recoup fixed costs on a per-customer basis through subscription fees. Welfare is maximised by setting the access price equal to the marginal cost of access (Laffont and Tirole, 2000, ch. 5).

An interesting observation in the literature is that, at least under certain circumstances, terminating access prices in a mature and symmetric market can be used in support of tacit collusion. On the other hand, this result depends heavily on the assumptions of the model (for a more extensive discussion, see Armstrong, 2001, and Laffont and Tirole, 2001, ch. 5). Overall, the risk of tacit collusion should not be neglected in a mature market, but it is not necessarily the most pressing problem facing policy makers in infant markets such as the current one for fixed telecommunications, given the assumptions under which the result is derived (in particular, the twin assumptions of symmetric operators and linear pricing). In the mobile telephony market, where operators are more symmetric and linear pricing schemes may be observed in practice, tacit collusion may be a policy concern.

In the fixed telephony market's current transition, most entrants are still relatively small when measured against the incumbents. A newcomer in the market, starting from scratch and slowly building up its market share, finds itself in a much more difficult position with regard to the incumbent than vice versa. In such a situation, asymmetric access regulation can be useful in the short term to protect small entrants as well as to prevent a reduction in consumer surpluses (see De Bijl and Peitz, 2002a, 2002b).[8] In particular, if an entrant is initially disadvantaged against an incumbent (e.g. in terms of market shares while there is customer lock-in, or because it lacks a quality track record and a reputation for quality), as is apparent in immature markets, it may be optimal to regulate the incumbent's access price at the marginal cost level (possibly allowing for a mark-up that yields a reasonable return), while allowing the entrant to charge a

substantial access mark-up. Such an asymmetric access mark-up increases the entrant's profits as well as consumers' surpluses. Welfare is barely affected, but the incumbent's profits are obviously reduced. This latter effect is undesirable in principle and should therefore not loom too large, since otherwise the incumbent may lose its incentives to invest. But within safe margins it is outweighed by the positive effects on consumer surpluses and the entrant's profits. Thus, non-reciprocal access prices may be temporarily useful to stimulate competition in an immature (and asymmetric) market. In the longer term, when the entrant has built up market share, cost-based access prices for both operators are optimal for consumer surpluses and welfare.

In summary: optimal policy with regard to terminating access prices is fundamentally different in infant and mature markets, a difference which is especially important in the light of the current fixed voice telephony transition phase. In an infant market, generally characterised by asymmetry between the incumbent and entrants, asymmetric access prices which favour entrants stimulate competition and consumer surpluses. In a mature and symmetric market however, asymmetric access regulation should be lifted, while oversight may be necessary to prevent tacit collusion through the access price.

The preceding discussion assumes that the 'calling party pays' principle (CPP) applies. And indeed CPP does apply in all European countries. Recent literature (Laffont et al., 2001) has shown that most of the problems mentioned above may disappear when the CPP principle is abolished. The reason is that a price is 'missing' under CPP, namely a price for receiving a call. As a consequence, all costs have to be incurred by the caller, while technically the costs

should somehow be split. One of the consequences is that operators have monopoly power over their incoming calls, which generates problems, e.g. in the fixed-mobile termination traffic (see section 4.1). Laffont et al. (2001) show that interconnection charges will be set at the competitive level if the missing price is recovered, i.e. when mobile operators are able to charge their customers for receiving calls. In contrast with the CPP case, perfect retail competition can exist and is viable. In the case of perfect competition, the access charge serves to determine how the cost of communication is split between the caller and the recipient, with the operators demonstrating indifference (since they achieve zero profits anyway). However, whenever there is imperfect competition (market power, albeit limited), operators' favoured access charges need not coincide with the socially optimal ones; unfortunately, the nature of the operators' bias is not that easy to predict.

3.3 Dynamic considerations

There are two issues relevant to the discussion between the short and the long term. First, since it has been pointed out that asymmetric access regulation is optimal in an infant market, an important issue queries the point at which the market can be considered to be sufficiently mature for asymmetric regulation to be lifted. Secondly, as with one-way access situations, any access price affects operators' profit levels, and hence their incentives to invest.

With regard to the first issue, if the regulator imposes asymmetric access prices to stimulate competition in the short term, a crucial question is: when does the transition phase end – in other words when are entrants sufficiently large so

that there is effective competition? For policy purposes, this is largely an empirical question. Clearly, the 'stopping rule' or 'sunset clause' should be based on observable market outcomes. In practice, policy makers can look at various indicators, such as the growth of entrants' market shares, reductions in retail prices, quality/price ratios and variety for end-users. However, there is a risk involved in making asymmetric regulation lifting dependent on market outcomes.[9] For instance, a criterion based on market share furnishes both the incumbent and entrants with incentives to compete less aggressively by increasing retail prices. By way of illustration, consider that the incumbent wants the entrant to gain market share more quickly so that it no longer incurs the access mark-up charged by the entrant, while the entrant wants to keep enjoying the access mark-up for as long as possible. To this end various regulators have been investigating the criteria for sunset clauses (e.g. Canadian Radio-Television and Telecommunications Commission, 2000, or OFTEL, 1999). However, it is not yet clear what all these dynamic regulation methods have yielded.

On the second issue, note that the policy prescription of asymmetric access regulation (as discussed above) does not conflict with the possibility that the regulator may want to introduce a bias towards facilities-based entry. Nevertheless, if the regulator explicitly wants to stimulate entrants into rolling out networks themselves, it is more direct to make other types of entry (Carrier Select, unbundled access to the local loop) increasingly unattractive. This may be sufficient to induce entrants who don't have their own local loops but who do have experience and market share, to start building them.

4 Policy Issues

This section covers four policy issues in more detail. These issues are fixed-mobile termination, margin squeeze, the allocation of common costs, and static versus dynamic efficiency. The issues are highly relevant for regulators at this stage of the telecom market's development. Considerable policy attention concerns the issues of fixed-mobile termination and margin squeeze. The allocation of common costs, and static versus dynamic efficiency issues, are both important general considerations in regulation.

This chapter is structured as follows. Each policy issue is preceded by a discussion of the policy's relevance, followed by an overview of what economic theory has to say. Occasionally this entails a summary from earlier chapters. The policy issues proceed with practical experiences and end with some policy conclusions. The policy conclusions combine insights from the earlier chapters with the practical experiences.

4.1 Fixed-mobile termination

4.1.1 Policy relevance

Fixed-mobile termination garners a lot of policy attention. It is an issue, which is the subject of a number of fixed network operators' complaints, while politicians also discuss the high fixed-mobile termination tariffs. In practice, one can concede

that these access prices are generally very high throughout Europe, compared to the cost of access. There is also a price asymmetry, i.e. the cost of being called is higher than that of making a call (although they place similar demands on the network). Once multiple mobile operators entered the market, the retail cost of making a mobile call decreased whereas the terminating access price decreased much more slowly, and price asymmetry increased. One might conclude that the fixed networks are subsidising the mobile networks. OPTA (2002) calculates that the revenue of fixed-mobile termination for 2000 and 2001 was about € 270 million above revenue based on cost orientation, i.e. a cross-subsidy of about € 270 million from the fixed networks to mobile networks. Furthermore, termination revenues are highly significant. For Telefonica, 75% of mobile revenue originated from termination and roaming services (OECD, 2000). This might give cause for regulators to intervene.

What intervention options are available? Regulators can decide not to intervene. One can expect the situation to continue unchanged because no alternatives for mobile termination are anticipated in the short term. Fixed subscribers subsidise mobile subscribers.

Regulators may decide to intervene to treat fixed-mobile termination as a competitive bottleneck and to regulate the terminating access tariffs. Different pricing principles can be applied to calculate cost-oriented tariffs (see section 2). These principles have to allocate the common costs across the different services, as the common costs need to be recovered somewhere. Ramsey pricing and LRIC are the two polar cases. When Ramsey pricing is applied, the problem may not be solved in a way, which satisfies the regulators (excessive prices).[10] Common costs

will be allocated based on the demand (elasticities). Demand in the retail market is more elastic than in the fixed-mobile terminating market. This implies that the largest segment of the common costs will be recovered in the terminating market, which itself implies that the current price asymmetry might be in line with Ramsey pricing. However, the low elasticity of the terminating access service would appear to be the result of market power. Fixed subscribers generally have no choice when they want to dial a mobile customer, so that the mobile operator on which these calls terminate possesses substantial market power. The regulator may wish to protect consumers from mobile operator market power abuse by regulating the terminating access prices (see e.g. Armstrong, 2001, section three). In other words, with given levels of market power and the opportunities for abusing market power, Ramsey pricing is a useful welfare notion. However, the regulator may have an interest in changing market power, which would also change elasticities and thereby Ramsey prices. The alternative to Ramsey is LRIC, commonly used by regulators in this situation. To recover the common costs, a fair share of these costs is often added to the price based on LRIC. The calculation of this fair share is somewhat arbitrary (see also section 4.3.2 on common costs).

4.1.2 Economic theory

When fixed subscribers want to reach mobile users, they have no choice but to pay the high access fee (as part of the overall per-minute price). The level of the fixed-mobile termination access price does not directly influence the level of the usage-based mobile retail prices, but forms part of the fixed cost related to the

mobile subscribers (e.g. lower subscription fee or subsidies on handsets). In essence, high access mark-ups imply that fixed customers subsidise mobile users. Mobile users benefit from high fixed-mobile access charges, to the extent that retail prices are subsidised.

Standard assumptions can show that welfare is maximised by making the terminating access prices equal to marginal costs (see Armstrong, 2001). However, it appears that no effective competitive pressure exists to set terminating prices at marginal costs. The policy discussion focuses on the extent to which different mechanisms restrict the MNO's ability to behave independently of its competitors, customers and ultimate consumers to an appreciable extent. In other words, does the MNO enjoy a dominant position or not? If the MNO is dominant and (also) sets excessive terminating access prices, intervention would seem to be called for.

How much do mobile consumers care about what it costs others to call them? One might argue that mobile users could also be contacted in other ways, for instance via a fixed connection. However this argument misses the point, which is that one generally tries to reach mobile users at their mobile phone when they are mobile, that is, away from home or office where there is a fixed line with a known phone number (or away from other contact alternatives). Even where there are substitution possibilities, they may not be significant enough to result in substantial downward access charges pressure.

One might also argue that market forces already exert sufficient discipline on mobile operators. For example, suppose that mobile customers derive benefit from being called (which is undoubtedly true in many cases). If consumers want

to receive more calls, they may care about the termination access price charged by their operator, and hence they may take the costs of inbound calls into account when choosing a mobile subscription. As a consequence, the presence of 'call externalities' gives mobile operators an incentive to reduce fixed-mobile termination access prices, although it is unclear how strong this effect is in practice. Armstrong (2001) actually shows that if mobile subscribers derive benefit from incoming calls, then welfare is maximised by setting termination charges below the marginal cost of access. This is because access below cost encourages fixed subscribers to make calls to mobile subscribers. Another way in which consumers may take the costs of inbound calls into account occurs, if one assumes that they care about the costs incurred by people calling them. In theory, if mobile customers allow for the welfare of people who will call them when choosing a mobile operator, this again provides the incentive for mobile operators to reduce termination fees (Armstrong, 2001). This scenario does have some grounding in reality, since people most often receive calls from family members or direct colleagues.

Ultimately, it is empirical as to how much mobile customers care about the costs of incoming calls. Existing evidence (e.g. OFTEL 2001) together with casual observation (the rates are indeed very high throughout Europe) suggest that it is most likely that mobile customers don't take much notice of the costs incurred by people calling them when choosing a subscription.

As the current system of 'Calling Party Pays' (CPP) and customer pressure does not appear to exert enough downward pressure on the terminating access prices, one might choose to change the payment principle. The alternative is to let

mobile operators charge their customers for incoming calls. In other words, call recipients pay a per-minute price to receive calls from fixed subscribers (instead of the CPP principle). This remedy removes the bottleneck problem (given that the mobile operator no longer charges the fixed operator for the termination), since customers who have to pay for incoming calls will be more sensitive to the prices levied for that service.[11] Changing the system may create a distortion, i.e. subscribers pay too little for calls to mobile users if the costs between callers and recipients are split for fixed-mobile but not for fixed-fixed. It is not clear how serious this distortion would be, but to solve it one might also consider introducing the split for other services.

The conclusion is that policy intervention appears to be necessary. There might indeed be some mechanisms available to restrict MNO behaviour, but the effect is too small to conclude that there is effective competition in the fixed-mobile termination market. In the short term, most research indicates that these mechanisms will not increase sufficiently in strength to create an effective competition situation (see e.g. Armstrong 2001).

4.1.3 Practical experiences

Central to the debate on fixed-mobile termination is whether fixed-mobile call termination is a ‘bottleneck’ and whether Mobile Network Operators (MNOs) can exercise market power. There is significant policy discussion between regulators and market players on this very point. The principle that the ‘calling party pays’ is central to this discussion. Arguments can be given as to why there is no (or a limited) effect on the extent of the terminating access prices versus arguments as

to why there is enough pressure on MNOs to compete on the terminating access prices (see e.g. OFTEL, 2001, IRG, 2002). Despite the results of this discussion, Armstrong showed that in most cases the pressure from these mechanisms is too limited to restrict MNO behaviour in setting their termination access prices. The high terminating access prices appear to support this argument. Most regulators also follow this line of reasoning, and intervene.

OFTEL (2001) discusses different mechanisms, which might constrain MNO behaviour. The first mechanism is that mobile subscribers might care about the costs of calling them and take the termination prices into account when choosing an MNO. The conclusion was that mobile callers pay little attention to terminating prices (this might be different for different groups). Secondly, the effect of closed user groups[12] generates a limited pressure on termination prices. In addition, OFTEL concludes that there are no real alternatives and that buyers have no countervailing power. The overall conclusion is that there is insufficient competitive pressure to constrain the terminating prices, and that pressure is expected to remain insufficient.

OPTA uses more or less the same arguments as OFTEL, and reaches the conclusion that mobile call termination is a bottleneck facility. OPTA decided to intervene and to set a timeframe to implement cost-oriented terminating access tariffs in 2003. In the meantime, OPTA suggested what it considered to be fair terminating prices.

In calculating terminating access prices, regulators use the LRIC pricing principle, or are planning to use it in the near future. LRIC is preferred over Ramsey pricing because Ramsey pricing might be difficult to calculate and might

result in an unfair allocation of common cost (cross-subsidisation and/or major price asymmetries). Common costs also seem to be a relatively small problem (see section 4.3 on common costs).

4.1.4 Policy conclusions

The policy conclusions combine insights from the earlier chapters with the practical experiences from the above.

- Fixed-mobile termination is a hot topic in the telecom sector and is the focus of much regulator attention. So far, the conclusion of most regulators is that fixed-mobile termination is (virtually) a bottleneck facility. There are too few competitive pressures to restrain MNO behaviour, thus regulation is considered to be necessary. However, it is important to review the development of potential mechanisms which might bring about effective competition. For example, it is unclear how closed user groups will develop and to what extent they might restrict MNO behaviour.
- Another important issue is the potential of alternatives to the 'calling party pays' (CPP) principle. CPP is common practice in Europe. CPP could be exchanged for 'receiving party pays' (RPP) or a combination of both. RPP can put effective pressure on the call termination prices because the individual who chooses the network operator is charged for both outgoing and incoming calls. The individual has the option of switching to a different network operator offering better charges. Countries with CPP have significantly higher terminating access prices than countries with RPP (OECD, 2000: 50). Thus

RPP might bring about lower terminating access prices. On the other hand, RPP might also have disadvantages. Prepaid could become less attractive. The subscriber has to pay not only for outgoing calls, but for incoming ones as well – calls over which he or she has no control (other than disconnecting the mobile). It will be more difficult for a prepaid subscriber to keep a rein on telephone expenses. It is also argued that CPP exhibits more potential for helping the mobile market to grow (OECD, 2000). Under CPP it becomes increasingly desirable to join a network, because receiving calls is 'free', while prepaid cards are more attractive in markets with CPP. Finally, with the new proposed 'open network provision' (ONP), it will be easier to designate MNOs with significant market power in narrower markets (e.g. mobile termination) and not to intervene in the more competitive retail market.

- Competitive pressures might also be introduced by offering consumers more options. For instance, it might be possible to offer consumers the possibility of choosing a different network for receiving calls than the one used for making calls. The mobile subscriber could select the network operator offering the lowest originating charge and another with the lowest terminating charge. This implies the creation of two separate markets. It is not clear to what extent operators focusing on a separate market (e.g. only terminating) can compete with operators providing a package of originating and terminating services. For this option, it would also be a requirement that SIM cards could communicate with all mobile networks. To be effective, it would also be necessary for the mobile subscriber to take the costs of the calling party into account.

- The calling party might also be able to call a mobile subscriber by using different MNOs. The mobile subscriber might have multiple mobile phones (not very realistic) or the mobile subscriber could be reached via various MNOs. This offers MNOs an incentive to compete on terminating access prices. This last option also assumes that the SIM cards could communicate across different mobile networks. It implies that the mobile subscriber owns the SIM card instead of the MNO, and that the networks need to be highly standardised. However it is unclear whether this could be either possible, or desirable, in the future.
- Increasing mobile subscribers' awareness of terminating access costs might increase their willingness to take these terminating costs into account. Transparency on the pricing structures is a tool to this end. To be effective, fixed network operators need to define their retail tariffs for the various MNOs.

4.2 Margin squeeze

4.2.1 Policy relevance

Telecom operators or independent service providers may rely on the key input (upstream market) of the vertically integrated operator, while at the same time competing with that operator in the downstream market. If the vertically integrated operator is dominant in the upstream market, there could be scope for it to leverage its dominant position into the downstream market, resulting in a margin or price squeeze.

A margin squeeze occurs if the retail prices charged by the dominant operator are so close to the wholesale prices of services offered to competitors, that even a more efficient competitor cannot enter into or survive in that market. The vertically integrated operator can subject its competitors to high access prices (raising the cost of the key input) and/or lower its prices in the downstream market. The vertically integrated operator's total revenue may remain unchanged. In extreme cases, the access price may even be higher than the incumbents' retail prices. In a margin squeeze situation, there is too little room for competitors to compete. Thus a margin squeeze can harm the development of effective competition over the long term (relatively high prices, low quality, etc).

The lack of (efficient) entrants or competitors does not necessarily reduce welfare. If the incumbent operates more efficiently than its (potential) competitors and the customers reap the benefits, the lack of entrants/competitors is not an issue. If this is not the case, then a margin squeeze is indeed a problem, because efficient competitors do not enter into or stay in the market and competitive forces cannot do their job.

More formally, a margin squeeze can be defined thus (given that access and retail services are strictly comparable):[13]

$$a + c > p$$

where:

a = access price charged by incumbent to entrants;

p = incumbent's corresponding retail price;

c = incumbent's cost of delivering the retail service, on top of costs already included in access price a.

This implies that:
the incumbent would make losses if it buys access at price a, given retail price p; p – a does not reflect the incumbent's own downstream cost c; the 'notional downstream margin' on the incumbent's retail activities is negative, $p – (a+c) < 0$.

Is the possibility of a margin squeeze sufficient reason to intervene as a regulator? Is a margin squeeze part of the process of becoming a competitive industry? Is a margin squeeze the result of normal business practices or anti-competitive behaviour? To assess the need for (potential) intervention, it is worthwhile to look at what economic theory has to say about margin squeeze.

4.2.2 Economic theory

The critical point in a margin-squeeze case is that an efficient (not vertically integrated) competitor cannot enter into or stay in the market and earn a normal profit because of the pricing behaviour of a (dominant) vertically integrated incumbent. In order to understand any margin squeeze case and the potential remedies, one should first consider its causes. There can be three different instances, which may give rise to a margin squeeze:

a. access price is too high (and competitive retail price); risk of exploitative abuse

b. retail price is too low (and cost-based access price); possibly predatory pricing;
c. margin squeeze between the prices above; possibly cross-subsidisation.

These issues will be discussed in greater detail.

Access price is too high: risk of exploitative abuse

When it comes to excessive access prices, it is vital to consider the reason for the level of the access price and its relation to the underlying costs. Excessive access prices are characterised by the absence of a reasonable relationship between the price and the economic value of the product or service supplied. If there is a large gap between the price and underlying costs, profits are higher than might be anticipated in a competitive setting. There may be various reasons for excessive access prices. Access prices may be too high as the result of a dominant position. From an access perspective, this is a one-way access situation (see section 2). On the other hand, access prices may be too high because of conflicting regulation practice objectives.

In theory, excessive access prices at the upstream market are not a problem if this attracts new entrants into the network market. This depends on this market's entry barriers. Particularly in the local loop, entry barriers are high and no disciplinary forces can be expected in the short term. Regulators might thus opt to intervene.

Access prices may also be too high even though this is not an intentional abuse by the incumbent. Price-regulation rules set by the regulators could result in excessive access prices (in relation to retail prices). This could result from

conflicting objectives and/or different pricing principles for the access and retail markets. The objective of efficient investment (access price based on cost) might contradict the objective of creating effective entry and competition at the retail level (retail minus). In such a situation, the regulator has to set priorities: not all objectives can be achieved simultaneously and some objectives are more important than others.

Retail price is too low: possibly predatory pricing

Various causes can result in retail prices being too low. Retail prices might be too low because of different pricing rules charged by regulators on the wholesale and retail market, (compared to the access price). A uniform national retail price might also be too low compared to cost-based access prices in a rural area. On the other hand, low retail prices can also result from an intentional abuse by the incumbent, i.e. a special case of predatory pricing.

A margin squeeze as a particular form of predatory pricing implies that an efficient entrant or competitor cannot redeem its downstream costs, and that it will not enter or leave the market. The incumbent opts for a short-term loss in order to make long-term extra profits thanks to a dominant position.

Predatory pricing implies that there is a price reduction which is profitable only because of the added market power the predator gains from eliminating, disciplining or otherwise inhibiting the competitive conduct of a rival or potential rival (Bolton et al., 2000). In the short term customers may benefit from lower prices, but over a longer period weakened competition will lead to higher prices, lower quality or less choice. The fact that an activity is being run at a loss, is not

sufficient to establish a case of predatory pricing; the question is whether it has an anti-competitive effect. In order to prove the anti-competitive effect of predatory pricing, Bolton et al. (2000) propose a five-criteria rule:

1. a facilitating market structure,
2. a scheme of predation and supporting evidence,
3. probable recoupment,
4. price below cost and
5. the absence of efficiencies or business justification defence.

subnote 1) The market structure must make predation a feasible strategy. A company must have the power to raise prices (or to otherwise exploit consumers or suppliers) over some significant period of time (dominant firm or small group of jointly acting firms, entry and re-entry barriers).

subnote 2) Predation pricing and recoupment require that predation is plausible ex ante (i.e. based on prediction and extrapolation) and probable ex post (i.e. retrospectively). This means there must be a predatory scheme ex ante under which the predator can expect to recoup its initial losses. Using the tools of applied game theory can assist in identifying economic conditions under which predation is rational profit-seeking conduct on the part of a dominant firm. Ex post probability is shown by the subsequent exclusion of rivals and post-predation market conditions making future recoupment likely.

subnote 3) At the very least, losses incurred in a predation strategy must be recouped somehow. Should the operator be unable to recoup the losses, because of competition from existing or potential competitors, the predation strategy is not viable. Recoupment is only possible if there is an exclusionary effect on

(potential) rivals, or through disciplining the rival's competitive conduct. The most common and straightforward recoupment occurs when prices rise above the predatory market's competitive level. In more complex settings, recoupment can occur through other channels, e.g. by raising the prices of complementary or closely-related services. It is essential that these latter price increases should be explained unambiguously by the earlier predatory pricing (see also Cabral and Riordan 1997).

subnote 4) In the predatory period, prices should be below average variable cost, although prices which are also above average variable cost but below average total cost might be predatory and injurious to competition[14]. The most-used cost standards are average total cost (ATC) and average variable cost (AVC) (OFT, 1999) or long-run average incremental cost (LRAIC) as a substitute for ATC and average avoidable cost as a substitute for AVC (Bolton et al., 2000). If prices are above ATC, there is no problem. If prices are below AVC, predation can be assumed. A price between ATC and AVC is either presumptively or conclusively legal. If the price is presumptively legal, there is a need for evidence that the operator intends to eliminate or to discipline a competitor.

subnote 5) Finally, cases can arise where below-cost pricing by a dominant operator might be efficiency-enhancing rather than predatory. However in such cases one has to consider very carefully whether the efficiency enhancement is also to the benefit of consumers in the long term. Otherwise the argument could be abused to foreclose a market on the grounds that it is 'efficient' to do so.

The five-criteria rule provides a clear procedure for handling a potential predatory pricing case. However predatory pricing might be difficult to prove,

particularly the intentional aspect. The section on policy conclusions will include a more detailed discussion of this rule's practical value.

Margin squeeze between access price and retail price: possibly cross-subsidisation

Cross-subsidisation occurs where an operator uses revenues from one market to subsidise losses in another market. If the revenues derive from a market in which the operator has a dominant position and are used to subsidise losses in a competitive market, this hampers competition and might be illegal. In the case of cross-subsidisation, losses in one market are redeemed by another market. Cross-subsidisation is particularly viable if the market is characterised by economies of both scale and scope, that is to say, scale and scope economies broaden the opportunities for incumbents to use part of their business to cross-subsidise other parts. For example, causally attributable and avoidable service costs can be labelled as common costs resulting in too low incremental costs. If the price in such a situation is set according to incremental cost, then a case of cross-subsidisation exists (see also section 4.3 on common costs).

There is cross-subsidisation if revenues are expected not to cover the cost of the associated activities over their economic lifetime. If the costs are indeed covered over the service's economic lifetime, relatively low prices when introducing the service could be a normal business strategy (penetration strategy), and thus do not constitute anti-competitive behaviour. For example, if the production of a service is characterised by economies of scale or learning effects, an appropriate business strategy might be to offer the initial service below cost price (for initial customers). As demand grows significantly because of the low

price, this will yield a sharp decrease in production costs. This option might yield a higher profit over the service's economic lifetime (low price, large volumes and substantially lower costs) than the alternative of a high price and a slower drop in costs.

Assessing cross-subsidisation relates to predatory pricing assessment. If revenues exceed the long-term incremental cost (including the cost of capital), the service is sustainable in the long term and there is no cross-subsidisation (OFTEL, 2000). If services share common costs, the evaluation is more complex (see also section 4.4). In this instance, total costs consist of the long-term incremental costs of all services and the common costs. In the case of cross-subsidisation, a dominant operator might charge a higher price to cover its common costs in the market in which it enjoys dominance, using a price, which equals the incremental costs in the competitive segment. Although this is not illegal in terms of predation, it might hamper competition. A competitor who does not offer all services available from the dominant operator might be at a cost disadvantage in the competitive market (i.e. it cannot recover the common cost from another market).

Economic literature focuses on two tests for cross-subsidisation: the incremental cost test and the stand-alone test (Braeutigam, 1989: 1337-1342, Sharkey, 1982, Faulhaber, 1975, as cited in Nicolaides and Polmans, 1999). The incremental cost test requires that the revenues from each subset S at least cover the increment to the total costs, which occurs when S is produced, as opposed to not being produced at all. If the revenues of S do not cover the incremental cost of S, then service S is subsidised. The incremental cost test sets the lower boundary

on the subsidy-free price range. The stand-alone test argues that if the revenues of S exceed the costs of providing service S alone, then users of S are subsidising other services. The stand-alone test sets an upper level on the revenues generated by S. The tests must be done for all possible subsets[15] (often referred to as the combinatorial cost test). If the price is between the average incremental cost and the average stand-alone cost, then the activity is neither a subsidy source nor recipient.

Economic theory produces important insights into margin squeeze, its causes and ways of assessing different forms. How can such theoretical insights be interpreted in real-life cases?

4.2.3 Practical experiences

Until recently, regulation and intervention by regulators focused on preventing excessive prices (e.g. price caps). Over the past two to three years, regulators have devoted more attention to the potential harmful effects of low prices. As a result of complaints from entrants and new competitors, it became more obvious that low competitive prices have a hampering effect. However, the distinction must be drawn between low prices resulting from predatory behaviour, and low prices resulting from legitimate competitive behaviour. In case law, [Dutch manufacturer] Akzo versus the European Commission is the leading example (ECJ, 1991). The European Court of Justice argued that “prices below average variable costs (that is those which vary depending on the quantities produced) by means of which a dominant undertaking seeks to eliminate a competitor, must be regarded as abusive” (paragraph 71), and “Moreover, prices below average total

cost, that is to say, fixed cost plus variable costs, but above average costs, must be regarded as abusive if they are determined as part of a plan to eliminate a competitor" (paragraph 72). In the first situation, the ECJ found the conduct of a dominant company to be an abuse without explicitly considering whether recouping losses would be feasible. Thus if the accused company is dominant, a case of predation does not require explicit analysis of the predatory scheme (see ad 2 and 3, page 31).

Interest in margin squeeze increased as a result of the potential problems incurred in Local Loop Unbundling (European Commission, 2001). It must be emphasised that margin squeeze does not necessarily imply intentional abuse. Margin squeeze can be the result of an incomplete price rebalancing. If prices are not completely rebalanced on a cost-oriented basis, such prices do not adequately represent the underlying costs. This may result in overcharging or underpricing. The Commission can launch infringement proceedings against member states to counter this problem.

Accounting inconsistencies may also result in margin squeeze. Certain pricing principles focus on static efficiency (e.g. Ramsey pricing), whereas others focus on dynamic efficiency (e.g. LRIC). National uniform retail prices could conflict with differences in the regional costs of providing access. Pricing principles may conflict and may result in margin squeeze.

Most regulators studied the effects of price squeezing and suggested price-squeeze tests when approving the incumbent's interconnection tariffs (e.g. OFTEL, 2000, OPTA and NMa, 2001). This ex ante regulation can be interpreted as part of the regulators' cost-orientation check. In a price-squeeze test, the costs

of the incumbent are studied, more specifically the costs that the incumbent incurs to provide the service internally. The extra costs the incumbent incurs to deliver the access service (e.g. billing) should be omitted from the test. In determining the access tariffs, most regulators use LRIC, or plan to use LRIC in the near future.

A problem could arise when using LRIC in a sector with high common costs – the telecommunications sector being just such a one. A multi-product supplier may choose the market in which to recover these common costs, provided it does not abuse its dominant position. If all services are priced at LRIC, the common costs are not recovered. To test that there is no case of predation, OFTEL tests whether the service prices are at least at LRIC level plus a fair share of the common costs. If a dominant operator is pricing below LRIC, predation is presumed, unless the operator provides solid arguments to the contrary. In addition, all the relevant services should be able to recover all common costs together (combinatorial test, OFTEL, 2000). If the dominant operator prices above LRIC, but overall revenue fails to cover total costs and the dominant operator has the intention of eliminating a competitor, predatory pricing is presumed.

OPTA and NMa (2001) published price-squeeze guidelines. These guidelines explain how the price-squeeze test is implemented. The test could be applied to all the dominant operator's retail products, and will be applied to the various retail price components (set-up cost, peak, off-peak, etc.). It will be difficult for entrants to cherry-pick as a result of this detailed approach..

Another issue needing attention is the allocation of costs (costs related to access versus those related to retail activities, see also section 4.3 on common costs). An incumbent might allocate costs related to retail activities, to access

activities. As a result, access costs are too high and retail costs are too low. If access prices are cost-oriented based on these (high) access costs, then the entrant is paying too much (i.e. part of the incumbent's retail costs). Thus when setting access prices based on an incumbent's information, the cost relevance must be addressed.

In conclusion: most regulators recognise the price squeeze problem. They have developed or are planning to develop a price-squeeze test. This instrument is used as part of the cost-orientation check. As this instrument is relatively new, it is too soon to draw conclusions on the effectiveness of price-squeeze tests.

4.2.4 Policy conclusions

Theoretical insights require a practical interpretation. A margin squeeze can be difficult to assess, proper data is difficult to obtain and interventions can also come at a cost. Mistaking competitive pricing for a margin squeeze will tend to distort the market. On the other hand, mistaking predation for competition may result in less competition and higher prices over the longer term.

It is vital to look at the margin squeeze cause in order to minimise these mistakes in practice. Depending on the cause and the effect on competition, regulators might decide to intervene. Thus the regulators should first determine the actual cause of the margin squeeze. The discussion in the theoretical section can be an important starting point. Some issues do need special attention. These are: the time dimension, allocation of (common) costs, information asymmetry, and ex ante regulation versus ex post intervention.

- Time is an important issue in margin-squeeze cases caused by predatory pricing and cross-subsidisation. In a predatory pricing case, time refers to the period required for recoupment. The period in which recoupment is possible is likely to be short, particularly in the telecom sector with its rapid technological development. On the other hand, path dependence might imply that once a company has missed a stage, it will be difficult to make up the disadvantage. Entry barriers for new technologies might also increase the period for recoupment. In a cross-subsidisation case, time is important to evaluate whether certain business practices constitute anti-competitive behaviour or not. For normal business practice, revenues over the economic lifetime of the service should cover the costs. If the costs are covered, the operator uses a penetration price to generate a large custom base quickly. If costs are not covered over the service's economic lifetime, there might be a case of cross-subsidisation.[16]
- A margin squeeze case can be very complex if common costs are involved. If the revenues of a service are higher than the LRIC plus the common costs, then there might be a case of cross-subsidisation. By using the incremental cost test and the stand-alone test one can identify whether there is indeed a case of cross-subsidisation. Predatory pricing is also more difficult to prove when common costs are involved. Part of the problem is related to the allocation of the common costs across the relevant services.
- It is clear that a margin squeeze case caused by predation or cross-subsidization is difficult to prove. It is difficult to obtain all the necessary cost information, and to allocate the costs to the relevant activities. If a regulator

cannot perfectly oversee the incumbent's cost levels, the incumbent may try to hide such a margin squeeze by allocating costs to its wholesale activities, which are actually incurred as a result of its retail activities. Accordingly, in practice, the risk of a margin squeeze is closely linked to information asymmetry. Secondly, it is difficult to prove operators' intentions. The likelihood of a specific case satisfying all criteria is limited .

- Finally, there is a trade-off between ex ante regulation versus ex post intervention. An ex ante margin-squeeze test clearly helps to make regulation transparent, and creates clarity for market parties. It can stimulate more market competition. On the other hand, it can also result in over-regulation or needless regulation, particularly if the price-squeeze test is very detailed. Margin-squeeze cases can also be handled ex post, with the risk that intervention is too late (complaints procedures take a very long time). Regulators have to make a trade-off between possibly needless regulation and the risk of intervention, which is too late.

4.3 The allocation of common costs

4.3.1 Policy relevance

One of the toughest regulation problems concerns the allocation of common costs. Technically speaking, entrants only use part of the infrastructure when leasing a line or using only parts of the spectrum, for example. However, access would be of no use if the other parts were not there (i.e. the parts of the infrastructure it does not directly use). In other words, there are parts of the infrastructure, which are 'used' for various services, but it is unclear in just what proportion the costs of that part of the infrastructure should be allocated across those services. A related but slightly different issue occurs when the costs do not depend on the amount of the bandwidth actually used or the number of services, which run on it. Finally, there are also often provisions for public services to be shared among the various services. An intervening regulator has to come up with an answer by deciding on some method of allocating the common costs. Misallocating these common costs can distort competition.

4.3.2 Economic theory

The necessity of financing part of the common costs is a problem reasonably well covered in the economic literature (see e.g. Armstrong 2001 or Laffont and Tirole 2000). At the same time, it is acknowledged there is no easy way of doing it. Determining access charges consists typically of two stages. In stage one, the regulator determines the causally attributable costs (see section 2 for various

methods). Stage two then tackles the common costs. The lack of allocating appropriate costs in stage one implies that the residual would also contain costs which might otherwise have been allocated as causally attributable costs, but which will be allocated as if they were common costs. The question then becomes: how does one allocate these common costs?

Theoretically there is an optimal way (Ramsey pricing, see section 2.2.2). In short, common costs are allocated inversely proportional to the service's price elasticity, i.e. the service with the lowest price elasticity carries the highest burden. It will be recalled that there are three reasons why Ramsey prices are rarely used in practice. The first is political: Ramsey prices often involve a (very) skewed distribution of prices across services and may imply sky-high prices for some services, which may be politically unattractive. The second is a practical one: to calculate Ramsey prices, one may need very detailed information, which is often unavailable. The third is that Ramsey prices do not allow for clear comparisons between countries or operators. Such comparisons could be useful for benchmarking (see e.g. Jullien 2001).

Because of these practical difficulties, regulators had to devise alternative approaches (see DTe 2000, Europe Economics 2001, OFTEL 2001). Four are detailed below:

- Equal proportionate mark-up (EPMU)

The mark-up for common costs is proportional to the incremental costs of the service provided. It differs from Ramsey prices in that the willingness to pay does

not count – only the costs. Thus this method neglects demand-side factors and scores worse than Ramsey pricing on allocative efficiency. In cases where common costs are small relative to incremental costs, this distortion is modest.

- Pro-rata appointment

The mark-up on common costs is decided as a fixed proportion between various services. The ratio can be decided by the regulator. This method's advantage is that it is very simple to implement, and that it can easily be tailored to a specific goal by the regulator. However, unless the regulator comes up with plausible arguments, the proportions are set in an arbitrary manner.

- Incremental and stand-alone costs

This method is based on the calculation of two extremes. The boundaries of the costs to be allocated to any service are determined by:

- a floor, the incremental cost — the cost which would be avoided were that service not provided, and
- a ceiling, the stand-alone cost — the cost that would be incurred if that service were provided in isolation.

The difference between these two boundaries represents the level of common costs. With pricing at incremental cost, the service concerned makes no

contribution to common costs; with pricing at stand-alone cost, the service concerned bears the totality of common costs. At intermediate prices, the service makes some contribution towards common costs. This method is particularly useful in casting light on what causes costs. The method provides an estimate of the magnitude of common costs, which is independent of the accounting data and thus could be applied as a cross-check in identifying attributable and common costs. As a method of determining some intermediate level, the same criticism is applicable as to the pro-rata method.

- Commercial negotiation

The regulator sets prices, which are consistent with a hypothetical commercial negotiation between two independent parties in advance of the expenditure being incurred (Europe Economics 2001). The advantage of this method is that it takes demand-side considerations into account, but is not likely to lead to the extreme Ramsey outcomes, nor does it require such a high information burden. The downside is that the outcome of this method is not clear a priori. At worst, the method is as arbitrary as pro-rata appointment.

4.3.3 Practical experiences

Before considering the main telecommunication experiences, it should be observed that other utilities face quite similar problems. Worthy of mention below are the examples of an airport and electricity supply.

- Airports UK

In a number of European airports, regulators are considering a so-called ‘dual-till’ system. Under a dual till, airport activities would be divided into ‘aeronautical’ and ‘non-aeronautical’ activities, and only costs and revenues in the former category, termed the ‘regulatory till’, would be taken into account when setting the price control for regulated charges (Europe Economics, 2001). The UK has the longest and best-documented experience with this, hence the UK focus here.

Europe Economics (2001), in advising the regulator on matters such as common costs, rejects Ramsey pricing for practical and political reasons, such as those discussed above. It rejects the EPMU, because it would cause profitable activities to cease (or prevent others from being introduced). Here the absence of demand-side considerations takes its toll. As a compromise, Europe Economics (2001) suggests the commercial-negotiation principle. In practice, the method boils down to a variant of EPMU, but corrected for demand-side factors.

- Electricity in the Netherlands

The Dutch regulator considers ‘the local electricity loop’ as regional natural monopolies and regulates access to the local loop via yardstick competition. Naturally the regional structure of the electricity networks helps the regulator in applying such yardstick competition. But even TenneT, the national grid company, is benchmarked against best-practice grids in the USA and Europe (DTe, 2000). Yardstick competition implies that there is no relationship between

access charges and the costs of the network. Access charges are set based on the best-practice performance in the country (or elsewhere). Yardstick competition obviously avoids any discussion on common costs. The applicability of yardstick competition to telecommunications depends on just how comparable accounting and regulation principles are between European countries. If they are indeed comparable, then yardstick competition could be a useful alternative to other methods of price-cap regulation. A recent example is the Local Loop Unbundling in Ireland, where tariffs are based on international comparisons (Office of the Director of Telecommunication Regulation Ireland, 2001).

Moving on to the two main telecommunications applications:

- Unbundled Local Loop/shared access in the UK

Unbundled Local Loop (ULL) enables competing operators to lease the fixed-line incumbents' local loop and to install equipment at their exchange sites. The main implication is that it enables these operators to offer high bandwidth directly to consumers and thus fosters competition for high-bandwidth services. ULL is a typical one-way access situation. Section 2 has already discussed access charges with ULL, while leaving the common-cost aspect until this section. One way of reducing additional consumer costs for hiring these new operators with their high bandwidth is to introduce shared access (see e.g., EC Commission Recommendation C2000 1059, OFTEL). An issue specific to shared access is how to split the common costs between high and low frequencies.

OFTEL has suggested a method for sharing common costs in the case of shared access (OFTEL, 2000). It first notes that shared loops differ from fully unbundled local loops, in that the competing firms only lease a portion of the local loop – the high frequency one. So after applying the attributable costs in step one (LRIC is used in this instance), in step two it must be determined just what to do with the costs of the loop which are invariant with the bandwidth used.

OFTEL rejects Ramsey pricing for the usual reason: the information burden is too high. It goes on to state that its objective is to adopt a method, which ensures that the take-up of higher-bandwidth services is neither discouraged nor deterred, i.e. allocative efficiency in DSL services appears to be the main goal. OFTEL concludes that the easiest way to achieve these goals is to let the local loop provider (LLP) set the access charges itself so long as there is non-discrimination, i.e. the LLP cannot charge competitors a different amount to that which it itself charges. Since the LLPs do not charge any of the common costs to their own downstream higher-bandwidth business, OFTEL suggests that the total of common costs will be allocated to voice.

- Fixed-mobile termination

It has been concluded from sections 3 and 4.1 that fixed to mobile termination is a service which should be regulated since mobile network operators (MNOs) have a near monopoly position in this service. Indeed, a number of regulators have adopted RPI-X regulation to prevent MNOs from exploiting their near monopoly

position (see e.g. OFTEL, Review of the Charge Control on Calls to Mobiles – 26 September 2001).

OFTEL opts for the EPMU system for the common costs (Kobold and Maldoom, 2001). It endorses this choice since the recovery of these costs is a relatively minor issue in the context of mobile termination, because a mobile network's common costs are relatively small. In OFTEL's cost model, common costs amount to only about 3%-5% of the total costs of a 900 MHz and 1800 MHz network. The common costs comprise the cost of a network-management system plus the fixed costs of coverage, i.e. the cost of acquiring, renting and operating the number of sites required to provide coverage. The characteristics of a 1800 MHz spectrum mean that more base stations are required to provide the same degree of coverage, which is why common costs are a slightly higher proportion of total costs than in a 900 MHz network. As a consequence, EPMU distortion is modest. Since EPMU is easy to implement, the choice is clear.

4.3.4 Policy conclusions

- Ramsey pricing, though theoretically appealing, is often discarded as being impractical or politically unfeasible. Nonetheless, Ramsey pricing's message is that demand-side factors can be important, in particular in situations where the common costs are fairly high relative to incremental costs.
- When common costs are high, Ramsey prices, or more generally demand-side factors, should not be readily dismissed. Difficult data requirements can sometimes be resolved. Imperfect demand estimations can be better than no

demand-side estimations at all. Political (income-distribution) problems can also often be resolved..

- In cases where Ramsey pricing is not feasible, the regulators will find themselves at a crossroads since none of the options can be deemed to be superior to the other, a priori.
- The two leading questions in determining the best available alternative are: (i) to what type and level of distortion does the method lead? (ii) which are the main objectives the regulator wants to achieve with the access charges?
- The first question demonstrates sensitivity to the relative damage that the method inflicts: in the case of fixed mobile termination, for example, we saw that the common costs were relatively small, implying that an alternative such as EPMU only causes a small distortion.
- The second question gives weight to the distortions: one distortion is not as bad as the other. In the airline example, we saw that demand-side considerations were considered to be very important, which made EPMU unappealing.
- A totally different perspective emanated from the electricity market. Indeed, one can wonder how international comparison could be a feasible future option in the telecommunications field, whether by benchmarking or even by the more formal companion to benchmarking, yardstick competition.

4.4 Static versus dynamic efficiency

4.4.1 Policy relevance

Sections 2 and 3 devoted space to the dynamic consequences of access prices. Because of the policy relevance in virtually all access problems, and the lack of a theoretical consensus in general, as well as the lack of a well developed analysis of the linkage between access pricing and dynamic efficiency, this section reverts to the issue in more detail. The notions of static and dynamic efficiency are used throughout this section; see the box below. This enables the possible trade-off between short- and long-term policy choice implications to be addressed.[17]

Static and dynamic efficiency

Roughly speaking, static efficiency is high if competition is sufficiently intense, if there is downward pressure on prices, if consumers can choose between several suppliers, and if they get good value for money. Some static efficiency drivers are: low entry barriers, no collusion and no substantial consumer switching costs. Unfortunately one cannot measure the static efficiency of a market, but in practice it can be roughly assessed by consulting industry experts, antitrust economists and consumer organisations which may each have specific views on certain indicators related to static efficiency drivers.

Dynamic efficiency is high if there are low entry barriers for new technologies developed by rival firms, and if firms have incentives to invest in R&D and innovation. Some dynamic efficiency drivers are: low entry barriers for players with new technologies, the possibility of recouping R&D investments, little regulatory uncertainty, and effective standardisation. Again, it is very difficult to measure dynamic efficiency, but in practice one can look at the levels of investment in product and process innovation, R&D budgets, the number of patents, the prospects for new players introducing new technologies, and consumer take-up of new products. Such indicators may provide a good approximation of dynamic efficiency. It is important to stress that a high level of innovation does not always mean that dynamic efficiency is high. There can also be too much innovation in a market, for example in a situation of planned obsolescence.

Why do regulators struggle with this issue? There are several explanations. First, theory is well developed for static access rules, not for dynamic ones. Second, while the emphasis was on developing competition for services in the early stages of liberalisation, attention has shifted (at least to a certain extent) to competition for infrastructures. Third, there could be trade-offs between static and dynamic efficiency. Regulating access prices, which foster service competition, does not automatically foster infrastructure competition. The well-known example here is low one-way access prices: entrants who can ‘free-ride’ on the incumbent’s network do not have incentives to build their own networks. Fourth, partly because of the low market sentiments for telecommunication, investments

in new infrastructure are not as booming as expected (with some notable exceptions). Fifth, there is considerable political pressure on regulators to go easy on telecommunications operators because tight regulation is allegedly harmful to investment, particularly given current market sentiments. Finally, the emphasis on dynamic considerations cannot be overstated. The telecommunications industry is one, which is driven largely by innovation and investments in new infrastructure. So putting right the investment and innovation incentives can easily override any of the static considerations. This notion is further reinforced when considering a regulatory time path. Bennett et al. (2001) showed that if investment and/or innovation can create a breakthrough in market structure, creating a market-driven transition from low efficiency to high efficiency, this route should not be obstructed (see also Leo et al., 2002). Since market-driven routes are often preferable to regulation-driven ones, the likelihood of this route should be assessed first before considering some form of regulation.

4.4.2 Economic theory

Summarising the theoretical insights of sections 2-3, six conclusions can be drawn:

1. Forward-looking access rules (e.g. LRIC) may offer better incentives for investment than backward-looking ones, although the empirical evidence is somewhat mixed (Cave et al., 2001).
2. However, by implicitly assuming that markets are contestable, LRIC rules can have adverse dynamic consequences, which can completely offset the

positive incentives mentioned above. Typically, telecommunication markets are not contestable: there are substantial non-recoverable costs as well as many unsuccessful investments (stranded assets). Failing to take this into account can hamper investments and innovation.

3. Providing ex ante clarity on the leading regulation principles reduces uncertainty and thereby contributes to dynamic efficiency.
4. Uncertainty can further be reduced by defining 'sunset clauses', i.e. regulators predetermine conditions under which regulation can be lifted or softened.
5. Dynamic efficiency is not one-dimensional: access charges which serve well as investment incentives for incumbents are not always as good for entrants, and investment is not equivalent to innovation.
6. A simple focus on incentives for investments by the entrants is too short-sighted: investments can be wasteful. It is only useful to provide incentives for entrants if the incumbent's infrastructure is replicable.

4.4.3 Practical experiences

There are a number of practical experiences with access in a dynamic context. While any number of practical experiences could be listed here, the focus has been laid instead on a couple of instances where the dynamic nature was most apparent. Much of this subsection on practical experience is taken from Cave et al., (2001) and Bennett et al., (2001).

- Country experiences

Three country studies can be mentioned, taken from Cave et al. (2001). The fourth is taken from Bennett et al. (2001).

In the UK, OFTEL realised in 1996 that its interconnection regime should be revised, because insufficient attention had been given to dynamic considerations. The new regime was characterised by low (forward-looking) interconnection charges for entrants. Entrants obtained a special status (Relevant Connectable Systems) which allowed them these low charges. At the same time, OFTEL made it clear that operators could not assume they would maintain that status unless they were prepared to invest.

For the USA, there is certainly empirical evidence that access charges influence investment levels, to the extent that lower access charges have promoted a greater deployment of digital technology among US incumbent local-exchange carriers.

For the Netherlands, empirical evidence suggests that interconnection policies have influenced the level and structure of investment, most notably in relation to ULL. A new group of DSL providers has entered the market, boosting investment in fibre optic networks.

There has been a Dutch policy debate as to whether cable companies should be forced to provide access to their cable infrastructure. Mandated cable access is good for the television market, where cable companies hold dominance. It can also be good for static efficiency in the Internet market. The effects on dynamic efficiency are unclear, since it may provide disincentives for cable companies to invest. Potential harm to dynamic efficiency can be reduced by

time-dependent access regulation, taking into consideration the stranded assets discussion (see section 2.3.2).

The conclusion drawn from these national scenarios is that various countries do certainly have experience with the dynamic consequences of access charges. While these consequences are considered important, it is not yet clear what lessons to draw from these experiences.

- Unbundling the local loop

The discussion on ULL also yields interesting results for the discussion on static versus dynamic efficiency. Consider various types of entry, where entrants adopt different strategies or where the maturity of the market differs (quoted from Cave et al., 2001; the current authors have added some notes in italics).

"First, a cable operator like UPC [a Dutch cable operator] will already have replicated some aspects of the local loop, although further investments are required both for telephony and for the development of Internet services. In this respect, local access pricing makes little difference. The cable operator has to buy call termination from the incumbent, probably in respect of the majority of calls. This service is wholly non-replicable. But reciprocal pricing between operator and incumbent neutralises this factor. Finally, the cable operator needs access to long-distance conveyance, which is replicable, either by the cable operator or by other entrants and which should, with time, be competitively priced.

"Second, consider the case of Tele2 [the largest Dutch Carrier select player], whose strategy consists of targeting a mass market, involving marketing and advertising expenditure, on the basis of – initially at least – a minimal investment in infrastructure. As time passes, Tele2 makes further investments in switching and conveyance at the national level, but its investments are limited (possibly confined to marketing costs).

"Third, consider the case of non-cable entry into the high-bandwidth markets. This is a new market, in which the incumbent has no historic market share, although clearly it has the advantage of providing the related service of basic telecommunications to the vast preponderance of domestic and business customers. Incumbents and entrants are under the same necessity to make the appropriate investments in servers. In this instance, the key non-replicable resort for entrants is the local loop, or access to part of the bandwidth provided by the loop. Setting on one side wireless technology, which is essentially untried, unbundling of the loop is a necessity.

"This raises the key question of how the unbundled loop should be priced [see also section 2]. Commitment to a low regulated rental for the unbundled loop would clearly encourage a competitor's complementary investment. If, however, the regulator were concerned about the entrant's short-term cash position, it might choose to alleviate its difficulties in the early phase of entry by a lower rental charge, which would then rise to above cost. The choice of a final level on the price curve at which the price would be stabilised would be influenced by the regulators' preference for network duplication. If the preference were slight, then a level-pricing policy would be preferable (or one which levelled off at cost). If

the preference for duplication was strong, then this could be reflected in a pricing strategy, which rose to above cost. A time limitation on mandated access to the local loop might also be appropriate. But this last policy, which has been adopted in certain areas in Canada, appears to carry the risk that entrants' fears about what might happen towards the end of the mandated access period might discourage investment."

This all leads to the conclusion that regulators have two instruments with which to influence investment decisions. The first relates to whether incumbents' assets are replicable or not. Clearly, entrants are not sensitive about access charges in parts of the network they cannot replicate. Second, access charges can have a dynamic nature. Most entrants choose a gradual investment strategy, leasing lines from the incumbent in the initial stages and meanwhile investing in their own networks. A dynamic access policy takes this into consideration: access charges can rise over time, or the level can be made contingent on the level of investments.

4.4.4 Policy conclusions

A number of policy conclusions emerge from the above.

- Universally high access prices are not smart

A possible implication of static analysis is that the best way to stimulate infrastructure investment is to have universally high access prices, i.e. access

prices which are high and which remain high over time. The argument is that high access prices provide an investment incentive for incumbents (risk premiums, stranded asset discussion) as well as for entrants (who can avoid high access prices by investing themselves). This is not necessarily true when taking a dynamic view. There is neither theoretical nor empirical evidence to support high-cost access pricing as a means of encouraging infrastructure competition. This is because future profit opportunities are more important than current price levels when contemplating an investment decision. Admittedly, if entrants are certain that future access levels will be at such low levels that it will always be cheaper to obtain access than to invest, then (future) high access levels are problematic. But there are many reasons why that picture may look different. The investments can yield other types of returns (unrelated to access) and there can be a belief that regulators will reward investments in infrastructure.

- Eligibility and replicability

Access charges, which explicitly take entrants' investments and investment possibilities into consideration, do well on dynamic efficiency. Think of OFTEL's strategy of conditioning charges, or the status of operators on investment decisions. The same applies to replicability: on which parts of the network is it reasonable to provide investment incentives? This is a tough question for regulators. A few years ago local loop competition seemed to be the name of the game, with wireless, cable and Internet coming up. Nowadays, over-investment

predominates throughout Europe. Perhaps tomorrow – who knows – everyone will quietly forget it when fibre to the house is installed.

- Sunset clauses and other dynamic regulation

Regulatory uncertainty can be lowered by pre-specifying sunset clauses, not only (or rather: not necessarily) purely the date, but perhaps the market condition under which some specific regulation can be removed. Access charges which rise over time or which contain some other temporal element, enable the fulfilment of both static and dynamic goals.

- A dynamic focus can reduce the need for future regulation

Setting appropriate access charges is difficult. Stimulating investments in new infrastructure provides the basis for removing access regulation and replacing it with bilateral agreements on interconnection charges.

5 Conclusions

This chapter provided an overview of the literature on access regulation and noted some lessons for regulation practice.

A common element in all this chapter's discussions is that access prices can never be neutral by definition. Access prices affect competition – or more

precisely, operators' profits, market shares, and retail prices. Much can already be gained by acknowledging this and by using economic theory to analyse the various links between markets and prices.

A second element is that access charges often perform too many tasks. Access charges are used to stimulate static as well as dynamic efficiency while at the same time serving equity goals such as universal service obligation. This is a bit much for an access charge which, in its simplest form, is a one-dimensional price. Several of these goals can easily conflict, so choices have to be made and trade-offs persist.

A third element is that the focus has shifted from competition in service to competition in infrastructure, or, more generally, attention has shifted from static access prices to dynamic notions. The emphasis on dynamic considerations cannot be overstated. The telecommunications industry is one, which is largely driven by innovation and investment in new infrastructure. So putting right the incentives for investments and innovation can easily dwarf any of the static considerations. Successful competition for infrastructure can reduce the need for access regulation and does not preclude competition for services taking place at the same time. Dynamic access policies, e.g. access charges that rise over time, enable the fulfilment of both static and dynamic goals. It is important to reduce uncertainty by providing regulatory transparency on these time-related issues.

A fourth element concerns the various methods of determining access charges. No one method is superior to another, a priori. It is both the relative levels and weighting of distortions, which determine which method is best. Some methods distort dynamic efficiency, but if the focus is on static efficiency, such a

method could be adequate. Other methods ignore demand-side considerations, but if the distortion is small, such a method could indeed be used. Ramsey methods require a high information burden, but sometimes this burden can and should be overcome. LRIC, when used too simplistically, can produce economically perverse outcomes, reducing incentives to invest instead of achieving the intended incentive boost .

A final element is the potential of an alternative to the ‘calling party pays’ (CPP) principle. CPP is common practice in Europe (with the notable exception of international roaming). The CPP principle implies that a receiver obtains a service (receiving a call) but doesn’t actually pay for it. Some distortions take place as a result, e.g. in interconnection tariffs and fixed-mobile termination, while Internet charges may be set at sub-optimal levels – where a balanced cost allocation would lead to incentives to set the charges at competitive levels (Laffont et al., 2001), thus reducing the need to regulate services which are currently under regulator scrutiny.

References

Armstrong, M. (1998), 'Network interconnection in telecommunications,' Economic Journal, (108), pp. 545-564

Armstrong, M. (2001), 'The theory of access pricing and interconnection,' in M. Cave, S. Majumdar and I. Vogelsang (ed.), Handbook of Telecommunications Economics, North-Holland.
http://www.nuff.ox.ac.uk/economics/people/armstrong/acc-int-num.pdf

Bennett, M., P. de Bijl and M. Canoy (2001), 'Future policy in telecommunications: an analytical framework,' CPB Document (005), CPB Netherlands Bureau for Economic Policy Analysis, The Hague.,
http://www.cpb.nl/eng/pub/document/5/doc5.pdf

Bolton, P., J.F. Brodley and M.H. Riordan (2000), 'Predatory pricing: Strategic theory and legal policy', The Georgetown Law Journal, (88), pp. 2239-2330

Braeutigam, R.R. (1989), 'Optimal policies for natural monopolies', in R. Schmalensee and R.D. Willig (eds), Handbook of Industrial Organization, Volume II, Elsevier Science Publishers: Amsterdam, the Netherlands

Buigues, Pierre (2002), Latest progress in LLU as reflected by the sector enquiry, LLU hearing 8 July 2002, downloadable from EC website

Cabral, Luis M. B. and Riordan, Michael H. (1997) 'The Learning Curve, Predation, Antitrust, and Welfare', Journal of Industrial Economics; 45 (2), pages 155-69

Carter, M. and J. Wright (1999), 'Interconnection in network industries', Review of Industrial Organization, 14, 1-15

Cave, M., S. Majumdar, H. Rood, T. Valletti and I. Vogelsang (2001), 'The relationship between access pricing regulation and infrastructure competition,' Report to OPTA and DG Telecommunications and Post, Brunel University

De Bijl, P.W.J., and M. Peitz (2000), 'Competition and regulation in telecommunications markets,' Special Study, CPB Netherlands Bureau for Economic Policy Analysis, The Hague., http://www.cpb.nl/eng/pub/bijzonder/26/bijz26.pdf

De Bijl, P. and M. Peitz (2002a), 'New competition in telecommunications markets: regulatory pricing principles' (2002), forthcoming in IFO Studies, also published as CESifo Working Paper No. 678 (9), CESifo, München

De Bijl, P. and M. Peitz, (2002b), Regulation and Entry into Telecommunications Markets, Cambridge University Press, Cambridge, UK.

DTe (2000), Guidelines for price cap regulation of the Dutch electricity sector, (http://www.nma-dte.nl/pdf/guidelines9mei.PDF)

European Commission (2001), Pricing issues in relation to unbundled access to the local loop, ONPCOM 01-17

Europe Economics (2001), Airport Cost Allocation Report for the CAA

IRG (2002), Working paper on market definition for mobile termination, February 2002, IRGs Significant Market Power – Working Group

Koboldt, Ch., and D. Maldoom (2001), Optimal fixed-to-mobile interconnection charges, paper presented at the 12th European regional ITS conference, Dublin, 2-3 September 2001

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

Laffont, J.-J., S. Marcus, P. Rey and J. Tirole (2001), Internet Peering, American Economic Review.

Laffont, J.-J., and J. Tirole (2000), Competition in Telecommunications, MIT Press

OECD (2000), Cellular mobile pricing structures and trends, OECD: Paris, DSTI/ICCP/TISP(99)11/final

Leo, H., M. Pfaffenmayr and G. Schwarz, (2002) Innovation und regulierung im Telecomsektor, (in German) Wifo study

Nicolaides, P., and R. Polmans (1999), Competition in EC Telecommunications: Cross-subsidisation, access and predatory pricing, World Competition, Vol. 22 (4), pp. 21-43

OFT (1999), Assessment of individual agreements and conduct, OFT 414

Office of the Director of telecommunication regulation Ireland (2001), Local Loop Unbundling – Eircom's Access Reference Offer (ARO), Decision Notice D8/01, September 2001

OFTEL (1999), Ideas for relevant markets/Sunset criteria for access market http://www.OFTEL.gov.uk/ind_info/international/irgtf99.htm

OFTEL (2000), The Application of the Competition Act in the Telecommunications Sector, London. http://www.OFTEL.gov.uk/publications/ind_guidelines/cact0100.htm

OFTEL (2001), Review of the price control on calls to mobiles

OPTA and NMa (2001), Richtsnoeren prijssqueeze

OPTA (2002), Beleidsregels inzake de regulering van mobiele terminating tarieven, 28 maart 2002

Sharkey, W.W. (1982), The theory of natural monopoly, Cambridge: Cambridge University Press

Telecom Public Notice CRTC 2000-96, Local competition: Proceeding to consider extending the sunset rule for near essential facilities and to treat copper loops as essential facilities, July 11, 2000

Tirole, J. (1988), The Theory of Industrial Organization, MIT Press

Notes

[1] Armstrong (2001) also mentions competitive bottlenecks as a separate category, but the current authors prefer to discuss competitive bottlenecks in the policy chapter.
[2] Entrants without their own backbone may lease long-distance capacity from the incumbent. Since the incumbent may have spare capacity, this may be in its own interest.
[3] Only the leasing part of ULL is one-way access. The interconnection part is two-way access. It is the leasing aspect which is the most attractive here.
[4] There need not be a conflict of interest if the incumbent wants to increase traffic on its network, (compare with virtual mobile operators).
[5] See Tirole (1988), chapter 8.
[6] Applying LRIC to certain services (e.g., wholesale services) and backward-looking prices to others (e.g., retail services) could lead to inconsistencies.
[7] Or at least less directly competing operators.
[8] Armstrong, 1998, and Laffont et al., 1998, also address asymmetric competition between an incumbent and an entrant, but focusing on different issues, e.g. the entrant's optimal network coverage.
[9] This is pointed out by De Bijl and Peitz (2002b).
[10] It is not clear, however, whether this would constitute a problem in welfare terms.
[11] However, it does not solve the *originating* access from fixed to mobile. These charges should still be regulated.
[12] Closed user groups are groups which have an interest in how much it costs to call each other and therefore all subscribe to the same network (e.g. family members or a company). Special offers are often made for these closed user groups.
[13] If different retail services use the same access service, the allocation of common costs to the different retail services will come into play; see section 4.3.
[14] If prices are below cost but stay at that level 'for ever' then there is no predation. This can be a viable business strategy in industries characterised by network externalities or learning by doing. The period of below cost pricing extends no longer than necessary to achieve the network economics, or until the cumulative production experiences result in lower costs. See also the next section on cross-subsidisation.
[15] This also solves the discussion as to which subset to start with.
[16] A practical issue is how to determine the economic lifetime. This is an accounting matter.
[17] See also Bennett et al. (2001) and Leo et al. (2002), in which the notions of static and dynamic efficiency play a central role in addressing various telecommunications market policy issues.