CPB Document

No 005

Future Policy in Telecommunications

An Analytical Framework

Matthew Bennett, Paul de Bijl, Marcel Canoy

Centraal Planbureau
Van Stolkweg 14
Postbus 80510
2508 GM Den Haag

Telefoon (070) 338 33 80
Telefax (070) 338 33 50
Internetadres: www.cpb.nl

ISBN 90-5833-072-9

Table of Contents

Samenvatting (Summary in Dutch)

De toekomstige ontwikkelingen van telecommunicatie zijn uitermate onzeker. Beleidsmakers worden geconfronteerd met allerlei elkaar in snel tempo opvolgende ontwikkelingen. Hoe kunnen beleidsmakers anticiperen op toekomstige beleidsvragen?

Dit werkdocument levert een analytische bijdrage aan dit probleem. Beleidsmakers staan regelmatig voor het dilemma dat beleidskeuzes goed voor de korte termijn kunnen zijn (statische efficiëntie), maar minder goed voor de langere termijn (dynamische efficiëntie) en omgekeerd. In welke situatie is zo'n dilemma te verwachten en waarom?

Bovenstaande vragen worden geconcretiseerd door drie ontwikkelingen in de markt te bezien die exogeen voor beleidsmakers zijn. De ontwikkelingen genereren allerlei nieuwe beleidsuitdagingen. De gekozen ontwikkelingen zijn: Internet via mobiele telefonie (ontwikkeling I), de toenemende vraag naar breedbandcapaciteit (ontwikkeling II) en consolidering en allianties (ontwikkeling III).

Deze samenvatting is als volgt georganiseerd: eerst beschrijven we de methodologie, daarna bespreken we de drie ontwikkelingen, tot slot trekken we algemene lessen uit de drie ontwikkelingen.

Methodologie

Het is allereerst van belang om te bepalen welke factoren ertoe bijdragen dat een markt zich in een situatie van lage of hoge efficiëntie bevindt. Er blijken vier groepen van zulke factoren ('drivers') te bestaan, te weten:

- marktstructuur en toetreding
- anti-competitieve acties
- regulering en mededingingsbeleid
- technologie

Combinaties van deze drivers kunnen leiden tot een efficiëntieniveau waarop de markt zich bevindt. Een voorbeeld is “lage statische hoge dynamische efficiëntie”. Zo kunnen technologische ontwikkelingen in combinatie met anti-competitieve acties ertoe leiden dat bedrijven goed geïnvesteerd hebben in netwerken (hoge dynamische efficiëntie) maar wel hoge tarieven in rekening brengen voor hun diensten (lage statische efficiëntie).

Na het vaststellen van welke driver tot welk efficiëntieniveau leidt, is het zaak te bepalen wat de prikkels van marktspelers zijn om een hoger efficiëntieniveau te bereiken. Indien die prikkels niet of slechts beperkt aanwezig zijn, ontstaat een rol voor beleid om de markt richting een hoger efficiëntieniveau te bewegen.

In de studie worden drie ontwikkelingen bestudeerd, Internet via mobiele telefonie (ontwikkeling I), de toenemende vraag naar breedbandcapaciteit (ontwikkeling II) en consolidering en allianties (ontwikkeling III). Bij elk van deze ontwikkelingen vragen we ons af waar een afruil tussen statische en dynamische efficiëntie dreigt, welke prikkels partijen hebben om een hogere efficiëntie te bereiken en wat de rol van beleid daarbij is.

Internet via mobiele telefonie (ontwikkeling I)

Kenmerkend voor de ontwikkeling van UMTS (in Nederland) is dat de markt nog niet bestaat, dat de spelers die daarop actief zullen zijn dezelfde zijn als in de huidige markt voor mobiele telefonie, dat de verdeling van licenties via een veiling niet bepaald soepel verliep en dat er grote onzekerheden zijn. De startsituatie in de markt is er zeer waarschijnlijk een van lage statische en dynamische efficiëntie.

De vraag is of marktpartijen een prikkel hebben om naar een hoger efficiëntieniveau te gaan. De route die begaan kan worden zonder al te veel overheidssturing loopt via investeringen en innovaties die de markt openbreken. Diensteninnovaties kunnen deze markt concurrerend maken, zonder dat dit al te nadrukkelijke regulering behoeft. Of het zover zal komen is gezien de 'stroeve start van de markt' (de veiling) echter maar de vraag.

De volgende conclusies kunnen uit deze ontwikkeling getrokken worden.

- Omdat de overheid niet kan inschatten wat de toekomstige baten van internet via mobiel zijn en omdat frequenties schaars zijn, is het logisch de licenties te verdelen via een veiling (en niet via een schoonheidswedstrijd).
- De verdeling van licenties heeft ook langere termijn doorwerkingen in de markt. Slecht ontworpen veilingen kunnen toetreders afschrikken en investeringen ontmoedigen, hetgeen dan weer achteraf door regulering gerepareerd moet worden.
- De mate van regulering hangt af van de kwaliteit van de marktopzet. Een goede marktopzet vergemakkelijkt de taak van de toezichthouder en stelt hem in staat meer op de mededingingswet te vertrouwen. Een minder goede marktopzet vergt meer van de toezichthouder.

Zo moet hij met de delicate vraag omspringen welke mate van netwerkcoöperatie gewenst is. Vragen daarbij zijn: (i) Wie is er verantwoordelijk voor congestie op het netwerk? (ii) Kunnen investeringen op het netwerk en onderhoud gecoördineerd worden? (iii)Ontstaat er wel genoeg concurrentie in netwerken zelf ? (iv) Beperkt samenwerking concurrentie in diensten?

Een andere belangrijke vraag is in welke mate de toezichthouder concurrentie aan kan wakkeren door spelers toe te laten zonder netwerk.

De toenemende vraag naar breedbandcapaciteit (Ontwikkeling II)

De huidige capaciteit is niet voldoende om de toekomstige vraag naar breedbanddiensten te accommoderen. Er is een aantal mogelijkheden om capaciteit uit te breiden (via telefoonlijnen, kabel of 'fiber to the house'). De vraag is hoe de diverse marktpartijen een optimale prikkel krijgen om de juiste keuzes te maken. Keuzes omtrent de mate van openstelling van netwerken en de voorwaarden daarvoor kunnen een afruil impliceren tussen concurrentie in diensten nu (statische efficiëntie) of netwerkconcurrentie (dynamische efficiëntie).

De volgende conclusies kunnen getrokken worden:

- De overheid weet niet beter dan de markt wat de beste technologische keuze is om breedbandcapaciteit uit te breiden. Beleid dient derhalve in aanleg technologie-neutraal te zijn. Dat is niet altijd makkelijk omdat keuzes voor bijvoorbeeld de toegangsvoorwaarden van netwerken impliciete keuzes tussen de diverse varianten kunnen betekenen. Bij keuzes tussen statische en dynamische efficiëntie ligt de nadruk bij dit beleidsprobleem eerder bij dynamisch dan bij statisch: zelfs een suboptimale uitbreiding van capaciteit kan immers beter zijn dan een stagnatie in investeringen.
- Het bepalen van toegangsvoorwaarden (third party access) is het belangrijkste reguleringsinstrument van de toezichthouder. Een manier om recht te doen aan huidige en toekomstige belangen is om de voorwaarden te laten variëren in de tijd: lage tarieven nu zijn goed voor concurrentie in diensten nu, maar kunnen de prikkel tot investeren verminderen. De toezichthouder kan van te voren aankondigen dat de tarieven in de toekomst stijgen. Dit gaat alleen op voor die delen van het netwerk die dupliceerbaar of vernieuwbaar zijn.
- Het openstellen van de kabel (mandated access) is goed voor de televisie- en internetmarkten. Het heeft hooguit enige nadelen voor breedbandcapaciteit als de kabelexploitanten geen redelijk rendement op toegang krijgen èn de kabel de toekomst voor breedband is (hetgeen lang niet zeker is).
- Er moet aandacht zijn voor zogeheten 'prioritization technologies'. Deze technologie stelt netwerkspelers in staat eigen klanten te bevoordelen ten opzichte van andere klanten, bijvoorbeeld met snelheid en toegang. Bezien moet worden of dit de concurrentie niet teveel belemmert.

Consolidering en allianties (ontwikkeling III)

Getriggerd door technologie en internationalisering veranderen martkstructuren snel in telecomland. Een afruil kan bijvoorbeeld ontstaan doordat fusies marktmacht vergroten (lagere statische efficiëntie) maar grootschalige investeringen vergemakkelijken (hogere dynamische efficiëntie). Is dat reden voor beleid?

We trekken drie conclusies:

- Horinzontale fusies en allianties zijn in de regel geen reden voor speciale aandacht (anders dan via de Mededingingswet). Er zijn twee uitzonderingen: (i) het is opvallend dat er in Mededingingsland zo weinig aandacht is voor deelnemingen (cross shareholdings). Deelnemingen tussen concurrerende partijen reduceren de intensiteit van concurrentie en verdienen dus aandacht. (ii) Indien fusies leiden tot 'gesloten oligopolies' (tight oligopolies), waarbij een beperkt aantal spelers opereren op een markt die in aanzienlijke mate afgesloten is voor toetreders, is extra waakzaamheid vereist omdat de Mededingingswet niet goed toegerust is op het oplossen van problemen bij dergelijke gesloten oligopolies.
- Bij telecommunicatie is speciale aandacht nodig voor verticale constructies, zoals verticale integratie of verticale afspraken. De Mededingingswet geeft meer ruimte voor verticale acties dan voor horizontale. Daar is op zich een goede reden voor maar in telecomland zijn verticale acties net zo gevaarlijk als horizontale, omdat ze de concurrentie ernstig kunnen beperken. Zo kan het gevaarlijk zijn als één partij zowel het netwerk als de 'content' als de diensten in één hand heeft.
- Bij het beoordelen van gedrag en fusies is het vaak belangrijker naar de gevolgen voor dynamische efficiëntie te kijken dan voor statische efficiëntie.

Conclusie

Deze studie analyseert toekomstige beleidsvragen in telecomland waarbij de afruil tussen statische en dynamische efficiëntie centraal staat.

We trekken vijf algemene beleidsconclusies:

1. De opzet van de markt is van groot belang voor zowel statische als dynamische efficiëntie.
2. Bij keuzes tussen statische en dynamische efficiëntie is het vaak beter de nadruk op dynamische efficiëntie te leggen.
3. Het bepalen van toegangsvoorwaarden is een belangrijk beleidsinstrument. Het geloofwaardig aankondigen dat tarieven in de toekomst gaan stijgen kan goed zijn voor zowel statische als dynamische efficiëntie.
4. Het verlagen van toetredingsdrempels is vaak gemakkelijker en effectiever dan directere vormen van regulering.
5. Speciale mededingingsaandacht is vereist voor deelnemingen tussen concurrerende partijen, verticale relaties en technologieën waarbij een netwerkspeler zijn klanten kan bevoordelen.

Preface

The future of telecommunications is notoriously hard to foresee. Who would have predicted the mass mobile consumption? Yet, some developments are foreseeable and policymakers need to be prepared. This CPB Document helps them to anticipate new policy challenges related to exogenous developments, such as the increasing demand for broadband capacity.

This work is co-financed by OPTA. The authors appreciate feedback from OPTA's Luc Donners, Mindert Mulder, Robert Stil, as well as from Eric van Damme (KUB), Patrick van der Duin (KPN Research and TU Delft), Peter Vis (KPN), colleagues and workshop & conference participants. Kathy Schuitemaker did her usual good job in editing the manuscript.

Henk Don, Director CPB

1 Introduction

The Dutch telecom regulator Opta has asked the CPB to create a framework in which to analyze future policy issues in the telecommunication sector. The aim of the framework is to analyze in a consistent way market developments that are *exogenous to policymakers*. Although exogenous to policymakers, the development potentially creates lots of policy questions, and a framework is needed to analyze these questions and to anticipate problems. An example of an exogenous development is e.g. fixed-mobile convergence. It is a development that is largely exogenous to policymakers, but the convergence can create new policy issues. The policy issues that are analysed in this study focus on the relationship between so-called static and dynamic efficiency. An example: policy instruments such as Open Network Provision can be very effective for enhancing competition for services on the short term (static efficiency), but are potentially bad for the development of new infrastructure (dynamic efficiency).

The value of this study mainly lies in the framework in which to examine different types of exogenous developments. This framework is new and opens up future research possibilities. The three selected developments are: third generation mobile telephony, excess demand for broadband capacity, and tendency towards consolidation and alliances. It is important to note that the three developments undertaken in this document are not comprehensive studies of each individual development. Regulatory success or failure often depends on details. Clearly in a document of this size it is impossible to do three such large and important questions of policy and exogenous market development full justice. However these three developments serve the purpose in showing exactly how the methodology proposed may be applied to specific but different developments. For this reason, the policy implications are reasonably general in nature.

For each development this document identifies the relationship between static and dynamic efficiency. By doing so, the robustness of available policy tools can be checked and when necessary, new tools can be identified. These developments are not forecasts: no probabilities are attached to each of them. On the other hand: developments must be probable in some sense.

Structure of this report
Chapter 1 states the report aim, structure and gives a brief description of the Dutch Telecom industry. In particular it discusses the boundary of policy and exactly what is within the policy framework. Chapter 2 introduces the methodology and discusses the difference between the exogenous developments and the states within each development.

Chapters 3, 4 and 5 analyze the three exogenous developments. We conclude each exogenous development with the main policy message of the development. Finally, in Chapter 6 we conclude by bringing to light common elements and trade-offs across the different developments.

1.1 Background of the Dutch telecommunications industry

Telecommunications markets are going through a period of fundamental change, ignited by privatizations, liberalization of national markets, technological innovations, and, in Europe, convergence of economic policy and institutional structures. From 1998 onwards, the markets for networks and services in almost all EU countries were liberalized. Market liberalizations have been accompanied by partial or full privatization of incumbent operators. Historically, operators in Europe were state-owned, vertically integrated monopolists. Because of the large fixed cost of building a network, telecommunications networks providing voice telephony were viewed as natural monopolies. Technological change and innovation have generated new transmission systems and decreased the cost of building infrastructure, so that the idea of a natural monopoly is no longer seen as valid. Moreover, the current view is that public ownership does not provide strong incentives to decrease costs, resulting in inefficiencies. Therefore, most operators have been or are subject to plans to be privatized. For instance, the former national incumbent's in the UK, Italy and Spain are fully privately owned, while the Dutch former monopolist is still partly in public ownership.

Market liberalizations have caused large scale entry by all sorts of network operators and providers of telecommunications services. The rules of competition are being changed rapidly by these entrants, who may be less hampered by possibly obsolete infrastructure and associated cost structure, or by bureaucratic organizations and associated working practice (although this is not true for all entrants, as former monopolists are entering other counties). Hence, the former incumbents are facing a profound challenge to redefine their value propositions and strategies, and to reorganize their business practice. Incumbent operators are coping with the drastic changes in the industry in several ways:

- internal restructuring to increase efficiency, flexibility, quality of service, and responsiveness to customers' wishes;
- diversification (e.g. mobile telephony, Internet access provision) and internationalization (e.g. foreign acquisitions, international alliances and mergers) to compensate for the expected loss of domestic market share in the market for fixed voice telephony;
- innovation (e.g. high-bandwidth transmission systems, intelligent applications in the network).

The fundamentals of deregulation in the EU are defined in the Open Network Provision (ONP) framework. The ONP rules aim to ensure interconnection of networks, and to allow entrants access to various elements of networks of dominant operators at cost-based prices. In the language of the European Commission, a dominant operator is called an operator with "significant market power," and is usually defined by a market share above 25%. The incumbent operators naturally fall into this category.

The European Commission recently proposed a new framework with the purpose to accelerate the process of economic and structural reform in the EU by lowering access prices and reinforcing competition (the "Review" by European Commission, 1999). According to the Commission, the European market is fragmented and dominated by incumbent operators, even though entry has been substantial and prices have decreased. While the existing framework was designed to cope with the transition to competition, the new framework seeks to reinforce competition, especially in market segments at the local level. Moreover, it must cater for the rapid technological change and unpredictability of telecommunications markets.

The Dutch former "PTT," nowadays called KPN Telecom, was a state monopolist until 1989, when it became a public limited liability company, although still fully owned by the state at that moment. The government sold 30% of its shares in 1994, and another 25% in 1995. When competition was introduced in the Dutch market in 1996 on the basis of the so-called "interim legislation," two operators, Enertel and Telfort obtained a license for national, fixed telecommunications infrastructure. The Dutch market was completely liberalized in 1998 when the new Telecommunications Act came into force, in line with EU policy on liberalization.

Recently, KPN Telecom has been active with reorganizations, foreign acquisitions (in Eastern Europe), diversification (e.g. mobile telephony, Internet provision, Internet banking), and joint ventures (with US firm Qwest, building a pan-European, IP-based fibre network). Although entry in the Dutch market has been substantial, KPN has large market shares in several markets (e.g. in 1999: 80-95% in the markets for national telephony and for calls from fixed phones to mobile phones, 90-99% in the market for local telephony, and 85% in the consumer market for international telephony, according to Opta, 2000, p. 16).

2 Methodology

This section introduces our methodology to analyze the exogenous developments. We argue that common to all the developments, and most of the current policy problems, is a relationship, possibly a tradeoff, between dynamic and static efficiency. The existence of such a potential trade-off is contentious to many economists. Our methodology identifies under which circumstances these two efficiencies co-existent or mutually exclude each other. For this reason we split the efficiency states into high and low. This enables us to examine each combination of the two efficiency states and allows us to analyze what type of policy (if at all) is needed to move to the most desirable state (i.e. the state where both static and dynamic efficiency is high).

The three selected developments that will be analyzed with the methodology sketched out below are: (1) Third Generation mobile telephony, (2) Excess demand for broadband capacity, and (3) Tendency towards consolidation and alliances. We selected these three developments for the following reasons:

- Firstly, there is a large enough probability of them being realized and thus being directly useful to policy makers.
- Secondly, they are sufficiently narrowly defined and are suited for analyzing the trade-off between static and dynamic efficiency (not all policy questions deal with this trade-off).
- Thirdly, following on from the second point each of the three looks at a different area of interest. The first looks at changes in technology, the second development looks at changes in consumer demand whilst the last one looks at changes in the market structure.

We use these three developments in chapters 3, 4 and 5 merely as examples that illustrate our methodology. They should give sufficient confidence to analyze any future policy question where relationships or tradeoffs between static and dynamic efficiency play a crucial role.

2.1 Exogenous developments and efficiency states

The segregation of developments and states allows us to separate events that are exogenous to policy and those that are endogenously determined. For example, a regulator is hardly able to determine that Internet will develop over mobiles. Once we assume that Internet is developing over mobiles we find that a regulator may be able to determine to some degree both the level of static efficiency in this market and also the speed at which the market is taken up and new products are developed (dynamic efficiency). Thus if the industry starts in the low static and low dynamic efficiency state, policy may influence the market not only whether it moves to a new state, but also which state to move towards and how fast.

The methodology can be illustrated as follows:

Figure 2.1 Methodology

Exogenous Development 1

Efficiency States	**Dynamic Efficiency**		
Static Efficiency		*Low*	*High*
	Low		
	High		

For each development we look at the combinations of static and dynamic efficiency, called **efficiency states** (or simply "states"), that the market may be in. An example of an efficiency state is "high static, low dynamic".

Exogenous Developments

As stated above, the exogenous environments itself are beyond the control of policy. This does not mean that all environments should be treated with the same policy tools. The ultimate goal of policy is to ensure that the market moves to (or stays in) the most desirable state (high static and dynamic efficiency). There can be various reasons why the market is in a particular state. Different reasons require different policy cures. Suppose the market is in the state with low static and dynamic efficiency. If the market is inefficiently set-up , e.g. if licenses are allocated through a failed auction, then the policy cure should address repairing the damage of inefficient licencing. However if collusion drives the "low" state, competition policy is called for.

2.2 Defining efficiency states

This section describes the various efficiency states, first in general terms to define what we mean by the high and low efficiency states, and then in more detail by listing the elements that push a market towards one state or another. We term these elements "drivers". Each driver is first explained in general. These broad drivers will then be made concrete for each individual development. For example, whilst the set up of a new industry is clearly an important driver to both static and dynamic efficiency, the means to set up the industry will be quite different between the developments. In development 1 (3G mobile telephony), the set up issue may be more specifically described as the choice between auctions or a beauty contest, whilst in development 2 (excess demand for bandwidth), the choice is between the type of access created.

2.2.1 Static Efficiency

We define what we mean by “low” and “high” levels of static efficiency. Static efficiency is a combination of allocative and productive efficiency, keeping investments in product or process innovation constant. Roughly it means that short-term combined welfare of consumers and producers is maximized, while production takes place at the lowest possible cost using available technology.

In practice, it is hard to assess static efficiency on basis of these notions. A reasonable interpretation is to see static efficiency as good value for money for the consumers while firms achieve normal profits.[1] The next question is how to assess what is good value for money, as well as normal profit levels.

Instead of a directly assessing static efficiency, an indirect way can be followed by considering the level of competition. Intense competition will, under a set of well-specified conditions, lead to static efficiency. It is possible approximate the level of static efficiency by defining these conditions and analyze the extent of competition. For instance, given the level of fixed costs, a certain intensity of competition is compatible with normal profits. If firms make supranormal profits and entry barriers are high, arguably static efficiency can be enhanced by intensifying competition.

An example clarifies the complexity of the relationship. The situation with five mobile telephony players can be consistent with both high and low levels of competition and static efficiency. Even if five players is the optimal number of firms, five players can compete fiercely or softly. In the latter case supranormal profits and stable market shares can sustain over time, despite cost differences. If, however, market shares of the five players move up and down, corresponding with the relative cost levels of the firms, competition is arguably intense. Even for a fixed number of players more intense competition is not equivalent with more welfare. More intense competition can lead to the bankruptcy of the least efficient firm. This is only good for welfare if the firm is replaced by another firm. Therefore, intense competition should be accompanied by sufficiently low entry barriers.

To approximate static efficiency by considering competition is worthwhile when there is clarity about the relevant market, there are enough data etc. For the more general questions of this study it seems better to use a set of proxies for static efficiency, such as:

- entry barriers
- number of firms
- level of fixed costs
- intensity of competition

[1] The “normal” profit level includes a fair rate of return on capital invested, and hence include some degree of “profits” as they are commonly perceived.

A market is then considered to be (relatively) statically efficient if sufficient proxies indicate it. Admittedly, this approach is not very precise, albeit practical. For the purposes of this study, however, a more elaborate approach was not feasible.

2.2.2 Dynamic Efficiency

Similar to static efficiency we need to define what we mean by "low" and "high" levels of dynamic efficiency. Firstly we define the dynamic efficiency in a market as the extent to which the present value of a stream of static total welfare can be maximized over long periods of time (long enough to allow for investments in product and process innovation). Whilst the measurement of total welfare is difficult it is possible to proxy dynamic efficiency using other similar measurements. Examples of these proxies (some of them overlap) are:

- introduction of new products and services ("product innovation");
- improvement of production technologies ("process innovation");
- increases in choices for consumers;
- increases in quality of products and services.

The level of dynamic efficiency is said to be high if over time these measurable proxies for dynamic efficiency also are high. For example, high dynamic efficiency is characterized by a fast product roll out, with substantial increases in quality and rapid introduction of new services. Similar to static efficiency we believe that these proxies are sufficient for our purpose.

What further complicates matters is that there is no one-to-one correlation between competition and dynamic efficiency. We assume that there is a certain link between incentives to innovate and the level of profits that a firm makes. For instance, a firm with higher profits may have lower incentives to innovate, or reversely, higher incentives. The nature of this relationship will typically depend on the intensity of competition. This inverse U-shaped relationship between competition and innovation is further discussed in section ??.

2.2.3 Drivers for Efficiency

There are several potential reasons why a market is characterized by high or low efficiency levels. The following is a list of the most important drivers of efficiency.

Driver 1 Market structure and entry

- In telecoms the number of firms is often determined by licenses. One reason is that frequencies are scarce and have to be allocated to a limited number of firms. Both the means of allocating licenses, and the type of access on networks that should be permitted, will have an impact on efficiency. Increasing the number of firms in the industry can drive down prices and profitability, can increase output which in turn increases static efficiency. This general relationship is more complicated in industries with a substantial fixed cost element such as the UMTS and broadband market. In this case continual entry in the industry results in the duplication of fixed costs. At some point the increase in static efficiency through greater competition is outweighed by the duplication of fixed costs.

- Market structure not only impacts static efficiency but also dynamic efficiency. For instance, both the number and type of licenses or access will have an impact on both the profitability and the competitiveness of the industry and hence impact dynamic efficiency. Conditions which are too favorable to new entrants may reduce incentives to invest, whereas a lack of competition between firms may reduce firms" aggregate incentives to invest. Newcomers may introduce more variety for consumers than incumbents, and also increase the pressure on incumbents to invest in product and process innovation.

Driver 2 Anti-competitive practices

- Firms have an incentive to collude or to engage in other anti-competitive actions, because profits are strictly increasing with collusive prices until they reach monopolistic levels. Tacit or explicit collusion increases prices and decreases output and static efficiency. For this reason it is vital to ensure that measures are taken to ensure that anti-competitive practices are unattractive.
- The relationship between anti-competitive practices and the level of investment or innovation (and therefore with dynamic efficiency) is more complicated than with static efficiency. In general, more intense competition can either increase or decrease incentives to innovate. Recent theoretical and empirical work has tended to side with the view that monopolies have a reduced incentive to increase dynamic efficiency. However this relationship is highly complicated and is studied in much greater depth in the technical appendix and in the following subsection.

Driver 3 Regulation and competition policy

- Regulatory interventions may directly affect static efficiency. For instance, access price regulation that makes third-party access cheap and easy, intensifies competition and may therefore increase static efficiency.
- Regulatory interventions may also directly affect dynamic efficiency. Coming back to the previous example, access price regulation that makes third-party access cheap and easy, may reduce network operators' incentives to invest in maintaining and upgrading their networks.
- Uncertain regulation increases risk and hence the costs of capital to compensate for this increased risk. These higher costs reduce the level of investment. Explicit regulation especially increases uncertainty where the regulator is not consistent in its decisions of acceptable rates of returns allowed to firms to ensure that their investments can be recouped. Even where certainty exists, but the regulator decides on a very low allowance for rate of return, the incentive to invest will be reduced.

- Competition policy is obviously related to the previous driver. The effectiveness of reducing anti-competitive practices strongly affects efficiency.

Driver 4 Technology

- Technological uncertainty (e.g. about the success of innovations and adoption in the market) can drive a reluctance to invest. Uncertainty may be derived from many areas, some within policy control such as regulation (see previous driver), and some outside such as price and demand uncertainties. If the future profitability of a market is very uncertain, then firms may under-invest, thereby reducing dynamic efficiency.
- Compatibility and standardization: New technology industries such as the third generation mobiles are susceptible to problems of compatibility and standardization. The current WAP and Bluetooth standards are examples of successful coordination on a standard, as are GSM and UMTS. If there is more than one standard, mobile network operators may be wondering which to set up on their new services. Clearly there is an incentive for firms to wait to see which compatible standard will be set up rather than potentially committing themselves to a standard that becomes obsolete in a few years.
- Technological lock-ins in previous generations or old technology, for example because of a failure to coordinate on a certain, already existing technological standard, can reduce static efficiency. Similarly, lock-in effects may make it simply unattractive to develop new technologies, thereby reducing dynamic efficiency.

2.2.4 Interaction between Efficiency States

From above it follows that although one may have developed ideas about static and dynamic efficiency, these notions have to be made more concrete in individual developments. For example, high static efficiency for the first exogenous development (third generation mobile telephony) may not be exactly the same as high static efficiency for the second one (excess demand for broadband capacity). Also, we simplify in the sense that only high and low efficiency levels are considered, whereas in reality efficiency is a continuous variable.

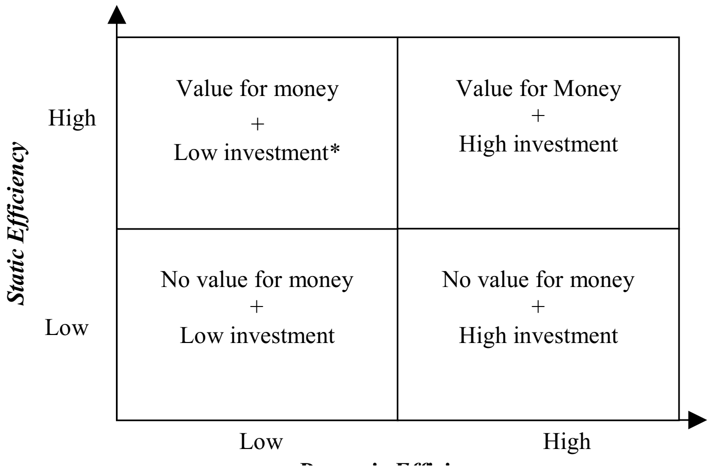
In Figure 2.2, we illustrate the four possible efficiency states that the market may be in. For simplicity we called “high static” value for money and “high dynamic” high investment.

Appendix A.1 discusses theoretical and empirical links between static and dynamic efficiency. A conclusion drawn in the appendix is that there may be an inversed U-shaped relationship between the intensity of competition and the rate of innovation in an industry.[2] The relationship only focuses on one proxy of static efficiency (competition) and one of dynamic efficiency (innovation), but is still highly instructive for our purpose. For the sake of exposition,

[2] See Scherer and Ross, 1990, p. 645.

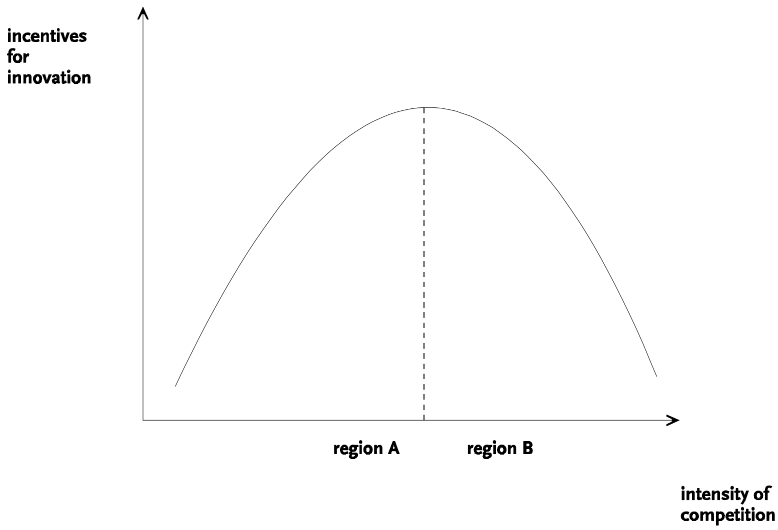
we therefore adopt this relationship, keeping in mind its limitation. See figure 2.3 for an illustration.

Figure 2.2 Diagrammatic Relationship between Static and Dynamic Efficiency



**investments in product and process innovation*

Figure 2.3 Competition and innovation



Region A, where innovation and competition move together, is characterized by a market structure in which firms have a considerable level of market power. Hence, competition is not very intense. Increasing competition, for instance by lowering entry barriers, then encourages innovation. At some point, when competition becomes more intense, innovating becomes less worthwhile, because there are less rents to be taken from innovation. Region B, where there is a tradeoff between competition and innovation, is characterized by a market structure in which firms compete intensely and face strong price pressure (e.g., Bertrand-style competition). Firms engaged in this type of competition are typically unable to support the investment costs needed for innovation.[3]

The drivers of efficiency as described above, may or may not be compatible with each other. Within one efficiency state the market may be in region A or region B. For this reason it is vital to look at each efficiency state in turn to determine the relationship between the two types of efficiencies and the reasons for this relationship and hence efficiency state emerging.

Given the exogenous development there are four different efficiency states. Obviously the most attractive efficiency state is where both static and dynamic efficiency is relatively high, but, as previously mentioned, this will not always be feasible for all developments.

Low Static and Low Dynamic Efficiency

In many cases this may be the start up of a particular environment: new technologies still have to be tested in practice, while the profitability of a market is uncertain.

High Static, Low Dynamic Efficiency

Think, for instance, of cut-throat competition: firms are unwilling to invest in innovation because the possibilities to recover their fixed costs through healthy profit margins. There may be intense competition, but this comes at the cost of little investment in R&D.

Low Static, High Dynamic Efficiency

Suppose that more intense competition reduces firms' incentives to invest in innovation. Hence, if competition is not very intense, there may be a situation in which low static efficiency is combined with high dynamic efficiency. Firms are not competing strongly, perhaps they are colluding, to ensure prices substantially above cost. This leads to large profits and a relatively statically inefficient market. These large profits not only accrue to the managers and shareholders but also provide the firms with substantial capital to reinvest and innovate to differentiate their products.

[3] This argument depends on the assumption that capital markets are imperfect.

High Static, High Dynamic Efficiency

This is the optimal state, although as mentioned not always feasible. Suppose that more intense competition increases firms' incentives to invest in innovation. Then a situation may arise in which competition is intense, so that static efficiency is high, while firms have strong incentives to leap ahead of competitors. If innovations succeed, they are able to price to some degree above cost, making it possible to recoup their investments.

Since both static and dynamic efficiency can be influenced by different drivers, there can be various reasons for the market for being in a state. The identification of the reasons why the market is in a state enables the assessment of the incentives for market players to leave the state. If the incentives do not point at the "high-high" state, policy tools can be discussed that help the market to move in the right direction.

2.3 Summarizing the methodology

Our methodology can be summarized by the following steps:

1. For each development, four drivers of static and dynamic efficiency are identified. Various combinations of these drivers can lead to various "states". An example of a state is "low dynamic high static efficiency".
2. After identifying the most relevant combinations of drivers, we determine the incentives of the market players to move to another state.
3. If incentives point in the "wrong" direction, determine what policymakers can do to move the market in the "right" direction.

3 Development One: Third Generation Mobile Telephony

3.1 Introduction

In this first exogenous development, we concentrate on the provision of “3G” mobile telephony, which is the next generation of mobile telephony with the possibility of mobile internet access. Its underlying standard is called UMTS. The current (“2G”) generation, GSM, is mainly used for voice telephony and short messages (SMS services). It has limited internet-type functionality through the so-called WAP and GPRS technologies (see the next section for background information). Nevertheless, the general expectation is that the next generation will make mobile internet access widely available. Throughout this chapter, the emergence of the UMTS market for mobile internet is considered to be an exogenous development.

Policy question

We focus on one general policy question. Within this general policy question more specific subquestions can be asked.

What is the optimal policy to ensure both effective competition and take-up of new services in the new generation of mobile telephony?

This general question may encompass many issues such as:

- Which incentives can be given such that technology is rolled out in the maximum speed?
- Should mobile licensee firms see their networks open to virtual competitors?
- Should mobile firms have an explicit form of regulation?
- How does the set-up of the market, in particular the allocation of licenses, affect efficiency tradeoffs?

3.2 Background

This section contains two boxes, providing some background for the rest of this chapter, both on the Netherlands as well as on mobile technologies and standards. The first box describes the current situation in the Dutch market for mobile telephony, which is strongly influenced by the license allocation in the year 2000.

The mobile markets in the Netherlands

In the Dutch market for mobile telephony, there are currently five operators active: KPN Telecom, Libertel (owned mainly by Vodafone), Dutchtone (owned by France Telecom), Telfort (owned by BT and Dutch Railways), and Ben (initially owned by Belgacom and TeleDanmark). KPN has teamed up with Japan's NTT DoCoMo to introduce i-mode (a mobile Internet technology; see the box below) in Europe.

Five licenses for the 3G (third generation) market were auctioned in the summer of 2000, and were won by the incumbent operators KPN, Libertel, Ben as a part of the 3G Blue consortium (Belgacom, TeleDanmark, and Deutsche Telekom), Telfort, and Dutchtone. Note that the number of licenses equalled the number of incumbent operators. During the auction, Telfort sent a letter to entrant Versatel with the apparent purpose to exert pressure to Versatel to stop bidding and quit the auction. Many involved parties are now challenging the outcome of the auction, either in court or in an investigation (e.g. the Parliament).

The second box describes the different generations of technologies for mobile telephony, and the associated terminology. Distinctions between current and future generations are made.

3.3 Efficiency drivers

We now move on to discuss the drivers of the various efficiency states in detail. Before going into the drivers, it is important to recall that a description of dynamic efficiency can approached be several indicators. These include both the length of time for the rollout of coverage and the increases in the quality of service taking place. In particular, high dynamic efficiency is characterized by a rollout of coverage at the fastest possible level, with substantial increases in call quality and network capacity, facilitating a rapid introduction of Internet-based services.

Driver 1 Market structure and entry

In the mobile market, entry is legally restricted by a licensing system. The licenses must be correctly allocated to guarantee high static efficiency. This involves ensuring that the most efficient firms, or the firms for whom it is most profitable to enter, will be operating within the market. Where licenses are not allocated to the most efficient firms, or firms engage in significant amounts of lobbying for the licenses, static efficiency will be reduced. Auctions give no guarantee either: an auction has to be well designed to ensure no overbidding or underbidding take place. Auctions typically perform better than beauty contests in terms of efficient allocations of spectrum rights.

Overview of Mobile Technology

First generation "1G"
The first generation of mobile voice telephony, which is no longer operational in most European countries, used analog technology. The standard is known as NMT.

Current generation "2G"
The second (and current) generation of mobile telephony is based on digital technology. Voice and data are carried by signals in the radio spectrum within certain frequency bands (the 900 and 1800 MHZ bands). The underlying standard in western Europe is called GSM (Global Standard for Mobile telephony). There are two important 2G data services:

- SMS (Short Message Services) allows consumers to send short text messages at a relatively slow speed; messages are transmitted whenever a network has some spare capacity. SMS is (unexpectedly) popular because of its simplicity and straightforward functionality.
- WAP (Wireless Application Protocol) is a standard to adapt Internet-based services to mobile phones. Because of the small screens of WAP phones and the slow speed of access, WAP cannot be used for Internet surfing and using web-based services, although it is sometimes advertised as such. Users have to connect every time that they call a new "page" (it is not "always-on"). Unlike the Internet, WAP is (or at least started as) a closed platform: users initially only had access to content supplied by mobile operators themselves, who act as gatekeepers. There has been a lot of hype about WAP, without generating a strong consumer response (The Economist, 2000).

Intermediate generation "2.5G"
By upgrading software of the base stations of GSM operators, better data-handling capabilities can be obtained than with 2G. The main underlying technology is GPRS (General Packet Radio Services), which denotes packet-switched communications that run over GSM networks. Some view GPRS as a fix of the limitations of WAP (e.g., it has the always-on feature). In Japan, there is I-mode, a mobile technology offered by NTT DoCoMo. There is a relatively large customer base, not only because it was introduced earlier, but presumably also because it does not share the disadvantages of WAP: it is "always-on", it is an open platform; usage prices are much lower. KPN Telecom has teamed up with NTT DoCoMo to introduce i-mode in Europe.

Next generation "3G"
Mobile phone services offering high-speed Internet access, supposedly much faster than previous generations. In the UK and in the Netherlands, there have already been auctions for spectrum rights for the third generation of mobile telephony. Voice and data, especially video and multimedia applications, are carried by radio signals in the spectrum within certain frequency bands (higher frequencies than the 2G bands). The standard for 3G in western Europe is called UMTS (Universal Mobile Telecommunications System). The main difference with 2G and 2.5G is that a larger amount of information can be conveyed, which makes it suitable for regular Internet access and multimedia applications. 3G mobile telephony is not yet available.

A necessary condition for high dynamic efficiency is that firms bid at correct price levels for licenses. Overbidding increases the cost of capital (as markets discount the return on the infrastructure investment due to the decreased profit streams), which increases the marginal cost of an investment and hence reduces the level of investment or innovation. Also, overbidding may increase the incentives to collude later. The effects of underbidding are mixed: whereas the windfall profit of low bids may be neutral, it is likely that an inefficient allocation results, leading to low static efficiency. On the other hand, firms do not face the increased cost of capital and high indebtedness, so that funds to invest in innovation are more readily accessible.

The number of network operators is typically fixed because of the way the market was set up. Increasing the number of licenses drives down price and profits, increasing the degree of static efficiency. This relationship holds until the point where the benefits achieved via a lower price are outweighed by the duplication of fixed costs incurred through entry. Where these fixed costs are large, the point of excessive fixed cost duplication may be reached at a relatively low number of competitors. One possible way around this trade-off is the allowance of virtual competitors (MVNOs) which offer services over incumbents' networks. See the following bullet point.

In a situation of high static efficiency, several MVNOs may exist in the market. Too many may not be optimal, though, as it may lead to cut-throat competition and degenerated quality offerings. Also, introducing too many MVNOs too early reduces network operators' incentives to invest in upgrading and maintaining their networks. However, introducing MVNOs at a fair access price may encourage both product and further infrastructure development, increasing the level of dynamic efficiency. By changing the access price over time, a regulator can shift priorities.

Driver 2 Anti-competitive practices

In the absence of competition policy, firms might have incentives to collude based on their profit maximization behavior, although the stability of collusive pacts will depend upon the market conditions. Prices above the competitive level resulting from collusion decrease static efficiency.

Cross-shareholdings, no matter how small, directly reduce the intensity of competition and inflate prices—a point that does not seem to be commonly recognized. In the short run, cross shareholdings decrease static efficiency, while in the longer run, they may also reduce firms' incentives to invest in innovation (e.g. if the market is in region A in figure 2.2 in chapter 2).

Driver 3 Regulation and competition policy

Regulation may have a strong impact on the possibilities for MVNOs, that is, operators offering Carrier Select-type services on mobile networks. This affects both static and dynamic efficiency: cheap and easy access for MVNOs increases the intensity of competition, but may reduce network operators' incentives to maintain and upgrade their networks. The reverse is also possible, namely that MVNOs impose greater pressure on incumbents to innovate. See also driver 1.

Regulatory uncertainty, implying that firms are uncertain about regulatory measures, may effectively increase their costs of capital to compensate for the increased risk. For instance, ambiguity about access for MVNOs implies uncertainty about the profitability of network operators. The higher costs of capital reduces investment levels, and may therefore reduce dynamic efficiency. Even where certainty exists, but the regulator decides on relatively low allowances for rate of return, this decision will drive a reduction in the incentive to invest.

On the European level, licensing conditions across member states of the EC are not identical. For example, some countries have used auctions, whereas others have used beauty contests, the numbers of licenses vary across countries, and licenses apply to different time periods (length and starting date). This fragmentation makes it more difficult for operators to operate in more than one country, and is likely to reduce static as well as dynamic efficiency.

Driver 4 Technology

The future success and profitability of new technologies (including new services) is very difficult to predict. Therefore, technological uncertainty may lead to underinvestment, which may reduce dynamic efficiency. Conversely, if firms are over-optimistic, they may over-invest, which also harms dynamic efficiency. The future profitability of UMTS services is quite uncertain. Another example is "2G" WAP services. Note that this uncertainty not only affects network operators, but also equipment suppliers. For example, the timely availability of handsets for a new generation of mobile telephony is not evident (recall the joke that WAP stands for "Where Are the Phones?").

New technology industries such as telecommunications are susceptible to problems of compatibility and standardization. New technologies such as WAP and UMTS require standards and common communication protocols. Any uncertainty with regard to standards may induce firms to wait; they have an incentive to wait and see which standard will emerge rather than potentially committing themselves to a standard that becomes obsolete in a few years. This waiting will reduce dynamic efficiency relative to the case where standards are already determined. In Europe, common standards such as GSM and UMTS have been developed in a successful way, so this issue is of limited relevance in practice.

Lock-in effects may prevent operators from switching to an existing superior technology, which reduces static efficiency, or from investing in new technologies, which reduces dynamic efficiency. For example, if operators would fail to agree on the standard of a next-generation technology, under-investment is likely to occur. This may reduce static and dynamic efficiency.

Contrary to lock-in, firms may also switch to a next-generation technology too soon (excessive technological change). This may, for instance, happen in situations of "winner-takes-all" technologies, where firms have strong incentives to be the first one introducing a new technology. This is likely to result in over-investment, reducing dynamic efficiency.

Given the recent history of license allocation in the Netherlands, the most relevant efficiency drivers are market set-up and regulatory ways to deal with its consequences (e.g. regulation of third-party access). The next section contains an elaborate discussion of these issues. Before going into the analysis, we recapitulate the various drivers and their potential effects on static and dynamic efficiency.

Table 3.1 Overview of efficiency drivers

Driver	Static efficiency	Dynamic efficiency
1. market structure and entry	• license allocation –/+ • number of licenses –/+ • entry by MVNOs +	• license allocation –/+ • number of licenses –/+ • entry by MVNOs –/+ • pre-competitive alliances +
2. anti-comp. practices	• collusion – • cross-ownerships – • high interconnection prices –	• collusion – • cross-ownerships –
3. regulation and comp. policy	• effective oversight + • easy/cheap access MVNOs +	• effective oversight + • regulatory uncertainty – • easy/cheap access MVNOs –/+
4. technology	• technological lock-in –	• uncertainty about new technology – • standardization and compatibility problems – • technological lock-in – • excessive technological change –

By combining different drivers in the table above, one can see how certain states can arise. In principle, any combination of low or high static, and low or high dynamic efficiency is possible. Among the drivers analyzed here, there are no incompatibilities that may prevent certain states from prevailing. It is beyond the scope of this study to discuss all possible combinations, though. For the sake of illustration, consider the following two examples:

- The low-static/low-dynamic state may arise due to an inefficient license allocation coupled with ineffective competitive oversight. For instance, incumbent operators heavily overbid in an auction and won all the licenses, increasing their cost of capital as well as their sunk cost level. They try to recover fixed costs through tacitly colluding, which may be feasible if competitive oversight is weak. Since no newcomers managed to win a license, the level of innovation is lower than it could have been. Hence dynamic efficiency is relatively low.

- The high-static/low-dynamic state may arise due to access price regulation (for the sake of simplicity we ignore other drivers here). If MVNOs have cheap and easy access to networks, then of course the intensity of competition increases. However, it is possible that network operators' incentives to maintain and upgrade their networks are reduced, leading to low dynamic efficiency.

Relationship between static and dynamic efficiency

It is important to acknowledge that the relationship between static and dynamic efficiency is difficult to predict in practice, since in theory, there is no uniform conclusion. Because in this development, the set-up of the market takes place at a moment that is separated from actual competition in the market, both regions A and B in figure 2.2 in chapter 2 are possible. Accordingly, either more intense competition increases firms' incentives to invest in innovation (region A), or it decreases innovation (region B in figure 2.2). The occurrence of any of these regions depends on the actual intensity of competition. If competition is relatively mild, region A is more likely, whereas relatively intense competition is more likely to give rise to region B.

3.4 Moving between states

This section discusses in detail how the market may exhibit tendencies to move between efficiency states. For each state, we discuss a certain number of reasons for being in that state, and address whether the market can move towards a better state by itself or whether regulation is needed. Implicitly, we make a distinction between the current situation, in which the market is not yet operational, and the future situation in which networks have been rolled out and operators offer UMTS services.

Given that the market for 3G mobile telephony does not yet exist, it is crucial to discuss how pre-market phenomena (e.g. auctions, pre-competitive collaboration among firms) influence the initial state of the market, and how policy may try to prevent the low/low state from arising. In particular, we will consider how the allocation of licenses may be the principal cause of a certain market state, and discuss how inferior states can be prevented from arising.

Additionally, although the 3G market does currently not yet exist, it is important to consider policy questions that may become relevant when firms start offering 3G services. These policy questions naturally depend on the causes that gave rise to the efficiency state at hand.

3.4.1 Low static, low dynamic efficiency

Market set-up

Wrongly allocating licenses, or allocating the wrong number of licenses, introduces a tendency towards this inferior efficiency state right from the moment of the set-up of the market. An efficient allocation of licenses is a necessary, but not sufficient, condition for the high/high state to prevail. Arguably, in the Netherlands the set-up of the 3G auction was not optimal because of several reasons. The set-up gave the five incumbents in the 2G market (KPN, Libertel/Vodafone, Telfort, Dutchtone, Ben) an advantage: five licenses were auctioned without reserving slots for entrants. Most potential entrants dropped out before the auction started, while the single participating entrant quit during the auction. Accordingly, it was relatively easy for the five incumbents to obtain licenses—in fact, it is likely that the bidding during the auction was not as intense as it could have been, so that questions about the efficiency of the license allocation arise. An issue of particular importance is that consumers will not be able to benefit from increased competition owing to newcomers trying to gain market share (bad for static efficiency) as well as from innovative services introduced by newcomers in the market, since there are no newcomers (bad for dynamic efficiency).

The main policy tool to prevent this problem is to customize the design of license allocation to the market environment. It is not sufficient to use an auction design that was successful in the past in a changed environment. As Klemperer (2000) writes "Auction design is a matter of "horses for courses", not "one size fits all"". Whilst a detailed analysis of auction theory is beyond the scope of this paper, different types of auctions and more entry encouragements will increase the efficiency over more simple designs.

Uncertain technologies

Since UMTS services have not been established yet, it is hard to predict the future attractiveness of the 3G market. Whereas in general technological uncertainty may lead to under-investment by firms, in some countries in Europe operators may have been over-optimistic, leading them to place high bids in auctions. The rather slow introduction and consumer pick-up of WAP services may be one of the reasons why financial analysts have downwardly adjusted their optimistic beliefs about UMTS. Banks have reduced the credit rating of some telecommunications operators, for instance KPN, thereby increasing these operators' cost of acquiring new capital. Accordingly, the uncertainty about profitability of new services may create a tendency towards low dynamic efficiency due to initial over-investments.

It is likely that the future profitability of new markets is even more difficult to predict for regulators and policy makers than for firms. Therefore, even though technological lock-in and excessive technological change may occur, it is best to leave investment decisions to private parties as much as possible. Nevertheless, policy can play an important role in trying to prevent problems of coordination and compatibility by facilitating standardization for new technologies.

The market does not necessarily have a natural tendency to avoid this state. If one expects the 3G market to start in the low-static/low-dynamic state because of technological uncertainty, it is best to aim at increasing dynamic efficiency only, although to a limited extent, it may be possible to increase static efficiency. More importantly, high dynamic efficiency can be ensured before the market comes into existence, and may lead to a breakthrough in market structure which makes more intense competition possible (and thus take care of static efficiency as well). Given that licenses have already been allocated, there are several ways to make it more likely that the market will start with high dynamic efficiency:

- Stimulate a quick network rollout through antenna site sharing.
- Allow operators to cooperate in network rollout, so that they can save fixed costs of networks and have more funds to direct towards investment in innovation. Overall, duplication of fixed costs is avoided. Apart from the straightforward cost argument, it also enables more intense competition for customers while the threat that operators go bankrupt is greatly reduced. Therefore, cooperation has a direct impact on productive efficiency and an indirect effect on allocative efficiency (through increased competition). Hence a positive side-effect is that static efficiency is increased as well, while the risk of collusion can be kept at an acceptable level by effective competition oversight (e.g., by implementing certain conditions for infrastructure sharing). A risk is that by making the network more homogeneous, services offered over the network become more homogeneous as well, increasing the intensity of competition without generating more variety for consumers. Hence static efficiency may be harmed to a certain extent.
- Eliminate regulatory uncertainty (if there is any), so that operators are not unnecessarily discouraged from investing in new technologies.

Thus, given the right circumstances, the market itself may prevent a situation of low dynamic efficiency, while making high static efficiency more likely as well in the somewhat longer run.

The previous drivers applied to pre-market conditions. We now move to drivers in an existing markets. Two remarks are important before starting. Firstly, one can expect that when operators start offering UMTS services in a couple of years, that the uncertainty about the profitability of the market, and about so-called "killer applications," will be greatly reduced. Hence, technological uncertainty is less likely to drive low dynamic efficiency at that moment (i.e., when the market has come to existence). Other reasons that may drive low dynamic efficiency, such as the license allocation, are now bygones. Therefore, when the 3G market has started, increasing dynamic efficiency typically takes places through influencing the intensity of competition or by stimulating entry by newcomers (e.g. MVNOs).

Ineffective third-party access

A reason for being in this state is that static efficiency is low due to little competition among network operators while there are no MVNOs. If the market is in situation A (figure 2.2), then this goes together with low dynamic efficiency.

It is, in general, unlikely that firms will deliberately intensify competition. In a low static-efficiency state, policy can accomplish this by, for instance, increasing the pressure from competition authorities. An example is forbidding cross-shareholdings, as they tend to reduce the intensity of competition and static efficiency. Another example is regulating interconnection tariffs at marginal-cost levels, if the reason of low static efficiency is that there are high mark-ups in terminating access fees.

More directly, the regulator can facilitate access for MVNOs, but this introduces the risk that dynamic efficiency is further reduced. This risk, however, is likely to be small, for a low intensity of competition is most likely to give rise to region A (figure 2.2). Indeed, to a certain extent it should be possible to give MVNOs access to networks at fair access prices (i.e. sufficiently high), which do not reduce network operators' incentives to maintain and upgrade their networks, and at the same time introduce some pressure in the market to increase the level of innovation. Increasing the intensity of competition by facilitating entry for MVNOs may introduce new services from entrants as well as increase the pressure on incumbents to remain innovative themselves (especially if the market is in region A in figure 2.2).

3.4.2 Low static, high dynamic efficiency

The state of low static and high dynamic efficiency can come into existence not only due to "exogenous" efficiency drivers, but also as an intermediate state between the low-low and high-high state (see the discussions in subsection 3.4.1). In the most relevant cases, the remedy to get out of this state and move towards the high/high state depends on regulatory measures with regard to third-party access.

Market set-up

If one expects the 3G market to start in the low-static/high-dynamic state due to an ineffective license allocation mechanism, the focus should obviously be on increasing static efficiency.

Given a fixed number of licensed operators, it is unlikely that the operators themselves will take care of the problem. Hence, the regulator can aim at introducing third-party access. Just as discussed in the previous state, it is important that access prices provide a fair rate of return to network operators, in order to maintain high dynamic efficiency.

A caveat is that opening networks for MVNOs without pre-announcement before the auction, may be seen as unreliable government interference (harming the government's reputation). Hence, while a bad market set-up will almost certainly deteriorate efficiency when the market comes into existence, and therefore naturally calls for added regulation or policy, it is costly for policy makers and regulators to introduce measures that were not announced before the auction and not taken into account by the bidding parties.

Ineffective third-party access

If the market is in region B in figure 2.2, then a low intensity of competition is coupled with high incentives for innovation. Hence it is possible that a situation where third parties do not have access to incumbents' networks, induces network operators to invest in networks and services. In such a situation, the market is not likely to change by itself, since incumbents are not likely to open up their networks by themselves. Similar as in the low-low state, the regulator can facilitate access for MVNOs while taking care that incumbents' incentives to invest are not eroded (see the previous state). Hence, access should not become too cheap. See the previous state for a more elaborate discussion.

3.4.3 High static, low dynamic efficiency

This state, characterized by high static and low dynamic efficiency, is unfortunate in the sense that it is not part of the natural path towards the high-high state. It is most likely to arise because of an unsuccessful license allocation. To avoid too much repetition, we will not repeat that here (see the low-low state). Instead we only discuss a reason that can arise in an existing market.

Ineffective third-party access

An important reason for being in this state is that third-party access leads to intense competition but erodes network operators' incentives to invest in infrastructure (region B in figure 2.2; notice that there is a subtle difference with the previous state). Vigorous competition from MVNOs can occur if access prices are too low.

It is very difficult for firms to get out of this situation themselves, since presumably, regulatory intervention lead to this state. In such a situation, access regulation should aim at increasing access prices, although not to levels that prevent access altogether. This eases the pressure on incumbents that prevented them from investing in networks and new services. Care should be taken that the intensity of competition does not suddenly drop too much, as this would reduce static efficiency.

3.4.4 High static, high dynamic efficiency

Obviously, this is the state that policy is aiming at for the longer run, as well of a feasible follow-up of the low-static high-dynamic state. It is unlikely that any new market will start in the high/high state. Typically, competition does not seem to be very intense in newly-born, fast-growing markets, although the intensity of competition may quickly increase over time as firms try to gain market share fast. If the market happens to be in the high/high state, it may not be easy to remain there, although this depends on how the state came about (see below).

Market set-up

An important beneficial factor for this state to arise is an effective license allocation. See the previous states for more elaborate discussions. A side-effect of an effective allocation is that less regulation will be needed later on. Hence, if the market moved towards this stated without heavy-handed regulatory interventions, then light-handed but effective competition policy is required to ensure that the market remains in this state.

Effective third-party access

Access for MVNOs is well-balanced: it increases competition up to the level where firms keep their incentives to invest in innovation. If extensive access regulation has lead to this state, then it is unlikely to be stable in the longer run, and trying to remain in this state is, unfortunately, a very fine line to walk. The situation is more fortunate if third-party access has come about without imposing it, but by mutually beneficial agreements between network operators and MVNOs.

3.5 Conclusions

In general the 3G telecommunication market players, left by itself, are unlikely to have the proper incentives to move to the high-high efficiency state. The reason for that is the fact that a tight oligopoly has incentives to behave anti-competitively. The only market driven route to the high-high state is via increasing dynamic efficiency. If the incentives are such that operators are investing in technology and networks, and these investments imply a breakthrough in market structure, then the route to high/high might be open. In the Netherlands this route seems to be blocked, given the design of the UMTS auction. Without a clear market driven route to the high efficiency state, there is a role for policy in influencing the market parties to move to the high-high state.

Because of the delicacies involved in the set-up of the market and the presence of technological restrictions related to spectrum scarcity, the low/low state is the most likely place to start and the easiest to regress back to. Therefore it is vital that the market is set up with both the correct infrastructure and the correct regulatory structure. Given that the UMTS auction in the Netherlands is not considered to have been successful, regulation seems to be called for.

The general policy lessons for this development can be summarized as follows.

Lesson 1 Market set-up

(a) Optimal license allocation is vital. Badly designed auctions or inefficient beauty contests are likely to lead to inferior efficiency states. An optimal auction not only selects efficient players (static efficiency), but also ensures an optimal number of players and adds to proper incentives to invest in product and process innovations (dynamic efficiency).

(b) The choice for competition policy versus regulation depends on the history of the market set-up. When license allocation is optimal, competition law can deal with most (but not all) problems. Market incentives are proper, so roughly speaking, competition law should be enough to keep the players in line. Some additional regulation on compatibility, interconnection or other co-ordination problems can be used additionally, but the overall picture is light-weight regulation. When license allocation is not optimal, additional regulation is likely to be necessary, since many problems can occur. Zooming in on the actual situation in the Netherlands, where the set-up of the auction may have led to underbidding and a selection of same players as in the 2G market, what can be done? If the market starts in a situation of low static and low dynamic efficiency then action is needed, since it is not expected that investment and network rollout will create a market structure breakthrough. Supposing that investments in networks come to a standstill because of capital market problems, allowing the operators to selectively cooperate with building and maintaining networks may be a good solution to increase dynamic efficiency. These advantages should be weighed with possible new cooperation and competition problems. Question are: (i) Who is responsible for congestion? (ii) How to coordinate investments in the network? (iii) Will there be enough competition in services after sharing?

A possible way to repair market set-up damages is by introducing MVNO's. Opening up networks for MVNOs without pre-announcing before the auction is damaging the government's reputation. The only way out seems to specify a set of conditions on the level of competition. If there is too little competition, resulting in low static efficiency, the networks will be opened for MVNOs under conditions that were specified in advance. If the market starts in the low/high or high/low state, then the market itself if likely to solve the problems at hand.

Lesson 2 Importance of technology.

Dynamic efficiency is enhanced by stimulating compatibility and standardization, issues the market does not always handle effectively. It is, however, more important to stimulate standardization per se than to direct the market towards a certain standard. In the same manner, it is best to leave the assessment of future profitability of new-generation technologies (such as UMTS) to market players, in order to prevent unnecessary technological lock-in or excessive technological change. This can be done by auctioning off licenses rather than organizing a beauty contest.

4 Development Two: Excess Demand for Broadband Capacity

4.1 Introduction

In this development we discuss the case where the demand for broadband services exceeds supply. "Broadband" refers to high speed transmission of data, and allows internet connections to be "always on", contrary to narrowband, dial-up access associated with traditional telephony lines. Broadband is necessary, for instance, to view interactive, graphic-rich web-pages, and to enjoy innovative services such as audio content of digital quality, streaming video, video-on-demand, and video-conferencing.

Currently, there is a lot of capacity in the long-distance networks and the fibre rings surrounding the big cities in Europe. However, there is a capacity bottleneck within the local connections (the "last mile") to customers outside these areas. The cost of connecting local users has turned out to be very high, giving incumbent telcos a big advantage, and cable operators a somewhat smaller advantage (because they have to invest in upgrading their networks), over entrants. The demand for capacity is larger than supply, creating relatively high prices that slow down the growth and use of the Internet in many countries (although prices for, e.g., ADSL are decreasing). Moreover, the widely predicted quick and smooth rollout of broadband service hasn't happened yet for a variety of reasons. There have been technical glitches, financial problems, declining market sentiments and regulatory difficulties. At the same time, increasing competition may be seen as a means of bringing high bandwidth capacity prices down to more widely affordable levels. Throughout this chapter, the current situation of limited last-mile capacity coupled with increasing demand is considered to be exogenous. The road ahead and the possible destinations are partly endogenous, depending to a certain extent on policy.

Policy question

Clearly, the current situation is highly immature. Hence, one has to address whether the market is likely to move to a better situation, either by itself or guided by policy, and what a mature market may look like. The central question in this chapter, which will be addressed both from the current as well as from a future perspective (mature market) is the following:

How can policy and regulation either promote or not stand in the way of growth and investments in broadband capacity, and assure that consumers benefit from competition?

Overview of broadband technologies

At present, incumbent operators mostly use the **copper local loop** (i.e., copper-based customer access networks), carrying analog signals, to transport voice and to provide Internet access. Ordinary telephone lines deliver Internet access at a relatively slow speed, and are hardly or not capable of delivering video services. This is sometimes referred to as narrowband, characterized by dial-up access and slow content delivery. **Broadband** stands for the capacity to deliver "high-bandwidth" services, such as fast Internet access (interactive, graphic-rich web-pages) and multimedia applications (e.g. television, audio content of digital quality, video conferencing).
The most important technologies that allow for broadband access to end-users are:

- ISDN (Integrated Services Digital Network): A digital technology (and internationally agreed standard) that was originally developed to transmit voice, but is now adapted to carry data and video. It is a reliable technology, although slow for broadband, and can simultaneously transmit voice and data.
- DSL or xDSL (Digital Subscriber Line): A family of technologies that increase the bandwidth of existing copper wire networks. DSL transforms ordinary telephone lines into high-speed digital lines that can support fast Internet access and video-on-demand. Perhaps most important in the short term for the residential market is ADSL (Asymmetric Digital Subscriber Line). "Asymmetric" refers to different upstream and downstream bandwidths: it makes it possible to send a large quantity of data (e.g. video) to consumers, and send a small quantity (e.g. voice, control data) back to the operator.
- Cable modem: Cable operators can upgrade their existing cable TV networks so that they are capable of providing two-way traffic and interactive broadband services. Consumers then have to install cable modems. Whereas ordinary telephony lines consist of twisted copper pairs, cable networks often use coaxial cabling (i.e., cable with a metal core and insulated sheath) or hybrid fibre/coaxial cabling. A drawback of cable technology is that an increase in the number of active users in the same neighborhood substantially reduces the speed of each of their connections.
- Broadband wireless access (broadband wireless local loop): Wireless link that beams voice and data over the airwaves and allows entrants (1) to bypass incumbent operators' access networks, and (2) to avoid the large fixed cost of laying cables in the ground. Most systems are still in trial.
- Fibre: Besides wireless, entrants can also use optical fibre to bypass copper local loops, although they then incur the fixed cost of network rollout. It is not very likely that there will be large-scale rollout of "fibre-to-the-home" in the short run.

Broadband transmission can be asymmetric. That means that internet pages load quickly (transmission to the end-user), but that sending e-mails and clicking (transmission from the end-user) is much more slow.

Netherlands

Broadband Internet access is available on a limited scale in the Netherlands. KPN offers DSL services under the brand name Mxstream. UPC offers, in some parts of the country, Internet via cable modems under the brand name Chello. Other cable operators also offer Internet access, but not yet in the whole country. Spectrum auctions for wireless local loop have been postponed; it is unclear when they will take place. There are no signs that fibre to the home will be rolled out in the near future.

4.2 Background

Before going into efficiency states and policy cures, the following box provides a nutshell overview of broad technologies.

Next, it may be helpful to gain more understanding about access to local networks for third, independent parties. The following box provides a brief introduction.

Local access and third-party access

Local access
Local access is the connection between the customer's premises and a local switch in a telecommunications network. Entrants without their own local access network can, in principle, have access to end-users by using existing networks of incumbent operators. This situation is usually referred to as "third-party access."

Third-party access
In the case of traditional telephony networks, third-party access typically occurs by unbundling an incumbent's local network, implying that an entrant can lease or own the incumbent's local connection and directly serve end-users. For unbundling to be possible, "colocation," that is, opening up facilities (e.g., space and technical equipment) for entrants, is necessary.
Local connections in cable networks cannot be unbundled. Nevertheless, third-party ISPs can offer high-speed Internet access to cable customers if cable operators have upgraded their networks to support two-way broadband traffic.

Scarcity problems
Scarcity seems to be a larger problem in cable access networks, due to their serial nature, than in regular, upgraded telephony networks. For example, ADSL may be a better solution for scarcity problems: the copper wire has turned out to be more "upgradable" than cable, in the sense of adapting it for higher speeds of access. In the Netherlands, this may give KPN's ADSL offerings (brandname Mxstream) an advantage over cable operators such as UPC. Overall though, it is still uncertain to what extent scarcity will occur.

In the next sections, when we present the drivers of the efficiency states and discuss policy options, we will regularly use terminology introduced in the boxes above.

4.3 Efficiency drivers

Given that this development and the previous one have in common that networks have to be upgraded or newly rolled out, the drivers for static and dynamic efficiency are roughly the same. A major difference is that in this development, there is no history of a license allocation. Instead, the rise and fall of old and new players in the market is much more fluid. Entrants are more or less free to roll out new networks. A common theme is that third-party access (and its regulation) strongly affects both static and dynamic efficiency, in roughly similar ways.

Dynamic efficiency in the context of providing broadband capacity may be measured through a number of ways. The most important proxies for dynamic efficiency in this development are: investment in broadband capacity, introduction of new, innovative services and content that ask for broadband connections. High dynamic efficiency is characterized by the sufficiently rapid introduction of sufficient capacity, whether competing networks or a natural monopoly situation, coupled with the rapid introduction of new services.

Driver 1 Market structure and entry

The number of suppliers with their own local infrastructure and the number of suppliers without local connections to end-users directly affects static and dynamic efficiency. Let us look at static efficiency first. At least in the short run there is only a small number of alternatives for local networks available. Therefore, the intensity of competition typically is low if entrants without networks do not have access to incumbents' networks, which is likely to happen if third-party access is unregulated. Thus static efficiency is typically low in the current situation, and will remain so in the short run.

Static efficiency can be increased by opening networks to entrants and regulating the terms of access (see also driver 3 on the terms of access, and see also the discussions on MVNOs in the previous chapter). That is, to the extent that it increases competition within feasible bounds. If the bottleneck capacity of the last mile deteriorates due to third-party access, then static efficiency may actually decrease. Also important is the ease, speed, and terms at which network operators and cable providers are able to ink deals with multiple internet service providers and other content providers.

Depending on whether there is a positive or negative correlation between the intensity of competition and the incentives to innovate (region A or region B in figure 2.2, chapter 2), dynamic efficiency will then be increased respectively reduced. A priori we allow for both possibilities. In addition, one should not rule out that third-party access increases incumbents' incentives to maintain and upgrade their networks because of the increased possibilities of generating revenues, or that limited possibilities for third-party access stimulate entrants to bypass existing networks. Both possibilities are good for dynamic efficiency.

Competition will be increased if network rollout occurs. Whether the benefit of increased infrastructure competition will outweigh the cost of duplicating the local loop will depend on the degree of resulting competition. Between the two extremes of full infrastructure competition and third-party access in a monopoly situation, there exist a host of variations where entrants use more or less parts of incumbents' networks. At the limit of high static efficiency, several operators may exist in the market using existing pipes, or several firms may be competing in offering access. A priori it is difficult to tell which possibility yields the highest level of static efficiency, given that there is incomplete information about costs and benefits of laying new networks.

Driver 2 Anti-competitive practices

In a situation where entrants wish to have access to incumbents' networks, for instance through local-loop unbundling, incumbents may try to delay colocation and unbundling requests from entrants.[4] Also, they have strong incentives to charge inflated access prices for third-party access. This directly harms static efficiency, since it delays and stifles competition. It also slows down entrants' ability to offer innovative services over existing pipelines and hence harms dynamic efficiency (if the market is in region A in figure 2.2). This in turn makes incumbents "lazy" in the sense that it reduces the pressure on them to remain innovative themselves. Arguably, it increases entrants' incentives to invest and build networks themselves. In that respect it need not always be bad for dynamic efficiency. A report by Analysys argues that by making unbundling and colocations more difficult, incumbents reduce entrants' incentives to invest in DSL technology, while it makes them consider alternative access technologies, such as fixed wireless, cable modems and fiber to the home.[5]

Standard anti-competitive practices, such as collusion and cross-ownerships, may occur. We will not discuss these here. See also the previous chapter.

Driver 3 Regulation and competition policy

Access for entrants without networks to incumbents' networks may strongly affect efficiency.[6] Regulation that facilitates third-party access stimulates competition and hence static efficiency. However, cheap and easy access may reduce network operators' incentives to maintain and upgrade their networks. The reverse is also possible, namely that it imposes greater pressure on network operators to introduce new services themselves (region A in figure 2.2). Also, consumers miss services by new firms, which might have resulted in more innovation compared to a monopoly situation. On the other hand, if access is difficult or expensive, entrants have stronger incentives to bypass incumbents' networks (region B in figure 2.2). Third-party access at an price that is not too low (taking fixed investment costs of network operators into

[4] For instance, in Germany, Colt has filed a legal complaint with regulator RegTP on unbundling against incumbent Deutsche Telekom. In the UK, several operators have filed a joint complaint with regulator Oftel, charging incumbent BT with discriminatory behavior in allocating colocation space in local switches. Also in the Netherlands, entrants have been complaining about the incumbent's charges for colocation and delay tactics.

[5] Report by Analysis is discussed in "German, UK operators in bid to force DSL unbundling progress", *Communications Week International*, 13 November 2000, www.totaltele.com.

[6] Potentially the type of universal service that voice telephony currently provides will never be economically feasible for high bandwidth access. Earlier competition through third-party access will speed up the introduction of broadband services in the most economically viable areas (such as cities) but may erode the cross-subsidies that would have allowed a universal service to be developed.

account) may encourage both services and infrastructure investments, thereby increasing dynamic efficiency. Also, access regulation can be time-dependent: by gradually making access for entrants less attractive over time, the incumbent's incentives to invest in the local loop need not be reduced.

Dynamic incentives critically depend on what Cave et al. (2001) call eligibility (see box in section 4.1), that is, the extent to which entrants face access terms that depend on their own investment levels. Accordingly, access prices can be used to guarantee that dynamic efficiency is not negatively affected.

An important issue in the short run concerns the conditions of local-loop unbundling (LLU), that is, unbundled access to the incumbent's local loop and local switches (this is important for entrants aiming to offer ADSL through LLU). It is important that regulators keep a close eye on collocation and unbundling requests, in order to prevent delay tactics by incumbents that harm static and possibly also dynamic efficiency (see driver 2).

Finally, similar to development 1 (chapter 3), uncertainty in regulation will increases the costs of capital to compensate for the increased risk.

Driver 4 Technology

Prioritization technologies: Network operators can use technology to build new differentiated levels of service for internet content, and mechanisms to control and manage internet data. For instance, traffic-type identification (into e-mail, Web-content, voice, data) allows ISPs to isolate different traffic types in networks by, for example, the type of application, brand, interface used, user identification, and site address. According to Cisco, a supplier of internet infrastructure technology, this enables an ISP to implicitly discourage the use of certain sites by slowing down the speed of access, or to promote and offer the ISP's own or its partners' services with full-speed features.[7] In a similar vein, cable modems can be deployed to force consumers to use proprietary software to use internet services over cable, and to restrict consumers' access to internet sites. While such technology may actually increase network efficiency, clearly there is a risk that static efficiency is be reduced. Also, to the extent that restrictive technologies slow down consumers' exposure to innovative services introduced by outsiders, it reduces suppliers' incentives to develop new services in the first place. This is bad for dynamic efficiency.

There may be uncertainty about the nature of the technology best suited for broadband access. For instance, what type of xDSL is best? Is cable better than xDSL? How costly will the modifications needed to make phone lines run at DSL speeds actually be? Technological uncertainty reduces firms' incentives to invest, leading to less dynamic efficiency than in a situation of complete information. Also important here is the state of market at the demand side. Whether there is an immature or a mature market, uncertainty about size of the market,

[7] See Cisco (1999).

uncertainty about profitability of market, all influence dynamic efficiency. Compare this to the 3G mobile market: also here no-one really knows what consumers are willing to pay for broadband applications, and how much time they want to be on-line, and what types of services will be most valued. Demand uncertainty is typically bad for dynamic efficiency, as it makes it difficult to know whether investments can be recovered. Conversely, a large, expected demand is good for dynamic efficiency. Typically, strong demand growth gives competing firms room to grow without competing intensively, which encourages investments.

An existing technology may prevent switching to a new technology (lock-in). Suppose that there is a fine mazed cable infrastructure which is, for the most part, not yet adapted to two-way broadband access. For cable operators it is much easier and less costly to enter the broadband market than for new facilities-based entrants (who have to roll out their own networks). Suppose that from a social point of view, adding new networks (e.g. fibre to the home) generates a larger surplus than only upgrading existing networks. Nevertheless, the presence and potential competition from cable operators may prevent newcomers from rolling out networks. In such a situation, lock-in effects reduce dynamic efficiency. Conversely, firms may adopt new technologies too soon (excessive technological change). This is a risk in the case of "winner-takes-all" technologies, which may lead to over-investment and reductions in dynamic efficiency.

Most relevant for the Netherlands is the current discussion about opening the national cable access market to competing Internet Service Providers (ISPs). Consumer pressure driven by complaints about service failures on UPC's cable network, has lead to a debate about stripping cable operators of their monopoly position earlier than originally planned.

The following table summarizes the efficiency drivers and depicts their potential, qualitative effects on static and dynamic efficiency.

Table 4.1 Overview of efficiency drivers

Driver	Static efficiency	Dynamic efficiency
1. market structure and entry	• nature of entry and competition –/+ • third-party access –/+	• nature of entry and competition –/+ • third-party access –/+
2. anti-comp. practices	• incumbent's tactics to delay network access – • collusion – • cross-ownerships –	• incumbent's tactics to delay network access –/+ • collusion – • cross-ownerships –
3. regulation and comp. policy	• effective oversight + • cheap/easy access third-parties –/+	• regulatory uncertainty – • cheap/easy access third-parties –/+
4. technology	• technological lock-in – • prioritization technology –/+	• prioritization technology – • technological lock-in – • excessive technological change – • uncertainty about success of new technology –

By combining drivers from the table above, potential reasons for each of the possible states can be constructed. Obviously it is not very useful to discuss all possible combinations. Instead, we will concentrate the discussion to the most relevant ones (see the following section). For the sake of illustration, two examples are the following:

- The low/low state may arise due to making third-party access cheap and easy. Suppose that although competition becomes more intense, the capacity bottleneck becomes more stringent (leading to quality degeneration), so that static efficiency is reduced. Dynamic efficiency is low because network operators lose their incentives to maintain and upgrade their networks.
- The low-static/high-dynamic state may arise if incumbents successfully ward off entrants that want to have access to their local networks, giving entrants incentives to bypass incumbents' networks. Of course, whether this is good for dynamic efficiency crucially depends on the network improvements and services that entrants are able to offer.

In the following section, we will discuss the most important reasons for all possible states in more detail. Before that, we provide some further comments on the relationship between competition and innovation.

Relationship between static and dynamic efficiency

It is important to notice that static and dynamic efficiency may be related. As in the previous development, we do not make predictions about this relationship. Instead, we allow for opposing situations (i.e., the relationship is either increasing or decreasing, as depicted in figure 2.2) and use them throughout the analysis. We will briefly recapitulate the possibilities that we allow for.

Incumbents may protest against forced third-party access, e.g. through unbundling of their last mile. While third-party access may increase the level of competition and hence static efficiency (at least if capacity constraints do not become overwhelmingly stringent), it is possible that lower expected profits for incumbents reduce their incentives to invest in the network (region B in figure 2.2). If this is the case, there is a trade-off between static and dynamic efficiency.

Conversely, there is also the possibility that increased competition spurs development of innovative services and new means to offer bandwidth (region A). Note that in this case ,there is no tradeoff between static and dynamic efficiency. A priori, we allow for both possibilities (i.e., either one of the regions A and B may apply).

To conclude this section, we discuss recent work that introduces interesting and new perspectives to the discussion about the effects of competition on innovation.

Cave et al. (2001) explore the relationship between access price regulation and infrastructure competition. Their main conclusion is that it is crucial to look at the dynamics of competition between incumbent and entrants and among entrants, and the dynamic process through which an entrant establishes itself in the market. Concerning the first aspect, they find no evidence that high access prices encourage infrastructure competition. Concerning the second aspect, they argue and motivate that “facilities-based entrants” (rolling out their own networks) are heavily dependent on incumbent’s wholesale services early on, and gradually become self-reliant as they roll out their networks. Therefore it is important that access prices should not be prohibitively high initially, but may be gradually increased later on.
Cave et al. jointly consider access prices and “eligibility”. Wide eligibility means that entrants have access to wholesale prices irrespective of their own investments in networks, whereas narrow eligibility means that investing in networks is a necessary condition to have access to wholesale prices. The optimal access price regime depends on the eligibility. For the purpose of this study the main findings are:

- Wide eligibility and low access charges are good for static and bad for dynamic efficiency.
- Wide eligibility in early days of competition is bad for dynamic efficiency.
- Effects on dynamic efficiency critically depend on the extent to which assets by the incumbent are replicable.
- Time-dependent access pricing (rising) implies that in early days (when most assets of the incumbent are not replicable) easy access is possible, but firms have a rising incentive to replicate the incumbents assets whenever feasible.

4.4 Moving between states

In this section, we explore the tendencies in the market to move from one state to another one, and how policy can affect this. We do this by discussing, for each possible state, a certain number of reasons for being in that state.

In the current situation of an immature market, one cannot have high static and high dynamic efficiency at the same time. Neither is the high-static/low-dynamic state likely to occur. Hence, either the low-low state or the low-static/high dynamic state is the most likely state to occur in the situation we are currently facing.

4.4.1 Low static, low dynamic

Technological uncertainty

Typically, in an immature market static efficiency is low, since there are very little players (and hardly any competition), there is scarcity in capacity, and demand is still undeveloped or uncertain. Dynamic efficiency, on the other hand, may be low or high. To see this, remark that it is likely that the current capacity shortages in broadband communications are of a temporary nature. Reasons for such a situation are related to technological uncertainty: there may not yet be a cost-effective technology that makes bypass economically feasible, or there is still uncertainty about the attractiveness of the market for broadband services. Firms may be waiting until they have more information and confidence in the profitability of the market. The only reasons for low dynamic efficiency in this case are technological and regulatory uncertainty leading to under-investment.

Firms themselves may find it difficult to overcome low dynamic efficiency. Policy can prevent this state as far as low dynamic efficiency is concerned. Firstly, its crucial role is to eliminate conditions that withhold firms from investing in innovation. Nevertheless, it is best not to interfere with firms' assessments of the attractiveness of their investments. That is, policy should be technologically neutral. Secondly, third-party access at fair prices to incumbents' networks, possibly for some predetermined period coupled with pre-stipulated investment conditions, may be useful to speed up consumers exposure to newcomers and their services, without harming incentives to invest in infrastructure (both by incumbents and entrants). Clearly, access policy has to be carefully balanced. On the one hand, access prices should be low enough to make entry feasible. On the other hand:

- Access regulation should allow for access at wholesale prices that do not discourage network upgrading and maintenance by incumbents. Nor should third-party access make capacity problems worse. Hence, access prices should be sufficiently high.
- Access regulation should not discourage bypass investments, if entrants consider such investments to be worthwhile. Optimal access policy is likely to include a rising scheme, at least for that part of access where entrants can replicate the assets of the incumbent. Extra dynamic incentives can be created by letting access terms depend on investment levels of entrants ("eligibility", Cave et al 2001; see the previous box).

Pre-announcing policies concerning third-party access in a timely fashion will prevent regulatory uncertainty. In addition, competition policy should prevent that incumbents stifle innovation (e.g. through delaying access or unbundling agreements). There must be regulatory frameworks that clearly state the requirements on incumbents.

The most relevant possibility for the Netherlands is an immature and uncertain market. There is uncertainty about demand (experience is still limited), the nature of access conditions related to local-loop unbundling is uncertain and the auction for wireless local loop spectrum has been postponed. There is currently no mandated access to the cable and there will be increasing scarcity in cable networks if demand grows (which is very likely to happen). Therefore, mandating third-party access is an obvious candidate to help the market to overcome the current situation, if one takes the necessary precautions so that dynamic efficiency is not reduced (e.g., sufficiently high access prices, no regulatory uncertainty). In a report paid by Dutch cable companies, Nyfer suggests that mandated access is bad for dynamic efficiency (Nyfer 2000). We have seen that this crucially depends on the terms of access and economic theory is ambiguous (see also the appendix). It is always possible to impose third-party access and structure the requirements on incumbents in such a way that they provide a fair price to cable operators.

Market structure and excessive technological change

Now suppose that the uncertain situation has crystallized, and there is more information available about the attractiveness of rolling out new networks. However, alternative local networks have already been rolled out, perhaps too soon, or worse, based on overly optimistic expectations. Accordingly, the low/low state may arise in a mature market if there has been inefficient duplication of networks, which increases fixed costs and in turn may reduce funds available for investments in innovative services.

Will the market leave this state on its own, given those reasons for being in this state? In the situation sketched above, industry consolidation is likely to occur, so that the market is likely to move out of this state sooner or later. If that happens, the remaining firms will find themselves in a healthier environment, with better possibilities to invest (such as in the low-static/high-dynamic case). The main role for policy is to monitor consolidation moves, such as mergers and acquisitions, to make sure that increased concentration does not prevent effective competition.

Ineffective third-party access

We distinguish to more cases for the low/low state in a mature market. Firstly, competition may be very intense, leading to degenerated quality offerings and lower static efficiency, and also further reducing dynamic efficiency (as in region B in figure 2.2). Entrants have no incentives to bypass existing networks; there are many entrants without their own networks, the physical bottleneck in the local network is strongly experienced, there is cut-throat price competition, network operators do not have incentives to upgrade and maintain their networks. Secondly, there may be too little competition, providing too much comfort for incumbents and taking away their incentives to remain innovative (region A, figure 2.2). There is no entry in the market, and incumbents face little downward price pressure. Because of high fixed costs, entrants have no incentives to bypass existing networks. Incumbent network operators enjoy entry barriers and have weak incentives to offer new services, to upgrade and maintain their networks.

In both cases, policy is desirable, for the firms may have difficulty to overcome the inefficient state themselves. In the former case, if intense competition is due to cheap and easy third-party access, regulation is apparently too harsh on incumbents. A reassessment of the terms of access is needed then. In the latter case, competition can be intensified by stimulating third-party access at fair access prices.

Prioritization technologies

Prioritization technologies may reduce static as well as dynamic efficiency, and hence may give rise to various states. Here we single it out as a reason for being in the low-low state. For example, network operators offer differentiated service levels for internet content, effectively limiting consumers' exposure to third-party content. In turn, the incentives of newcomers to introduce new content may be reduced. Given that these technologies are most likely to be introduced and applied by incumbent operators (and not by entrants), it is important that policy tries to prevent that such technologies are used to implicitly introduce cable-operator imposed "censorship" or discrimination in speed of content delivery.

4.4.2 Low static, high dynamic

Before turning to reasons for being in this state, recall that this state may arise as an intermediary state while the market is moving from the low-low state to the high-high state. Compared to the high-static/low dynamic state (discussed later), this state provides better possibilities to ameliorate efficiency.

Technological uncertainty

Recall from the discussion of the low-low state that static efficiency is low in an immature market. Dynamic efficiency is not necessarily low, though, because firms may already be working on new access technologies. However, it may be hard to observe the actual level of dynamic efficiency, given that firms may be waiting and innovations still have to materialize. Firms typically have incentives to break out of the initial situation by investing in innovations.

Therefore, if initially the market is in the low-static/high-dynamic state, it is likely to move to a better state by itself in the short-to-medium run: introduction of broadband access technologies and network rollout by new players will introduce competition and increase static efficiency in the somewhat longer run. Regulation and policy should not introduce biases towards or against certain technologies, and neither should they introduce uncertainty. An immature market in the low-static/high-dynamic state should mature on its own as much as possible.

Ineffective third-party access

As we have seen earlier in similar cases, and also in the previous chapter, the low-static/high-dynamic may arise if a low intensity of competition prevents that incumbents' incentives to invest are not eroded. If the reasons for being in these states are related to third-party access (whether it is too easy, too difficult, or delayed by incumbents), then policy cures are rather obvious and we will not repeat them here (see the discussion above).

Market structure

An alternative reason for being in this state is that an entrant has rolled out fiber to the house while that technology has the characteristics of a natural monopoly. For example, the entrant has been able to overtake incumbents, effectively inducing them to exit the market for local broadband access. If that happens, by definition consumers do not benefit from competition, although they may strongly benefit from a new technology that does not exhibit capacity restrictions.

A natural monopoly typically arises for exogenous reasons, so that firms cannot take care for a move to the high-high state. Sector-specific regulation (e.g., retail price regulation) may then be desirable, at least to rule out excessive prices. It is important, though, that the possibility of this type of regulation is known in advance by firms that are considering to invest in fiber networks. Notice that at present, it is unclear (i) whether fiber-to-the-home will be economically viable, and (ii) whether it will allow for competition (i.e., it is not a natural monopoly).

4.4.3 High static, low dynamic

The state of high static and low dynamic is rather unfortunate, since it is not part of the natural path towards the high-high state. See also the discussion on this state in the previous chapter. For the market itself, it may be difficult to get out of this state, because a low level of innovation typically precludes a drastic shift in market structure. Hence there is a role for policy or regulation.

Market structure and ineffective third party access

The high-static/low-dynamic can occur if intense competition erodes the incentives of network operators to invest in their networks and new services. Hence, the structure of the market does not allow for market power that makers Schumpeterian innovation possible. A likely reason for this state is based on ineffective third-party access. For instance, access is too easy and cheap for entrants without networks, which erodes incumbent's incentives to maintain and upgrade their networks, and entrant's incentives to bypass existing networks (region A in figure 2.3). In this case, earlier discussed policy remedies on third-party access apply (see the discussion on this topic in the low-low state).

4.4.4 High static, high dynamic

The high/high state is obviously the most desirable state to be in, although it is unclear to what extent it is feasible, especially in the short to medium run. This state is not feasible in the short run because the market is still immature (which is currently the case). Nevertheless it is interesting to discuss how this state can come about and what can be done to remain in this state. Notice that it is most likely that the high-high state is precessed by the low-static/high dynamic state, since high levels of investments in innovation are likely to lead to a breakthrough in market structure, making more intense competition feasible.

Market structure: network competition

The most natural way to end up in this state is through the development of network competition, at least if the market is not a natural monopoly. If this happens, light-handed regulation and competitive oversight are likely to be sufficient to remain in this state, since access regulation becomes much less relevant.

Market structure: natural monopoly

Another reason for the high-high state is delicate balancing of access regulation, if the market happens to be a natural monopoly. In this case, however, it will not be easy to remain in this state, because regulation of third-party access will remain a necessity, giving rise to the continuing need of balancing static and dynamic efficiency. Accordingly, compared to the previous reason to be in this state, it is more likely that the market will slip back into the low-static/high-dynamic or high-static/low dynamic state.

4.5 Conclusions

Is there a role for policy in influencing the market to move to the high-high efficiency state? Similar to the previous development, the only market driven route from the low-low to the high-high state is via low-static/high-dynamic. If the incentives are such that market parties are investing in technology and network and these investments imply a breakthrough in market structure, the route to high-high is open. Cable companies or other market parties who invest in broadband such as e.g. fiber-to-the-home, can create a breakthrough in market structure, which in turn may imply a route to the high-high state.

The general policy lessons for this development can be summarized as follows.

Lesson 1 Third party access

Optimal access policy is vital. It can cater for both static and dynamic efficiency. To accomplish that, (i) tougher access terms are needed for those parts of the incumbent's assets that are replicable and (ii) differentiation of access terms for entrants that invest and entrants that do not invest in infrastructure is desirable.

Low access charges are good for static efficiency. Pre-announcing that the charges will rise gives proper investment incentives provided that the incumbent's assets at stake are replicable. Extra incentives for network rollout are possible if the regulator treats entrants that invest more favorably.

Lesson 2 Technology-neutral policy

A key question is how to set-up a regulatory regime that is not biased towards or away from certain technologies. Typically, the government is even more uncertain than the market about the optimal technology: cable, ADSL, WLL or fibre to the house. So it does not want to manipulate technology races by their regulatory choices. Yet, this is quite hard, since the government has to decide e.g. on mandated access for cable. Whatever it decides, it influences the competitive relationship between the various infrastructures. The principles should therefore read: whatever the decision is, it should not restrict the build up of (various types of) new infrastructure, but not push the market in some direction. Take mandated access to the cable as an example. Will that manipulate infrastructure choices or hamper dynamic efficiency? This is not easy to say. Lesson 2 further develops this point.

If the starting point is low static efficiency, regulation aimed at stimulating investments tends to be more successful if these investments create a breakthrough in market structure. A breakthrough in market structure (for the better) implies that there is a possible way out of the low-low state that is largely market driven, which is preferable for obvious reasons.

Lesson 3 Mandatory access

Mandated access to the cable is good for the television market and for static efficiency on the Internet market. The effects on dynamic efficiency are unclear. Potential harm to dynamic efficiency can be reduced by time-dependent access regulation. The argument runs as follows:

- Suppose that cable is the most attractive future option for broadband. Mandated access to the cable is then unlikely to deter profitable investments by cable operators
- Suppose now that cable is not the most attractive future option for broadband. Now mandated access yields ambiguous results. It depends how important competing infrastructures are. Suppose that fibre to the house is not realistic in the coming five years, and infrastructure competition is the best way of guaranteeing future static efficiency on the Internet broadband market. Access policy that focuses on dynamic incentives (investment in cable), along the lines discussed in lesson 1 above, is called for.
- Without mandated access on the other hand, cable operators might leverage their telecom and Internet activities to tighten their grip on the television distribution bottleneck. So if regulation of cable operators for their television activities is hard (and the UPC example shows that this might indeed be the case), mandated access can be good for efficiency on the television market, possibly somewhat at the expense of the broadband developments.

Lesson 4 Risk of prioritization technologies

Notwithstanding the intentions of suppliers of network infrastructure and switching/routing equipment to increase the internal efficiency of networks, it is conceivable that prioritization technologies are used to restrict consumers' access to Internet content or to discriminate in the speed of content delivery. Policy makers should keep a close eye on such restrictions, since they tend to harm static and dynamic efficiency.

5 Development Three: Tendency towards Consolidation and Alliances

5.1 Introduction

This chapter applies our methodology to a development that is quite different in nature from the previous ones. The first two developments were technological developments in a specific market. The third development relates to the exogenous trend that market structures change. The principal aim of the study is to show how various future policy issues can be analysed. By taking a completely different development this chapter adds to the richness of the methodology. At the same time, because this development is quite general (i.e. not related to a specific market), the insights are also fairly general in nature.

In this chapter we discuss the effects of horizontal and vertical consolidation and alliances on telecommunications markets. For the purposes of this study, we consider all tendencies toward a more concentrated industry, such as horizontal and vertical acquisitions, mergers, joint ventures and alliances.[8] In particular, we address how these developments affect static and dynamic efficiency, and how policy can try to enhance efficiency.

In the 1990s, there has been a lot of discussion about "convergence" which can be loosely described as the increasingly intertwined nature of, for instance, telephony, cable, TV, publishing, Internet, and entertainment. Mergers and alliances are sometimes motivated by convergence, which makes consolidation either a necessity or a vehicle to realize substantial "synergies."[9]

Telecommunications industries are consolidating both on a national and international scale. These integration developments extend vertically to cover the markets for the provision of "content" and Internet services.

Convergence can be seen as a technology-related background. Another background, which is not directly related to technology, is the trend towards internationalization, implying that the geographical scale of markets changes, leading to cross-border activities. As an illustration, consider AOL Time Warner, which consists of the US's biggest Internet company (America Online) and its largest media firm (Time Warner). The company's chairman and chief executive recently announced that they plan to rapidly expand in Europe, seeking to increase international revenues from 17 to 50% during the next ten years.[10]

[8] An analysis of the potentially different effects of these types of consolidation is beyond the scope of this study.

[9] The difficulty of predicting how convergence will take place is illustrated by the failure of WebTV, a home TV set to access the Internet. See "Perspective – AOL/Time Warner and the Convergence Myth," by M.A. Noll, published on www.totaltele.com, 5 February 2001.

[10] "AOL TW plans expansion," FT.com, 8 March 2001.

Policy Question:
The general question is: to what extent and how, should policy makers be concerned with consolidation and alliances?

This question has two important parts to it. Firstly, there is the fundamental question of whether policy makers should worry about increased consolidation of firms. Secondly, if it is perceived as a problem then what should be done about it? Specific issues to focus on are: horizontal and vertical consolidation, cross-border consolidation (e.g., are there problems with a foreign firm owning the network, and should one be concerned that firms operate in many markets as long as the domestic market can be separated and controlled?), and pre-market consolidation (e.g., R&D joint ventures).

Given that consolidation creates some issues that need to be addressed, whether the benefits of a regulatory body are outweighed by its regulatory burden remains to be seen. Clearly here is the tradeoff, introducing regulatory checks and balances may well increase allocative efficiency, but it may also reduce firms ability to create pan-European networks, seamless mobile and data services. This may not be an issue for some services whilst the lack of large scale and international consolidation may create problems.

5.2 Background

The exogenous development explored in this chapter is characterized by the fact that a "traditional" way of market definition is perhaps less useful, given that telecommunications services, infrastructure, equipment, and content markets are all part of the same consolidation trend. The traditional notion of market definition in economics seems too narrow to deal with multi-market competition and cross-market mergers and acquisitions. In particular, it does not take into account the possibility that a firm "leverages" its dominant position in one market into another market (see e.g. Hobbs, 2000). Leveraging of market power can take place if markets are "connected" in some way. In CPB (2000), markets are said to be connected if there exists a link between them (e.g. supply-side link, demand-side link, vertical link), while that link is rooted in firms' business operations or strategy, or in the market's demand side. The relevant market cluster is the set of markets that is obtained by adding all connected markets that affect the possibilities for entry on the narrowly defined relevant market where competition problems are surfacing (the particular market under scrutiny).[11] Concepts such as connected markets and leveraging are very relevant here. Nevertheless we will not formally define the relevant market cluster in this exogenous development. Instead we indicate that when analyzing mergers and

[11] The importance of an integrated approach to markets is also acknowledged in a recent working document by the EC DG IV (EC 2001).

alliances in telecommunication markets, an integrated approach to relevant markets is called for, taking into account services, content, infrastructure, and distribution.

The next two text boxes describe recent developments with respect to horizontal and vertical consolidation.

Horizontal consolidation

The telecom sector might be heading towards a market with a relatively small number of global operators, along with a large number of smaller operators with a narrower focus (Goldman Sachs, 2000). Currently observed takeovers may just be the start of an intensifying battle for growth, driven by the urge to satisfy shareholders' expectations. According to Balsinde et al. (2000), three groups of companies are striving for European leadership:

- **Large integrated corporations**, such as BT, Deutsche Telekom, France Telecom, Telecom Italia, Telefonica: they have scale, large customer base, shareholders with low expectations for earnings growth
- **Focused mobile and data communications companies**, such as Vodafone AirTouch, Colt Telecom: they have relatively small customer bases, shareholders with high expectations for earnings growth (most of the large, focused European firms are either mobile or data operators).
- **Small integrated companies** or the former ex-monopolists in the smaller countries, such as KPN Telecom, Belgacom, Portugal Telecom, Telia, Telenor, and Swisscom: they have relatively small customer bases, they are attractive acquisition targets for companies of types (1) and (2).

The European industry is partly moving away from integrated corporations to more specialized and focused companies. Examples include: KPN Telecom's plans to float its mobile division KPN Mobile on the stock market, and the division of Telefonica in focused operating units. According to Balsinde et al. (2000), the scale needed to increase market power, to set industry standards, and to shape regulatory agendas, can only be obtained by crossing national boundaries. The following events can be observed, although the scale and importance are still unknown:

- Cross-border consolidation in the form of mergers between large integrated corporations, acquisitions of small by large corporations, due to the drive for efficiency and growth (Balsinde et al, 2000).
- Breaking up of integrated companies and desegregation into units offering specific services or targeting specific market segments; small, independent integrated operators may vanish (Balsinde et al, 2000).
- Consolidation in backbones (because of increased competition to fill underused capacity), in local loop, although it is still unclear whether operators or content providers will see value accrue to themselves (Beardsley et al., 2000).

Vertical consolidation

On the vertical dimension, there are many, and there will be many more partnerships and joint ventures between companies in telecoms, services, media, software, and hardware industries (Balsinde et al, 2000). One can observe some consolidation in content, portals and e-commerce in need of broadband infrastructure (Beardsley et al. 2000), but the scale and importance are still uncertain. Coordination of content management and broadband services is difficult: companies face the complex task of providing platform-independent services. Companies are focusing on removing bottlenecks in infrastructure, billing, and customization of content. Some important competition issues that may arise in relation to vertical consolidation are:

- duration and scope of exclusivity embodied in intellectual property rights and programming rights related to content;
- bottlenecks and capacity shortages in transmission networks, in particular in customer access networks;
- access to installed customer bases for operators and content providers without direct customer access;
- interoperability of proprietary technology (e.g. digital set top boxes).

Examples 1 Takeover of Endemol by Telefónica
An recent example of vertical integration is the takeover of the Dutch content producer Endemol by Telefónica from Spain. Endemol is a well-established television producer and the creator of popular program formats such as "Big Brother." According to Telefónica's press release (August 2, 2000) Endemol will become one of the principal providers of content for Telefónica, supplying television and Internet productions for Telefónica's different platforms: television (Telefónica Media); Internet (Terra) and third-generation mobile telephony (Telefónica Móviles), as well call centers and future broad-band initiatives. According to the Presentation to Analysists (Telefónica/Endemol Entertainment, 2000), Telefónica aims to distribute content over the different platforms, such as fixed telephony, mobile telephony, Internet, and cable, in order to "leverage the multi-platform network of the Group to move towards convergence".

Example 2 KPN Telecom's activities in e-commerce and content provision
KPN Telecom aims to offer networks, access to Internet, and content to end-users. Its content provision activities are focused in Planet Internet (an ISP and content provider), which collaborates with Disney and Endemol. With respect to the combination of content provision and e-commerce, KPN has teamed up with the ABN-AMRO bank, to develop a joint on-line financial services provider, named MoneyPlanet. This provider will be integrated in KPN's Internet portals (KPN/ABN-AMRO, 2000).

Example 3 UPC's activities in content provision
Cable operator UPC has recently (February 2001) merged its ISP, Chello Broadband, with its TV broadcasting operations, into a new division UPC Media. The restructuring is based on the view that "television and Internet are coming together so fast that is makes a compelling case for this integration of resources," according to chief executive Mark Schneider, and on the hope to reduce operational costs. UPC Media will consist of four units: Chello Broadband Internet access, Chello interactive services, transactional TV services, pay TV channels.
Source: Total Telecom, 13 February 2001 (http://www.totaltele.com)

5.3 Efficiency Drivers

The central economic question for this development is: how does consolidation (e.g., a merger, takeover) affect static and dynamic efficiency? Again there are four efficiency drivers.

Driver 1 Economies of scale and scope

Consolidation may lead to realization of economies of scale and scope.[12] Examples in telecommunications are: the linking of network operators in order to save interconnection tariffs or to reduce interoperability problems, and the offering of different services over a single network. An example of firms trying to create synergies is KPNQwest, where KPN Telecom provides experience in Europe and Qwest has a presence in the US. Another example is that mobile operators may share antenna sites or other attempts to avoid duplication of fixed or R&D costs.

Vertical integration may be helpful to solve coordination problems, e.g. by making integration of content and access possible that are otherwise harder to achieve. Also, innovations and investments levels can be slowed down by problems with differing technology standards and compatibility. Research alliances or strategic cross-ownership networks that reduce the number of standards and increases cooperation in determining the standards will thus increase dynamic efficiency. For this reason, some level of research cooperation and sharing may be desirable. Thus dynamic efficiency may be increased, as new services can more easily be developed.

Driver 2 Market power

The number of firms determines the level of competition and hence (more often than not) static efficiency. Horizontal consolidation typically reduces the number of firms and hence reduces the level of competition. Therefore cost reductions might be outweighed by the detrimental effects of increased market power, thereby reducing static efficiency. If there is a large number of firms within both industries, the adverse effects of an alliance will be relatively small. Again, the effect of increased market power on dynamic efficiency is ambiguous (see appendix).

Another way in which consolidation may affect static efficiency is through takeover threats, which can impose discipline on management to act in shareholders' interest. Hence, (potential) takeovers may increase efficiency.

[12] Roughly speaking, economies of scale exist when the unit production cost decreases with the number of units produced. Economies of scope are cost-saving externalities between product lines. If there exists a substantial potential for economies of scale and scope, then there will also be potential for significant static efficiency gains post-merger or alliance.

A special but important case of horizontal consolidation is formed by cross-shareholdings. Cross-shareholdings constitute an important way to set up an anti-competitive way of horizontal consolidation. It directly reduces the intensity of competition. Ceteris paribus, cross-shareholdings do not have positive effects for consumers. Cross-ownership is a form of consolidation that does not generate economies of scale and scope. A special case is a situation where the state owns (part of) a company (see box). Also important are multi-market contact and cross-border consolidation. Horizontal alliances or mergers across borders may create problems due to their impact on potential competition. In general, multi-market contact may lead to tacit collusion. A firm with significant alliances or cross-ownerships has to weigh the negative impact that entering will have upon its business via these alliances or ownerships. Economies of scope and scale may be non-existent where two companies are based and operate in separate markets.

State shareholding

Some former national monopolists have not yet been privatized completely, such as KPN in the Netherlands. State shareholdings can be problematic, as they may impede takeovers and mergers. Consequently, the disciplinary force of takeover threats is reduced, whereas internationally, there is an unlevel playing field. The Dutch government recently announced plans to dispose of its stake in the incumbent. To the extent that a takeover increases static efficiency, a government shareholding may be potentially harmful.

Vertical integration can be driven by the wish to obtain control over end-users (i.e., by limiting consumers' choice in content and by favoring their own programming at the expense of others' content). The aim may be to use network bottlenecks to give consumers access to preferred content suppliers (i.e., reduce their choice). Hence vertical integration can increase possibilities to leverage market power, thereby reducing static efficiency. For dynamic efficiency it is ambiguous again: stronger control over the customer interface makes it possible to reap greater gains from innovation (good for dynamics). On the other hand, the scope for innovation from newcomers may be reduced, as access to end-users becomes more difficult (bad for dynamics). Overall, the total effect of vertical consolidation on dynamic efficiency seems more likely to be balanced towards a reduction: integration, by its very nature, seems to make an industry more static, since the ties are stronger (an integrated industry is less open to independent entrants). Network operators in content markets may use their gatekeeper position to extract anti-competitive terms and disfavor competitive content providers. This makes it less attractive for content providers to invest in new content and services, reducing the variety of future supply.

Driver 3 Entry and exit barriers

Low entry barriers are often good for efficiency, since it disciplines incumbents. This holds until the point where competition becomes too intense to recover fixed investment costs or wasteful duplication of fixed costs occurs. As long as there are no significant entry barriers, consolidation does not need to pose problems. In practice, the big investments, the need for a customer base, and the sunk nature of investments, needed for infrastructure may pose entry and exit barriers. Entry and exit barriers typically reduce dynamic efficiency as well, since new firms may introduce more variety for consumers than incumbents, or, more generally, increase the pressure for product and process innovation.

Telecommunication markets are often characterized by regulatory entry barriers (e.g. a license is required). This may lead to "tight oligopoly" markets in which players jointly have ample possibilities to behave anti-competitively without being picked up by competition law. The most straightforward examples occur on the vertical side, where strategic alliances, tie ins, exclusive agreements and other such contracts tend to make entry in downstream markets more difficult. Network operators can use technology to build new differentiated levels of service for Internet content, and mechanisms to control and manage Internet data (see also chapter 4). Similarly, cable modems can force consumers to use proprietary software to use Internet services over cable and to restrict access to Internet sites. Deploying such technologies increases network operators vertical control, while reducing consumers' effective choice and hence static efficiency. Network operators may grant preferential treatment to company-owned and affiliated programming, while reducing the transmission speeds for competitors' content.

An important issue related to vertical integration is whether technology will allow closed, controlled networks to gain vertical control, stifling innovation and competition along the way. As discussed in the previous exogenous development, if a network operator controlling a bottleneck expects that his market power will not be affected for some time and mandated access is not allowed, it may either have stronger incentives to invest (because it will reap all the profits) or lower incentives to invest because it gets lazy (the situation is too comfortable). Also, on the one hand consumers miss variety and new services by entrants, which might have resulted in more innovation compared to a monopoly situation, but on the other hand entrants have stronger incentives to bypass the incumbent's network by rolling out local access networks themselves. Vertical integration may tend to reduce innovation (and hence choice and variety for end-users) in the short run, but it may be beneficial in the longer run.

Driver 4 Technology

Technological developments often create new opportunities (or even strong incentives) to change the market structure, hence leading to new mergers, alliances and vertical integration. Quite often these type of mergers and alliances are survival strategies. If mobile and fixed markets converge it can be inefficient to have separate firms for both activities.

As introduced in the previous chapter, network operators can use technology to build new differentiated levels of service for internet content, and mechanisms to control and manage internet data. This vertical link might enable an ISP to implicitly discourage the use of certain sites by slowing down the speed of access, or to promote and offer the ISP's own or its partners' services with full-speed features (prioritization technology) . This is in fact the technology variant of exclusive deals and can damage both static and dynamic efficiency.

Note that regulation or competition policy is not mentioned as a separate driver for this development. It goes without saying that the quality of competition policy influences efficiency when mergers are at stake. Mentioning this point on each occasion does not generate special insights here, so that we refrained from it.

Table 5.1 Overview of efficiency drivers

Driver	Static efficiency	Dynamic efficiency
1. economies of scale and scope	avoid duplication of fixed cost + synergy +	research alliances + network sharing + standardization +
2. market power	increased concentration – cross-ownerships – vertical integration –	increased concentration +/– vertical integration –
3. entry and exit	barriers –	barriers –
4. technology	breakthrough + prioritization technology -	breakthrough + prioritization technology -

If one combines drivers from the table above, one can construct potential reasons for each of the possible four efficiency states. There is a long list of possible combinations, and not all of them can be discussed here. The next sections discuss the most relevant combinations for each state. To illustrate, two examples are the following:

- The low/low state may arise if vertical integration and exclusive deals foreclose markets, reduces options for consumers and increase prices.
- The low-static/high-dynamic state may arise if increased concentration reduces consumer surplus in the short term, but create a platform for innovations and investments (the Schumpeter story, see appendix).

5.4 Movement between states

Combinations of the specific drivers lead to "states". Suppose e.g. there is a technological breakthrough. Small firms share networks and R&D and thereby avoid duplication of fixed costs, without creating market power or vertical problems, this can lead to a high dynamic- high static efficiency state. We start with combinations of drivers that lead to low static low dynamic efficiency. We verify if firms have incentives to leave this low-low state, in if so in which direction (towards high static or high dynamic?). If firms do not have an incentive to leave a state with low efficiency which policy is called for? We repeat the procedure for the other states, each time limiting the analysis to just a couple of combinations of drivers (to avoid lengthy run-downs of cases).

5.4.1 Low Static, Low Dynamic Efficiency

Working from Table 5.1, the most important reasons to be in the low-low state have to do with increased concentration and vertical integration in combination with entry and exit barriers.

Increased concentration

One reason for the low-low state could well be due to too many alliances and mergers, resulting in a small number of large integrated corporations, while effective, beneficial alliances have failed to come about. With an uncontrolled market, it may happen that a few large multinational firms, such as BT and Deutsche Telekom, emerge, controlling more than one country's voice, data and mobile networks and services. The resulting possibilities for abusing market power are not offset by important economies of scale or scope. Dynamic efficiency can be low due to the lack of competition in the market (region A of Figure 2.2).

Do firms have incentives to leave this state? It depends. Large incumbents hardly have an incentive to leave this state, but the process towards this state may involve quite a number of "losers". Whether the losers and potential entrants are successful in forming countervailing power and activate competition policy preventing bulky mergers is an open question. Without that, sector specific regulation is clearly called for. In case competition policy is activated, firms might have an incentive to leave this state. This is typically done by innovation. Thus the natural tendency is toward the low static / high dynamic state. Consumers do not immediately benefit from innovation, since time is needed to bring innovation to the market. The market does not have a natural tendency toward the high static / low dynamic state.

Vertical Relations and Entry/Exit barriers

Vertical integration can reduce both static and dynamic efficiency. Think e.g. of cable. If a cable operator vertically integrates into content, it might leverage its market power on distribution into the content market. As vertically related firms integrate or sign contracts aiming at controlling end-users, this results in lower allocative efficiency and reduced incentives to innovate. Thus, in this state vertical consolidation does not lead to technology enhancement, but to more control over the customer interface between infrastructure and content, while effectively limiting consumer benefits. Vertical integration may also be driven by the intention to create bottleneck positions in order to create control over scarce capacity and to restrict entrants' access to end-users.

Do firms have incentives to leave this state? The answer is often "no" here. Once vertical integration leads to low static and dynamic efficiency sector-specific regulation is called for to repair the damage, or, alternatively, harsh US-style break-up policies (Microsoft). The only market driven way out of this state is when the vertically integrated firm can be replaced altogether by firms using an entirely new technology.

In general, vertical integration, in particular by firms having some market power, should be treated with suspicion in telecommunication markets, as they may restrict variety for end-users and access to end-users by other service/content providers. Even less far-reaching vertical relations between firms that do *not* possess significant market power can damage efficiency. Think of exclusive deals between operators and retailers. More often than not these deals are not in the interest of consumers (who benefit from independent retailers), but the problem seems difficult to tackle with competition law.

5.4.2 Low Static, High Dynamic Efficiency

This state is characterized by low levels of static efficiency but high levels of dynamic efficiency. As observed above, the reason for being in this state can be a market-based move out of the low-low state, driven by innovation. Are there other reasons for being in this state? The main reason is a Schumpeterian story (region B of Figure 2.2), where mergers lead to relatively high margins which are needed to finance innovation. Another explanation occurs when vertically integrated firms have realized new services but at the expense of static efficiency.

Increased Concentration

This state may correspond to a situation in which large integrated corporations and strategic alliances turn out to be the more successful in realizing investment projects. Smaller players cannot match this and are taken over or merge to mimic the strategy of larger firms. Accordingly, the intensity of competition is reduced. Horizontal integration makes it more feasible to reap the gains of innovation. Joint structures enable firms to invest and consequently dynamic efficiency relative to not having joint structures will be high. Thus, where there is a tendency for firms to consolidate, this may potentially increase the level of dynamic efficiency but reduce the level of static. Assuming that alliances and mergers bring previously unviable investments or innovations into possibility or help to define industry standards. In this case dynamic efficiency may increase. An example of a consolidation move that is potentially beneficial for dynamic efficiency is the partnership between KPN Mobile and Japanese mobile operator NTT DoCoMo. The latter firm's experience with the i-Mode business model (2.5G mobile telephony; see the chapter on mobile Internet) may help KPN to avoid costly mistakes in the development of UMTS mobile telephony.

Vertical integration and new services

Where there is the possibility to force either the consumer or downstream company to only accept the upstream firm's products, static efficiency is reduced. A reason to be in this state related to technology is that vertical integration gives firms larger incentives to invest in new services. Note, however, a downward effect on dynamic efficiency. An integrated industry is less open to independent entrants who introduce innovations themselves and also provide incumbent firms with stronger incentives to innovate. Moreover, applying technology that restricts consumers' access to independent services and content, generally stifles innovation. So high dynamic efficiency through vertical integration can only be realized under specific technological conditions, i.e. the integration should be a necessity to realize substantial benefits.

5.4.3 High Static, Low Dynamic Efficiency

In the short run consumers benefit from high levels of static efficiency. In the longer run, the low level of dynamic efficiency, implies that these benefits erode. The main reason for being in the state are unjustified merger blocks.

Increased Concentration: blocked mergers

This can, for instance, occur when regulation or competition law is overly harsh to mergers and alliances. Under the European Competition Law, efficiency criteria cannot be used in a merger decision. Hypothetically, if mergers lead to market power in the short run, but are necessary in order to invest successfully in new technology, the current law can block the merger, leading to high static but low dynamic efficiency. Clearly there is no market driven way out of this state. European Competition Law should be modernized, in particular for a sector where technology and scale is important. Efficiency criteria and dynamic considerations should play an important part in merger decisions.

5.4.4 High Static, High Dynamic Efficiency

Lastly, this is the most desirable stage. If large scale and integrated structures are necessary for investments, consolidation may have come about without harming static efficiency. This can be the case when large scale is not only needed for investments but is also the most efficient scale of production.

Economies of scale and scope

The level of concentration and consolidation in the market is such that it contributes to efficiency up to the point where it would lead to reductions in the intensity of competition and incentives to invest in innovation. Cooperation on technology takes place to the extent that it helps firms to overcome problems of standardization and compatibility. The market has reached a certain level of maturity, with the number of firms down to a level where the impact of an anti-competitive alliance would be strongly felt on both static and dynamic efficiency. For this reason competition policy has been developed that is consistent, fair and sufficiently selective. Alliances that create net benefits are allowed whilst alliances that have significant problems are either passed with restructuring of the problematic element or banned. Cross-ownerships, either horizontally or vertically, are generally banned thus reducing collusive incentives.

Will the market, left on its own, remain in this state? There are two options here. The market driven option is when firms are checked by innovation threats. There are only a few firms, but they cannot afford to be slow or set too high prices because they will be replaced by new firms with superior technology. Some software markets have this characteristic. Within the telecommunication markets, one can think of the market for business systems. The regulation driven option is where market check is too weak and regulation manages to walk on the thin line that ensures both static and dynamic efficiency. Clearly, in the case that a small number of firms is needed for dynamic efficiency, it will be hard to maintain static efficiency for regulators.

Technology

Technological developments can lead to the high-high state if they lead to a change in market structure. New mergers, alliances and vertical integration can be a necessary response to the emergence of a new technology, and can hence lead to high static and dynamic efficiency. To link up with the previous developments: think, hypothetically, of a successful auction of UMTS licenses. There are mergers and alliances between foreign and national firms, such that there is new entry and quick roll out. Think alternatively of fiber to the house. Under appropriate conditions, this leads to new mergers and alliances realizing quick roll out and new broadband services. Consumers benefit from competition and new services. This situation is clearly desirable and accentuates the importance of a regulatory environment that has a sharp eye for technological breakthroughs. A successful regulatory decision in the early stages and set-up of new markets can prevent a lot of difficult regulation in the future. It also accentuates that in telecommunication markets, a regulatory focus on dynamic efficiency seems more fruitful that a (too narrow) focus on static efficiency.

5.5 Conclusion

Consolidation does not often lead to the high-high efficiency state. The only market driven route from low to high efficiency is when technological developments lead to new scale and scope opportunities.

Lesson 1: All forms of horizontal and vertical concentration and consolidation (including cross-ownerships and alliances) should be subject to general competition policy, taking an integrated approach to relevant markets (including all linked markets). No special regulation seems required with regard to horizontal alliances, mergers and international take-overs. More than in other markets, competition policy may need to pay special attention to efficiency and dynamic considerations when assessing mergers.

An noticeable exception to this principle emerges when the market digresses to a tight oligopoly as a result of a merger or take-over. Competition law is notoriously vulnerable to tight oligopolies. Among other things, it is very hard to prevent or cure anti-competitive practices when no individual firm has a dominant position. In those circumstances, light-weight ex ante regulation may be preferable to ex post competition policy.

A special case of horizontal consolidation is cross-shareholdings. In the US competing firms cannot have shares in each other. We do not see why in Europe cross-shareholdings (in competitors) are allowed, since it is does not seem to have any other purpose than to reduce competition, while competition law cannot do much about it. Also, a too narrow focus on relevant markets and market shares misses some vital possibilities that market players have to behave anti-competitively.

Lesson 2: Vertical integration and vertical contracts, although they may help to circumvent coordination problems, should be treated with suspicion, as they may lead to tight control over end-users, make entry more difficult, and stifle innovation. Competition law may even be often too late to correct for it ex-post, such that a merger type of control would be preferable. Also, policy makers should keep a close eye on network traffic prioritization technologies, as they may harm static and dynamic efficiency.

Lesson 3: Given the choice, regulation should focus on dynamic rather than static efficiency. Firms will try to get out of the low-low state by innovating and differentiating themselves from competitors. Thus the natural tendency from low-low is toward the low static / high dynamic state, which over time may lead to the high-high state. Consumers do not immediately benefit from innovation, since time is needed to bring innovation to the market. Focusing on dynamic efficiency is particularly important with technological breakthroughs.

6 Conclusion

This study discusses future policy issues related to the trade-off between static and dynamic efficiency. This chapter concludes by identifying some policy principles.

Policy principle 1

The set up of a market is vital.

In a number of instances new markets need to be set up by the government of regulator. Think of allocating frequencies or licenses. Getting it exactly right is very hard, but getting it wrong often implies more need for regulation and a longer path towards high static and dynamic efficiency. This holds in particular when the new market involves a technological breakthrough.

Policy principle 2

Announcing upfront that some specific regulatory step is changed at some pre-specified future date can often be good for both static and dynamic efficiency

Suppose regulation has an impact on both static and dynamic efficiency, such as regulation of access prices. A policy principle that can cater for both types of efficiency often involves a phasing in time. Think of setting temporary low access prices to stimulate competition in services, with the pre-announcement that access prices will increase in due time, to stimulate investment in infrastructure. Nevertheless, access prices should increase only if one can reasonably expect that entrants consider investments to be viable.

Policy principle 3

Ex-ante clarity about the policy principles whenever possible.

Ex-ante clarity about the policy principles reduces policy uncertainty and therefore increases the willingness to invest. This also holds for so-called sunset clauses, i.e. the regulator specifies ex-ante under which conditions of the market it will decide to withdraw regulation and rely on competition law. It should be mentioned, though, that clarity cannot be placed above all.

Sometimes the future is simply too uncertain to provide clarity about the principles in an early stage.

Policy principle 4

Focus on entry barriers.

Reducing entry barriers is often more efficient than direct regulation of prices. Entry barrier reduction tends to be good for static efficiency since it enhances competition, but also for dynamic efficiency because it increases the willingness to invest in new markets and stimulates entry by innovative newcomers. It is also the most “market-oriented” way of regulating a market.

Policy principle 5

The regulatory regime should be as technology neutral as possible.

A regulatory regime that favors one technology (e.g.cable) over another (e.g.WLL) manipulates the market. In general there must be good reason to exercise such influence. The market knows best and decides itself which types of investments are optimal. However, it will not always be possible to let the market do the job. Access needs to be regulated, and by doing so, future investment opportunities are influenced.

Policy principle 6

A limited sharing of joint facilities should be stimulated.

A trade-off presents itself when entry is good for competition but involves duplication of fixed costs. These fixed costs can often be (substantially) lowered if players can share joint facilities. An important example is the possibility to jointly build the 3G mobile telephony network. This sometimes involves coordination problems or explicit permission by (local) government (e.g. mobile telephony antenna site sharing). It is therefore not clear that the market generates sharing by itself. However too much sharing can lead to competition and coordination problems on its own. Questions are: (i) Who is responsible for congestion? (ii) How to coordinate investments in the network? (iii) Will there be enough competition in services after sharing?

Policy principle 7

No special regulation seems required with regard to horizontal alliances, mergers and international take-overs. Vertical integration and cross-ownerships may deserve extra scrutiny, though.

There does not seems to be a need for special (i.e. other than regular competition law) attention to international take-overs or mergers and horizontal alliances. Special care is required when vertical integration is at stake, because vertical integration can restrict end-user choice, create effective entry barriers and can be a means to leverage market power from one market into the other. Cross-ownerships between competing firms also deserves special attention since it does not seem to serve any purpose other than reducing competition.

Policy principle 8

The only market drivenroute towards the state with high static and dynamic efficiency, is via dynamic efficiency. If e.g. investments in new technology create a market structure breakthrough, the market can lead itself to the high efficiency state (provided competition law is effective).

There is no market driven route towards the state with high static and dynamic efficiency, since market parties typically have little incentives to solve static efficiency problems themselves. Conversely, there exist situations where market parties need some market power to have incentives to invest in innovation. If investment or innovation create a breakthrough in market structure, future static efficiency is conceivable. Hence a transition from low efficiency to high efficiency via the low static / high dynamic state, is only possible if market parties have incentives to invest in innovation. If this happens, innovating firms gain and maintain market power. Although to some extent this may be harmful for current-generation customers, later generations will benefit from it. Since market driven routes are often preferable to regulation driven routes, the likelihood of this route should be assessed first before considering some form of regulation.

Policy principle 9

Technologies that allow network operators to prioritize network traffic (e.g. internet access) may increase static efficiency, but at the risk of limiting consumers' choice and access to competing content providers or ISPs. Hence policy makers should carefully check that prioritization technologies do not stifle competition or block the introduction of innovative services by newcomers.

The importance of this last policy principle is, at this moment, not yet clear. Nevertheless, it is important to be aware of the potential risks involved. One can currently observe the emergence of prioritization technologies, which are developed by strong players such as Cisco, with the purpose to be employed by, for instance, cable operators providing Internet access services.

Appendix A.1 Theoretical and empirical links between static and dynamic efficiency

Introduction

This section explores the theoretical and empirical links between dynamic and static efficiency. In particular it looks at the possible trade off between the level of competition (one proxy for static efficiency) and the incentive to innovate or invest in a future investment (one proxy for dynamic efficiency). Section A.1.1 concentrates on the most important models that have defined the way economists view the relationship between dynamic and static efficiency. As the section progresses we introduce more complicated relationships and attempt to identify broad patterns within the literature.

Section A1.2 considers the empirical side of the link between innovation and static efficiency. It should be noted that whilst the theoretical literature provides often relatively precise predictions these have been harder to translate into empirical proofs. Whilst there has been a large literature on the empirical links between firm size and the level of innovation, the literature on the degree of competition or the degree of profits has been much smaller. Consequently this section is more brief than the theoretical review.

A1.1 Theoretical relationships to innovation

A1.1.1 Defining the Relationship between Innovation and Competition: The Core Models

This trade-off has been covered in some detail by economic literature, generally falling into one of two categories: Those that believe increased levels of competition reduces the level of innovation or investment, and those that believe increases in competition increase the level of innovation or investment. We study these two arguments firstly in the simplistic case of symmetrical costs, where all firms within the market have identical costs before the decision to innovate or invest. We then look at the more realistic assumption that firms within the market have different costs of productions and see what the impact the relaxation of symmetry has upon the initial conclusions.

Innovation requires Market Power (Schumpeter)

The first school of thought was introduced by Schumpeter (1943). His negative relationship between the level of competition and innovation was based upon three arguments. First there are difficulties for firms in securing capital for innovation or research and development that have significant risks attached to their successes. Schumpeter argued that market power and the substantial profits that are associated with this, are prerequisites to financing innovation independent of external sources. Today, with developed venture capital markets in many countries and a willingness to invest in companies that will not make profits for some time, this argument is less apt than in Schumpeter's time. The second argument, that the incentive to innovate is greater for monopolies, is still highly relevant. Market power and the reduction of competition increases the level of future profits. The increase in future profits increases the return on research and development or innovation of a successful product thus increasing the incentive to innovate at the firm level. We will return in more depth to this argument to see under what conditions this link can be reversed and most importantly will the overall incentive for the industry increase. Lastly within research and development there are often significant economies of scale that may be exploited by a firm of significant size. From these three arguments he concluded:

"the trail [to progress] leads not to the doors of those firms that work under conditions of comparatively free competition but precisely to the doors of the large concerns...big business may have had more to do with creating that standard of life than with keeping it down" [13]

The belief in the destructiveness of competition and the necessity of market power to ensure innovative gains has driven a whole literature both theoretical and empirical in support of and against this controversial proposition.

Competition leads to Innovation (Arrow)

The second school of thought may be traced back to Arrow (1962) in his pioneering article to determine the pure innovation incentives. Arrow concluded that competition enhances the incentive to innovate as the value a competitive firm places on an innovative process will be greater than the monopoly's value. This finding is important and is worth explaining. Consider some innovation that enables the producer, be it the monopoly or socially optimal planner, to secure the lower marginal cost c_1 on successful completion. Relative to the social optimum it is easy to see that the monopoly's innovative incentive is lower as the benefit of moving from c_0 to c_1 is only realised on the much smaller monopoly level of output. The monopolist is pricing where MR = c_0, whilst the social planner prices equal to c_0. Thus output under the monopolist is smaller and hence the benefit/incentive for the monopolist to innovate is smaller than the social optimum even though both face identical (and hence symmetrical) reductions to their costs.

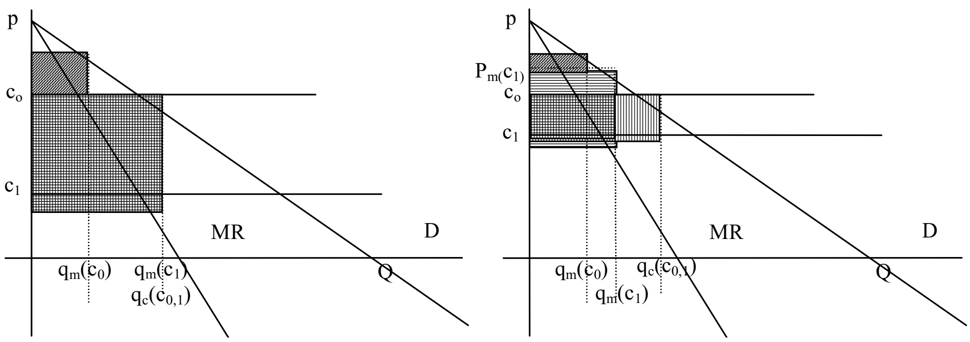
Looking next at a competitive market, consider a large number of firms that produce a homogeneous good at price equal to marginal cost c_0. The firm that obtains a new technology is awarded with a patent allowing it to charge the monopoly price. We assume that the innovation

[13] Schumpeter (1943) page 82.

is purely in process merely creating a cost advantage rather than a separate product. Correspondingly, it still is faced by competition from the remaining competitive firms. Whether the firm charges this monopoly price depends on the level of innovation. Firstly assume that the innovation is “drastic”, that is the cost advantage realised by the new innovation is so great that even when pricing at the monopoly level, the resulting price is still less than the competitive firm's price (c_o). In this case the new firm prices at the monopoly level (lower than c_o) and all remaining firms are forced to leave the market. In the second case, assume the innovation is “non-drastic” thus the cost advantage of the new innovation is not great enough to allow unconstrained monopoly pricing lower than the competitive level. In this instance the new monopolist takes the entire market by pricing at just below the competitive price level (c_o-ϵ) making a profit, but not the monopoly profit of drastic innovation.

Figure 6.1 “Drastic” Innovation and “Non-Drastic” Innovation

Fig x.x ‘Drastic’ Innovation *Fig x.x ‘Non-Drastic’ Innovation*



In the “drastic” case the monopolies incentive to innovate is characterised by the loss of the top diagonally shaded box and the gain of the cross hatched box, resulting that net profit from innovation is less than the total cross hatched area. The competitive firm initially prices at c_o and hence starts without the diagonal profit made by the monopolist. Hence innovation for the competitive firm moves it from 0 profit to the entire cross hatched area. The conclusion within this “drastic” case then is that the competitive firm always gains by a greater amount relative the monopolist, and hence has a larger incentive to invest. Naturally this greater incentive to invest does not necessarily translate into greater social welfare, and we return to this point within the study of patent races.

The “non-drastic” case is a more complicated, the monopolist loses the diagonally shaded box as before, but gains the horizontally shaded area plus the cross hatched. The competitive firm once again does not have this loss of profit and is able to appropriate the vertically shaded plus the cross hatched area. From the diagram it is not clear which structure creates the biggest incentive, as the monopolist gains the horizontal but loses the diagonal areas relative to the competitive industry which gains the vertical shading. To resolve this we turn to algebra: By our assumption and demonstrated graphically c_o is less than the monopoly pricing at new lower level $p_m(c_1)$, which in turn is of course less than the monopoly price at any other cost level greater than c1, ie $p_m(c)$. This implies that $q(c_o) > q(p_m(c_1)) > p_m(c)$. From the innovatory value functions it is then possible to show that for both the drastic and non-drastic innovations the competitive industry has a greater incentive to innovate than the monopolist.

It is relatively easy to see that both of these incentives are lower than the incentive for an optimal social planner, where the post innovation price is set to the new marginal cost. This is because the gain in efficiency is utilised for a much greater quantity than that of the monopolies. Thus the total gain is not subject to any dead-weight loss and will always be greater than either the original monopoly or the new monopoly. In conclusion, the benefit from innovation for the competitive firm is strictly greater than that of the monopolistic firm. The monopolist simply becomes another monopolist at a slightly greater profit level, he merely replaces himself (hence Arrow's term the "replacement effect"), whilst the competitive firm becomes a monopolist from nothing, all the gain is new profits for the innovating competitive firm. Although Arrow does not explicitly analyse the impact of research cost, it is relatively easy to extend this basic model using some form of fixed research cost.

It is difficult to sum up exactly why the differences between Schumpeter and Arrow arise because of the more philosophical tone that Schumpeter adopts. This philosophical tone does not set out explicitly why a monopoly will be more likely to innovate, hence the different interpretations of his work. However one may think of the main argument as an appropriability argument without patents. The market power of the monopoly allows it to gain from or appropriate the entire benefit of the innovation even without the existence of a patent. However a competitive market without a patent only gains from the innovation until the rival firms are able to imitate it, thus reducing the benefit and hence the innovatory firm's valuation of the innovation. In summary, one may think of the Arrow model as a model where there is perfect patent protection and the competitive firm goes from zero to monopoly profit. Schumpeter's work however, implies that the competitive firm will be unable to realise this monopoly profit due to problem associated with lack of market power and appropriability.

A1.1.2 Innovation Incentives with Strategic Interaction Between Firms

Whilst very instructive, Arrow's model only provides a partial story of innovation. Implicit in the basic model is the assumption that firms do not compete with each other through innovation. Literature subsequent to Arrow has concentrated more on an oligopolistic industry with two firms competing for a patent on some form of product or process innovation.

Innovation with Certainty and a Fixed Cost

Needham (1975) developed one of the earliest attempts at capturing competing firm's reactions. Focusing on a product rather than process innovation, he postulated that after a firm has introduced an innovation, its rivals retaliate some time afterwards with a competing product of their own. Thus the value of the innovation can be split into three elements. Firstly there is the negative effect due to the one off cost of research. Secondly the high degree of profits before imitation, and lastly the lower degree of profits after imitation. Needham showed that an individual firm's incentive to innovate depends largely on the degree of response from rivals. If no retaliation is expected then a firm's incentive to innovate is simply the difference between additional expected revenue and research cost. With increasing rival response, both in severity and time, the expected profit and hence incentive to innovate falls.

This model supposes however that there is a large degree of certainty within the innovation process. The profit from the innovation is known, as is the response time and severity from rivals. Hall (1990) proposes that these conditions are only likely to be met within a settled mature market. Using this, one can interpret increasing maturity as being likely to decrease the level of innovation. This is due to two reasons: Firstly an immature market has relatively few products. Thus an increase in market share will be relatively large and more profitable with the introduction of a new product that has few substitutes. As more products are introduced the marginal revenue from innovation will fall. Secondly the length of time that it takes rivals to imitate will probably be longer as the innovation is larger and less is known about a rivals products, also patent protection will be more difficult to obtain with small variations in product as in a mature market.

Increasing Costs of Research, Uncertainty

The important but unrealistic assumption in Needham's model, that research always leads to developments within a known time and before rivals, is relaxed by Kamien and Schwartz (1982). They imagine some future date at which the development will take place that decreases as the intensity and hence cost of research rises. Similar to Needham, they separate the net value of innovation into four terms. First the discounted value of the profits that the firm expects from continuing to sell the old product, this of course depends negatively on the rivals ability to innovate first. The second element represents the profit that the innovation (if first to the market) will produce along with the reduced degree of profit from subsequent imitation. This second term will fall with an increased probability of imitation. The third term is the profit the firm realises if a rival innovates before it. This third term incorporates the idea of the firm as the imitator and thus the later the imitation date the lower this third term. The last term represents the cost of research as discussed previously.

Under the assumption that it is profitable to innovate, there are two main results. Firstly the greater the current present value of profits the lower the incentive to innovate (ie delay). This follows as delaying the time to innovate results in lower research costs whilst existing profits remain high. Secondly, the higher the expectation of rival innovation, the greater the incentive to innovate. This follows as the expectations of the first element of profit from using old technology falls as the likelihood of being beaten to the innovation by a rival, increases. This model shows that the absence of rivalry favors fast innovation when the increased returns from innovation are modest. However if the increased returns are large the pressure of rivalry will increase the speed of innovation toward its maximum.

One Time Research Expenditure Patent Race

A well known patent race model is Dasgupta and Stiglitz's 1980 paper. To simplify the analysis they assume that the firm's probability of discovering a useful innovation at any point of time depends only on the firm's current research expenditure and not on past experience. They then return to the issue of whether a monopolist in the product market is more likely to innovate than an entrant. Suppose that the first firm to innovate will obtain and exploit an infinite patent. Assume the entrant earns nothing via the industry initially but after the innovation will earn monopoly profits.

Which of the two firms spend more on research depends upon the relative weight of two effects:

The efficiency effect suggests that the monopolist has a higher incentive to innovate and therefore will spend more on R&D. This is because the difference between monopoly and duopoly profit will always be greater than duopoly profit. The monopolist gains a greater flow of profit by pre-empting the entrant and maintaining its monopoly profit, than the entrant can by entering and becoming a duopoly. Note that under drastic innovation the entrant becomes an instant monopoly (rather than a duopoly) hence both the entrant and monopoly have equal incentives to innovate.

The replacement effect suggests that the monopolist will have a lower incentive to innovate and therefore spend less on R&D. This is because the marginal productivity of R&D expenditure for the monopolist decreases with its initial profit. As the monopolist increases R&D it hastens the replacement of its existing profit with new (albeit greater) profit. In contrast, the entrant does not forgo a profit after a successful innovation as it does not have an initial presence in the market. Thus the incentive to innovate measured by the difference between the two profit states will be greater in the competitive than the monopolist firm.

Looking at the extreme case of drastic innovation, as the entrant becomes a de facto monopolist there is no dissipation of rent and hence no efficiency effect. Thus the replacement effect must dominate. Reinganum (1983) derives this proof. Thus in drastic innovation there is a tendency toward innovatory entry into the output market. For non-drastic innovations creating small benefits, there is a tendency for the monopoly to persist because the established firm has a higher probability of obtaining the patent. Innovation is obtained early as the monopolist is much more concerned with the possibility of innovation by the entrant than by the date of it's replacement. Note that this analysis does not say how the market structure effects innovation but analyses how innovation defines the market structure.

Martin (1993) returns to this initial question of market structure impact via a similar model by Lee and Wilde (1980). He finds like Dasgupta and Stiglitz that the incentive to innovate per firm may increase or decrease with greater numbers of firms in the market. Most importantly he shows through numerical simulations that regardless of whether the innovation incentive per firm increases or decreases, the aggregate innovatory intensity across all firms always increases. This increase in aggregate intensity may not translate into an actual increase in social welfare where firms are forced through competition to innovate at a higher than optimal speed. Where the innovation is drastic increasing the number of firms translates into a strict increase social welfare. However where the innovation is more modest such as within the telecoms industry the welfare gains through increasing the number of firms are much more modest, and at the limit of entry may tend to fall.

A1.1.3 Innovation and Accumulated Research Experience

In the last twenty years the main developments of innovation theory has sprung from the introduction of game theory into the analysis. Fudenberg et al (1983) and Harris and Vickers (1985) use a game theoretic model to show that if the probability of discovery per unit of time depends not on current R&D but on experience accumulated to date, the conclusions may strongly differ from Dasgupta and Stiglitz's. They consider a patent race between two firms in which both firms make negative expected profits if neither of them drop out of the race but positive if becoming a monopoly. Under this scenario the firm that drops out, is the firm with the least research experience to date (i.e. the follower). Assuming rationality and research being a costly good, backward deduction shows that it is optimal for the follower to drop out a short time after the beginning of the race. They prove that even if the monopolist is able to enter some small time (e) before the entrant, the entrant will rapidly leave.

However this simple equilibrium may not be the only equilibrium. If there exists information lags, i.e. firms can only observe their rivals R&D efforts after some delay, the follower may be able to leapfrog the leader in the time lag between being noted as increasing its research rate.

These results can be relatively easily demonstrated by supposing that each firm can accumulate 0, 1 or 2 unit of experience per period at respective costs of 0, c_1 and c_2 where $c_2>2(c_1)$. Thus an R&D monopoly without the threat of entry prefers to accumulate only one unit of experience per period. In the equilibrium, competition is very intense with each firm accumulating 2 units per period. When the leader is one unit ahead the optimal strategy is to randomise between 1 and 2 whilst the follower's is to randomise between 0 and 2 thus even with a small lead there is still some probability of competition within innovation. However when the leaders lead is at 2 or more points ahead the follower is forced to drop out of the race and the leader proceeds at the monopoly pace of 1.

Whilst Dasgupta and Stiglitz predict strong competition within innovation until the time that the innovation is completed, Fudenberg et al predict that even with competition, as soon as one firm becomes too far ahead of the other, competition breaks down and the follower leaves the market. Thus this game theoretic approach seems to draw the pessimistic conclusion that competition will have little lasting impact on the rate of innovation development. However where the switching costs of reducing the level of innovation of the follower are significant or there is some uncertainty about how the leader is doing, both of these factors will increase the value of continuing to compete in innovation. This will increase the length and degree of competition over the time of the patent race. The model also makes the fundamental assumption that were both firms to create the innovation both firms profits would be negative. Relaxing this by assuming both firms make some degree of profit via product differentiate and introduces the new result that both firms continue to compete for the innovation but differentiate the innovation slightly to ensure both firms can make a profit.

A1.1.4 Other Effects On Innovation

The previous section looked at the perhaps the four most important developments in innovation theory. The initial questioning of the relationship between monopolies and innovation by Schumpeter, the reply by Arrow, Dasgupta & Stiglitz's introduction of strategic interaction and finally Fudenberg's introduction of game theory into the analysis. All of these four models look to determine the core underlying relationship between market structure and innovation. This next section will look specifically at the relationship between other fundemental characteristics of an industry and innovation. The section is designed to show the reader that whilst the core relationship may point to one competition effecting innovation in one way, this may be effected by other characteristics of the industry.

Imitation, Spillovers, Innovation and Diffusion

The degree to which a firm innovates or to which increases in firm numbers have on innovation will be a function of how easy the innovation is to imitate and the degree of spill-overs created.

Spill-overs are the degree to which a firms innovation can be copied by the competing firm. Industries with high spill-overs are ones in which patent protection is not feasible, firms are able to relatively easily reverse engineer existing innovations given some degree of time. One might expect that because of these spill-overs the degree of innovation will be lower, but it is not cleat that it will change the relationship between innovation and competition.

Fudenberg and Tirole (1985), (1987) consider a Bertrand duopoly in which two firms initially have constant unit cost c and neither firm makes a profit. Adopting the innovation reduces the unit cost so that the firm that adopts first makes some degree of profit. Imitation never occurs in this extreme model because Bertrand competition with identical costs yields 0 flow profit and the adoption incurs a cost at all times. The payoff for the leader then increases as time increases until at some point it starts to drop off as the discount rates becomes higher. Meanwhile the payoff of the follower is 0. This is because once the profits have accrued to the leader there are no other profits for the second firm on imitation of the innovation. The optimum time for the leader to adopt is when the payoff is at the maximum, however if it chooses to adopt at this time the other firm can still profitable pre-empt and seize all the profits. This undercutting in time continues until all profits are taken up by the cost of researching. This zero profit equilibrium ensures that monopoly rent is totally dissipated and secondly this dissipation is completely wasteful from a social viewpoint.

This model shows that increasing the degree of competition to the extremes of Bertrand results in welfare losses. The absence of imitation hinges on the Bertrand nature of competition and the fact that any fixed costs can never be sustained under such a competitive market structure. However when we relax this competition assumption and assume another means of competition such as product differentiation or Cournot type competition we find that the pre-empted firms will end up adopting the technology in the long run only if the cost of imitation falls sufficiently.

Entry Barriers and Innovation

Note that many of the models when finding that the incentive to innovate is greater in the monopolist case, relies entirely on the ability of the entrant to enter to derive any positive welfare effects. Where entry is barred such as in a limited license network or the requirement to invest in large infrastructure costs, this result will be invalid. A firm that is forced to overcome a large entry barrier will need to take this into account when determining whether it should invest in an innovation. The monopolist is then able to use this barrier as a means to allow it to introduce innovations at a slower rate than previously without allowing the new firm to enter and innovate.

In the case of high barriers to entry, whilst the end result that the monopolist is in many cases still able to maintain its monopoly by being the first to innovate, the rate of which it is forced to innovate will be lower. These lower rates of innovation may push the rate of dynamic change within the industry far below the optimal welfare maximising levels.

However, the other side of the argument is the possibility of innovation being more valuable to the monopoly simply because it is a means to create these high entry barriers. As we saw from the game theoretic treatment of the patent race, a firm that is able to stay at least one innovatory step ahead of its competitor may effectively be able to create an entry barrier. This would seem to apply that the monopoly has a higher incentive to innovate where there is the additional benefit to profit of preventing future entry.

How can these two seemingly opposite effects be squared? One area of innovation research that helps us reach a solution, is the finding that monopoly firms have an incentive to innovate, but not necessarily to introduce those innovations. Several studies have found that monopolies have a tendency to develop their innovation up to the patentable level but never introduce the innovation. This practice of "patent shelving" is looked at by Gilbert and Newbery (1982) who illustrate many cases of patents being introduced purely as a means to deter entry. These shelved patents give the firm the option of picking them up and developing them should a new firm threaten its market with entry. Both IBM and Microsoft have been accused of using these types of tactics to prevent entry into their markets. In conclusion then it is important not only to look at the number of patents but the number of patents that are developed into successful innovations.

Joint Research and Innovation

Joint research or innovation is one of the few policy areas where there may be a case to encourage firms to work together. Joint research may take place in many ways, it may be deliberate sharing may be via a joint research venture in which they both finance and share all findings, or one firm can pay another for the right to know anything that emerges from the first firms research. In practice the most common agreements are for technological information supplied on an ex post basis with firm A providing a license to firm B permitting B to use a specified innovation of A's. In many cases the agreements are totally informal.

Recently the area of joint research has received a lot of literature and attention with its increasing practice in countries. Shapiro and Willig (1990) look at the relationship between innovation and joint ventures. They begin by asserting the basic belief that there are benefits both socially and private in collaborating within innovation. These include the sharing of risks associated with uncertain demand, synergies arising when firms share complementary skills or assets and the attainment of economies of scale and scope in production or logistics. These benefits appear to be backed up firstly by the fact that joint ventures are quite common. Secondly the accumulation of market power does not appear to be the motivation as the parents are generally not horizontal competitors or in cases where they are, the relevant market is relatively unconcentrated or not closed to new competition. Thirdly the ex-ante sharing arrangements offer some degree of insurance to each participant against the risk that its own innovation does not bear fruit as cartel members receive access to all research. Lastly the costs of maintaining an innovation joint agreement will generally be much less than an output agreement. Members of an innovation cartel have a huge competitive advantage over non-members. Those outside the agreement don't have access to shared information and thus will have a rate of innovation very much lower than those inside. For this reason it is likely that the firms outside the agreement will either join or may eventually be forced out of business.

However joint agreements can also cause problems to both individual firms and society. Firstly decision making may be slow due to disagreeing parents and hence innovation slow. Secondly one firm may try to free ride on the other as neither firm can truly see the innovatory effort put in by the other. Thirdly the co-operation required for a joint venture may limit the ability to develop other opportunities that would arise for one firm to expand profitably at the expense of another. An example of this third problem is where assigning key personnel to a joint venture limits possibilities in others areas. Ordover and Willig (1985) point out that a research joint venture by replacing competitive independent decision making with a joint decision structure, may prevent a patent race. This decreases the pace of R&D if the joint research does not sufficiently lower the R&D costs or if the agreement combines the R&D activities of principle competitors. Lastly from societies view the formulation of a joint venture that significantly reduces competition (through placing industry capacity under joint control) derives private benefits at the expense of social welfare.

Williamson (1968) showed that even with relatively small unit cost savings the gains in welfare may still be significant as long as firms are not able to extend the joint agreements into the second period. Thus joint ventures with independent competition are what the competition authorities should strive to achieve. However policy makers may find it difficult to substantiate the efficiencies claimed by parties and may equally question whether it may be possible to realise the cost benefits without the deficiencies via the joint output. An example of how difficult it is to tell if a consortium truly is competing is where there are several firms who join the venture and after being sold the product via the venture then compete in the product market. If the joint venture then prices the product to be sold to the members at the monopoly level, then regardless of the competition between the members, society pays the welfare price of a monopoly structure. The members realise monopoly profits via their ownership of the consortium rather than via the product market. To avoid this the consortium must be forced to price at marginal cost and allow the firms to compete in the profit. This solution corresponds to the case where a joint venture produces royalty free innovations to all the consortium.

When attempting to investigate whether such conditions in a joint research structure are met, competition authorities must ensure that the fixed costs of the venture are also covered or the venture will collapse. However like within all regulation optimal tariffs rely on the ability of the regulator to know exactly what the fixed, and marginal cost of the innovation is, whilst firms within the joint structure clearly have an incentive to over report their costs.

The above paragraphs demonstrate that like the core relationship between innovation and competition, the relationship between innovation and joint research produces mixed theoretical evidence. Like innovation and competition we turn to numerical solutions of theoretical models to attempt to shine some light on the welfare effects. Comparing an oligopolistic industry that is allowed to make joint innovatory decisions with the same industry but restricted from joint decisions provides some interesting welfare results. Martin (1991) presents a model that allows a common per firm research intensity and allows members of the group to share in all innovations. He finds that whilst per firm and aggregate research intensity are lower relative to the patent races of Dasgupta and Stiglitz and hence time to innovation longer, welfare may increase due to the enforced competition in the second stage. This contrasts with Dasgupta and Stiglitz where a single innovating firm becomes a monopoly within the innovation for at least some time. Thus he concludes that whilst joint ventures are socially desirable in this model, the desirability has little to do with the impact of the joint venture on R & D. Clearly if joint research encourages collusion in the downstream market (rather than the competitive assumption made) then these agreements will certainly lower social welfare further.

Horizontal Collusion and Innovation

It is generally agreed that co-ordination on prices or outputs by firms reduces social welfare, indeed it is often argued that collusive agreements that push the price to monopoly levels reduce welfare to levels below that of monopoly. However, recent literature has shown that there may be some arenas in which horizontal collaboration can be beneficial to society. Despite tendency to yield a less that optimal allocation of resources in stationary equilibrium, collusion may speed productivity and output growth and even reduce the costs.

Baumol (1992) discusses the pro and cons of collusion and concludes that there are the following main results. Firstly, ceteris paribus, the effects of dissemination are unambiguously beneficial. Secondly, horizontal collusion will sometime stimulate R&D and innovation and sometimes depress them, Thirdly even if collusion restrains innovation the net consequence may be beneficial because of costs savings and faster dissemination. Finally the core problem with technology collusion is that it often leads to a monopolistic agreement to keep prices high and avoid expense of rapid innovation.

Even though the degree of concentration may not be a good approximation to the degree of competition in a market (two firms competing using Bertrand strategies will secure lower prices than 10 using Cournot strategies) it is a good approximation for the degree of ease in setting up and maintaining a cartel. In a few cases price fixing may create benefits to consumers i.e. where vertical constraints are used to prevent the double mark-up that a downstream industry may desire to place. However generally price fixing in the product market will result in lower welfare. Increasing the degree of competition within the industry will not harm the technology cartel stability but will make it more difficult to sustain a cartel within the product market.

Network Compatibility and Innovation

The last additional effect we consider is Network Compatibility. Economides and Flyer (2000) show the trade off between wanting to make products network compatible so the number of users increase and the firm can capture the value added by being a member of a large network and the desire to product differentiate to decrease the level of competition. They show that because this trade-off exists, the optimal market structure for the firms is to create coalition groups of compatible platform. As society generally desires to maximise the size of the network and hence capture the largest possible value from being a member of the network, the market structure results in a socially sub optimal degree of compatibility.

A1.1.4 Conclusion and Implications for Telecoms.

The telecommunications industry consists of several vertically separated markets aside from the actual infrastructure that this paper discusses in most depth. Unlike other research dominated industries like pharmaceuticals, a single firm does not does not both undertake research and production of a complete final product to the customer. The product that arises from the research is simply one of the many inputs into the final good, for example being able to receive the internet over a mobile requires many innovations not all undertaken by one company. Most of the research in telecoms is done by specific research orientated companies such as Bell-Lucent Laboratories. These independent research institutions develop the product and sell it to the telecommunication companies at a price so that they may improve either their quality or price via the innovation. For this reason there is some degree of separation between the researching institution and the implementing or network institution. This is important when considering whether the infrastructure companies require a certain market power, or profitability to adopt these innovations.

Much of the literature discussed within this paper has concentrated on the impact of a completely new innovation, which is not directly applicable to any of the three exogenous developments discussed previously. However, as long as there is some degree of risk and uncertainty in implementing the innovation, or more importantly some competitive gain which can be appropriated for a single firm at the expense of another, then existing literature has an important bearing. Although we are primarily interested in the development of the infrastructure industry, where ties are developed with other areas of telecoms, (such as mobile producers and network tie ins) then the analysis of this becomes important to the third exogenous policy development. For example the tie between third generation mobile phone manufacture and the development of the infrastructure creates a link between the successfulness of the infrastructure and the ability of the mobile company to innovate. For this reason, and the more general reason that tie ins are often a form of vertical constraint designed to curtail competitors, it is generally undesirable for such tie ins to occur unless investment would not take place without them.

Even though the infrastructure development that is so crucial to the telecom industry may not be directly mapped onto the more mainstream innovation literature, there is still some benefit in being the first to develop some new quality or product. The key question for these markets is one of appropriability: Can a firm that introduces some new innovation or increase in infrastructure quality appropriate sufficient gain before its competitors are able to imitate and reduce the benefit to unprofitable levels? Perhaps the most directly relevant area of literature for telecom infrastructure innovations is the analysis of a market in which innovations are quickly imitated by competitors. Models of imitation and adoption such Fudenberg and Tirole's become particularly relevant although their conclusions hinge on the Bertrand assumption of competition not particularly realistic for telecommunications.

With generally high consumer switching costs, one may consider that telecom infrastructure markets have a form of patent protection. The first mover has a strong advantage as long as its product innovation is sufficient to lure consumers away from its competitors. Once these competitors have caught up or imitated the innovation, unless they have managed to sufficiently improve the benefit above the level of the innovator, the cost savings will not be sufficient to lure the customers back to themselves. This lack of ability to recapture customers effectively creates a form of patent protection for the developing innovator. In this case the innovation patent race literature becomes once again highly relevant for telecom. We summarise exactly which areas of the literature in the conclusion.

A1.2 Empirical evidence on innovation

Introduction

As stated at the start of this appendix, direct and relevant empirical literature that addresses the relationship between dynamic and static efficiency is more limited than the theoretical literature. For this reason, and due to the necessity to conduct only a brief overview, this section is more succinct than the theoretical.

Profits and Innovation

Schmookler has been one of the main researchers within the empirical field of innovation. His 1966 book advanced the idea that resources are allocated to innovative activity in proportion to the expected profit flow from this activity. Scherer (1982) looked at this hypothesis and an alternative hypothesis that innovation occurred because some industries have lower costs of innovation than others. His study found that there is some support for both these hypotheses coming to the conclusion that the two factors should not be thought of substitutes but complements.

Many of the recent studies, and perhaps more appropriate to this paper, is the empirical work on the relationship between firm's cash flow and the level of innovation. This is in line with the Schumpeterian belief that large profitable firms will have easier internal access to finance than other firms. Most of the studies support this belief. Particularly relevant is Halls (1990) work which finds that firms that increase their leverage reduce their R&D intensities.

Size and Innovation

Perhaps the earliest empirical research was based on National Science Foundation data from the 1950's and 60's. Villiard (1958) and subsequent researchers established that the likelihood of a firm conducting research increases with firm size. Across later multi-industry studies it was concluded that R&D rose somewhat more than proportionately with firm size, more detailed studies such as Scherer (1965) found that past some threshold the increases in R&D levelled off to proportional levels. More recent studies such as Scherer and Ross (1990) that take account of industry effects in cross industry samples have redefined the earlier finding and conclude that the *degree* of innovation rises proportionately with firm size. This offers little in the way of support for the Schumpterian notion that the size of a firm results in a disproportionate increase in R&D.

Concentration/Market Power and Innovation

Farber (1981) examined the relationship between concentration and innovation finding that the higher the concentration in the product market the greater number of research personnel would be hired. This was seen as evidence for the belief that firms need market power to be secure of appropriating some of the gains from innovation before committing themselves. However the presence of research personnel may not be evidence of actual innovation where firms have incentives to patent and shelve their research.

Scherer (1967) was one of the earliest to test the hypothesis that Dasgupta and Stiglitz later theoretically gave strength to, that R&D intensity appears to increase with increases in industrial concentration but only up to a rather modest value (a CR4 of about 55%) after which R&D efforts begin to decline. Later empirical work has replicated and hence strengthened this finding.

A1.3 Conclusions

The reader to this brief literature survey will have grasped the difficulties in making clear-cut conclusions from the main texts within this field. Some of the models predict that competition will increase the rate of innovation per firm (Arrow), some predict that it will generally be unchanged (Fudenberg) and others predict that it will decline (Schumpeter). However it is generally true to say that the rate of innovation per firm will increase with competition when the degree of competition is not already too severe. Once an industry is within the Bertrand sphere of competition increases in the number of firms only serve to increase the race for patents, generally beyond the socially optimum levels. Firms who are engaged in competition of this type will be unable to support the fixed costs necessary to ensure innovation. However, increasing the number of firms when the underlying industry structure is already monopolistic generally increases the rate of innovation both at the firm level and at the industry level. Increasing the number of firms encourages firms to race for new patents. Secondly it reduces the degree of loss that an imperfectly competitive firm experiences via the replacement effect and hence increase the likelihood that the efficiency effect will dominate.

This leads us to think of a relationship between the degree competition and innovation that is initially increasing with the degree of competition but at some point when the degree of competition is close to Bertrand, further increases in competition or number of firms will reduce the degree of innovation. This fits in with Scherer (1967) and later empirical findings of an inverted U-shaped relationship between competition and the degree of innovation.

The other main conclusion is that looking merely at the level of concentration or degree of competition is not a sufficient guide to the relationship between the two. Other factors may well have significant elements that outweigh the initial relationship. The degree of entry into an industry is a characteristic which appears important in determining that competition will reduce innovation. Where some barriers to entry exist, the monopoly is able to reduce the level of innovation in a limit pricing manner to where the cost of entering and innovating is just above that possible. Collusion within an industry may well increase the degree of benefits that accrue from innovation but only if the collusion does not carry on to the second stage of product competition. Lastly spill-overs and the ability to imitate between competitive firms reduce the appropriability of innovations and hence will reduce the incentive to innovate. In this case a firm structure that can capture all the externalities of the innovation (such as the monopolistic structure) may benefit society to a larger extent.

More importantly for this paper is how do these theories translate to the telecom industry. At the end of section A1.1 we concluded that although the telecom industry is not the same as more classical research industries such as pharmaceuticals, it may still borrow heavily from the existing literature to make both theoretical and empirical predictions. The key conclusions taken from these two areas of literature are as follows:

Current Concentration: It is clear that the telecom industry within the Netherlands is not currently intensely competitive, for this reason increases in the degree of competition is likely to increase the degree of innovation, if not on a firm level than on the aggregated industry level.

Limitations to Competition: Whilst the first point demonstrates that competition could well be increased substantially without damaging dynamic incentives this does mean that this relationship always holds constant. At some point the level of competition will be sufficiently intense so that an industry without patents will no longer be able to sustain the risk and costs of research.

Joint Research: Where there are areas requiring substantial new technology investment and hence associated risk and costs, there are significant gains to be made from joint research. For this reason joint research in areas such as infrastructure development (particulary applicable to 3G infrastructure) or new ways of increasing bandwidth, should be encouraged.
Limitation of Joint Research. Once again like increasing competition, the degree of joint research has its limitations, the joint agreements should only be encouraged as far as the innovation stage. All innovatory developments resulting from the agreement should be distributed to the members of the agreement at 0 or cost price. Without this it is too easy and too tempting for a telecom joint research agreement to become a telecom joint pricing agreement.

The first two conclusions are especially applicable to the first exogenous policy development, concluding that generally increasing the degree of competition in relatively concentrated industry such as the new 3G mobile industry will increase innovation. The last two conclusions have particular relevance to the third policy development, where by joint agreements or convergence should be allowed but only if they are to facilitate research and not the vertical constraint of competition. The second exogenous policy development, the need for new capacity, may be thought of a high bred of the two. Firstly it takes the innovatory conclusions from the first showing that increasing competition will increase the degree of innovation. Secondly it also borrows from the last two conclusion to show that potentially increasing joint research agreements to develop new ways of introducing higher degrees of capacity will bring further welfare gains.

To conclude then, this survey is not intended to produce a definitive guide on which way competition will effect innovation in every instance but is designed to show the different effects it may have. For this reason it would be naive to make generic policy recommendations within telecommunications, as each policy development will develop different states and hence different policy recommendations.

References

Arrow K.J. (1962), Economic Welfare and the Allocation of Resources for Invention, in: R.R. Nelson (ed.), The Rate and Direction of Inventive Activity, Princeton University Press.

Balsinde, R., C. Böhmer, T. Calleja, H.-P. Kaukonen, and R. Perttunen (2000), McKinsey Quarterly Nr 2, p. 20-29.

Baumol, W.J., (1992) “Horizontal Collusion and Innovation” *The Economic Journal*, Vol. 102 pg 129-137

Beardsley, S.C., R. Raghunath and M.J. Wilshire (2000), "The emergence of broadband in Europe," The McKinsey Quarterly, Nr. 2, p. 39-42.

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, draft Brunel University.

Cisco (1999), White Paper: Controlling Your Network – A Must for Cable Operators, www.cptech.org/ecom/openaccess/cisco1.html

CPB (2000), Publishers Caught in the Web? Strategies, performance and public policy: General Framework and Conclusion, Working Paper 119, The Hague.

Dasgupta P. And Stiglitz J. 1980. Uncertainty, Industrial Structure and the Speed of R&D. Bell Journal of Economics 11: p. 1-28.

Economides, N., and F. Flyer. (2000) “Equilibrium Coalition Structures in Markets for Network Goods”. In D. Encaoua et al. (Eds.), The Economics and Econometrics of Innovation pg 339-358. Kluwer Academic Publishers, Netherlands.

European Commission (1999), Towards a new framework for Electronic Communications infrastructure and associated services: The 1999 Communications Review, Communication from the Commission to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions, COM(1999) 539, Brussels.

European Commission (2001), Commission Working Document on Proposed New Regulatory Framework for Electronic Communications Networks and Services, COM (2001) 175, Brussels.

Farber, S. (1981) "Buyer Market Structure and R&D Effort: A simultaneous equlations model", Review of Economics and Statistics 62, pg. 336-45.

Fudenberg, D., R. Gilbert, J. Stiglitz, and J. Tirole (1983), Preemption, Leapfrogging and Competition in Patent Races, European Economic Review, 22, p.3-31.

Fudenberg, D., and Tirole, J., (1985). "Preemption and Rent Equalisation in the Adoption of New Technology", Volume 3 of "Fundementals of Pure and Applied Economics" ed. J. Lesourne and H. Sonnenschein. London: Harwood

Fudenberg, D., and Tirole, J., (1987). "Understanding Rent Dissipation, On the Use of Game Theory in Industrial Organisation" American Economic Review: Papers and Proceedings 77: pp 176-183.

Gilbert R. and D. Newbery (1982), Preemptive Patenting and the Persistence of Monopoly, American Economic Review 72, 514-526.

Goldman Sachs Global Equity Research (2000), Telecom Equipment: Europe, London.

Hall, B., H. (1990) "The Impact of Corporate Restructuring on Industrial Research and Development", Brookings Papers on Economic Activity, Washington D.C., The Brookings Institution.

Harris, C. and J. Vickers (1985), Perfect Equilibrium in a Model of a Race, Review of Economic Studies 52, p. 193-209.

Hobbs, A. (2000), "The Commission's assessment of the participation of BSKYB in the pay-tv operator in Germany", *Competition Policy Newsletter*, EC, No. 3, October, 56-60.

KPN/ABN-AMRO (2000), *Launch on-line financial services partnership*, July 18. http://www.kpn.com/common/downloads/final.ppt

Lee T. and L. Wilde (1980), Market Structure and Innovation: A Reformulation, Quarterly Journal of Economics 194, p. 429-436.

Martin, S. (1991), Private and Social Incentives to Form R&D Joint Ventures, working paper 91/35, European University Institute, Florence.

Needham, D. (1975), Market Structure and Firms" R&D Behavior, Journal of Industrial Economics 23, p. 241-255.

Opta (2000), Jaarverslag 1999, The Hague.

Ordover, J. and R. Willig (1985) "Anti trust for High Technology Industries: Assessing Research Joint Ventures and Mergers". Journal of Law and Economics, May. Vol.75, pg 25-30

Reinganum, J. (1983), Uncertain Innovation and the Persistence of Monopoly, American Economic Review, 73, p. 741-748.

Scherer, F. M., (1965), "Firm Size, Market Structure, Opportunity and the Output of Patented Inventions", American Economic Review 55, pg. 1097-1125

Scherer, F. M., (1967), "Research and Development Resource Allocation Under Rivalry", Quarterly Journal of Economics, Vol. 81, No. 3. pp. 359-394.

Scherer, F. M., and Ross D., (1990) "Industrial Market Structure and Economic Performance" Boston, Houghton Mifflin Company.

Schumpeter, J. (1943) "Capitalism, Socialism and Democracy". London: Urwin University Books

Shapiro, C. and R. D. Willig (1990) "On the Antitrust Treatment of Productino Join Ventures". Journal of Economic Perspectives. Vol 4, Number 3, p 113-130.

Telefónica/Endemol Entertainment (2000), *Presentation to Analysts*, March 17. http://www.telefonica.com/index/investor_relac.html

Villiard, H. (1958) "Competition, Oligopoly and Research", Journal of Political Economy 66, p. 483-97.

Abstract

CPB investigates future policy issues within the telecommunication sector with the aid of so-called "exogenous developments". The aim is to analyze a limited number of developments that are *exogenous* to policymakers. The policy issues that emerge depend mainly on the trade-off between so-called static and dynamic efficiency. The three exogenous developments we consider are: growth of internet through mobile, the outpacing of demand for supply of broad band internet access, and the concentration and consolidation of the European mobile industry. This paper describes the policy trade-offs for each of the exogenous developments and charts both the likely development of the industry given certain key drivers and suggests ways in which the market can be pushed towards the optimum state of high dynamic and high static efficiency.
This work is co-financed by OPTA.