THE ACTIVITIES OF A MONOPOLY FIRM IN ADJACENT COMPETITIVE MARKETS:

Economic Consequences And Implications For Competition Policy

Patrick Rey - Paul Seabright - Jean Tirole[1]
Institut d'Economie Industrielle, Université de Toulouse-1
September 21, 2001

[1] We thank Christian Ahlborn, David Evans and Albert Nichols for valuable comments. This survey was supported by National Economic Research Associates, whose clients include Microsoft. However the responsibility for all views expressed remains that of the authors alone.

ABSTRACT

The activities of firms with market power are the staple fare of anti-trust policy and have been extensively analysed in the legal and economic literature. Most of that literature is concerned with activities that take place directly in the markets in which the monopoly power is actually or potentially exercised. But some writers and competition policy enforcers have expressed concern of a different kind: namely, that a monopoly firm's activities are to be feared even when they take place in other markets (usually called adjacent markets) in which that firm does not currently have a monopoly, nor even substantial market power.

In contrast to the literature on the direct exercise of market power the literature on adjacent markets has reached much less of a consensus as to when and why such indirect exercise of market power is to be feared. This paper seeks to assess where the debate now stands. The comparative confusion over the issue of adjacent markets, we argue, may be due to a failure to appreciate two important points.

(1) Firms exercising direct market power typically have both the *ability* and the *incentive* to inflict damage on consumers: actions that raise prices or lower product quality often also increase firm profits. By contrast, in adjacent markets firms with an ability to inflict damage on consumers do not necessarily have an incentive to do so: actions that raise prices compared to a competitive level in adjacent markets may lower overall firm profits.

(2) Very widespread and commonplace synergies may exist between activities in nominally different markets, either because costs and knowledge can be shared across activities, or because the assembling of component goods and services into new combinations creates real value for consumers.

These two points do not imply that concern about a monopoly firm's behaviour in adjacent markets is always misplaced. In fact, in this paper we identify a range of circumstances in which such behaviour may indeed raise a legitimate concern. But we emphasize that it is important to identify these circumstances fairly precisely: behaviour in adjacent markets *is* different from the direct exercise of monopoly power, and involvement in multiple markets *is* a normal and often innocent form of industrial activity, and neither of these intrinsically justifies a presumption of suspicion.

1. Introduction

The activities of firms with market power are the staple fare of anti-trust policy and have been extensively analysed in the legal and economic literature. Most of that literature is concerned with activities that take place directly in the markets in which the monopoly power is actually or potentially exercised. But some writers and competition policy enforcers have expressed concern of a different kind: namely, that a monopoly firm's activities are to be feared even when they take place in other markets (usually called adjacent markets) in which that firm does not currently have monopoly or even substantial market power.

In contrast to the literature on the direct exercise of market power the literature on adjacent markets has reached much less of a consensus as to when and why such indirect exercise of market power is to be feared. This paper seeks to assess where the debate now stands. The comparative confusion over the issue of adjacent markets, we argue, may be due to a failure to appreciate two important points.

(1) Firms exercising direct market power typically have both the *ability* and the *incentive* to inflict damage on consumers: actions that raise prices or lower product quality often also increase firm profits. By contrast, in adjacent markets firms with an ability to inflict damage on consumers do not necessarily have an incentive to do so: actions that raise prices compared to a competitive level in adjacent markets may lower overall firm profits.

(2) There is a common suspicion that a monopoly firm's activities in adjacent markets must be a signal that it is up to no good – why else, it might be thought, would a firm venture away from the deep peace of monopoly power and into the hurly-burly of the competitive market place? This suspicion ignores, we suggest, the very widespread and commonplace synergies that exist between activities in nominally different markets, either because costs and knowledge can be shared across activities, or because the assembling of component goods and services into new combinations creates real value for consumers. The typical modern firm is intrinsically a flexible, multi-tasking organization, and even if it exercises

> market power in some of its activities it is unlikely that it exercises market power in them all. The fact that it is active in adjacent markets is often a consequence of the complex nature of modern industrial organization.

These two points do not imply that concern about a monopoly firm's behaviour in adjacent markets is always misplaced. In fact, in this paper we identify a range of circumstances in which such behaviour may indeed raise a legitimate concern. But we emphasize that it is important to identify these circumstances fairly precisely: behaviour in adjacent markets *is* different from the direct exercise of monopoly power, and involvement in multiple markets *is* a normal and often innocent form of industrial activity, and neither of these intrinsically justifies a presumption of suspicion.

In this paper we shall spend much of our time discussing adjacent markets of a particular kind, namely those for goods that are useless unless combined with the monopolized good. These are typically the kinds of market that have featured most often in relevant anti-trust decisions, since they are the markets in which the monopoly firm appears to have the greatest capacity for exploiting the "essential" nature of its own product.[2] Indeed the legal doctrine of "essential facilities" seeks to address the competition problems inherent in this situation. However, we shall also look at the more general case in which the two goods are complementary products, meaning that their value is greater when they are consumed in combination, even if neither of them is strictly useless without the other.

In either case, whether the monopoly good is essential or merely complementary to those produced in adjacent markets, we shall be concerned greatly with the questions of how and by whom these different goods are combined. As we shall discuss in more detail in

[2]Such markets have been the object of much antitrust scrutiny at least since the celebrated *Terminal Railroad* case in 1912. We can now avail ourselves of a wide range of decisions regarding the vertical behaviour of owners of electricity, railroad, telecommunications, and transportation infrastructures, of stadiums, computer reservation systems, wholesale markets, software and intellectual property, equipment producer, delivery networks, and franchises, not to mention decisions related to aftermarkets. Examples include, respectively, *Otter Tail* (1973); *Terminal Railroad* (1982); *Sealink* (1992, 1994), *Port de Roscoff* (1995), *Rodby* (1993), *Frankfurt Airport* (1998), *Eurotunnel* (1994); *Hecht* (1978); *Sabena* (1998), *CAB* (1984); *Gamco* (1952), *New York Stock Exchange* (1963); *IBM* (1984), *Magill* (1995); *Tetrapak* (1994), *International Salt* (1947); *Oscar Bronner* (1998); *Chicken Delight* (1971); *Renault* (1988), *Volvo* (1988), *Kodak* (1992).

section 2, the creation of economic value takes place as much in the combination of components as in the production of the components themselves. Whether we think of a car- or television-assembly plant putting together components that have been manufactured elsewhere, or a chef mixing together the ingredients of a legendary sauce, the activity of assembly creates real value for which consumers are typically prepared to pay, and sometimes to pay a lot. Shops do sell furniture that you put together yourself, and there are restaurants where you create your own salads from ingredients supplied. But these are typically bottom-of-the-range outlets whose products trade at much lower prices than the norm. Such price differentials may sometimes reflect a strategy of price discrimination, or just a need to cut the cost of shelf space in the store, but they are often based upon a significant perception of added value in assembly. Combining components often requires not just labour but skill – skill at selection, skill at knowing the kinds of combination that will work, and skill at managing the often fluctuating stocks of components that the assembly process requires. We shall therefore consider to what extent the fact of possessing monopoly power in the production of a particular component influences not only the production of other components, but also the decision by the monopoly firm to assemble the components itself, to allow other firms to assemble components, or to allow consumers to buy components separately for assembly by themselves. Each of these decisions will have a potential economic rationale in its own right, and any potential anti-competitive consequences need to be set against an understanding of that rationale.

We begin in section 2 by setting out more precisely what we mean by adjacent markets, and by asking two questions about the firm's involvement in these markets. First, what determines whether a monopoly firm will wish to be active in adjacent markets in the first place? Secondly, what determines whether it will undertake the assembly of the monopoly product and the competitive product itself, do so in competition with others, or leave assembly entirely to other firms or to consumers? We then consider the various strategies that the monopoly firm can pursue with respect to the goods concerned, including its pricing decisions, the nature of the contracts that it signs with other firms, and the kinds of innovation it undertakes. We emphasize the need to consider the underlying economic rationale for each of these strategies as a prelude to understanding whether they may also have possible anti-competitive effects.

2. Adjacent markets: a framework for analysis

2.1 What are adjacent markets?

We are interested in situations (depicted in figure 1) in which

a) A firm *M* is present in two adjacent segments,[3] *A* and *B*, such as local and long distance telecom services, rail infrastructure and travel services, operating systems and application software, client and server operating systems, or equipment production and maintenance.

b) Firm *M* has substantial market power in market *A*, called for simplicity the "monopoly segment" (the local loop, the transportation infrastructure, etc.) and faces actual or potential competition in market *B*, labelled the "competitive segment" (long distance telecommunications, train operating services, etc.);

c) Product *B* is less valuable, and may even be useless, unless combined with product *A*; in contrast, product *A* may or may not have a life on its own. For example, in the case of telecommunications, the local loop (product *A*) can be used for long distance calls (product *B*) but can also be used for local calls or Internet access. In the *Magill* case, TV offerings (product *A*) had a life of their own, independently of the production of generalist TV guides (product *B*), which did not even exist. As a last example, the Windows client operating system (product *A*) has many other usages than those related to the network functionalities provided by servers (product *B*). In contrast, stadium seats or railroad tracks (product *A*) are useless unless they are combined with an entertainment offering or train service (product *B*).

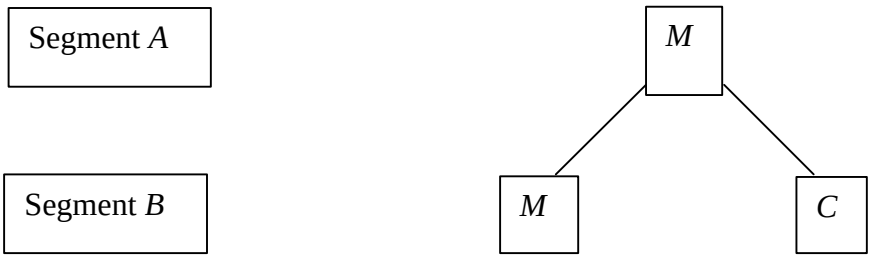

Figure 1: Adjacent markets

[3] Alternatively, *M* could be present in segment *B* through "privileged partners".

The monopoly firm faces two fundamental decisions about the kinds of activity it will undertake. First, should it be active in market *B* at all? And secondly, should it combine (or assemble) the two goods for sale to consumers or leave this to consumers themselves or to other firms? These two decisions are distinct, and give rise to four possible outcomes, depicted in Figure 2. However, the considerations influencing the two decisions will typically be interrelated, in that a firm that wishes to assemble the goods itself may be more likely also to wish to produce good *B* than one that has no wish to do so. We therefore begin by outlining the considerations that will influence the assembly decision.

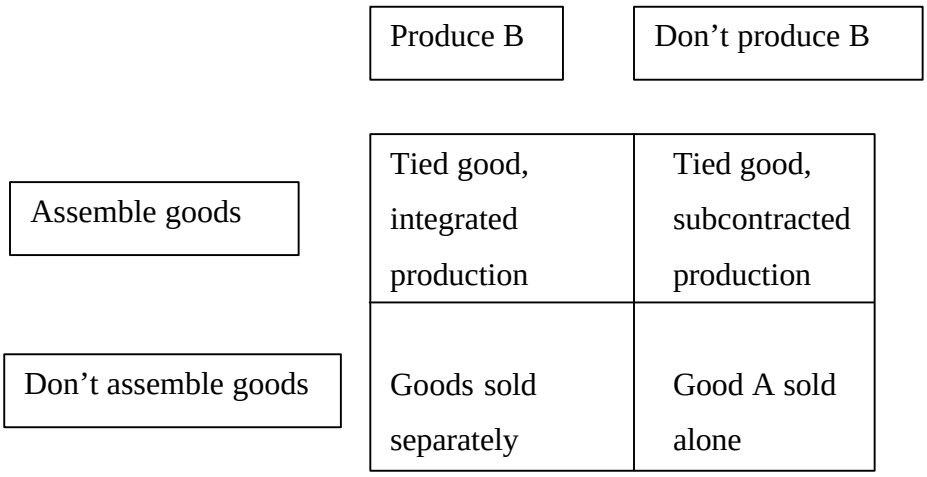

Figure 2: The production and assembly decisions of the monopoly firm

2.2 The assembly decision: what are the issues?

2.2.1 Examples of assembly

Even though products *A* and *B* may be useless unless consumed together, the assembly may be performed by the consumer who purchases them in two separate transactions. For example, the client Windows operating system is sold separately from consumer support, applications, hardware, or server operating systems. Similarly, the sale of primary goods may not include customers' rights to maintenance services. Local telephone services may be marketed with handset equipment included (as used to be the case in most countries until recently); or else they may be sold separately with the handset equipment marketed (and priced) on a stand-alone basis. Your local restaurant may or may not allow you to bring in a bottle of wine of your own choice.

Alternatively, product *A* may not be sold separately to the consumer; it may instead be used internally or sold externally to product *B* firms which incorporate the *A* input into the final product. Both may even occur simultaneously. Competitors in market *B* are then said to pay an "access charge" or "access fee" to *M*. Examples include the licensing of biotechnology patents to pharmaceuticals,[4] regulated access to tracks, high-voltage grid, pipeline or local loop facilities, the access to harbour, airport or tunnel infrastructures, the access to computer reservation systems or stadiums, and the licensing of code to be incorporated into larger software programs. Alternatively, firms in market *B* may act as component suppliers to the monopoly firm, possibly in competition with some of the monopolist's own in-house component production ("make or buy" does not have be an all-or-nothing decision).

What are the benefits that are created for the consumer when the assembly of the two products is performed on her behalf? In ascending order of complexity we can think of such benefits as consisting of saving labour, exploiting economies of scale in specialization, and overcoming information problems. We deal with these in turn.

[4] The "access charge" is then called a "licensing fee".

2.2.2 Saving labour

Saving labour is the simplest to understand: I may wish someone else to put my furniture together just as I may wish someone else to mow my lawn. This may be reasonable even if putting the furniture together involves no skill whatever, just so long as I have enough skill at some other task for my time to be better spent doing that. As trade theory teaches us, it is not the absolute but the relative advantage of the furniture assembler that determines the efficient allocation of tasks. And another insight of trade theory, namely Adam Smith's famous observation that "the division of labour is limited by the extent of the market," leads us to expect that the demand for assembly should increase as economic development proceeds: the richer the country, the fewer the people who assemble their own furniture. Three hundred years ago in Europe and North America most families ate food they had grown themselves, used furniture they had made themselves and lived in houses they had built themselves. They were not self-sufficient in the strict sense since they usually relied on others for some things – metal for agricultural tools, for example. But changes in their links with the outside world would rarely threaten their ability to eat. Today in these countries most families who were prevented from exchanging with others would starve within a few weeks. Many of them also expect their furniture to be manufactured by others.

In general, though, there is no particular reason why sellers should not offer buyers a choice: you may have the assembled version or the disassembled version and you choose freely between them. Or rather, how much of a choice you are offered will depend on the transactions costs associated with the two alternatives. Twenty years ago in Europe most household appliances were sold without plugs, and consumers would buy the plugs separately; now, with increasing prosperity, most consumers value the convenience of buying the appliances already plugged. Theoretically stores could offer two versions of each appliance, one with a plug and one without, but this would involve additional costs of shelf space, sales and marketing, and stock control that would probably outweigh any gain to consumers from the greater choice. And plug manufacturers have not complained that they are being driven out of business; they have increasingly become sub-contractors to appliance manufacturers rather than sellers direct to the public. But it is a fair bet that if new types of plug were invented that had very different properties from existing ones, so that some consumers wanted the new plugs while others did not, appliance manufacturers would start to offer consumers a choice again.

Once the motivation for assembly becomes more complex than simple labour-saving, however, the options typically become substantially more restricted. Why should this be so? We can distinguish two kinds of tying of products: technical tying and commercial tying. The former occurs as part of the assembly process, while the latter involves offering goods in combination at the point of sale that it is not particularly advantageous to combine in the assembly process. As we shall see, while scale economies tend to be associated with technical tying, asymmetries of information may be an important motivation for tying of either kind.

2.2.3 Economies of scale

Consider assembly designed to exploit economies of scale in specialization. These economies of scale may arise, for example, through the use of capital equipment. A likely reason is that such equipment allows assembly to be automated, though capital equipment may be necessary even without automation. Most households own a hammer and a screwdriver, enough to assemble simple furniture, but assembling more complex equipment may require sophisticated and expensive tools, which it makes no sense for every household to own. Assembling electronic goods, for example, often requires an expensive dust-free environment and precision equipment. Hiring such equipment may be possible, though it may economise on transport costs to take the components to the equipment rather than vice versa. But hiring the knowledge of how to use the capital equipment is much harder. I could hire the equipment, then hire someone to train me in its use, but it will usually be more efficient for me simply to buy the assembled goods.

The real economies of scale, then, often arise through the acquisition of knowledge and skill. It is simply impossible to become skilled in everything; it takes time and practice as well as talent, and it would involve wasteful duplication of costs even to try. When assembly requires a lot of skill, it is therefore quite rare for consumers to be offered a choice of the components and the assembled product. Simple electronic goods represent an interesting intermediate case. This is partly because there are enough amateur hobbyists around to constitute a significant demand for the disassembled product. And partly it is because most goods contain some components (such as fuses) that can be easily replaced at home at less expense (and with less expertise) than is involved in repairing the goods at the factory. So

hardware stores usually contain a wide range of electronic components, and there exists a variety of products that come in kit form. An impressive if dwindling band of enthusiasts manufacture their own radios. But with most goods of any sophistication you simply cannot buy the kit.

Why not? In theory a manufacturer of (say) magnetic resonance imaging equipment could offer a kit version to those few consumers who claim to have the necessary expertise, and thereby allow them to substitute from other sources components they prefer to the ones supplied by the manufacturer. But in practice such manufacturers do not, and for a very good reason, which brings us to the third and much the most important motivation for assembly.

2.2.4 Overcoming information problems

There are significant informational obstacles to assessing the reliability of components, the quality of expertise and the soundness of assembly procedures. A consumer who believes herself to have the necessary skill to assemble magnetic resonance imaging equipment, but fails to produce a working product, may wrongly blame faulty components rather than her own defective expertise. Or she may mistakenly blame one of the manufacturer's own components rather than one of those she has purchased from elsewhere. Equipment manufacturers invest heavily in the acquisition of a reputation for reliability. This reputation rests on the soundness of judgement with which they select the types of component for their products, the accuracy with which they screen individual components for defects, the uniformity of the assembly procedures, and the thoroughness of the post-assembly testing. A manufacturer choosing whether to offer the consumer a choice of self-assembly has to trade off the benefits of satisfying a group who might prefer the kit, against the potential costs to its reputation of involving in the assembly consumers whose enthusiasm is not matched by the quality of their expertise. All manufacturers can and must choose among competing suppliers of parts, labour and skill. The inevitability of such choice is in no way diminished by the fact that the competing suppliers may include the eventual customers themselves.

The solution most assembly processes adopt is modularisation. When part of a finished product can be separated relatively easily (either physically or functionally) from the rest, the component can be considered a module. The transactions costs and, often, the

reputational risks of allowing the consumer to add it or subtract it are small. (It may be easier to tell whether a fault is in the detachable component or in the rest.) Similarly, consumers can easily judge the benefits of having more or different modules, and do so more easily than if the parts work together without clear boundaries between their functions. So consumers may typically be offered a choice of modules, but no choice about the components that enter into each module. A fuse is a good example of a module.

More generally, the nature of modularisation will change with changing technology, and integration of components that seems unnecessary at one time may become important as the character of technology develops. For instance, it is possible that satellite navigation systems for cars, which are currently available as optional modules, may become integrated seamlessly into the functioning of cars' on-board computers so that route choices can be made that integrate data from the car such as tyre wear and fuel consumption with data from the environment such as road quality and the distance to the nearest gas station. This might be simply in order to reduce the costs of manufacture. But even if this were not a motive, reputational considerations might play a role: a driver whose car runs out of gas would not know whether the fault lay in the navigation system or in the signal sent to that system from the car's fuel tank, so there would be an important reputational externality between the manufacturers of the two products. It is also interesting – and entirely explicable – that as more and more computing functions depend on continuous networking, more and more computers are being made with modems and Ethernet cards internally installed instead of being supplied as detachable components. What these examples show is that the boundaries between modules frequently change as rapidly and as unpredictably as the other characteristics of the products themselves. They do so in response to changing technological constraints and the pressures of consumer demand, and these shifting boundaries between modules in turn determine the extent of the assembly choices it is reasonable to offer to consumers.

Reputational considerations are not the only way in which asymmetric information can militate in favour of tying. When goods can be combined in variable proportions, the firm may wish to signal to consumers that it can determine these proportions more efficiently than the consumer herself can; the result will be commercial tying even in the absence of technical tying. This might be for two reasons: one is that the relative prices of different components may be distorted (because, for example, the monopolist charges a mark-up on its own

differentiated component that is different from the mark-up charged by rivals, reflecting a different degree of market power in component manufacture). The firm may then tie the goods together to eliminate inefficient substitution away from its own product[5]. It could simply have reduced its mark-up on good *B* while raising the price of *A* to compensate, but this might conflict with a desire to sell units of good *A* on a stand-alone basis.

A second reason relates to a wish to signal product quality.[6] Suppose that *A* is a durable good of quality unknown to the consumer; *M* would like to convince the consumer that he will enjoy using good *A*. One possible strategy is then to sell good *A* at a low price and tie the sale of complementary good *B*. The strategy amounts to telling the consumer that *M* is confident that the consumer will make much use of good *A* and therefore consume much of good *B*, enabling *M* to recoup the loss made on good *A*. This strategy *de facto* ensures the consumer against any attempt to cheat on the quality of good *A*.

However, information problems do *not* all militate in favour of prior assembly by the firm. For just as there may exist asymmetric information about products and their components, with firms typically better informed than consumers, so there may also exist asymmetric information about consumer preferences, with consumers better informed than firms[7]. This will be more common where the product *A* has a life of its own (as for local calls - as opposed to long distance calls - or TV channels - as opposed to TV guides). In such cases it is reasonable to think that consumers' preferences may vary significantly over the two components, and that it may make sense to sell the goods separately to allow consumers to express these different preferences in their purchasing decisions. By contrast, left shoes are not sold separately from right shoes because it is very rare for consumers to have preferences for left shoes that are distinct from those for right shoes (though fashions could change…).

Uncertainty over consumers' preferences can sometimes be dealt with by offering consumers a choice of bundles. Dell Computers, for example, allows buyers to design their

---

[5] See Vernon & Graham (1971), Schmalensee (1973), Warren-Boulton (1974), Tirole (1988, p.179).

[6] See Schwartz-Werden (1996).

[7] For example, computer reservation systems have been operating in the "input mode". The computer reservation system service is sold as an input to travel agent, and not priced directly to consumers. In contrast, search engines can charge the consumer (in money or in advertising) for being matched with ticket offers.

own computer prior to manufacture by specifying combinations of components required; this evidently depends on a successful process of component standardization as well as on flexible manufacturing and reduced transaction costs (e.g., through web ordering). This phenomenon encompasses such diverse examples as cars, where website ordering is increasingly allowing customized manufacture, and engagement rings, where the De Beers website allows customers to design their own offering to the beloved. Furthermore, the uncertainty on the consumers' preferences over the bundle of components must be compared with that relative to the consumers' preferences over each of the components. For example, when there are many components and consumers have independent preferences for these components, the law of large numbers makes it easier to "know" consumers' preferences for the bundle.[8]

Uncertainty over consumers' preferences might even provide a reason in favour of tying if consumption of one good is much easier to monitor than that of the other. The monopolist may tie goods *A* and *B* in order to meter the consumer's demand (willingness to pay) for good *A*, and charge more to those who have a higher willingness to pay. Such tie-ins are often efficient. Consider a monopolist supplying fibre-optic cables to consumers' homes. Suppose that there are two types of consumers in equal proportions: heavy users, interested in video streaming services, and light users, interested in telephony, email and similar communication services that are less bandwidth demanding. To fix ideas, suppose that heavy users consume twice as much bandwidth as the low users. The heavy users are willing to pay 10 and the low users 6. The cost of connecting an extra consumer's home is 4.

Suppose first that the operator only sells the fibre equipment to the consumers, who are then free to use the bandwidth as they wish (contract with content and phone operators). The fibre company then charges 10 to the consumers, as its profit per inhabitant is (10-4)/2=3, while the policy of reaching all inhabitants by charging 6 raises only 6-4=2. If instead the fibre company controls the use of the bandwidth (e.g., video versus telephony or email), it can then charge 10 to heavy users and 6 to light users (provided heavy users derive value no

[8] See Armstrong (1999) and Bakos-Brynjolfson (1999). Another classic example, due to Stigler (1963), is when the consumer has negatively correlated preferences over two components – in the extreme case of perfect negative correlation, there is no uncertainty over the consumer's preferences for the bundle, even though his or her preferences for each component are uncertain.

more than 6 from a light use). Profit has increased and so has consumer surplus and thus social welfare, since low-willingness-to-pay consumers are no longer excluded.

While this example is hypothetical, more concrete instances of the use of a tie-in as a metering strategy have been suggested. IBM used to tie punched cards with the rental of its computers in order to measure the use of the equipment.[9]

Of course there may be alternatives to tying that allows *M* to meter the usage. For example, in the fibre-optic illustration, *M* could just meter the use of bandwidth[10] rather than tying the content. This strategy however would not work if the heavy use of bandwidth could correspond to different services with very different willingness to pay. More generally, tie-ins may be necessary to measure the monetary value of consumption.

[9] A similar motivation has been invoked in *Chicken Delight* (1971) and *Tetrapak* (1994).

[10] In as similar way, a server software vendor may meter the usage of the software through the number of computers that access the server.

2.3 The production decision: what are the issues?

2.3.1 Technological considerations

What determines whether the monopoly firm will also produce good *B*? As we have indicated, the nature of modern production means there will often be straightforward technological reasons why such a decision makes sense. There may be economies of scope between the two segments (for example, R&D meant to improve the *A* product may have fortuitously generated a *B*-innovation). Alternatively, there may be cost advantages to developing components (such as pieces of software code) that are used in both *A* and *B*; similarly, the benefits of learning-by-doing can spill over from one segment to another. In principle, these technological reasons for producing good *B* may exist independently of whether the assembly of goods *A* and *B* is performed by the monopoly firm. In practice, though, it is likely that some kinds of synergy may be enhanced by the possibility of performing the production and the assembly of *B* in the same place (which may make just-in-time production easier, for instance).

However, there are three further reasons why the monopoly firm may wish to produce good *B* which rely intrinsically on its also undertaking the assembly of the two goods. These are vertical externalities, hold-up problems and restrictions on price discrimination.

2.3.2 Vertical externalities.

Vertical externalities refer to the uncoordinated choice of prices, investments, specifications or other attributes on the two segments. The most well-known such externality is double marginalization: in the presence of market power on the *B* segment, the price of the service obtained by combining products *A* and *B* embodies two margins, and so the price of the combination even exceeds the monopoly price. This is sometimes summarized by saying that there is nothing worse than a monopolist – except two successive monopolists.

Vertical integration is one of the ways in which the corresponding externality can be internalised. It is then socially beneficial because it eliminates one of the margins, and produces a situation in which both the industry and the consumers gain.[11]

Similarly, vertical integration can help solve other coordination problems. Suppose for example that the consumer can only assess the quality of the final product, and not the respective contributions of goods *A* and *B* to the quality of the final product. Then, in the absence of vertical integration, when deciding to invest or not in the quality of their components, the competitors in segment *B* will fail to take into account the impact of their quality efforts on *M*'s reputation – and likewise, *M* will not take into account the benefits, for *B* firms, of its own effort for maintaining the quality of the input *A*. Similar externalities arise in the case of advertising. For example, they are frequent in the production-retail chain: when deciding whether to launch a national advertising campaign, the manufacturer takes into account the impact of the additional sales on its own profits, but not necessarily on its retailers' profits; the same argument applies similarly, on the other side, for the local advertising campaigns run by the retailers.

Also, vertical integration solves the familiar chicken-and-egg problem that may prevent complementary products from being supplied: in the absence of vertical integration, product *A* (or an innovation thereupon) may not be supplied if product *B* does not exist, and conversely. As for other externalities, there may exist other ways to address the problem. For example, in the software industry, the early release of interface specifications, expenditures on evangelisation, and reputation effects may sometimes create enough commitment to prevent a coordination failure. Formal contracts may also be feasible. These are frequently used in the aerospace industry: Airbus Industrie sought advance orders for its A380 super-jumbo prior to taking an irrevocable decision to proceed with development (and in return, Airbus commitment to develop the A380 may encourage airlines to organize their routes and hubs so as to take advantage of large capacity aircraft). However, when contracting is complex, for instance because of the difficulty involved in specifying the end products, vertical integration may be a more effective alternative.

---

[11] There are other ways of addressing the double marginalization problem such as non-linear pricing of the input *A* to producers in segment *B*. Different approaches are not equivalent in general. On this, see Tirole (1988, chapter 4).

2.3.3 Hold-up problems.

Vertical integration may also be motivated by the desire to protect specific investments when effective long-term contracts are hard to write.[12] For example, suppose that a new automobile model can run only if some specifically designed one-Euro-a-piece component is included. While in principle many suppliers may be able to design and produce this component, in the absence of formal agreement, the one that is selected may have enormous hold-up power when the car is ready for the assembly line. Indeed, it can expropriate part of the automobile manufacturer's investment by threatening not to supply the component and thereby block the production of the car.

Three solutions can be contemplated in order to prevent such opportunistic behaviour. First, the automobile manufacturer can encourage dual sourcing to create ex post competition between rival suppliers of the component. However, component producers may not want to invest in designing the component if they anticipate that they will wage tough competition and therefore will not recoup their investment; more generally, this solution may exacerbate hold-up concerns for the suppliers. Second, the automobile manufacturer and the component supplier can enter into a long-term contract that sets the specifications and the price of the component; this approach requires being able to set the specifications precisely at the date of the agreement.[13] It may also require the posting of collateral (in money or in reputation) by the component supplier to avoid any temptation to force ex post renegotiation.[14] Third, the automobile manufacturer can produce the component itself; it is then able to take over the blueprints and machinery employed to produce the component and to replace the division manager if the latter attempts to behave opportunistically. To sum up, vertical integration

[12] See Grossman-Hart (1986) and Williamson (1975, 1985).

[13] This does not necessarily prevent the automobile manufacturer and the component supplier from adapting ex post the contract to the circumstances. What matters, however, is that the initial contract (i) be indexed to verifiable shocks, (ii) provides parties with flexibility to react to unverifiable shocks through options, and (iii) defines an appropriate starting point for the ex post bilateral renegotiation – see Aghion-Dewatripont-Rey (1994).

[14] On this, see Hart-Moore (1994).

may substitute for long-term contracts to protect specific investment when the latter prove hard to write.

2.3.4 Restrictions on price discrimination.

Economic efficiency and profit maximization both recommend offering low prices to consumers with low willingness to pay (the high-elasticity-demand consumers) and higher prices to consumers with higher willingness to pay (the low-elasticity-demand consumers).

Suppose, in our framework, that firm *M* produces only good *A*, which is used as an input by good-*B* producers who then sell the combined product to the final consumers. A standard example of "good *B*" is the retailing or after-sales service activity, but the scope of the analysis is obviously not confined to these cases. Other examples include pharmaceuticals and medical equipment (which are combined with medical services), and books and IT equipment (which may be combined with educational services). "Downstream" competition among *B*-producers prevents price discrimination, and in many cases results in an inefficient outcome.

Suppose for example that *M* considers developing a good to be distributed by independent retailers in two equal-sized geographic markets, a "low-price" region one in which consumers are willing to pay 10 for the final good, and a "high price" region in which consumers are willing to pay 20. To fix ideas, suppose that all production and retail costs are zero, but that the development of the good involves a fixed investment cost which, *per capita* (that is, divided by the total number of potential consumers), is lower than the average willingness to pay (that is, $I < (10+20)/2 = 15$), so that it is efficient to develop the good. In that case, *M* would indeed choose to develop the good if it can sell at 10 in one area and at 20 in the other. However, competition in the high-price region would in that case lead the retailers to buy at 10 in the other region and to sell at that same price in the high-price region. *M* is then left with two options: either it sells at 10 in both regions, or it withdraws entirely from the low-price region in order to keep selling at 20 in the high-price region; in both cases, the average per capita revenue drops down to 10. As a result, the good will not be developed, even though it would be efficient to do so, whenever the per capita cost *I* lies below 15 (the average willingness to pay). The same argument applies to any situation where consumers

differ in their valuations for the good, and applies more forcefully when consumer heterogeneity increases.

Two solutions are then at hand (although as it happens both conflict with European Union antitrust rules). First, *M* can preserve downstream competition, but delineate the market available to different groups of retailers. An example of this strategy is the prohibition of parallel imports by one geographical area from another. It is an important strategy used by manufacturers of branded products seeking to supply low-income markets, which would probably not be supplied at all in the absence of some way to prevent the low-price goods from being re-exported back to the high-price market. There has been much controversy about such cases (branded jeans, compact discs, sunglasses, and prescription medicines are leading examples). Sunglasses were the focus of the 1998 Silhouette case that upheld a manufacturer's right to prevent parallel imports from outside the EU (in this case from Bulgaria into Austria) – a right that does not extend to intra-EU trade.[15]

Second, *M* can enter the *B* market, charge a high intermediate price to the independent *B*-producers who then supply the low-elasticity (high-price) segment, and leave its internal *B*-division to focus on the high-elasticity (low-price) segment. This division of labour requires discriminating between external buyers and the internal buyers of product *A*,[16] although it could be argued that there is no price squeeze as long as *M*'s *B*-division does not go after low-elasticity-demand consumers (which *M* has no incentive to do if independent *B*-companies are competitive).

---

[15] For example, the Block Exemption Regulation CE 2790/1999, which reflects a generally more tolerant attitude towards vertical agreements, stills blacklists restraints on parallel imports. As for prescription medicines, which are subject to different price regulation systems across Europe, pharmaceutical companies have attempted to limit the parallel imports triggered by the disparity in these national regulations. The European Commission has so maintained its firm condemnation of these practices, although the Court of First Instance took a more open position in *Adalat* (2000).

[16] In the sense that the Efficient Component Pricing Rule is not satisfied (the intermediate price exceeds *M*'s *B*-division margin). Such price discrimination would be considered as an abuse of a dominant position under Article 82. This is for example a driving concern of the Commission in the telecommunications industry (see e.g. the *Regulation* EC 2887/2000, which entered into force beginning 2001, and the *Communication on unbundled access to the local loop* OJ C 272, 23.9.2000).

In conclusion, then, a decision by the monopolist to produce good *B* may arise out of natural technological complementarities with good *A*, or out of complementarities that are more strategic, arising from the ability to avoid various transactions costs or strategic constraints in the production of good *B* alone. However, the question must arise whether such strategic considerations may sometimes be deployed in an anti-competitive way. This is the subject of section 3.

3. Anti-competitive conduct

3.1 Incriminated practices

The literature and jurisprudence of competition policy have given rise to a large number of allegations that the firm that is dominant in market *A* may abuse this dominance in the adjacent market *B*. The practices that have been incriminated at various times all relate to the limitation of access to competitors in *B*, the dominant firm being accused of leveraging its market power on segment *A* to gain dominance of segment *B*. Such limitation of access is usually referred to as "foreclosure," and can aim at preventing entry into segment *B*, driving competitors out of the segment, or simply reducing their share of the segment without driving them out. The incriminated practices include:

- refusal to sell input *A* to competitors;
- incompatibility between product *A* and versions of product *B* sold by competitors;
- prohibitive access charges;
- tying products *A* and *B*, that is, offering the two goods as a bundle to consumers;[17]
- and if such foreclosure is not uniformly applied to all outsiders, discriminatory pricing and licensing.

[17] If product *A* is never offered as a stand alone product, this de facto excludes competitors in market *B*; these competitors are also *de facto* excluded if input *A* is only offered at a price higher than the price of the bundle, net of the cost of the other components (this situation is sometimes known as a "price squeeze").

Essentially these strategies fall into two categories: supply and demand-side strategies. Supply-side strategies (such as excessive access charges – of which refusal to supply is just an extreme instance) are those that make it more difficult for competitors in *B* to supply the combined product to consumers. Demand-side strategies (particularly tying) are those that make consumers less willing to purchase competitors' versions of good *B*. Technological incompatibility may be either a demand-side or a supply-side strategy, depending on whether assembly of the two goods is typically performed by consumers or by firms.

It has also been suggested that the presence of a link between the two markets gives the dominant firm an anti-competitive motive for innovation in market *B*, so that innovation (normally something the anti-trust authorities would not wish to discourage) should be regarded with particular suspicion when undertaken by a dominant firm even when the market in which it innovates is not the one in which it is dominant. We examine this argument further below.

In Europe, the competition rule that is most relevant for such practices is Article 82 (formerly 86) focusing on abuses having a potential appreciable effect on trade between Member States by a firm enjoying a dominant position in a relevant product and geographic market within the Union. Article 82 explicitly prohibits unfair prices and trading conditions, the limitations of production or markets, discrimination and tying, and can be used to incriminate other practices such as refusals to supply or grant a license, predation, or the use of loyalty-inducing schemes.

The remedies that may be applied when the dominant firm is found guilty of foreclosure are many:

- Structural remedies[18] include *M*'s divestiture of assets in segment *B* together with a line-of-business restriction preventing *M* from re-entering that segment (AT&T's 1984

[18] These include *de jure* restrictions, but also remedies that amount to *de facto* line-of-business restrictions. For example, a policy that would force *M* to share (with no or low compensation) its innovations on segment *B* with its competitors would discourage *M* from innovating in that segment and is thus similar in spirit to a line-of-business restriction.

breakup), and the obligation to share the assets in segment *A* (Terminal Railroads 1912). Although such remedies have not to date been applied by the European authorities they are in principle available to them.

- The control of access prices and conditions forces *M* to charge prices for access to *A* that do not discriminate among external units in the *B* segment (computer reservation systems) and do not discriminate between external buyers and *M*'s own division in the *B* segment[19] (US railroads, local loop resale).
- The control of the volume of access sets a capacity to be reserved for competitors (Eurotunnel).

3.2 Foreclosure: the Chicago critique

We have emphasized that both the monopolist's decision to enter the market for good *B* and its decision to tie the two products together may have a rationale grounded in the creation of genuine benefits to consumers, and it is important to bear these in mind when considering possible anti-competitive consequences. In addition, while it is certainly possible for dominant firms to use these strategies to engage in anti-competitive foreclosure of competitors in adjacent markets, this does not necessarily make it a rational strategy for them. The reason why it may not be rational is that when the second product is a complement to the first, then a monopolist that can exploit its market power for its own monopolised product has no interest in excluding low cost and high-quality varieties from the market since their presence makes its own product more attractive to consumers. This is true both when the monopoly product is sold separately on a stand-alone basis (consumers will be prepared to pay more for it when there are high-quality low-cost complementary goods also available in the market), and when it is only an input into a downstream product (the higher quality or lower cost of the downstream segment raises the overall willingness to pay for the final

[19] This is usually understood to mean that the access charge to external customers satisfies the Baumol-Willig Efficient Component Pricing Rule (ECPR), according to which the access charge should not exceed *M*'s opportunity cost in segment *B* (i.e., *M*'s price minus marginal cost in that segment). More generally (since the ECPR is applicable only in certain specific circumstances) this means that external customers should find it profitable to enter the market only when their costs in the competitive segment are lower than those of the integrated monopolist in that segment.

product, allowing the upstream firm to extract more rent for its input). The key point is that the monopolist can extract its profit through its pricing in the monopoly market rather than through seeking to exercise monopoly power in the second market. As we shall see in section 3.4, foreclosure may sometimes be driven by the monopolist's inability to exploit its hypothetical market power in its own market. But in circumstances when it can do so unhindered, the presence of high-quality and low-cost complementary goods is something it will wish to welcome.

The circumstances when it would make sense for a monopolist to engage in anti-competitive foreclosure are different. One possibility, discussed in section 3.3 below, is when the products are relatively independent and the monopolist has a realistic chance of driving competitors out of or of discouraging entry in the adjacent market. In this way it can establish a second source of monopoly power whose exercise (because of the independence assumption) does not damage the value of its original monopolised product. While we illustrate this argument in the case of tying, similar considerations apply to the remaining strategies of foreclosure.

This argument applies to the case of a monopoly good marketed separately on a stand-alone basis, but there is a second set of circumstances in which foreclosure may make sense in spite of complementarity, and which applies when the monopoly good is used solely as an input. While it remains true that the monopolist benefits from high-quality low-cost supply in the downstream segment, it nevertheless may have an interest in restricting the degree of competition among downstream firms. Here there is no question of tying (since the monopoly good is an input anyway), but other strategies of foreclosure may be attractive. We consider this situation in section 3.4.

Yet another possibility, discussed in section 3.5, is when entry in the adjacent market facilitates entry in the monopolised market. Then, the incumbent monopolist may be tempted to deter entry in the adjacent market (possibly through tie-ins when the monopolised good is sold on a stand-alone basis, or through vertical integration or similar exclusionary practices otherwise) in order to help prevent entry in the core market. We further discuss incentives to invest and innovate in the adjacent market in section 3.6.

Before going through economic analysis of foreclosure, it is useful to briefly recall the legal status of the various practices considered.

Tying has been the focus of a number of high-profile anti-trust cases on both sides of the Atlantic. For example, in its Statement of Objections of December 1980, the European Commission alleged that IBM abused of its dominant position in CPUs for large mainframe computers (good *A*) by tying other elements (good *B*) such as main memory or basic software. When competition in the latter markets became more intense, IBM first reduced the prices of its own products and then started to bundle them with CPUs. Another famous case of tie-in in Europe is the *Tetrapak* case (1994) where cartons were tied to the filling machines.

Tie-ins have for a long time been per se illegal in the US. Famous cases of tie-ins include *International Salt* (1947), in which the producer of salt dispenser equipment bundled salt with the equipment, and *Chicken Delight* (1971), in which the franchiser tied various inputs (ingredients, cooking equipment) with the franchising contract. Tie-ins are still currently *per se* illegal under certain conditions, although determining whether the conditions for a *per se* violation have been established frequently involves analyses of market conditions and effects.[20]

Tie-ins have also surfaced in the context of "after-markets." The *Renault* (1988) and *Volvo* (1988) cases involved a refusal by the car manufacturers to sell spare parts to independent repair shops.[21] The Court assessed that there was no abuse of dominant position. In the US, the standard case is *Kodak* (1992), in which the company refused to sell spare parts to Independent Service Organizations.

The leverage doctrine that inspired these cases was sharply criticized by the Chicago School in the 1970s. The Chicago School argued that when two goods are complementary,

---

[20] The decision of the Supreme Court in *Jefferson Parish* (1984) continued to treat tie-ins as *per se* illegal but the Court's review of the conditions that must be established for a *per se* violation increases the emphasis on market analysis, an analysis that has been prominent in a number of subsequent cases including the Appeals Court decision in the *Microsoft* (2001) case.

[21] See also Prime Computer, which involved tied maintenance.

there is really only one profit to be gained and so the idea of reaping a second profit is absurd. There are two versions of the Chicago School critique, depending on whether the monopolized good is an input or else is sold on a stand-alone basis:

- *Bottleneck good is an input*: The Chicago critique is most easily grasped when *A* is an input and is combined by *B*-producers to provide a single service to the final consumer. There is then a single final market and therefore only one profit to be reaped.[22] Foreclosure could at best influence the division of profit between different stages of production – but when the *B*-segment is competitive the monopolist already captures the full profit.

- *Bottleneck good is sold on a stand-alone basis*: When good *A* is sold separately, there are indeed two markets and two profits. Yet the leverage doctrine is not crystal-clear either in that case: If *M* forecloses the access to the producers of complementary good *B* and thereby reduces competition in market *B*, good *A* becomes less desirable to the consumer. To illustrate this, suppose that good *A* provides surplus 10 to the consumer, that *M*'s own version of the *B* good delivers an additional 5 to the consumer while several independent *B* producers produce a *B*-good yielding surplus 10 to the consumer. To simplify, suppose that marginal costs of production are equal to zero and that consumers want one unit of each good. If it forecloses access to its rivals, *M* can either sell good *A* at price 10 and good *B* at price 5 (being then a monopolist on that market), or the combination *A*-*B* at price 15; both options result in a profit of 15. In contrast, in the absence of foreclosure, competition among the independent *B* producers delivers an additional surplus of 10 to consumers who purchase good *A* (recall that good *B* is useless unless combined with good *A*); therefore, *M* can increase the price it charges for good *A* from 10 to up to 10+10=20 and realize a profit of 20. In other words, *M* loses from foreclosing access and becoming a *B*-monopolist.[23] The point is that any additional surplus provided by *B*-competitors

[22] For instance, Posner and Easterbrook in their antitrust treatise (1981, p870) argue that "there is only one monopoly profit to be made in a chain of production". See also the writings of Posner (1976), and Bork (1978), who argues that a "vertical merger does not create or increase the firm's power to restrict output. The ability to restrict output depends on the share of the market occupied by the firm. Horizontal mergers increase market share, but vertical mergers do not" (p231).

[23] When *M*'s version of the *B* good dominates its rivals', it does not matter whether *M* forecloses access or not.

increase consumers' valuation of good *A*, which *M* can then extract (at least partially) by exerting its market power on that segment.

3.3 When leverage is profitable

The circumstances under which tying can indeed be used profitably to extend market power from *A* to *B* are rather different. They have been analysed in Whinston (1990)'s classic paper, which considers two goods that could be sold separately. His idea is best illustrated in the case in which goods *A* and *B* are independent. Maintain our assumption that *M* is a monopoly in market *A*, but suppose that it faces potential competition in market *B*, from a rival who has not yet incurred a fixed cost of entry.[24] Whinston's key insight is that tying the two goods makes *M* more aggressive *de facto* in market *B*, and thus contributes to convincing the rival to stay out of the market. Of course *M* is not properly speaking in market *B* if it ties the two goods. But one way to think about this situation consists in interpreting the sale of the bundle as a situation in which *M* sells good *B*. For example, suppose that each good costs 4 per unit to produce and that good *A* brings surplus 6 to consumers; so if good *A* were sold separately, *M* would make profit 2 per consumer on the *A* segment. Suppose now that *M* sells the two goods bundled at price 11, say. This in fact amounts to selling *M*'s version of good *B* at price 11-6=5. The fictitious margin on good *B* however is not 1=5-4, but 3=5-(6-4). That is, the margin on good *B* should not be computed using *B*'s marginal cost of production, but rather *B*'s opportunity cost of production, which reflects the fact that *M* loses a sale on *A* (with value 2) every time it loses a sale of *B*. This generates a more aggressive behaviour by *M* and reduces the entrant's profit. A tie-in may therefore discourage entry and increase *M*'s overall profit.

To sum up, in this example the conditions for a tie-in to be profitable are: a) *M* must commit itself to a tie-in (for example through a technological choice implying incompatibility of *A* with competitive *B* versions), b) the strategy deters entry (or induces exit) of competitors in market *B*, and c) goods *A* and *B* are independent. The necessity of a) is obvious. That of b) has been demonstrated by Matutes and Regibeau (1988) and can be understood from the

[24] The theory obviously applies equally well to the case of a rival who decides whether to exit and must pay a fixed cost for operating in the market.

reasoning above. A tie-in is self-defeating if the competitor stays in the market, because it increases the intensity of price competition. Last, point c) is demonstrated by Whinston. When goods *A* and *B* become more complementary, the exit of competitors in market *B* mutilates good *A* (which it did not do under independent demands). Thus anticompetitive tie-ins are less likely for very complementary segments.

There is an alternative motive for tie-ins that is worth discussing even though it strictly speaking applies to oligopolistic rather than monopolistic upstream markets, and is used to weaken price competition rather than to drive competitors out. This motive applies to markets in which the goods are or may be purchased sequentially (examples include razors and blades, mobile telephones and accessories such as car chargers, new cars and spare parts, and computers and component upgrades). In many such markets there is fairly intense brand competition. Consumers are easily and plentifully informed about relative prices of the original purchase and display considerable price-sensitivity. By contrast there is often little commitment as to the future cost of additional equipment – consequently manufacturers charge lower mark-ups on the original purchase and higher mark-ups on the additional equipment, a strategy that requires incompatibility between the original equipment and the additional equipment of rival manufacturers in order to weaken price competition in the subsequent market. For example, purchasers of an automobile are usually aware of the fact that purchases of spare parts in the future will be costly. In these contexts the tie-in can be seen as a form of endogenous switching cost (see Klemperer, 1995). The welfare implications are complex, since the weakened ex post competition in the additional equipment market triggers stronger competition in the price-sensitive original purchase market; while the price structure may become distorted, the overall impact on price levels may be good for consumers and the society as a whole, making it difficult to identify in practice the particular cases in which anti-trust intervention might be desirable. In addition, this strategy is not really applicable to cases in which the original equipment market is already monopolised, since in this case the monopolist has no need to weaken price competition in that market.[25]

[25] Actually, the work of Farrell-Gallini (1988) and Shepard (1987) suggests that the monopolist may be better off committing to low prices though the absence of tie-in.

3.4 Foreclosure to protect the monopolised market

We now consider the second set of circumstances that may be favourable to foreclosure, namely those in which the monopoly good has no stand-alone value but is used purely as an input. A recent line of theoretical research (associated with Hart-Tirole 1990, O'Brien-Schaffer 1992, McAfee-Schwartz 1994 and Rey-Tirole 1997) provides a synthesis of the old leverage doctrine and the Chicago critique, and allows a reasonably sophisticated analysis of the impact of foreclosure practices on consumer surplus and social welfare. Recall that the case in which the monopolized good is an input includes such important instances as the bottleneck in the network industries, computer reservation systems, and the licensing of intellectual property.

The Chicago School is right in pointing out that there is then a single profit and therefore the idea that a monopolist in one market wants to *extend* its market power to an adjacent segment makes little sense as it stands. The new synthesis argues that the incentive for foreclosure is not to extend, but to *restore* market power in the monopoly segment. It therefore applies to circumstances where, for one reason or another, the monopolist's capacity to extract profit in its original market is insecure.

The basic idea behind the new theory is straightforward. Suppose that there are 10 final consumers. Each consumer can consume 0 or 1 unit of the combined good and production of *A* and *B* is costless; the *n*th consumer is willing to pay *n* €. The industry's monopoly price is then equal to 5 €, leading to sales to 5 consumers and profit equal to 5x5 = 25 €.[26]

Suppose first that *M* is not in market *B* and that this market is supplied by a competitive industry. To obtain the monopoly profit of 25 €, *M* may try to pick a *B*-supplier and sell 5 units at a unit price of 5 €. Once it has collected 5x5 = 25 €, though, *M* will be tempted to go to another *B*-supplier and sell additional units to this supplier. For example, even if it has observed the first transaction and thus knows that 5 units will already be "on the market," this

[26] This is the price that maximizes the total profit; for example, reducing the price down to 4 generates an additional sale but reduces profits because of the reduced margin (4x6=24<25); likewise an increase in the price, from 5 to 6, say, increases the margin but reduces both the sales (4 consumers at most are willing to buy at a price of 6) and the profit (6x4=24<25).

other supplier would perfectly be willing to buy 2 additional units at a price of 3 € (which is the market price for a total number of 5+2 = 7 units). Having done this, *M* will turn to a third *B*-producer (who would be willing to buy an additional unit at a price of 2 €), and so on. When dealing the input *A* to new *B*-producers, *M* exerts a negative externality on *B*-producers who have previously purchased the input *A*, because this intensifies competition in the downstream segment; in a sense, *M* expropriates those producers' previous investment in input *A*. However, the ability to "flood the market" with new units of *A* backfires. Anticipating this incentive, no *B* -firm will want to pay more than 1 per unit of *A*, resulting in total profit 10x1=10 € for *M*. With a continuous demand curve, or if *B*-producers do not purchase sequentially and do not observe other purchases, *M*'s profit is actually even smaller and converges to zero when the number of *B*-producers becomes large. This phenomenon is essentially an atemporal version of Coase's 1971 conjecture about the lack of market power of a monopolist producing a perfectly durable good.

This erosion-of-profit phenomenon has been understood by market participants, who have learned to restore their upstream market power in several ways. Let us just mention three frequently used strategies.[27] One consists in *M*'s tying its hands by offering an exclusive contract to an independent *B*-supplier (or, more generally, a contract which specifies the number of other suppliers or franchisees within a given area). Another has *M* integrating vertically and refusing to supply to independent *B*-suppliers. The *A*-division then internalises the negative externality imposed on the *B*-division when it sells to independent suppliers. In the example above, both strategies enable *M* to obtain the monopoly profit (25 €). A third strategy consists in developing a reputation for not acting opportunistically. This reputation may lead to one of several economically equivalent outcomes: a *de facto* exclusive license in which the monopolist refrains from supplying to the licensee's competitors although he has no written of formal obligation to do so; or the absence of discrimination under which the monopoly keeps selling at the same monopoly price to all downstream producers. Interestingly, while economically equivalent, the former outcome would probably be considered as resulting from a foreclosure strategy while the second would not.[28]

---

[27] See Rey-Tirole (1997) for a more comprehensive list of strategies.

[28] There is another interesting difference between the two. The reputation for abiding by a *de facto* exclusive arrangement may be easier to develop as the act of supplying a competitor may be widely observable. In contrast, in intermediate good markets, unit prices may not be transparent (unless the monopolist sets up a

More generally, a vertically integrated *M* may not refuse to supply, but limit instead the sales of its input *A* to independent *B*-producers. There are two reasons for this:

- First, independent producers may create value in that they offer either a higher quality or lower cost *B*-good than *M*'s *B*-division, or they offer a differentiated product. Foreclosure is then costly to *M* because it limits the implicit price at which *A* can be sold. This idea can be applied to cases in which there are several *B*-markets: *B1, B2, B3,...* Provided that it can prevent resale, *M* may then want to sell its input *A* to some independent companies (those in the *B*-markets that create value through vertical or horizontal differentiation) and not to others.

- Second, *M* may not be a pure monopolist upstream. It may face a competitive threat. To the extent that the alternative and inferior producers of input *A* stand ready to supply the input to independent *B*-producers, both profit maximization and social welfare command that the most efficient producer of input *A* (that is, *M*) supplies the input to the independent producers of good *B*. In this case, it can be shown that *M* restricts, but does not completely foreclose access to its input by the independent *B*-producers.[29] Note that in that case, vertical integration is more efficient than alternative profit-preserving means such as exclusive dealing: resorting to exclusive dealing would completely foreclose efficient *B*-producers and reduce consumer choices.

These considerations suggest that the task facing competition authorities investigating alleged foreclosure is a particularly delicate one. Even if some downstream firms complain about restricted supply on the part of the upstream firm, this might represent either foreclosure or the by-product of an entirely legitimate selection of *other* downstream firms whose quality or product differentiation enhance the value of the overall product. Telling these different circumstances apart may not be easy.

---

surveillance and disclosure apparatus) and so the reputation for never lowering the price may be hard to develop.

[29] For more details, see Hart-Tirole (1990) and Rey-Tirole (1997).

3.5 Foreclosure to reduce the threat of entry in the monopolised market

The chief insight of the theory we have just described, namely that foreclosure may be a means of restoring the firm's original monopoly power rather than extending it into an adjacent market, is not directly applicable to the case where the two goods are sold separately. This is because the separate sale of the upstream good itself acts as a commitment device against the erosion of profit. Nevertheless, the argument has an interesting dynamic analogue in the case where the two goods are sold separately. While it remains true that the monopolist would suffer at any one point in time from the exit of efficient producers of good *B*, this might make it easier to protect the market for good *A* against *future* threats. Carlton and Waldman (2000) argue that *M* may tie goods *A* and *B* not so much to make more profit on market *B* but to deter future entry in market *A*. In their model, *M* is initially (in period 1) the unique producer of good *A* and faces competition in market *B*. Later on (in period 2) an alternative supplier of good *A* may appear and decide whether to enter. Goods *A* and *B* are complements.

From our earlier arguments, we know that tying *A* and *B* results in less profit in period 1 for *M*. But suppose that the tie-in induces the competitive supplier to exit at the end of date 1 (the arrival of the alternative *A*-supplier is too uncertain). Then at date 2, the alternative supplier, if there is one, may not want to enter because it faces *M*'s *B*-division as its sole complementary good producer. Or, to put it differently, the date-2 entrant may need to enter both markets at the same time in order to compete effectively with *M*. This may prove too costly.[30]

The Carlton-Waldman theory is interesting, but its validity requires a number of assumptions.

- First, the entry in the *A* segment must be sufficiently uncertain that the independent *B*-vendor does not want to hold firm until its arrival.

[30] See also Choi & Stefanadis (2001) for a development of this argument. A recent paper by Fiumagalli & Motta (2000) makes a related point in relation to "buyer power" – coordination among buyers may enable entry to take place in circumstances where too much competition among them would keep the entrant out.

- Second, it may be the case that innovating in adjacent segments is not that costly.

- Third, the model does not fare as well for "asymmetric complementarities." Suppose that *A* is indispensable for operating *B*, but *A* has a life on its own. Then the profitability of entry in the *A* segment is much less affected by the monopolization of segment *B*.

- Fourth, and quite importantly, the model is not fully dynamic. If dominance in one market and the strategic use of the simultaneous-entry problem lead to high incumbent profits, then there is a high incentive to become the new incumbent. The alternative *A*-producer may therefore decide to enter, even if in the "short run" (period 2) the monopolization of the *B*-segment by *M* makes entry unprofitable. It is therefore important to add periods 3,4,... We return to more dynamic aspects of competition at the end of section 4.

3.6 Innovation by the monopoly firm in the competitive segment

One way in which foreclosure has sometimes been feared is through the use of innovation in market *B*, the effect of which might be to discourage competitive efforts (including innovation) by rival *B*-producers. On the face of it, this fear seems incompatible with our normal intuitions that innovation is desirable, that competition should apply to the innovation process as well as to normal processes of production, and that the results of this innovation should be given intellectual property protection. Forcing *M* to share its innovation with its *B*-competitors creates an asymmetry in the competitive process and is undesirable. First, sharing induces the independent *B*-suppliers to free ride, reducing their R&D effort and probably product diversity. Second, and more importantly, access policies of this type would imply a de facto line-of business restriction, as *M* would stop engaging in innovations that would be competed away. Both arguments advocate protecting *M*'s rights over its innovation.

While this simple analysis is broadly correct, there is a twist, though, that has been analysed by Choi (1996, 1998) and Farrell-Katz (2000): The R&D competition in market *B* is affected by the presence of one of the competitors, *M*, in the adjacent market *A*. Suppose for the sake of argument that *A* is sold on a stand-alone basis (a similar analysis applies to the case in which *M* produces an input that is then used internally by its *B*-division or sold to

independent producers of good *B*). Then the value of *A* is higher, the better the quality over price ratio of the product offered (by *M* or its competitors) in segment *B*. This implies that *M* benefits from innovation in market *B* in two ways: directly through sales of the component *B* if *M*'s innovation in the *B* market is superior to those of its rivals; and indirectly through the increase in demand for good *A*.

To fix ideas, suppose for example that all *B*-competitors produce the same good (no differentiation) and that innovations reduce production costs in that market. The indirect or spillover effect is then clearly identified when *M*'s innovation in market *B* is dominated by a rival's innovation, in which case *M* makes no profit in the *B*-segment. In that case, a small increase in the quality of *M*'s innovation in market *B* still leaves it dominated and thus does not generate any profit to *M*'s *B*-division. Yet it increases *M*'s profit if it forces the efficient *B*-rival to lower its price (squeezing quasi-rents from the independent *B*-producer), and thereby boosts the demand for good *A*. The indirect effect takes a different form when *M*'s innovation in the *B*-segment dominates its rivals'. Then a marginal increase in the quality of *M*'s innovation increases the demand for *M*'s integrated solution (this is an example of the vertical externality effect identified in section 2).

The spillover effect implies that *M*'s R&D efforts in segment *B* are higher than they would be if *M*'s R&D division did not internalise the profit in this *A* segment. In turn, *M*'s enhanced R&D effort reduces that of its rivals. As Farrell and Katz (2000) show, the overall welfare impact of *M*'s *B*-division internalising *M*'s *A*-division's interest is ambiguous.

As usual, we should clarify the nature of the policy intervention that is being contemplated. Short of imposing structural remedies, no antitrust decision will prevent *M*'s *B*-division from internalising the *A*-division's interests; hence, the above analysis may seem irrelevant as it stands. This analysis would however apply stricto sensu to the following situation: Suppose that the antitrust authorities mandate the sharing of *M*'s innovation in market *B*. As we already noted, this would basically induce *M* to quit that market. If *M* is replaced in market *B* by another firm with equivalent R&D expertise,[31] then the Farrell-Katz analysis and its conclusion that there are social costs and benefits to *M* being in the *B*-

[31] The new firm could be *M*'s divested *B*-division.

segment apply –but if *M* is not replaced, then the situation is more likely to be detrimental to R&D.

The ambiguity of the welfare analysis suggest antitrust involvement is unlikely to improve matters unless it is possible to demonstrate that a) the reduction in the independent *B*-producers' R&D effort due to *M* being vertically integrated offsets the increase in that of *M*, and b) *M*'s *B*-division can be effectively duplicated by entry in the *B*-market (e.g., through an effective divestiture and in the absence of economies of scope). It is therefore not surprising that antitrust authorities have traditionally shunned direct intervention in the competitive market.[32]

Even if an analysis of this kind were used in a particular case as the basis for anti-trust intervention, the resulting intervention would run counter to the tradition of intellectual property law. That tradition seeks to resolve the tension between the benefits of competition and the protection of innovation by protecting the innovation from direct imitation, while encouraging rival innovations. Indeed (as the analysis above makes clear), while the quality of the best innovation determines the gross benefits to consumers who purchase it, the price at which they buy (and therefore the *net* benefits of the purchase) is determined by the quality of the *second-best* innovation (this is the same phenomenon as the fact that the price paid by the winner in an auction is determined by the valuation of the second-highest bidder). Consequently innovation by rivals plays an important part in the process of keeping prices low, one that IP law has consistently sought to protect. In contrast, intervention to restrict innovation by M in the *B*-segment would essentially consist of removing one firm's IP protection in order to protect the innovation of another firm from post-innovation rivalry.

These issues become of particular importance once we consider the impact on competition in adjacent markets of innovation in the original monopolized market. This is the subject of section 4.

[32] When they do intervene in that market through a forced divestiture and a line-of-business restriction to prevent reintegration, the motivation given is the failure of effective access policies and relates to the analysis of section 4, and not to the analysis of this section.

4. Market power and innovation in the monopolized market.

4.1 Access regulation and incentives to innovate

Our discussion so far has emphasised that the traditional leverage theory of foreclosure – namely, that a monopoly firm might seek to extend its market power into competitive adjacent markets – ignores the fact that the monopoly firm benefits from the presence of low-price high-quality complements (whether these are sold separately or are part of the downstream production process), provided it can extract the rent from this through its monopoly pricing in the original market. However, besides the commitment problem identified in section 3.4 above, there are a number of circumstances in which the firm may be prevented from using its monopoly pricing in this way. For example, the monopoly good may be a telecommunications network subject to access pricing regulation. Under these conditions the monopoly firm may seek instead to recoup some of the lost opportunities for exercising market power through its strategies in adjacent markets. This implies that any public policy to restrict monopoly power in the original market needs to be implemented in conjunction with attention to the risks of foreclosure in the adjacent markets.

To simplify the exposition let us assume that *M* is present in the market *B* and forecloses access to input *A* to independent *B*-producers, and that none of the efficiency gains to vertical integration discussed in section 2 is present. Should access to the bottleneck be forced upon *M*? The theoretical answer to this question is rather straightforward, and involves a *basic trade-off between ex ante incentives and ex post competition*.

Taking the existence and specifications of good *A* as given, a policy of open access to *A* increases competition in the downstream segment and lowers the price to the consumers. Thus forced access increases consumer surplus and social welfare.

But open access also reduces *M*'s incentives to develop or enhance the input *A*. In the spectrum of access policies, the two polar cases are commercially determined access (laissez-faire), and regulated free access (the input *A* is freely shared by *B*-competitors). The commercially determined access policy allows *M* to reap the maximum fruits of its investment while free access leaves no profit on the table. In other words – an obvious, but

sometimes forgotten point – *the regulation of access is really a regulation of the rate of return on the bottleneck segment*. Maximal incentives to invest/innovate in segment *A* are provided by the non-intervention policy, and minimal incentives by forced sharing. Thus the choice of an optimal regulated access policy involves the same kind of considerations as the choice between price caps, earnings sharing schemes, and cost of service in utility regulation: it is really about the determination of the incentives that are required to bring about a proper level of investment and innovation.

While the theoretical framework is recent, its underlying ideas have been around and in part understood by market participants, courts and policymakers for centuries. Much of its philosophy is embodied in patent law. The desirability of avoiding restrictions on the contractual freedom of the innovating firm is illustrated by the following example. Suppose a biotechnology start-up develops a new compound, for which it obtains a patent. It licenses its patent for a lump sum fee to a pharmaceutical company without granting exclusivity. The biotechnology start-up will subsequently be tempted to license its innovation to a second pharmaceutical, and then a third and son on. Since drug production involves a very low marginal cost,[33] pharmaceutical companies will quickly see the price of their perfect substitutes driven to near zero. They will therefore not be willing to pay anything for a non-exclusive license. Patent law has long recognized this point and has granted patentees the right to license their innovation to whomever they want at mutually agreed upon terms. The same approach applies to franchising as well, as a franchiser must guard itself against adding more and more franchisees in the same neighbourhood or city.[34]

Patent law thus leaves the innovator free to licence its innovation and protects this innovation during a given time window. More generally, the optimal openness of access to a bottleneck depends on the incentive needed to induce investment/innovation in the bottleneck.[35] As a general presumption, the more the monopoly in market *A* is the result of

[33] The reasoning does not depend on the marginal cost being small, though.

[34] This is recognised by franchise law in the US and by a specific Block Exemption Regulation in the EU.

[35] This point was clearly stated by the Advocate General (§57 and §62) in *Oscar Bronner* (1998).

innovation the more important it is to preserve the commercial freedom of negotiated access, as is recognized by patent and franchise law[36].

Circumstances that militate in favour of relatively open access (though not free access) can therefore be set out as follows:

- The bottleneck position is the outcome of a legal monopoly (as is the case for example for an airport). The monopoly position has then been acquired independently of investment or innovation.[37] In *Oscar Bronner* (1998), the Advocate General argues in §66 that situations in which the monopoly good has been created in a non-competitive context, perhaps even with public subsidies, should be treated differently.

- The monopoly position is due more to economies of scale than to investment/innovation. For example, there may be limited scope for multiple wholesale markets in a medium size city.[38]

- Market power is due primarily to the existence of network externalities and lack of compatibility/interconnection rather than to investment/innovation. This applies in particular to telecommunications networks (e.g., AT&T at the beginning of the century;

[36] One exception to this concerns circumstances in which the diffusion of the monopoly good is likely to lead to many other small incremental innovations that are individually difficult to patent or otherwise profit from but are cumulatively very important. Then the transactions costs for the monopolist of obtaining sufficient license income from its patent may make it unwilling to diffuse the technology to anything like a socially optimal degree. A pertinent historical example may be Richard Arkwright's technology for spinning with the use of rollers, patented in 1775. As Hall (1998) writes, "once the Arkwright patent fell, its combination with Crompton's machine meant that every one who had the slightest talent for constructiveness or appropriativeness fell to work to improve this process", a cumulation of innovation that Arkwright had not been able or willing to encourage while the patent was still in force (p. 332).

[37] Similarly, and while for other reasons we disagree with the many calls made in France for the regulation of access to shelf space in supermarkets, a coherent case could be made that the 1993 and 1996 French Royer and Raffarin laws *de facto* prohibiting the construction of new supermarkets have created a legal scarcity of shelf space.

[38] As was recognized in *Gamco* (1952).

see also the European Commission's concerns about the Internet mergers involving WorldCom and MCI, and then MCIWorldCom and Sprint).[39]

4.2 Intellectual property law and the essential facilities doctrine

In practice, however, the policy of competition authorities has not always been consistent with these principles of patent law. The European Commission in 1998 defined an approach to the treatment of bottleneck goods (known as "essential facilities") that is similar to that adopted in the US in the case of AT&T vs. MCI (1983). An essential facility is an input that

- is controlled by a dominant firm,
- is essential in the sense that duplicating it would be seriously uneconomic (just "less costly" than an alternative does not suffice), and
- for which the dominant firm has no objective reason to deny access (no capacity constraint, no need to disclose information covered by intellectual property rights, no fundamental technological problem of providing interconnection).

The third condition can be relaxed to require that a bottleneck owner take steps not to block a potential new service in an adjacent market (such as broadband access for an incumbent telecom operator that does not provide such services itself).

After concluding the existence of an essential facility, courts can mandate that access be given on "reasonable and non-discriminatory" terms. Similarly, the European Union's access notice states that when the access is essential, the requesting party must pay a reasonable price and abide by non-discriminatory terms.

[39] However, when the scope for network externalities is endogenous (e.g., is affected by penetration pricing or by the favourable treatment or subsidisation of a complementary segment – e.g., developers), there may be a need to encourage choices leading to larger positive externalities.

Of course, everything depends on the interpretation of what a "reasonable" price means. A high price amounts to commercially-determined access and a low price to forced access. Some observers have therefore wondered whether we really need a separate concept of essential facility in antitrust enforcement. After all, the list of abuses of dominant position drawn in Article 82 is not exhaustive. And, as is the case with the market definition exercise, the definition of an essential facility poorly reflects the economic effects (here, of regulating access). In our view, the definition of an essential facility serves the same purpose as market definition, namely to screen cases so as to alleviate the burden on competition authorities and courts. For example, cases in which the input can be duplicated at a reasonable cost or in which there is no dominant position, as in the 1972 *Topco* case in the US,[40] can be readily dismissed. Once this screening function is performed, one should turn to the more serious economic effects analysis.

A quick look at the criteria defining an essential facility shows the tension between this criterion and that embodied in patent law. A piece of intellectual property protected by a patent may well be an essential facility according to these criteria, indeed the more so the more important the innovation is! Yet the European Court of Justice, its American counterparts or the antitrust authorities would usually not consider applying the essential facility doctrine to material protected by intellectual property rights.[41]

The temptation is strong to argue that intellectual property is somehow different from other goods and services. But consider the following example: suppose that a new telecommunications operator enters the market deploying fibre-optic cables widely to consumers' homes. In the near future, this fibre to the home will be very much like intellectual property in its cost structure: a large upfront sunk cost and virtually no marginal cost as there will be excess bandwidth from the curb to the home. Is the fibre-optic-to-the-home network different from intellectual property? One should consider whether the fibre optic deployment will occur unless its owner is allowed to commercially negotiate access. This type of reasoning is quite similar to that employed to determine whether an innovation can be patented: The Patent Office will grant a patent to an inventor only if the innovation is

[40] The Supreme Court struck down the territorial restrictions (exclusive territories assigned to independent supermarket chains) as *per se* violations in *Topco*; the chains averaged 6 percent market shares in their regions.

[41] Compulsory licensing is an exception to the rule.

minimally useful and if it is novel. The innovation must be non-obvious and cannot represent a trivial step forward in the art. The Patent Office therefore considers the level of incentives that are needed to bring about the innovation. While its set of possible policies is somewhat limited (basically, to deny or grant the patent application), whereas access regulation is more continuous (the access charge can take a continuum of values), the broad principles of economic analysis are similar in the two domains.

Let us now come to the most difficult question of all: how do we determine a reasonable access charge?

4.3 What is a reasonable access charge?

Economists now widely agree on the pricing principles that should guide industry supervisors,[42] and the only remaining debates relate to the (least imperfect) implementation of these principles – in that respect, we shall later emphasise that access regulated by a competition authority is not the same as access regulated by a regulatory agency These principles can be stated as follows:

- First, access charges should reflect the marginal cost for the bottleneck to give access.

- Second, access charges should include mark-ups so as to allow the bottleneck owner to recoup its fixed costs of investment.

- Third, cost recovery should be efficient. That is, it should minimize the distortion on the economy. Thus mark-ups should be lower on those segments where demand is highly elastic and therefore consumption would be severely distorted by a high price. Indeed, access charges should generate retail prices that obey the standard commercial principles: the so-called Ramsey mark-ups that minimise economic distortions are inversely related to demand elasticities.

[42]See Laffont-Tirole (1999) for an exposition of these principles and of policy issues related to access in the context of regulated industries.

In addition, it should be stressed that access charges that are purely marginal cost-based (and therefore open access at a low price) are inefficient for several reasons. First, they don't allow fixed cost recovery and therefore discourage initial investment. This problem is particularly serious in the presence of a large fixed cost, or more generally when there are important returns to scale (which is often the reason for the existence of a bottleneck in the first place). Returning to the hypothetical fibre-to-the-home example, the marginal cost of giving access to bandwidth is equal to zero, and an access charge based purely on marginal cost would never allow the operator to recoup its billion-Euro investment. The same applies to biotechnology or software investments, and is particularly serious in the presence of uncertainty, since charges that appear to allow a reasonable rate of return when a project is successful may not adequately compensate for the risk of failure[43]. Second, by preventing the bottleneck owner from profiting from the bottleneck segment, the cost-based policy strongly encourages this owner to foreclose access through non-price methods, and thus creates a need for a heavy-handed, intrusive regulation. For example, the implementation of the European Commission's directive mandating unbundling of the "local loop" between customer premises and the local exchange is being made very difficult by disputes between the loop owners and the would-be entrants as to what exactly constitutes "access" (can the employees of entrants have unrestricted right of entry to exchange premises, which security measures should then apply, and so on). And this is a context in which the definition of access would appear to be relatively straightforward compared with many areas of intellectual property.

Access charges that include a mark-up to address cost recovery but are non-discriminatory are also unsatisfactory. Consider the case of an airline company offering service between Toulouse and Nantes by charging 2000FF to businessmen and 500FF to students. Probably no Toulouse-Nantes service would exist if the airline were forced to refrain from price discrimination, or equivalently if the airline had to sell its seats to marketers at a non-discriminatory price. Similarly, the hypothetical fibre-optic operator may not want to install fibre-to-the-home unless it is allowed to charge per unit of bandwidth a higher access charge for premium services.

---

[43] This is offset to some extent by limited liability, as well as by the flexibility of regulatory authorities in practice in the face of adverse commercial conditions. Agreements in some European countries to allow competitors to collaborate to share costs in third-generation mobile telephony since the downward revision of market forecasts are a case in point.

A difficult question in regulating access is that of the "granularity" of access, that is the level at which access should be provided. This issue, which involves a similar trade-off between ex ante incentives and ex post competition, has been discussed at length in the context of local loop unbundling. A more detailed unbundling catalogue is good for competition because it allows competitors with very heterogeneous needs to have access to exactly what they need. However, a wide range of unbundled options reduces the profit the telecom incumbent can obtain by investing in, maintaining and upgrading its network. A larger catalogue also increases transaction costs (preparing for collocation) and broadens the scope for "regulatory arbitrage," in which the competitors buy access to those elements that have been mis-priced downwards by the regulator and supply those that have been mis-priced upwards.

4.4 The institutions of access regulation

While the theoretical principles are identical for both forms of industry oversight, the actual regulation of access prices and conditions is likely to differ, and is probably more arduous in the case of regulation by a competition authority.

First, competition authorities usually lack the large-size staff and continual involvement in the industry enjoyed by regulatory bodies. The frequent lack of reliable data explains why competition policy is often less at ease with cases requiring quantitative evidence (tacit collusion, predatory conduct access pricing) than with cases that don't (overt collusion, per se illegal restraints). And access pricing is an eminently quantitative issue.

Second, the scope of regulation by a regulatory agency is usually broader than just access pricing. For example, it may be able to regulate adjacent segments (even though the trend is toward the deregulation of potentially competitive segments). For example, the possibility of imposing caps on the vertically integrated firm's competitive activity reduces the incentive for non-price foreclosure.

Third, and partly endogenously, regulatory agencies tend to regulate relatively stable products. The regulation of access charges is facilitated when the definition of the bottleneck

good moves slowly over time. Fast technological progress creates serious issues about the ex ante regulation of access terms, which is why many observers feel that regulation will soon give way to antitrust enforcement in the telecommunications industry.

In fast-changing industries the fact that innovations succeed each other has an important impact on the incentive issues with which regulation must deal, as we discuss briefly in section 4.5.

4.5 Dynamic contestability

Once we take into account the dynamic nature of innovation, the considerations shaping access regulation can become very complex and call for a very careful analysis. In particular, industries with rapid technological progress are characterized by an interesting topsy-turvy principle: the higher the barriers to entry, the higher the incentives to enter! If barriers to entry are high, incumbents are highly profitable, and so everyone wants to be the next incumbent. In short, barriers to entry encourage entry.

In such a context, developing an appropriate analysis may be a delicate task. An illustration can be found in Fudenberg-Tirole (2000),[44] who analyse barriers to entry resulting from network externalities and incompatibility between the incumbent and the entrant's standards. They consider a multi-period framework where in each period a new competitive threat arises. Entrants face large fixed costs of entry on top of the handicap of not having an installed base and therefore of offering fewer network externalities to their consumers. But by entering an entrant may dislodge the incumbent and become the new incumbent.

The incumbents build on network externalities to deter entry. In each period, and despite having full monopoly power in the short run, the incumbent charges low prices in order to broaden its installed base and make entry more difficult.[45] Entry thus occurs only when the entrant has a sufficiently superior technology and/or faces a low-entry cost.

---

[44] Whose ideas were anticipated (although less formally expressed) in Reddy et al (1998).

[45] Other and similar strategies could include evangelization or improvements in quality.

The welfare analysis of these entry-deterrence strategies has interesting implications. The key point is that the threat of entry has a positive impact, since it keeps the incumbent on his toes, while actual entry may have either positive or negative effects. Indeed, a marginal increase in the probability of entry (say because of a small subsidy to entry) may reduce consumer surplus or social welfare because of two effects:

- first, entry may be motivated by the usual "business stealing effect:" the entrant just wants to become the next incumbent and, as a result, can have excessive incentives to enter compared with its social desirability (that is, entry may be good for the new competitor but not for the consumers);

- second, entry leaves the incumbent's installed base of consumers stranded: because of the network externalities, the coexistence of multiple standards reduces consumer surplus and social welfare.

These already intricate welfare implications are further blurred when the entrant's innovation spills over[46] to some degree to the rest of the industry. This analysis however clearly stresses the need carefully to develop an appropriate dynamic framework when looking at rapidly evolving industries.

5. Concluding remarks

We began our discussion by pointing out that the concern for the possible anti-competitive consequences of action by monopoly firms in adjacent markets needs to take into account two important considerations:

- first, behaviour in adjacent markets *is* different from the direct exercise of monopoly power; and
- secondly, involvement in multiple markets *is* a normal and usually innocent form of industrial activity that frequently creates real value for consumers; for example, using the knowledge and expertise acquired in one market to enter another market, or taking

[46] An entrant's technological advantage that is kept proprietary should not be accounted for in the welfare analysis because it is internalized by the entrant in its decision of whether to enter.

advantage of synergies and economies of scope between the production in the two markets, is both profitable for the firm and desirable for consumers and the society as a whole.

The second point applies more particularly to tying. Some phenomena are almost invisible because they are very rare, while others are almost invisible because they are so ubiquitous that we have ceased to notice them. Tying is a phenomenon of the second kind. Tying products together, sometimes literally, in the assembly process is an omnipresent feature of modern life. This is a fact that is often forgotten when particular tying cases are discussed, because tying is made to seem the bizarre exception rather than the utterly familiar rule. In this paper we have tried to draw attention to the pervasiveness of firms' involvement in multiple markets, and the pervasiveness of tying complementary goods together in the assembly process, because it is only by understanding the functions these activities fulfil in the normal course of events that we can properly diagnose the circumstances where they can be used to anti-competitive ends.

The analysis shows that, in some circumstances, an incumbent monopolist may engage in exclusionary practices to foreclose entry. We have pointed out, however, that such attempts to foreclose entry are much more likely to be motivated by the wish to protect and exploit monopoly power in the core market than to extend that monopoly power into adjacent markets. Sometimes the appropriate conclusion is therefore to use access regulation as a means of limiting the original monopoly power, but this raises well-known conflicts when that monopoly power is the result of prior innovation. It can, of course, make sense for the authorities in certain circumstances to seek to make a monopolised good available to downstream users on terms more favourable than the monopolist itself would determine. An obvious example is access to fixed telecom networks, where the bottleneck assets may be less the result of innovation than the accidental heritage of regulatory history. In markets where innovation is an important contributor to consumer surplus and social welfare, any benefits of post-innovation competition need to be set against the costs of discouraging the innovation itself. It is particularly difficult for anti-trust authorities to strike this balance. Furthermore, a firm's competitors may feel threatened by its innovations, and may thus seek to persuade the authorities that such innovations should be prevented or expropriated (an effective way of preventing future innovation). But society at large should be much more welcoming of innovation, and ensure that policies to speed up its diffusion do as little as possible to hinder its creation.

References

Aghion, P., M. Dewatripont and P. Rey (1994), "Renegotiation Design with Unverifiable Information," *Econometrica*, 62(2):257-282

Armstrong, M. (1999), "Price Discrimination by a Many-Product Firm," *Review of Economic Studies*, 66(1):151-168.

Bakos, Y., and Brynjolfson, E. (1999), "Bundling information goods: Pricing, profits and efficiency," *Management Science*.

Carlton, D., and M. Waldman (2000), "The Strategic Use of Tying to Preserve and Create Market Power in Evolving Industries," Working paper NBER DP #6831 (1998, revised 2000).

Choi, J.P. (1996), "Preemptive R&D, Rent Dissipation and the 'Leverage Theory'," *Quarterly Journal of Economics*, 110: 1153-1181.

----- (1998), "Tying and Innovation: A Dynamic Analysis of Tying Arrangements," mimeo, Columbia University.

Choi, J.P., and C. Stefanadis (2001), "Tying, Investment and the Dynamic Leverage Theory," *Rand Journal of Economics*, 32: 52-71.

Farrell, J., and M. Katz (2000), "Innovation, Rent Extraction, and Integration in Systems Markets," *The Journal of Industrial Economics*, 48(4): 413-432.

Fudenberg, D., and J. Tirole (2000), "Pricing a Network Good to Deter Entry," *The Journal of Industrial Economics*, 48(4): 373-390.

Fumagalli, C., and M. Motta (2000), "Buyers' Co-ordination," European University Institute, mimeo, downloadable at http://www.iue.it/Personal/Motta/Buyers6.pdf

Farrell, J., and N. Gallini (1988), "Second-Sourcing As A Commitment: Monoply Incentives To Attract Competition," *The Quarterly Journal of Economics*, 103:673-694.

Grossman, S., and O. Hart (1986), "The Costs and Benefits of Ownership: A Theory of Lateral and Vertical Integration," *Journal of Political Economy*, 94: 691-719.

Hall, P. (1998), *Cities in Civilization*, London, Weidenfeld & Nicolson.

Hart, O., and J. Moore (1994), "A Theory of Debt Based on the Inalienability of Human Capital," *Quarterly Journal of Economics*, 109: 841-880.

Hart, O., and J. Tirole (1990), "Vertical Integration and Market Foreclosure," *Brookings Papers on Economic Activity*, Microeconomics, 205-285.

Klemperer, P. (1995), "Competition when Consumers have Switching Costs," *Review of Economic Studies*, 62(4), 5-39.

Laffont, J.-J., and J. Tirole (1999), *Competition in Telecommunications*, Cambridge: MIT Press.

Matutes, C., and P. Regibeau (1988), "'Mix and Match:' Product Compatibility without Network Externalities," *Rand Journal of Economics*, 19: 221-234.

McAfee, P., and M. Schwartz (1994), "Opportunism in Multilateral Vertical Contracting: Nondiscrimination, Exclusivity, and Uniformity," *American Economic Review*, 84: 210-230.

O'Brien, D. P., and G. Shaffer (1992), "Vertical Control with Bilateral Contracts," Rand *Journal of Economics*, 23(3):299-308.

Posner, R. (1976), *Antitrust Law: An Economic Perspective*, University of Chicago Press.

Posner, R., and F. Easterbrook (1981), *Antitrust Cases, Economic Notes and Other Materials*, second edition. St Paul: West Publishing.

Reddy, B., D. Evans and A. Nichols (1998), "Why Does Microsoft Charge so Little for Windows?," mimeo, NERA.

Rey, P., and J. Tirole (1997), "A Primer on Foreclosure," mimeo, IDEI, Toulouse, forthcoming, Handbook *of Industrial Organization*, ed. by M. Armstrong and R.H. Porter.

Schmalensee, R. (1973), "A Note on the Theory of Vertical Integration," *Journal of Political Economy*, 81: 442-449.

Schwartz, M., and G. J. Werden (1996), "A Quality-Signaling Rationale for Aftermarket Tying," *Antitrust Law Journal*, 64:387-404.

Shapiro, C. (1995), "Aftermarkets and Consumer Welfare: Making Sense of Kodak," *Antitrust Law Journal*, 483-511.

Shepard, A. (1987), "Licensing to Enhance Demand for New Techonologies," *Rand Journal of Economics*, 18:360-368.

Stigler, G. (1963), "A Note on Block Booking," *Supreme Court Review*.

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

Vernon, J., and D. Graham (1974), "Profitability of Monopolization by Vertical Integration," *Journal of Political Economy*, 79: 924-925.

Warren-Boulton, F. (1974), "Vertical Control with Variable Proportions," *Journal of Political Economy*, 82: 783-802.

Whinston, M. (1990), "Tying, Foreclosure, and Exclusion," *American Economic Review*, 80: 837-860.

Williamson, O. (1975), *Markets and Hierarchies: Analysis of Antitrust Implications*. New York: Free Press.

----- (1985), *The Economic Institutions of Capitalism*. New York: Free Press.

List of Cases:

AT&T vs. MCI, 708 F.2d 1081, 1132-33 (7th Cir), cert. Denied, 464 US 891 (1982)

Eurotunnel (1994): Communication de la Commission du 30 juillet 1994 faite conformément à l'article 19, paragraphe 3, du Règlement no. 17/62 du Conseil et à l'article 26 paragraphe 3 du Règlement no. 1017/68 du Conseil, Eurotunnel, JOCE C 210/15 du 30 juillet 1994, point 24.

Frankfurt Airport (1998) IV/34.801 FAG – Flughafen Frankfurt/Main AG

Gamco Inc. V. Providence Fruit & Produce Bldg. Inc. 194 F 2d 484 (1st Cir) cert. Denied, 344 US 817 (1952)

Hecht v. Prof-Football Inc., 570 F 2d 982 (D.C. Cir 1977) cert. Denied, 436 US 956 (1978).

IBM OJEL L118 May 4, 1984.

International Salt Co. V. United States, 332 US 392 (1947)

Jefferson Parish Hospital District No. 2 v. Hyde, 466 U.S. 2 (1984),

Eastman Kodak Co. v. Image Technical Services Inc. 112 S. Ct 2072 (1992).

Silver v. New York Stock Exchange, 373 US 341 (1963).

Oscar Bronner GmbH & Co. KG contre Mediasprint Zeitungs- und Zeitschriftenverlag GmbH & Co. KG e.a. Affaire C-7/97.

Otter Tail Power Co. V. United States, 410 US 366 (1973).

Port de Roscoff: Vingt-cinquieme rapport sur la politique de la concurrence, 1995, p. 29, point 43

Rodby (1993): Décision de la Commission du 21 décembre 1993 relative au refus d'accès aux installations du port de Rodby (Danemark), JOCE L 55/52 du 26 février 1994, point 12.
Sabena, OJ L 317/47

Sealink v. B&I, COM (92) 123 final (1992, 1994)

Siegel v. Chicken Delight Inc., 448 F2d 43, 47 (9th Cir 1971), cert. Denied, 405 US 955 (1972)

Tetrapak v. Commission, Court of First Instance, judgment of 6 October 1994, case T83/91; see also the Commission Decision 92/163/EEC of 24 July 1991 and the Judgment of the European Court C-333/94 of 14 November 1996.

United States v. Terminal Railroad Association, 224 US 383 (1912).

United States v. Topco Assoc., Inc., 405 U.S. 596 (1972).