

COMPETITION LAW ISSUES IN THE COMPUTER INDUSTRY: AN ECONOMIC PERSPECTIVE

CENTO G VELJANOVSKI*

I INTRODUCTION

The computer industry has been the subject of some of the most intense, extended and controversial investigations in competition law. The *IBM*¹ case in the 1970s and the *Microsoft* case² have polarized debate over the role and enforcement of competition law. Some argue that the computer industry and related sectors require a new type of antitrust, by which they mean a more permissive regime. Others believe that the concerns raised by the computer sector are no different to those routinely dealt with by competition authorities, say in the telephone, media and pharmaceuticals sectors - so there is no need for a new competition law and new economics - just better analysis and more stringent enforcement to prevent irreversible monopoly positions being created.

There is no doubt that the computer industry poses a challenge to competition law and its economics. The industry raises traditional competition law concerns of monopoly, foreclosure and bundling together with more complex considerations of innovation, tipping, and the interplay between intellectual property rights and competition law. It raises the question whether the static definition of market power – usually cast in terms of the ability to profitably raise price above the competitive level – is an adequate benchmark in emerging industries where non-price competition, dynamic efficiency, price discrimination and prices below marginal costs are endemic. It is also an industry where market boundaries are changing and converging, and where often a few large firms supply the ‘building blocks’ for those down- and up- stream.

This article addresses some of the economic issues raised by the computer sector relevant to competition law analysis. The discussion is organized as follows. Section II identifies several basic economic concepts relevant to the competitive assessment of the computer industry, its structure and business practices. This is followed in Section III with a critical assessment of the implications of these concepts for policy and competition law enforcement, and the concerns over how dynamic competition factors should be incorporated into competition law (Section IV). The discussion then turns to the definition of relevant markets and in particular whether complementary products

* Managing Partner, Case Associates (London) and Associate Research Fellow, Institute of Advanced Legal Studies, University of London. Contact: centog@casecon.com.

¹ *IBM v Commission Case 60/81* [1981] ECR 2639.

² *United States v Microsoft Corporation* [2001] Trade Cases (CCH) 73,321.

should be regarded as one or two relevant product markets (Section V), often referred to as aftermarket. The discussion then looks at the economics of product bundling (Section VI) and a brief discussion of intellectual property rights (Section VII). The discussion for the most part makes reference to European law and cases, and in relation to bundling *United States v Microsoft Corporation*.

II SOME BASIC ECONOMICS

Several key economic concepts are useful in the competitive analysis of the computer industry – on the supply-side there are economies of scale/scope; and on the demand-side network effects, and public goods. The discussion below focuses on the demand side concepts which have excited most interest. The supply-side considerations while not discussed are nonetheless important, and where significant will tend to reinforce the considerations addressed below.

A Basic Concepts

The first concept is network effect or externality. This is where the utility that a user derives from consumption of good increases with the number of other users consuming the same good.³ The concept has also been called a demand-side economy of scale,⁴ because the value of a network as measured by the number of subscribers, increases with the number of other subscribers accessible by the network.⁵

Professors Shapiro and Katz⁶ identify two types of network effects – direct and indirect.⁷ *Direct network effects* have been defined above as occurring when the value of the service increases with the number of users/consumers. The more people with telephones, the more useful and the more valuable are telephones to the user. Direct

³ Professor Jean Tirole: 'Positive network externalities arise when a good is more valuable to a user the more users adopt the same good or compatible ones. The externality can be direct (a telephone user benefits from others being connected to the same network; computer software, if compatible, can be shared). It can also be indirect; because of increasing returns to scale in production, a greater number of complementary products can be supplied – and at a lower price – when the network grows (more programs are written for a popular computer; there are more video-cassettes compatible with a dominant video system; a popular automobile is serviced by more dealers)'. J Tirole, *The Theory of Industrial Organization* (1989) 405. Professors McKnight and Bailey offer a similar definition: 'A network externality is the benefit gained by incumbent users of a group when an additional user joins the group. The group can be thought of as a 'network' of users, hence the term network externality. When the economic benefit of an additional user positive [sic], it is a positive network externality.' L W McKnight and J P Bailey (eds), *Internet Economic* (1997) 6, note 3; W H Page and J E Lopatka, 'Network Externalities' in B Bouckaert and G De Geest (eds) *Encyclopaedia of Law and Economics* (2000).

⁴ C Shapiro and H R Varian, *Information Rules – A Strategic Guide to the Network Economy* (1998).

⁵ It has also been termed a 'bandwagon effect'. J H Rohlfs, *Bandwagon Effects in High-Technology Industries* (2001).

⁶ M L Katz and C Shapiro, 'Network Externalities, Competition and Compatibility' (1985) 75 *American Economic Review* 424.

⁷ N Economides, 'The Economics of Networks' (1996) 14 *International Journal of Industrial Economics* 675; W H Page and J E Lopatka, 'Network Externalities' in B Bouckaert and G De Geest (eds), *Encyclopaedia of Law and Economics* (2000); N Gandal, 'A Selective Survey of the Literature on Indirect Network Externalities' (1995) 17 *Research in Law and Economics* 23. See <<http://www.stern.nyu.edu/networks/site.html>>; <<http://wwwpub.utdallas.edu/~liebowitz/netpage.html>>; <<http://www.sims.berkeley.edu/resources/infoecon/Networks.html>>.

network effects are relevant for two-way telecommunications systems such as fixed and mobile networks, the Internet, Instant Messaging and other communications networks. The more relevant concept for the computer sector is *indirect network effects* which occur when the value a consumer derives from a good or service increases with the number of additional users of identical and/or interoperable complementary goods. These give rise to what some have referred to as ‘virtual’ or ‘hardware/software’ networks. An example of an indirect network effect is a computer operating system (OS). If only five people use the same operating system, few would write any programmes and applications for that system, which would limit its usefulness. But as more people purchase that same operating system, programmers will create more programs for that system, increasing its usefulness. This will attract more users and begin to generate positive feedback effects that increasingly make the OS more attractive to both programmers and users.

The other demand-side concept relevant to network industries is a ‘public good’. A public good exists where the consumption by one individual does not detract from the consumption of others. A film or an episode of *Friends* is ‘public’ in the sense that consumption is collective and indivisible – a broadcast can be viewed by all in the transmission area, without the viewing by any one individual detracting from others. It thus differs from an apple – a private good - which once grown and eaten, is unavailable to the rest of the world. Television programmes, computer software, music and most other media and information products have public good characteristics.

While the notion of a public good is a demand-side concept arising from the indivisibility of consumption, it has a supply-side counterpart. Public goods imply that the marginal costs of supplying the good or service to more users are low or negligible. An episode of *Friends* has a large sunk cost of production which does not vary with the number of times it is broadcast or the size of its audience. This, in turn, gives rise to the existence of large sunk costs and products where marginal costs are below average (total) costs.

B *Welfare Economics Considerations*

The proposition that markets in the computer sector are characterised by significant network effects and public good features has significant implications for competitive analysis even in a static setting. In particular, the usual optimality conditions that prices equal marginal costs no longer holds, and this in turn means that many of the usual tests for misuse of market power will need to be modified.

1 *Network effects*

The empirical proposition that consumers value a larger network, whether real or virtual, more than a smaller network leads to a number of modifications to market outcomes and competition policy.

First, network effects explain why and how network operators and producers seek to actively network growth. The existence of significant network effects will encourage network operators to increase the number of users and expand the network. That is, they seek to ‘internalise’ the network effects and to the extent they are successful markets do not fail. Indeed it is a misnomer to refer to many examples as ‘externalities’. This is

because an external benefit (positive externality) or cost (negative externality) does not influence the actions of those who make resource allocation decisions.⁸ As a result those activities which impose costs are over-expanded, and those generating benefits under-provided by markets. A network effect as defined above, on the other hand, is often taken into account by network operators who gain by internalising it, either by growing or interconnecting physical networks, or adopting pricing and marketing strategies which seek to rapidly grow the number of users and the products diffusion in the marketplace. Therefore it is wrong to characterise these network effects as externalities, and hence to imply market failure. This is why the neutral term network effects has now been adopted in much of the literature.⁹

Second, network growth driven by attempts to internalise network effects increase consumer welfare as measured by the economist's concept of consumers' surplus.¹⁰ The network operator has an incentive to grow the network, and in particular to encourage take-up in the early years, by so-called penetrative pricing such as subsidised handsets and lower on-net tariffs. However, since larger networks generate greater consumer benefits than smaller networks, this outcome is economically efficient, provided the larger network does not misuse any market power. Society is better off with these larger networks because the value to consumers is greater than would be the case if the same number of subscribers were spread among two or more incompatible/closed smaller networks. In addition the costs of developing the second (smaller) network are avoided. Thus, the 'internalisation' of network effects by operators or companies generates large consumer benefits, and therefore carries no pejorative connotations either generally, or in the enforcement of competition law. As Joel Klein stated when Assistant Attorney General for Antitrust in the US Department of Justice:

... there's nothing illegal or even undesirable about [network externalities]: they are an outgrowth of market forces and consumer choice and, so far as the antitrust laws are concerned, business which have the skill and foresight to understand and take advantage of these forces are entitled to enjoy the fruits of their effects.¹¹

Third, the number of viable networks will depend on the significance and nature of the network benefits function, all other things equal. If the network effects are exhausted at subscriber/user numbers which are far less than the potential market, then in the presence of network effects more than one network will be sustainable. Thus, there is no necessary reason to assume that network effects, even in the absence of anti-competitive

⁸ An externality is said to exist if a transaction imposes a cost or benefit on others not taken into account by the transacting parties. The classic case of an externality is pollution where the production of a good (for example, from a power generation plant) gives rise to a third party effect (polluted rivers) not directly priced in the market and therefore not affecting the economic decisions of producers and consumers. As a result the activity in question is over-expanded because society at large is effectively subsidising its production by not taking into account the costs it imposes on third parties (local residents) and others.

⁹ On the other hand it is true that activities with increasing returns do pose theoretical problems for the economists' general equilibrium model of markets since it implies that there is no competitive outcome but either a monopoly outcome or a 'corner solution' or multiple equilibria.

¹⁰ Consumers' surplus is the difference between what individuals pay for a good and their maximum willingness to pay at different levels of consumption. It gives a monetary value to consumers' welfare.

¹¹ J I Klein, 'Rethinking Antitrust Policies for the New Economy' (talk to New Economy Forum, Haas School of Business, University of Berkeley, 9 May 2000).

actions, will necessarily lead to a natural monopoly situation, or market tipping (see below).

Fourth, network effects can also be exploited by interconnection, interoperability and licensing agreements which give one network access to another's customer base. Network effects create an incentive for co-operative solutions by holding out benefits from interconnection of separate networks. When separate networks are interconnected then the sector as a whole and consumers and operators benefit from network effects – all have the gains associated with a larger network.

Fifth, network effects can encourage operators to adopt a non-cooperative approach, such as a 'winner takes all' strategy where they refusal to supply interconnect, and compete aggressively by offering an exclusive and incompatible network to consumers. This non-cooperative strategy may be pro-competitive resulting in vigorous network competition or alternatively an abuse of market power designed to foreclose the market to smaller networks.

2 *Public goods*

The computer sector has many products and services which exhibit public good features, especially for computer hardware standards, middleware, operating systems, and particularly software. Public goods raise two main complications for the competitive assessments of the computer sector.

First, there will be a trade-off between efficiency in consumption and production. As stated above, a public good once produced and developed can be made available to all consumers at very little cost. As a result society would benefit from the widest distribution and exploitation of the good since the costs of doing this are low. This implies that, say, software should only be priced at the direct incremental costs of making another copy of the application. To charge a price higher than these low marginal costs (even to cover the high fixed costs) would be inefficient because it would cause many consumers either to reduce their consumption, or not consume the good at all even though they would be willing to pay the marginal costs. However, clearly if price were set at marginal costs, the revenues to the producer would be insufficient to cover total costs plus a reasonable return on investment made. As a result very few public goods would be produced by the private sector.

Second, public goods give rise to a further potential under-provision because consumers have limited incentives to reveal their true preferences and declare their willingness to pay. This arises from the non-excludability or joint consumption aspect of a public good. Once a (pure) public good is produced it is available to all, so that the individual can consume it irrespective of the amount paid. Clearly, if this is the case individuals will prefer to pay a little rather than a lot, even if they value consumption of the public good highly. An individual will not regard this under-valuation as necessarily having an effect on the production and availability of the public good. If all reason in this myopic way, then society will be under-supplied with public goods. The price mechanism will only work if there is some form of exclusion, and/or consumption can be monitored. However, in general the weak excludability features of public goods means that they create an appropriability problem for producers, and the real possibility of under-production.

The third and related feature of public goods is an appropriability problem. For the reasons just given it is often difficult for the producer of a public good to reap the revenues generated from a public good; consumers are not willing to pay and other firms will seek to free ride on the investment of the original developer and producer of the public good. On the supply side where such competition takes place (in the absence of adequate copyright and legal protection) it will force the market price down to the negligible marginal costs of copying and distributing the good. Thus in the absence of legal protection, copying, counterfeiting and mimicking will lead to the market price failing to reflect and hence recover the fixed costs of developing the public good in the first place.

The reconciliation of some of these tensions, at least in theory, is fortunately at hand. The optimal price of a public good has often been mischaracterized because of the implicit assumption that the same price will be charged to all consumers. However, allocative efficiency requires only that the marginal consumer who is assumed to place a very low marginal evaluation on the good, should be charged the low marginal cost. Those who value the public good more highly can be charged higher prices without distorting consumption. Thus the optimal pricing structure which satisfies both consumption and production efficiency conditions is one where a different price is charged to each consumer based on his or her marginal evaluation of the public good. In this way the supplier of existing and prospective public goods knows and receives the correct price signals and consumption patterns are not distorted. That is, allocative efficiency requires highly discriminatory pricing structure which generates sufficient revenues to cover the fixed costs and give producers a measure of the total value to consumers place on the public good. In practice such a system of (third degree) price discrimination would not be practical and crude systems of price discrimination will be practised by customer group, usage and territory. However, from this brief description it is clear that the optimality conditions of a public good jar with competition law where often price discrimination (unrelated to cost differences) is treated as a misuse of market power.

III IMPACT OF NETWORK EFFECTS

Network effects in particular have played an increasing role in the application of competition law to the computer and communications sector. The economic literature has identified two alleged market failures which may have a bearing on competition - 'tipping' and inefficient standards. In this section the theory underpinning these two alleged market failures are examined and assessed.

A *Alleged Market Failures*

1 *Tipping*

A novel element introduced by the application of network effects to competition law is the idea that markets may 'tip' or 'snowball'. The theoretical literature suggests that under some circumstances *indirect* network effects can result in one standard or product, and by implication, one firm dominating the industry. This phenomenon relates mainly to technical standards and virtual networks, rather than physical telecommunications networks. It is highly relevant to the computer industry particularly for operating systems and software.

Tipping is best understood by looking at computer operating systems. In the beginning a number of computer operating systems may compete each with a small but growing base of dedicated users. Eventually, one operating system attracts more users, say, because it is bundled with PCs or, as in the case of Microsoft Windows, adopts an open licensing policy. The user base grows which makes it more attractive to programmers. More software is written for the operating system which makes it yet more attractive to users, resulting in increased demand, and a growing user base. That is networks lead to powerful 'positive feedback effects'. Eventually the demand for the operating system may 'tip' when consumers see it as a 'must have' because it has the most users and most applications. This can lead to the survival of one standard for the industry.

The literature suggests that tipping requires several necessary (but not sufficient) conditions. The first is some incompatibility between the different networks or products. This may arise because they are not physically interconnected or interoperable.¹² In the computer industry, an exclusive attribute such as a proprietary technology or software, which is incompatible with other potentially substitutable products, provides the conditions for tipping. This factor forces consumers to make an either/or choice.¹³

High switching costs have also been listed as essential to tipping because they lock customers into one network. This condition may be more important during the early phases of the tipping process, since after a point users may be induced to switch to competitors by the presence of a large installed subscriber base itself.

Tipping or snowballing may suggest that a high market share is potentially of more concern than in old economy industries because of the higher likelihood that the firm will grow especially when a merger is contemplated or rivals allege anticompetitive behaviour related especially to interoperability, exclusivity or predation. In the presence of network effects firms may be encouraged to adopt non-co-operative or 'winner-takes-most' strategies. Recognising that the value of a 'virtual' network increases with the number of users, an operator may seek to increase the number of users and knock out competitive products and/or firms.

2 *Inefficient standards*

Another concern found in the literature is that market forces may lead to the adoption of inefficient standards or products.¹⁴ It suggests that for a variety of reasons (see below) market forces may set an industry on a trajectory that does not maximise consumers' welfare/economic efficiency, that is that market failure is endemic especially in the computer and software industry. Claims are frequently made that the market leader's

¹² M L Mueller Jr, *Universal Service – Competition Interconnection and Monopoly in the Making of the American Telephone System* (1997).

¹³ 'Network effects also result in compatibility being a critical dimension of industry structure and conduct. ... In the absence of compatibility, markets may tip' M L Katz and C Shapiro 'Antitrust in Software Markets' in A Eisenach and T M Lenard (eds), *Competition, Innovation and the Microsoft Monopoly: Antitrust in the Digital Marketplace* (1999) 33.

¹⁴ J Tirole, *The Theory of Industrial Organization* (1989) Chapter 10; B Owen and S Wildman, *Video Economics* (1990) Chapter 7; S Besen and G Saloner, 'The Economics of Telecommunications Standards' in W Crandall and K Flamm (eds), *Changing the Rules: Technological Change, International Competition, and Regulation in Communications* (1989).

computer operating system or product is inferior to some other product or standard, and it is only by an accident of history, first mover advantage, or market clout that the product/standard has come to dominate the industry. This type of claim has often been made about the success of IBM over Apple Mac computers, and between MS-DOS and other OSs. Put more technically, virtual networks are said to be path dependent, and therefore early decisions and policies can have lasting irreversible effects.

Among the reasons given for market failure in this area are:

- excessive inertia arising from the fear of early adopters of a new technology of the risk of losing their initial investment because the technology is not adopted by a sufficiently large portion of the market. In this case the parties' fear of being 'stranded' with a low-value technology¹⁵ may result in deferring its adoption. If many in the market take this stance there is 'excess inertia' resulting in an efficient standard not being adopted, even though buyers would be better off if it were;
- inefficient bandwagon effect occurs when the decision by one group of users to convert to a new technology induces others to switch, even when the latter would have been satisfied with the old technology, and when the combined benefits to both groups are lower;¹⁶
- inefficient adoption where some users adopt the new technology, others do not, and the result is lower welfare than if none had switched;
- inadequate compatibility that may occur when some users do not find it in their interest to switch to a new incompatible technology unless everybody else also switched, even though aggregate benefits would be maximised if all users switched.

This does not seem to have any immediate competition law implications, since the efficiency of a standard or product over some other which was technically available but failed in the marketplace would be difficult to causally link to an anticompetitive abuse.

B *Assessment of Network Effects Theory*

The concepts of network externalities, public goods and natural monopoly have been around for a considerable time both in theory and practice. The telegraph, telephone, railway and many other non-utility industries/activities (distribution, banking, and the media) experience network or reinforcing positive (and negative) feedback effects.¹⁷ They may now be more pronounced than in the past, but this is a matter to be established rather than assumed. And it would be unusual if they are the only factor which explains firm growth and behaviour. Nonetheless, it is the case that where

¹⁵ J Farrell and G Saloner, 'Standardization, Compatibility and Innovation' (1985) 16 *Rand Journal of Economics* 70; S Besen and L Johnson, *Compatibility Standards, Competition, and Innovation in the Broadcasting Industry* Santa Monica: RAND Publication No. R-3453-NSF, November 1986.

¹⁶ J Farrell and G Saloner, 'Installed Base and Compatibility: Innovation, Product Pre-announcements, and Predation' (1986) 76 *American Economic Review* 940; J Katz and C Shapiro, 'Network Externalities, Competition, and Compatibility' (1985) 75 *American Economic Review* 424.

¹⁷ Shapiro and Varian, above n 4, still argue that 'many tried and true principles are still valid [...] what has changed is that the internet, and information technology in general, offers new opportunities and challenges in applying these principles' (83). However, 'large fixed costs and small incremental costs are [...] hardly unique to information goods' (22).

network effects (and the other factors identified above) are empirically significant, the sector will be subject to increasing returns to scale and outcomes which may not necessarily maximise consumer welfare. The question is whether the application of network effects theory offers additional insights and is a practical concept for identifying when network operators are likely to abuse any market power.

1 *Theory*

Much of the theoretical underpinning of network effects, and certainly the tipping and inefficiency versions, concerns virtual networks such as software and product standards, and therefore is particularly applicable to the computer industry. These economic models generally seek to explain why in software markets there appear forces leading to the dominance of one proprietary standard. As Professor John Sutton, a respected economist who specialises in the analysis of market structure and innovation, has pointed out ‘[e]ach standards battle is likely to have individual aspects that exert significant impact on outcomes, making any attempt to generalise across industries by reference to a manageably short list of observable industry characteristics hopelessly difficult.’¹⁸ That is the predictions or claims of the models are highly sensitive to their specific assumptions. And, he further notes that ‘battles between proprietary standards are the exception rather than the rule’. Thus the conditions which give rise to snowballing are perhaps a special not general case.

Tipping or ‘snowballing’ is an extreme outcome which is not pre-determined even in the presence of network effects. It requires, in addition, network incompatibility, bottlenecks and possibly customer lock-in. That is, a situation where customers face a choice between incompatible networks, and which, once made, is costly to reverse.

Further, the existence of networks effects, even if significant at some levels of network size, do not enable one to conclude that monopolisation is the outcome. There may also be costs associated with increased network size which reduce and render negative the net benefits of further growth.¹⁹ As suggested above, the network effects may be small or exhausted before single firm output supplies the entire market. As for most products there will be diminishing marginal returns. It is unlikely that individuals value ever-increasing size positively, and are unconcerned about the identity of the additional consumers who join. They will value those users who they most want to contact or be contacted by most highly, and others less so. As Professor Lemley notes:

The network effect grows from the positive value placed on the ability to contact other people via the Internet, and from the access to information from a wide variety of different sources. However, it isn’t the case that people on the Internet want to communicate with everyone else on the network. The death of Usenet as an effective means of communication resulted from an overabundance of participation coupled with the lack of limits on the relevance of that participation. In this sense, then, people want to interact with a subset of all those on the Internet, though they still may get a benefit from expanding the pool of people with whom they can potentially interact. Similarly, people may well want to access only a subset of information available on the Web, if that subset is filtered or tailored in such a way as to make it more useful to them. ‘Zoning’ the Net –

¹⁸ J Sutton, *Technology and Market Structure – Theory and History* (1998) 411.

¹⁹ S Liebowitz and H Margolis, ‘Are Network Externalities a Source of Market Failure’ (1995) 17 *Research in Law and Economics* 1.

either by setting up privatised areas within it or by sectioning off entire areas by government mandate – is not inconsistent with the idea that the Internet exhibits positive network effects. Rather one must distinguish between structural barriers to access set up by inconsistent market standards and contextual barriers set up deliberately by participants themselves.²⁰

It may also be the case that network effects give rise to ‘negative feedback effects’ and dynamic instability. This will be particularly the case in fast growing industries where the existence of network effects can unravel the growth strategy. For example, if a network operator with market power has built up a large user base by subsidising connections, and then seeks to exploit users at a later date by raising prices, the network may rapidly unravel. To quote Professors Lee and McKenzie:

Dominant producers of ‘network goods’ are even less likely to act like monopolies, because they fear that their networks will unravel, thus reducing their market share far more quickly than is likely the case with traditional products. If a dominant supplier of a ‘network good’ restricts output to raise its price, the immediate loss in market share reduces the value of the product to remaining users, which leads to a further loss in market share, and a vicious downward cycle. So dominant network companies have to worry more about the long-run price sensitivity of their users than do dominant traditional companies.²¹

The most striking example in the antitrust field is *IBM v Commission*²² in the early 1980s. Five years after the case, the alleged dominance of IBM sunk under a welter of new products, innovations, PCs, desktops and laptops. The current high volatility of dot.com stocks is another example where an adverse profit warning or delay, gets magnified into extreme share price fluctuations.

However, in those instances where the structural features of the market do make tipping likely it has significant competition law implications. It means that the computer industry, or parts of it, will be subject to an inherent instability and endemic monopoly, and more troublesome is that this monopolization can occur without the firm engaging in overt anticompetitive abuses.

C *Inefficiency Version*

The inefficiency version of the theory is even more contentious and difficult to verify. As several commentators have argued the theory assumes that both consumers and producers are ignorant, and act in a myopic and passive way. Yet inefficiency and the prospect of monopoly imply gains from taking actions to correct these. As Professors Liebowitz and Margolis²³ stress when a new and more efficient technology is developed able to compete with an existing technology, or, similarly, when two new incompatible technologies are introduced simultaneously, the existence of performance differences

²⁰ M Lemley and D McGowan, *Legal Implications of Network Economic Effects* <www.law.berkeley.edu/journals/clr/library/8603.html> (internet version with no pages numbers).

²¹ D R Lee and R B McKenzie, ‘Technology, Market Changes, and Antitrust Enforcement’ Sept/Oct (2000) *Society* 31, 33.

²² *IBM v Commission Case 60/81* [1981] ECR 2639.

²³ S Liebowitz and H Margolis, ‘The Fable of the Keys’ (1990) 33 *Journal of Law and Economics* 1. Also see S Liebowitz and H Margolis, ‘Lock and Key’ *The Economist*, 16 September 1999.

creates an opportunity for substantial economic gains if the owner of the efficient technology can induce users to switch to his product. Firms convinced of this will encourage adoption and overcome the 'failures' by offering large discounts to early adopters, renting to early adopters (thereby reducing the risk faced by new buyers), offering subsidises, and developing 'converters' to assure compatibility with existing products. Moreover, a standard may be licensed freely, and therefore the existence of a proprietary standard does not preclude competition.²⁴

D *Limited Evidence*

The network effects concept remains a largely untested theory, and the evidence to date supporting it somewhat meagre. Indeed, its use as a basis in business models of dot.com ventures of the late 1990s, and the subsequent collapse of the dot.com bubble, is a salutary reminder that the extrapolation of simple theoretical concepts is fraught with difficulties.

There is also the empirical issue of identifying inefficient outcomes. To assess whether a standard is 'inefficient' requires detailed knowledge of the technical, cost and demand conditions for the product and all its potential substitutes. It is not enough to show technical superiority since this may not be what the consumers want, it may not in the consumers' eyes be value for money or has attributes that are attractive.

There have been several recent statistical econometric studies of network effects,²⁵ and several case studies that purport to show that markets have selected inefficient standards such as the alleged choice of VHS over Betamax; QWERTY over Dvorak keyboards; MS-DOS over Apple.²⁶ These on closer inspection provide little evidence that the standards selected by the market were inefficient or inferior to the alternatives which failed.²⁷

²⁴ These models assume that one player selects a strategy while the others passively respond. Under these assumptions with a repeated game it is not surprising that the 'active' players end up winning. A fully dynamic model would introduce other relevant behaviour which could materially affect the outcome:

- Competitors' responses and strategies;
- U shaped cost functions to take account of transaction costs which increase with volumes;
- The relation between economies of scale on the demand side and network size; and
- The role of endogenous investment sunk costs whereby firms strategically invest in sunk costs (advertising and Research and Development).

²⁵ Y-M Wang, J Gallagher and S Vasudevan, 'The determinants of network growth: The case of commercial online information networks' Proceedings of the Seventeenth International Conference on Information Systems (1996); J Gallagher and Y-M Wang, 'Network effects and the impact of free goods: an analysis of the web server market' (1999) 3 *International Journal of Electronic Commerce* 67; E Brynjolfsson and C F Kemerer, 'Network externalities in microcomputer software: an econometric analysis of the spreadsheet market' (1996) 42 *Management Science* 1627.

²⁶ P A David, 'Clio and the Economics of QWERTY' (1985) 75 *American Economic Review - Papers and Proceedings* 332.

²⁷ S J Liebowitz and S E Margolis, 'The Fable of the Keys' (1999) 22 *Journal of Law and Economics* 1; S J Liebowitz and S E Margolis, 'Network Externality: An Uncommon Tragedy' (1994) 8 *Journal of Economic Perspectives* 133; S J Liebowitz and S E Margolis, 'Are Network Externalities a New Source of Market Failure?' (1995) 17 *Research in Law and Economics* 1; S J Liebowitz and S E Margolis, 'Path Dependence, Lock-in, and History' (1995) 11 *Journal of Law and Economics* 205; S J Liebowitz and S E Margolis, 'Should Technology Choice Be a Concern of Antitrust Policy' (1996) 9 *Harvard Journal of Law and Technology* 283.

It is not surprising that verifying the theory would be difficult. The theory posits a dynamic process of network development which would require historical evidence of how standards and networks have evolved in the past. This is a painstaking task requiring detailed knowledge of markets and technical factors, and even then the evidence would be subject to differing interpretations.²⁸ Even the classic examples purportedly supporting the snowballing version of network effects have offered limited verification.²⁹ Indeed, what will most likely emerge from this much needed empirical work, is that network effects if present are only one of a number of factors explaining the development of networks and their competitive responses.

IV DYNAMIC EFFICIENCY

A missing part of the economic jigsaw presented above is the role played by and impact of innovation and investment. In a competition law context this relates to whether the emphasis on price competition and static efficiency capture the essence of competition in new economy industries such as the computer sector.³⁰ Some argue that this static approach to market definition is wrong and does considerable harm; while others believe that the criticisms are exaggerated, and the adverse consequences small.

Underpinning these views are different ‘theories’ of the new economy. One school believes that new emerging markets should be tightly regulated because developments in the formative stages of the industry will determine its future market structure. If these create monopolies and tight oligopolies, it will be impossible to impose structural remedies to re-create a more competitive industry. First mover advantages, network effects and tipping all point to a heightened risk of monopolisation which needs to be prevented. According to this view, the largely static view of competition is adequate, and any harm it does is the lesser of two evils - the more serious danger is monopolisation.

The EC Commission has interpreted network effects as implying a greater correspondence between dominance, and its abuse and the creation of an essential facility. It has even gone further to suggest that networks which are not dominant but large, even in a relatively immature market, lead to snowballing.³¹ EC Competition

²⁸ J Sutton, *Technology and Market Structure – Theory and History* (1998) 410.

²⁹ The experience of AM radio in the US has also been given as an example of excess inertia. However, the Federal Trade Commission has questioned this interpretation suggesting that consumers did not value the technology as much as the costs. Federal Trade Commission, *Commentary on the Regulation of New Digital Audio Radio Services (GEN Docket No. 90-357)*, Staff of the Bureau of Economics and the San Francisco Regional Office of the Federal Trade Commission, January 25, 1991.

³⁰ C G Veljanovski, ‘EC Antitrust and the New Economy - Is the EC Commission’s view of the network economy right?’ (2001) 22 *European Competition Law Review* 115.

³¹ These differences are summarised by Pons, (then) Deputy DG of the Commission’s Competition Directorate:

The WorldCom/MCI case essentially revolved around two different views of the Internet. On the one hand there were those who said the Internet was the twentieth century equivalent of the Wild West, where there were no rules, where anyone could enter, and where the small and the big competed against each other on an equal footing.

In contrast to this there were many others who said that only a handful of companies, the equivalent in the voice telephone world of facilities based telephone operators, were truly capable of delivering Internet connectivity. Most of the smaller players relied on the services offered by these bigger players, and competed against them at the retail level, but depended on

Commissioner Mario Monti has claimed that this applies even if the monopoly problem ‘only lasts for a short time’.³² This effectively sets a lower threshold for triggering competition concerns in the new economy than the old.

The other school contends that this is a misreading of the economics and history of innovation. They argue that initially concentrated markets are common and critical for innovation to take place. These are emerging industries in a fluid state buffeted by technical change and the need to make massive and ongoing investment in software, hardware, and infrastructure. Where one or a few firms appear dominant, they will be rapidly displaced by the technical innovation and superior marketing of new entrants which may, in turn, appear ‘dominant’. They are, to use the words of Joseph Schumpeter, subject to the ‘perennial gales of creative destruction’.³³ That is, a process of serial monopolisation occurs which is short-lived. The standard approach results in excessively narrow product market definitions and too much regulatory intervention which may chill investment and stall innovation.³⁴ Market definition should use a wide lens to take into account longer term non-price factors.

The paradox that monopoly can be competitive was resolved by Joseph Schumpeter over half a century ago.³⁵ He suggested that maximising longterm consumer benefits require large firms who may have temporary market power. These ‘monopolists’ still compete vigorously not necessarily ‘in the market’ but ‘for the market’, that is for the right to be the monopolist. This competition places dominant firms under constant pressure from rivals offering better and cheaper products/services to displace them as the sole provider in the marketplace. Thus the competitive process leads a regime of serial monopolisation of markets consistent with competitive outcomes.

This view has attracted increasing attention and adherents especially given the very rapid technological change in the communications sector. For example, it was given a semi-official boost in 1995 with the publication of Richard Gilbert’s (then Assistant Attorney General for Antitrust) and Stephan Sunshine’s article arguing that traditional market definition procedure was flawed in innovative industries.³⁶ The authors argued

them for the supply of Internet connectivity, i.e. the ability to get messages anywhere on the Internet.

J P Pons, ‘The application of competition and anti-trust policy in media and telecommunications in the European Union’ (Paper to International Bar Association, September, Vancouver, 1998).

³² M Monti, ‘Competition and Information technologies’ (Address to ‘Barriers in Cyberspace’ – Kangaroo Group Brussels, 18 September 2000). The concept of a ‘temporary monopoly’ is also raised in *ACCC v Universal Music Australia Pty Ltd* (2002) ATPR 41-855; see S Coronos ‘Technological Tying in the Computer Industry: When Does it Contravene Section 46 of the Trade Practice Act?’ (2003) 3 *Queensland University of Technology Law and Justice Journal* 48.

³³ J Schumpeter, *Capitalism, Socialism and Democracy* (1943).

³⁴ Bittlingmayer’s study of 21 major industries in the US covering 1947-1991 investigated the statistical association between antitrust case filings and investment. Each antitrust filing was associated with a significant decline in investment in the respective industry: G Bittlingmayer ‘Regulatory Uncertainty and Investment: Evidence from Antitrust Enforcement’ (2001) 20 *Cato Journal* 295.

³⁵ While Schumpeter saw the pursuit of monopoly as the engine of innovation he also felt that if acquired it would be temporary – someone would always come along with a better mousetrap lured by the beacon of the monopolist’s high profits!

³⁶ R Gilbert and S Sunshine, ‘Incorporating Dynamic Efficiency Concerns in Merger Analysis – The Use of Innovation Markets’ (1995) 63 *Antitrust Law Journal* 569; Cf R Rapp, ‘The Misapplication of the Innovative Market Approach to Merger Analysis’ (1995) 64 *Antitrust Law Journal* 19.

that market definition ignores technological change, and hence the effects on future competition and market power. They proposed defining an innovative market as well as a product market. The former would include firms likely to innovate in the market whether they currently produce or not. The competition authority would then be required to determine the effect of a merger on concentration in that innovative market, and finally the effect of this on innovative activity.³⁷ While this has not been widely endorsed, there are signs that competition authorities are paying more attention to dynamic analysis and the impact of intervention on long-term competitive factors.

V MARKET DEFINITION PRINCIPLES

In the practical enforcement of competition rules, many of these considerations have not yet made much of a mark. However, even at the initial stage of a competition inquiry, the computer sector gives rise to complications in applying market definition principles. It is generally stated that markets consist of products which are close substitutes, yet in the computer industry competition authorities and economists have argued that markets consist of products that are not substitutes but complements. This has been referred to as the aftermarket problem, and as will be seen from the discussion below can have a profound impact on competition enforcement.

A *Markets for Substitutes*

Markets are generally regarded as consisting of products and services which are close substitutes. In identifying a market there are a range of standard concepts and procedures. Here we draw on European practice but this does not in broad principle differ significantly from the approach in Australia or the US.

The delineation of markets for the purposes of competition law is set out in the EC Commission's *Market Definition Notice* published in 1997,³⁸ the UK Office of Fair Trading's (OFT) guidelines issued under *Competition Act 1998*,³⁹ and those of other Member States.

A 'relevant product market' consists of products which are close substitutes in consumers' eyes in conditions of effective competition. Under the EC Commission's decisional practice, a multifaceted approach is used to determine substitution. The EC *Market Definition Notice*, and European case law, states that the relevant product 'market comprises all those products and/or services which are regarded as interchangeable or substitutable by the consumer, by reason of the products' characteristics, their prices and their intended use' (para 9). It continues that the relevant geographical 'market comprises the area in which the undertakings concerned are involved in the supply and demand of products or services, in which the conditions of

³⁷ OFT/Oftel, *Innovation and Competition Policy - Economic Discussion Paper 3* (2002); W J Baumol, *The Free Market Innovation Machine* (2002).

³⁸ EC Commission Notice on the definition of the relevant market for the purposes of Community competition law, OJ [1997] C 97/C 372/03, 5-13. Under the EC Framework Directive (Directive 2002/21/EC) which sets out the new regulation of the communications sector, market definition will be critical. EC Commission guidelines on market analysis and the assessment of significant market power under the Community regulatory framework for electronic communications networks and services, 2002/C 165/03.

³⁹ *Market Definition*, OFT 403, March 1999.

competition are sufficiently homogeneous and which can be distinguished from neighbouring areas because the conditions of competition are appreciably different in those areas' (para 8).

These definitions are imprecise, as they do not indicate the degree of substitution necessary to include or exclude products in a relevant product market. A more rigorous approach referred to in the EC *Market Definition Notice* is the hypothetical monopolist test. This looks at whether a hypothetical monopolist of a group of products or services can profitably raise relative prices by 5% to 10%. If it cannot, then the products are deemed part of the same relevant market and others are added to the product grouping until the point is reached where the hypothetical monopolist can profitably and permanently raise price. When this point is reached, there are no products outside the grouping which are sufficiently close substitutes to constrain the hypothetical monopolist from setting price above the competitive level.

There are several aspects of market definition in competition law that need to be clearly understood.

First, the relevant market is the narrowest grouping of products over which the hypothetical monopolist can raise price profitably above the competitive level. This is clear from the OFT's Guidelines which state that the market definition test must be applied to 'the products (or group of products) at the centre of the investigation...' (para. 2.8), and '[T]he process starts by looking at a relatively narrow potential definition' (para. 3.1).

Second, the test seeks to identify consumers' reactions which are fairly rapid. Under the OFT's Guidelines rapid is defined as at 'short notice' (para. 3.13) and 'in the short run (for example, within one year)' (para. 3.15).

Third, the appropriate definition of the market depends on the product or complaint under consideration, and varies on a case-by-case basis depending on the facts. For example, mainframe computers and PCs may at the same time be treated as one or two separate relevant product markets depending on the issue being considered. If one were examining market power in the mainframe sector, it may well be the case that PCs place a competitive constraint on the hypothetical monopolists' ability to raise the price of mainframe computers. However, the opposite may not be the case since those who buy PCs and are subject to some alleged monopoly abuse by PC suppliers, may not regard mainframe computers as close substitutes, and will therefore not be able to substitute away from PCs. In this case, PCs will be a separate market from mainframe computers. This is sometimes referred to as asymmetric market definition.

Fourth, price discrimination for the same or similar product can result in several markets being defined. If it is possible for a supplier to distinguish between classes of consumers (large corporate, SMEs, personal) and charge these different prices, then it may be possible to define different relevant markets for each. Switching costs in particular can generate a group of core captive customers which may form a separate market also. The hypothetical monopolist can exploit the fact that existing customers cannot switch easily to charge them a higher price than to its new customers. These will form separate relevant markets if the switching costs prevent captive customers from substituting in sufficient numbers in response to a price increase.

Fifth, and related to the previous point, price differences between the same or similar products are not evidence that they are separate relevant markets. Different types of computers, sold at different prices, even if these reflect different specifications and functionality, do not mean that they constitute separate markets. At the margin the reactions of some consumers may place a competitive constraint on PC manufacturers from departing from price differentials which reflect quality differences. This is referred to as the 'chain of substitution'. Thus markets with differentiated products sold at different prices may be horizontally linked by consumer reactions at the margin (see next point). Clearly, this implies that there are low switching costs, at least for some consumers.

Sixth, and critically, in determining product substitutability it is the reaction of the marginal customers that counts. The hypothetical monopolist test only requires that sufficient customers at the margin react to a price increase, not all customers or the average customer. Thus, if in response to a hypothetical 5% price increase, a subset of existing customers reduces the quantity purchased of the product by more than 5%, then this will eliminate the monopolist's ability and incentive to raise prices. The fact that the 95% of the remaining customers do not react is immaterial. It follows from this, that survey evidence that many or most customers are price insensitive or regard the product as having no close substitutes does not establish that the product is in the relevant market.

Seventh, markets may be linked vertically where products are complements. This is of particular importance in the analysis of the computer sector because the primary product (hardware) will need complementary products such as software, consumables, peripherals and maintenance to function. It is here that considerable controversy and attention has recently centered, since the decision as to whether the products constitute one or two markets, have been critical for the assessment of market power and its abuse as shown in the US antitrust proceedings in *Microsoft* (discussed in more detail below).

B *Establishing Market Definition*

Market definition should ideally be established on the basis of factual analysis. This is often difficult because in theory determining whether two products are close substitutes is based on consumer and firm reactions at the competitive price. Since the prevailing price will often not be the competitive price, especially where the firm has market power, the data necessary to carry out the market definition test will not be available. Moreover, as widely recognised, using the prevailing price may lead to the market being defined too broadly, a defect known as the *Cellophane Fallacy* after the celebrated 1956 *duPont* case.⁴⁰ At a practical level while a variety of rigorous statistical techniques are available often these cannot be used because data are unavailable, incomplete and/or not in the correct form. Thus the evidence used will usually come from the industry, marketing literature, industry reports and market surveys. For example in *ICL/Synstar*,⁴¹ the OFT based its view that there was one market for computer hardware and hardware maintenance on survey evidence, industry views and simple statistics backed up by analysis, inferences and opinion.

⁴⁰ *Du Pont de Nemours & Company* 351 US 377 (1956).

⁴¹ Case No CA98/6/201 (20 July 2001). The OFT stated that it preferred to follow the EC Decision in *Pelican v Kyocera*, *Competition Newsletter*, Vol 1.

Further, it is a common belief that rigorously testing for market definition is impossible - how does one test whether a hypothetical price increase is profitable? More rigour can, however, be given to the market definition using techniques which do not require vast amounts of data and statistical analysis, but which can provide a simple check on proposed or disputed market definitions.

One such technique is Critical Loss analysis.⁴² A Critical Loss is the percentage loss in output necessary to make a hypothetical price increase unprofitable.⁴³ All one needs to calculate the Critical Loss is data on price and the firm's variable costs. While the Critical Loss does not estimate the actual lost sales expected from a five percent price increase - this would require estimating the price elasticities and involve considerable data and econometric analysis - it does give a figure whose plausibility can be tested to see whether the supply response would be below or exceed the Critical Loss. If the evidence shows that the reaction of consumers would exceed the Critical Loss, then the market is wider than proposed; and if it falls below the Critical Loss, the market has been correctly defined or is narrower.

This simple technique is increasingly used in US antitrust but not yet in EC law. For example, Critical Loss analysis was relied on in part in *US v SunGard and Comdisco*,⁴⁴ where the US District Court ruled against the US Department of Justice's narrow market definition. The merging parties supplied disaster recovery services to business. The Department of Justice defined the relevant market as shared 'hot-site' recovery services. The Parties contended that the market was wider including hot-site disaster recovery, cold-site disaster recovery, vaulting, high availability disaster recovery and mobile solutions. Using data on price and the firm's variable costs, evidence was given that Sungard's Critical Loss was approximately 5%. This meant that Sungard could not profitably sustain a unilateral hypothetical price increase if it were to lose as little as 5% of its sales as a result of the price increase. The Court agreed with the Parties' expert economists' view that given the large number of competitors and alternative solutions to which customers could turn, that the supply response would exceed 5%, and a broader market definition was appropriate. The merger was cleared.

C *Markets for Complements - Aftermarkets*

In the analysis of the computer industry it is often necessary to decide whether there is one market which includes hardware and complementary services, or several separate markets - a primary market (mainframe computers) and a number of aftermarkets consisting of the complementary services/products supplied both by the computer manufacturer and third party suppliers. This issue arises regularly because firms in the primary market are usually vertically integrated supplying both the primary and

⁴² B C Harris and J J Simons, 'Focusing Market Definition: How Much Substitution is Necessary?' (1989) 12 *Research in Law and Economics* 207; B C Harris and C G Veljanovski, 'Critical Loss Analysis - Its Growing Use in Competition Law' (2003) 24 *European Competition Law Review* 213.

⁴³ Critical Loss is equal to $Y/(Y + CM) \times (100\%)$, where Y = the hypothesized price increase of 5 or 10 per cent expressed as a proportion (.05 or .10), and CM = the contribution margin defined as the difference between the original price and average variable cost stated as a proportion of the original price.

⁴⁴ 172 F Supp 2d 172, 182, 186-92 and n 21 (D DC 2001).

complementary products/services. Those independent downstream firms who supply, or wish to supply, only the latter, often regard the competition as unfair and it is frequently observed that the prices for complementary services are high giving rise to concerns that there is monopoly pricing.

Some theory is needed to examine the aftermarket issue.⁴⁵ A firm selling, say, both mainframe computer equipment (primary product) and computer maintenance (the aftermarket) faces a trade-off when it sets the price for maintenance services. Those who have purchased a computer will be 'locked-in' by the initial capital costs of buying and setting up the computer system, and any subsequent costs of moving to a new computer system. This provides the commercial temptation to the seller to exploit those who have already purchased his computers. That is, there is the possibility of what is sometimes called '*installed user opportunism*'. A higher price will allow it to earn increased profits on maintenance sales to consumers who have already purchased its computer equipment.

The incentive of the computer manufacturer to raise maintenance prices will have a feedback effect. Higher maintenance prices will reduce computer sales, because some or all potential buyers will take into account the higher expected costs of continuing maintenance services – so-called 'life cycle' or 'whole life' costs. If a high price in the aftermarket deters a sufficient number of potential buyers, then anticompetitive pricing is not a rational strategy. Here the word sufficient is to be understood as loss of computer sales which make a price hike in the aftermarket unprofitable. Thus if there is effective competition in the computer market this acts as an effective competitive constraint in the aftermarket.⁴⁶

If, on the other hand, the competitive constraints are weak, because buyers do not have adequate information, anticompetitive practices in the aftermarket may be profitable. Employing the EC Commission's 'five percent test', the integrated computer manufacturer and maintenance provider could raise maintenance prices by 5% without inducing losses in computer sales sufficient to wipe out the profits from 'overcharging on maintenance'. In the latter case, aftermarket services and hardware would delineate separate relevant markets.

In summary there are two views which ultimately depend on the specific facts in each case:

- View I holds that buyers of computers are sophisticated and have or should have adequate information to judge 'whole life costs'. They are, therefore, unlikely to make wrong decisions, especially if they are repeat purchasers. Moreover, given

⁴⁵ There is now literature on the economics of aftermarkets in the wake of the US *Kodak* and *Microsoft* decisions. See C Shapiro and D J Teece, 'Systems Competition and Aftermarkets' (1994) 39 *Antitrust Bulletin* 135; C Shapiro 'Aftermarkets and Consumer Welfare: Making Sense of Kodak' (1995) 63 *Antitrust Law Journal* 483; S Borenstein, *et al*, 'Antitrust Policy in Aftermarkets' (1995) 63 *Antitrust Law Journal* 455.

⁴⁶ Indeed recent theoretical works suggests that even if the primary market is effectively competitive, in the absence of perfectly contingent long-term contracts governing secondary/complementary services manufacturers are still able to price above the competitive price (greater than marginal costs) for aftermarket products and services. See S Borenstein, J K MacKie-Mason and J S Netz, 'Exercising Market Power in Proprietary Aftermarkets' (2000) 9 *Journal of Economic and Management Strategy* 157.

the interrelationship between hardware and maintenance, maintenance costs have an effect on new computer sales, which in turn limits excessive charges for maintenance services. Thus, hardware sales act as a continuing discipline on manufacturers not to hike maintenance prices. That is, competition in the primary market implies competition in the aftermarket, and hence both constitute the one relevant market for competition law purposes.

- View II casts doubt on the ability of buyers to foresee ‘whole life’ costs at the time of purchase, and subsequent high switching costs give the seller the ability to exploit captive customers. This ‘lock-in’ or ‘hold-up’ theory results in two (hardware and maintenance) or many separate (hardware and a different maintenance market for each type of hardware) relevant markets.

It should be appreciated that the above analysis does not do justice to the complexity of pricing in the computer sector. Pricing of computers and related services is the outcome of a number of conflicting considerations which alter over the life cycle of the product. For example, network effects will lead to penetrative pricing which entails lower hardware costs and higher prices for secondary products. The best example is mobiles where handset subsidies are given in order to lower the costs of access, and enhance the value of mobiles to users who can connect to more users. Such penetration pricing is noted (although not termed as such) in the OFT’s Guidelines at para. 5.11 which state that in such circumstance ‘it may be more appropriate to treat the primary and secondary products as separate markets ...’.

Second, customer ‘hold-up’ (View II above) may not be the only explanation for high aftermarket prices. High aftermarket prices may be a method of metering and hence a form of price discrimination used to identify consumers with intense demand. The implication is that high aftermarket prices are counterbalanced by low(er) hardware prices. This, argue some economists, is not anticompetitive in the same sense that a restaurant ‘overcharges’ for wine or a pub ‘undercharges’ for food in order to maximise profits.⁴⁷ What is at issue is whether the total price of the hardware and maintenance package is set at near competitive levels, not the individual prices. Moreover, it is not the role of competition authorities to regulate component prices but to ensure that the total price is set at a competitive level. Others argue that the practice is inefficient in the broader sense by a) creating excessive hardware purchases and therefore encouraging buyers to economise on ‘overpriced’ aftermarket services, and b) leading to excess potential entry of independent service providers, and obviously a spate of antitrust actions.

In contrast it has been argued that aftermarket prices may be set too low. This was the allegation in the US *Microsoft* case where it was alleged that Microsoft’s bundling of its Web browser free with its Window operating system was predatory. The Department of Justice argued that operating systems (OS) and Web browsers were two separate markets, and that the bundling of the two products by Microsoft was an attempt to leverage its market power in the OS market onto the Web browser market. Microsoft countered that the two products were in the same market, and that the free provision of Internet Explorer was not anticompetitive because the marginal cost of adding IE to Windows was negligible and most likely negative because it boosted the sales of

⁴⁷ B Klein, ‘Market Power in Aftermarkets’ in F S McChesney (ed), *Economic Inputs’ Legal Outputs – The Role of Economists in Modern Antitrust* (1998).

Windows. The case pointed both to the complexity of the competitive issues, and the critical role that market definition can play in the enforcement of competition law.

The discussion can be turned into a checklist of factors as has been done in some recent guidelines by competition authorities. It is likely that, say, hardware and computer maintenance would form one market if all or most of the following conditions existed:

- effective competition in the hardware market
- a high proportion of current to historical sales of hardware
- initial purchasers know and can predict lifetime service costs
- transparent service prices
- many repeat purchasers
- low switching costs
- open service markets
- high degree of technical change leading to short life of equipment
- business rationale supporting metering justification
- evidence that hardware discounting policy supports metering rationale

Many of these factors have obvious implications; others less so. For example, View I places considerable emphasis on effective competition in the hardware market to discipline the supplier in the aftermarket. This force may be diminished if the proportion of new sales to installed subscriber base of hardware is low, so that the anticipated loss of sales due to high maintenance costs is less. Lost hardware sales are also unlikely to make an impact where the manufacturer is making no sales, losing market share and/or in a declining market. A supplier in a declining or obsolete market is likely to 'harvest' its captive customers by charging supra-competitive prices for secondary products/services.⁴⁸

High switching costs or infrequent purchase may dampen the link between vertical products/markets. The OFT's Guidelines state (para. 5.10) that:

If replacement is very infrequent or switching costs are high there may be a significant number of secondary product customers who are captive. Depending on the relative size of the primary market, even if new customers whole-life cost, the supplier may find it profitable to exploit these captive customers, implying that the secondary products will be a separate market.

This allows for the possibility that while some customers may be able to switch between products, others cannot, and that this may mean that the latter constitute a separate relevant product market.

D *Recent Cases*

The case law has at different times adopted the two different views on aftermarkets. The European Court has held that there are two separate product markets, and that the hardware purchaser subsequently becomes captive irrespective of whether there was effective competition in the market for hardware. Thus, spare parts (*Hugin v*

⁴⁸ M L Katz and C Shapiro, 'Antitrust in Software Markets' in A Eisenach and T M Lenard (eds), *Competition, Innovation and the Microsoft Monopoly: Antitrust in the Digital Marketplace* (1999).

Commission)⁴⁹, consumables (*Hilti v Commission*)⁵⁰ and computer maintenance (*Digital Undertaking*)⁵¹ have been defined as separate relevant markets. The US Supreme Court in *Eastman Kodak*⁵² was more moderate in its analysis but came to the same conclusion. It held that if there was effective competition in the hardware market, there could only be an abuse in the aftermarket if switching costs were high and if the information the purchaser received when the initial purchase was made was inadequate.

Some more recent decisions have taken the opposite view. For example, the OFT in *ICL/Synstar*⁵³ held that mainframe computers and hardware maintenance were one market. This case concerned a complaint by Synstar that ICL was abusing its dominant position under the *Competition Act 1998* by not providing access to its full diagnostic service, and thus foreclosing the market. The OFT concluded that the relevant market was mainframe computers, that ICL was not dominant in that market, and that hardware maintenance was not a separate relevant market because buyers take into account whole life costs when purchasing a mainframe computer. This was based on the fact that buyers are sophisticated, the costs of maintenance relative to that of the primary product was small, and price information transparent. The OFT dismissed Synstar's view that there was a significant base of legacy systems with 'captive customers'.

VI BUNDLING AND *UNITED STATES V MICROSOFT CORPORATION*

Increasingly the computer, and other communications products and services, are bundled together. This can occur horizontally where applications software is bundled together (Microsoft Office) or vertically when say an OS is bundled with applications software (Windows with a Web browser). Often such bundling is pro-competitive because consumers either want the products together, and/or there are economies of scope so that the bundled product is cheaper to produce and distribute than supplying each separately. However, where the firm has market power, bundling and tying raise the spectre of monopoly abuse and foreclosure. It is argued that a firm which has market power in one market can leverage this onto another related/neighbouring, or up/downstream markets where it is not dominant in order to increase its profits in the long run but with the immediate intention/effect of foreclosing the market to its downstream competitors. This, in essence, was the accusation against Microsoft. The difficulty in dealing with bundling and leveraging is that it is consistent with pro-competitive and anti-competitive behaviour. Obviously if the practice reduces the price or improves the service to the consumer (even if it enhances the profits of the supplier) then it benefits the consumer and is pro-competitive. This is the case even if it harms competitors. Harm to competitors is not the same as harm to competition (or consumers), and preventing the latter is the only legitimate goal of competition law.

⁴⁹ Case 22/78 [1979] ECR 1869.

⁵⁰ [1991] ECR II-1439.

⁵¹ Commission Press Release IP/97/868.

⁵² 122 S C T (1992).

⁵³ Case No CA98/6/201 (20 July 2001). The OFT stated that it preferred to follow the EC Decision in *Pelican v Kyocera*, *Competition Newsletter*, Vol 1.

A *Theory of Leveraging*

There are a number of difficulties with the vertical leveraging argument which have been well rehearsed in the literature. The first, and most persuasive, is that it is an abuse searching for an explanation. That is, why would a firm with market power in one market seek to use it in another market where it has no market power. It could simply extract the monopoly profits in the market in which it is dominant!

This simple observation forms the basis of the Chicago School's critique of vertical restraints. The Chicago School position is that the vertically leveraging of market power is not undertaken for anti-competitive reasons.⁵⁴ This is because there is 'only one monopoly profit' in the vertical chain of product distribution which can be reaped by charging a monopoly price for the input. The upstream monopolist can not therefore generate even higher profits by leveraging its market power downstream, because if it is competitive (as implicitly assumed) there are no additional profits which can be extracted, except at the 'cost' of sacrificing upstream profits. If there is a monopoly profit it will extract this upstream.

This is even true when the downstream rivals are more efficient than the vertically integrated firm's downstream division. In this case it would not 'pay' the upstream monopolist to raise the essential input price to levels which would be exclusionary and drive its rivals out of business. The upstream monopolist would raise the input price to rivals above the (own) profit maximising level to extract the profits from their superior efficiency. However, the resulting squeeze on the rivals' margins would fall short of seeking to foreclose the market to efficient downstream rivals. Such an exclusionary abuse would decrease the monopolists' profits.

The post Chicago literature accepts the Chicago School's conclusions but argues that they cannot be generalised. It argues that under some circumstances the vertical leveraging of market power can increase overall profits. This is particularly the case where the downstream market is not effectively competitive. In this case the vertically integrated firm may have an incentive to exclude its downstream rivals.

B *United States v Microsoft Corporation*

The issue of bundling arose both in the *IBM* case of 1970,⁵⁵ and the *Microsoft* case initiated in 1998.⁵⁶ Although the issues raised in *Microsoft* were not exclusively vertical, the US Department of Justice's allegation that Microsoft sought to vertically foreclose the market for computer operating systems was a key plank of its case. It was alleged that Microsoft saw that the Web browser might enable Netscape to enter the OS market via the Internet and so-called 'middleware' applications. Specifically, a Web browser together with 'platform independent' software such as Sun Microsystem's Java, might allow Netscape to develop an alternative Internet based OS that would not require it to be purchased with a PC. Microsoft is alleged to have systematically sought to wipe

⁵⁴ See R H Bork, *The Antitrust Paradox: A Policy at War with Itself* (1995).

⁵⁵ *US v International Business Machines Corp* Docket no 69 Civ (D N E) Southern District of New York; F M Fisher, J J McGowan and J E Greenwood, *Folded, Spindled, Mutilated - Economic Analysis and US v IBM* (1983).

⁵⁶ *United States v Microsoft Corporation* [2001] Trade Cases CCH 73,321.

out Netscape's browser, and perhaps Netscape, by seeking to establish a dominant position in the provision of Internet browser software. As part of this alleged strategy, Microsoft distributed Internet Explorer browser free of charge and bundled with its OS. This bundling strategy was alleged by the Department of Justice (DOJ) as predatory.

In the Court proceedings expert economists on both sides presented very different explanations and assessments of Microsoft's practices. Professor Richard Schmalensee, an expert witness for Microsoft, argued that because the two products were complements, it made good business and economic sense to bundle these and give one away free to stimulate demand for the other. The proper assessment was therefore not the price of the individual components but that of the bundled product. Schmalensee argued that because the marginal cost of producing additional copies of Internet Explorer was near zero, and maybe negative because it generated additional revenue from sales of the OS, the zero price was above its (net) marginal costs and therefore not predatory.

The more serious accusation was that the bundling was an attempt to foreclose both the OS and Web browsing markets. The DOJ's case was built on network effects constituting a barrier to entry. As discussed above OSs and software form a 'hardware-software' system of complementary products. The more software there is for a specific OS, the higher the demand for that OS (indirect network effects). Thus the OS are characterised by network effects and software by economies of scale (public goods). The tens of thousands of applications written for Windows in the DOJ's view constituted what it called an 'applications programming barrier to entry'. Faced by firms trying to enter with new OSs, Microsoft allegedly bundled its Web browser to effectively squeeze downstream rivals to deter them from entering the OS market.

Microsoft countered this allegation by questioning the definition and existence of barriers to entry. It argued that a barrier to entry exists if and only if it prevents an equally or more efficient entrant from competing effectively with a less efficient incumbent. Under this definition, sunk costs, switching costs and network effects are not barriers to entry because they do not prevent an equally or more efficient downstream firm from entering the market.⁵⁷

As part of the analysis of bundling and foreclosure the parties' defined the relevant market differently. The DOJ's experts proposed that the relevant market was for OSs for Intel-compatible computers. This was based on the method for market definition set out in the 1992 *Merger Guidelines* (as discussed above). Since Microsoft had over 90% market share of this market, it was easy to argue it had considerable market power. Microsoft argued that the relevant market was a broader one known as the 'Wintel standard'. Its economic experts approached market definition by challenging the standard approach and advocating a behavioural approach based on observing whether a firm is behaving competitively or not. On this basis, Microsoft Windows software was part of a broader market for a number of reasons. First, price equal to marginal cost does not make sense for software, since this would mean pricing equal to small shipping costs, and hence insufficient to recover the vast development costs. Second, estimates of the short-run elasticity of demand for Microsoft Windows implied that a monopolist

⁵⁷ Cf G J Werden, 'Network Effects and the Conditions of Entry: Lessons from the Microsoft Case' (2001) 69 *Antitrust Law Journal* 87.

would have charged a price allegedly 16 times higher than its US price of about \$100.⁵⁸ Others have observed that while this may be correct in the short-term, Microsoft probably has the ability to charge prices that are extremely profitable. An alternative interpretation is that the low prices are based on penetrative pricing in network industry as well as concerns about entry.

VII INTELLECTUAL PROPERTY RIGHTS

No discussion of the computer industry would be complete without reference to the interrelation between intellectual property rights (IPRs) and competition law. This has been an area of increased activity by competition authorities and governments given that in most advanced countries the concept of property and indeed business is increasingly based on the exploitation of patents, copyright and contractual protections. In Australia the Federal Government has reviewed this area.⁵⁹

The discussion of IPRs is in most respects the institutional counterpart of the issues discussed above in relation to public goods and dynamic efficiency. It thus raises no new economic concepts apart from the role played by legally enforceable rights in improving or impairing market efficiency and competition.

The economics of public goods provide the best way of examining IPRs. IPRs are the institutional mechanism which seek to create exclusivity for public goods. It gives the rights owner legal protection which through legal enforcement forces users to pay, prevents free riding and gives the rights holder the ability to appropriate the benefits/revenue for the exploitation of their intellectual property. The economic and legal justification for IPRs is that in giving temporary exclusivity, they enable inventive and innovative activity to be rewarded, and in turn encourage future invention and innovation. The argument goes that in the absence of such legal protection, others would free ride on the often huge research and development costs, thus expropriating the value, and thereby reducing the incentive to innovate.

There are nonetheless concerns over the terms and duration of IPRs. Even among economists there is a strand of economic literature best associated most with the British economist Arnold Plant which questions the economic justification for patents and IPRs.⁶⁰ Plant argued that the large investment and lead time often gave the creator sufficient return without the need to extend this excessively by legal protection, and thereby turn the IPR into a genuine barrier to entry. While most would not agree with Arnold Plant that they are unnecessary; they would agree that there is a stage when the IPR changes from a legitimate protection of private property to a monopoly right which can act as a legal barrier to entry with offsetting adverse effects. The rights holder can abuse this monopoly by charging monopoly prices, although this effect will be ameliorated if price discrimination is permitted. Further, the rights holder can seek to foreclose the market by denying access which extends beyond that necessary to encourage inventive activity. When this happens there will have been welfare losses

⁵⁸ Testimony of R I Schmalensee, *United States v Microsoft Corporation*, CA No 98-1233 (TPJ) 163 (Jan 11, 1999).

⁵⁹ Intellectual Property and Competition Review Committee, *Review of intellectual property legislation under the Competition Principles Agreement, Final Report* (September 2000).

⁶⁰ A Plant, *Selected Economic Essays and Addresses* (1974).

because dissemination and use of the software will be restricted below the socially optimal.

VIII CONCLUDING REMARKS

The computer industry poses a challenge to and is challenging for the competition law. The characteristics of the products produced by the industry result in practices which can have depending on the facts alternatively pro- and anti-competitive effects. As a result, competition law intervention in the computer industry has often been a deeply controversial and contentious matter dividing lawyers, economists, regulators and politicians. As a number of commentators, and indeed courts and regulators, have noted the economic literature is not yet sufficiently developed to provide definitive answers, if indeed it can ever hope too. Yet, at this formative stage, it does point to the correct questions and the types of evidence and facts which would assist in selecting from the competing views.