Network industries in the new economy

In this paper we discuss two propositions: the supply and demand of knowledge, and network externalities. We outline the characteristics that distinguish knowledge-intensive industries from the general run of manufacturing and service businesses. Knowledge intensity and knowledge specialisation has developed as markets and globalisation have grown, leading to progressive incentives to outsource and for industries to deconstruct. The outcome has been more intensive competition. The paper looks at what is potentially the most powerful economic mechanism: positive feedback, alternatively known as demand-side increasing returns, network effects, or network externalities. We present alternative demand curves that incorporate positive feedback and discuss their potential economic and strategic consequences. We argue that knowledge supply and demand, and the dynamics of network externalities create new situations for our traditional industrial economy such that new types of economies of scale are emerging and 'winner takes all' strategies are having more influence.

This is the first of a pair of papers. A second paper will take the argument further and look at the nature of firms' strategies in the new world, arguing that technology standards, technical platforms, consumer networks, and supply chain strategies are making a significant contribution to relevant strategies within the new economy.

John McGee, Professor of Strategic Management, and Academic Director, Full Time MBA Programme, Warwick Business School, University of Warwick, Coventry

Tanya A. Sammut Bonnici, Researcher on Network Strategies and Corporate Evolution, Warwick Business School, University of Warwick, Coventry

Introduction

The popular business and financial press has produced endless discussions of the high-tech industries, the dot.com revolution, and the 'new economy'. The academic literature in turn has seen much discussion of 'increasing returns' and 'positive feedback effects', especially since Arthur's two seminal papers in 1989 and 1990. With the bursting of the dot.com bubble and apparent overselling of high-tech stocks, particularly telecommunications stocks, it may seem that a high-technology bubble has been and gone and that normal times have returned and we can revert to normal economics. This paper argues that this is too simplistic a view and too sanguine by half. We argue first that knowledge-intensive industries are a fact of life and that their economic characteristics have significant and enduring implications. We acknowledge that microelectronics has had a powerful impact on the nature of competition through commoditisation of products and innovation in distribution channels. However, the more powerful strategic impact has to be seen through the new industry structures now evident in the supply chains that run from telecommunications through to new dot.com companies. We put knowledge intensity and the new industries alongside the notion of positive feedback effects to argue that this is the platform from which major new competitive forces are emerging. These sit alongside traditional industry structures but also contain important new strategic lessons. It is from this vantage point that we can observe a new strategic logic at work that requires new approaches to the analysis of industries, a different and novel conception of generic strategies, and new ways of strategising in companies. This is not entirely new. The idea of positive feedback effects and increasing returns has been around for decades. However, its effects have been seen as derived from the supply-side increasing returns and natural monopoly characteristics of traditional industries. It is the conjunction of these effects with knowledge specialisation and knowledge intensity that marks a distinctive change.

This paper discusses the supply and demand of knowledge and network externalities. We argue that these two factors create new situations for our traditional industrial economy (the old economy) such that

new types of economies of scale are emerging and 'winner takes all' strategies are becoming prevalent.

Our second companion paper will take the argument further and look at the nature of firms' strategies in the new world (the new economy), arguing that attention to technology standards, technical platforms, consumer networks, and supply chain strategies have created a significant contribution to new strategies within the new economy.

Knowledge economics in the network economy

In the past fifty years, the world has witnessed the rise of information and communication technologies (ICT). It has changed the nature of the commercial environment and has increased the speed of operation, driven by the increase in connectivity between companies, and between companies and their customers. Network industries, such as telecommunications, the Internet, computing and software, have had a significant role in shaping the corporate environment into a new network economy. (The structure and nature of network industries and the network economy is discussed in our second paper.)

Do the same economic 'laws' hold for the new network economy as for the traditional industrial economy? This question has attracted much attention in business and in academia. The popular press has been excited by the possibility of a New World in which the old economics is somehow stood on its head. Academic economists have generally been more cautious. Shapiro and Varian (1999) in their excellent book on the information economy recognise these dynamics but assert 'Technology changes. Economic laws do not.' The impact of the Internet and computer networks nevertheless has created a new type of market failure driven by technology, high-risk network externalities (discussed later on), and zero marginal costs (Dumont, 2000).

We are witnessing a new commercial scenario: networks no longer belong to a single firm. Strategy goes beyond the efficient use of network structure and the relevant allocation of cost. Collaboration between firms has become mandatory for intra-network compatibility, a feature that is crucial in telecommunications. Collaboration and anti-monopoly pressures have led to a shift towards fragmented ownership and oligopolistic circumstances. Underpinning networks and the growth of collaboration are the following well-known characteristics of information goods, such as digitised information as in recorded music, software, football scores, encyclopaedias and telephone directories:

1. High fixed costs but low to vanishing marginal costs, thus high costs of creating intellectual property, but low costs of reproduction.
2. Low costs of copying intellectual property.
3. Information is an experience good every time it is consumed.
4. With easy access to information, there is information overload – value arises from location, filtering and communicating what is useful. Search engines facilitate this.
5. An extensive, expensive technology infrastructure is required to produce and distribute information.
6. Pricing is value-based, rather than cost-based.

However, these characteristics on their own do not produce exceptional results. They do result in an emphasis on volume and a tendency in certain circumstances, for example in fragmented markets, for price to fall precipitously towards marginal cost, i.e. to zero. But information products are susceptible to differentiation to convey quality signals and endorsement (e.g. the Encyclopaedia Britannica name). Such products are customised through timing (e.g. Amazon's mail shots through its Permission Marketing programmes) or individual customer data in the form of personalised web sites. At the same time the low marginal cost characteristics of information goods make it attractive to create dedicated distribution channels for information transfer (one-way or two-way) through which information differentiation can also be attempted. This differentiation has provided the basis for e-commerce and the growth of the Internet.

On the basis of these characteristics, there has been a systematic and ever-increasing shift from the traditional industrial economy to a knowledge-based or information economy. The next section outlines the basis for this evolution/revolution.

The demand and supply of knowledge

A key strategy in developing knowledge industries is to create a capability to analyse an organisation's current knowledge-processing environment, both on the 'supply side' of securing and providing existing knowledge and on the 'demand side' of knowledge creation. The core of the knowledge management challenge lies in creating a perfect balance of knowledge supply and demand. Google.com, the leading Internet search engine

company, is an excellent example of this dynamic. (For more details see the article by Fred Vogelstein in *Fortune*, 27 May 2002.) The company's founders, Brin and Page, who left Stanford's PhD programme to start the company, recognised that a network could be built from existing outsourced commodity hardware. In this case, knowledge was extracted from its original appropriators (server, software and telecommunication companies), and diffused by Google, who could bundle and replicate their knowledge. More capital-intensive processes followed simple replication of this kind. What came next was the combination with other knowledge, such as revenue earning advertising practices, to assemble a new goods package that came to be known as the Google market.

As in Google's case, an important point of change in the new economy is that knowledge is being made explicit, being appropriated by others, being diffused and replicated, and is becoming industrialised. Knowledge is produced in large volumes at lower cost. External knowledge providers (service companies such as IT, billing and consultancy providers) are changing the nature of the firm, as they provide a substitution for the tacit knowledge base previously controlled by the firm. The implication is that vertical integration and diversification patterns will become unsustainable. As knowledge changes, proprietary links give way to market relationships under labels such as outsourcing, deconstruction and the hollowing-out of the corporation.

As knowledge production has grown in response to market growth over the past 25 years, it has become more specialised and more diverse in its sources. It has moved from being a cottage industry to being an industrialised activity. Portions of knowledge are originally produced and appropriated for use by individuals or individual firms. How this outside knowledge is used and absorbed by a company has become an important success factor. Google's flexible system, based on the knowledge of several suppliers, reduced the risk of becoming outmoded by rapid developments. Its knowledge providers, or suppliers, handle the strain and expense of technological change. Google managed to build a business based on searches where Alta Vista, Yahoo, Lycos, Excite and Infoseek had failed.

Value chains: from integrated to reconstructed networks

The serial decisions of 'assembly' versus the 'purchase' of customised products, which is exemplified by the Google case, reflect changing economic circumstances and shifts in relative costs. There has been an extensive change in the nature of knowledge investment and production leading to a further change in the supply and demand for knowledge. Quinn (2001) describes these pervasive changes in six distinct phases.

In the first phase, economies of scale are created as large organisations capture key activities, leading to the demise of smaller enterprises that lack capital and expertise. In the second phase, economies of scope come about as the same technologies that created the economies of scale allow the handling of more data, more output functions, or more customers without corresponding cost increases.

However, changes in the fundamental conditions of demand and supply of knowledge can lead to the next phase – disintermediation. Disintermediation is the process by which proprietary links within the firm give way to the coordination mechanisms of markets. It refers to the piecemeal replacement of internal activities by external provision. Where this replacement is systematic and extensive it is known as deconstruction – the process of systematically undermining fundamental concepts (in this case the logic of vertical integration). Next, we enter a phase of deregulation in the sense that increased competition replaces the regulation imposed through vertical integration and semi-monopolistic competition. In this scenario, new competitors with new knowledge make cross-competition more possible.

Finally, there are rounds of redispersion and redecentralisation due to a reassertion of the need for more localised and personalised contact as new forms of broking, selling, and agencies emerge.

To summarise, the systematic accretion of knowledge and its diffusion around the world has had the effect of forcing a restructuring of industries even to the extent that entire economies have been fundamentally changed. The open standards and the universal connectivity inherent in information technology enable knowledge modules to be 'snapped together' similar to a Lego system, without any expensive customisation or reworking.

With the widespread acceptance of international division of labour in knowledge, there are further characteristics inherent in the new economy. The first is that knowledge-based functions are significant elements in most value chains and that these are significant in size and critical for competitive advantage. The second is that the pressures of competition mean that simple technical efficiency is

not enough. There are opportunity costs of not buying from the most efficient suppliers and outsourcing becomes not only attractive but also necessary. Third and most important, firms need to focus their strategic investments and their management attention on those knowledge-intensive activities that form the basis for sustainable competitive advantage. This means that it is no longer sufficient to maintain a portfolio of competences in which most of the competences are at best only co-specialised with the core competences. Historically firms have invested capital in those journeyman competences that were required to bring the truly distinctive assets, the core competences, to market. Now, however, they do have the prospect of replacing them with lower cost or more effective substitutes from external suppliers.

Through this mechanism we are witnessing the development of three new business models: the new competitor; the deconstructed value chain; and the reconstructed value chain (Evans and Wurster, 2000).

The *new competitor* mounts direct attacks on established businesses by splitting the information flows from the physical flows. Thus, the success of business hinges on two levels of transaction: the traditional flow of products to the supply channels, and the flow of information from the company to the customer, via channels such as the Internet. Figure 1 illustrates these two levels.

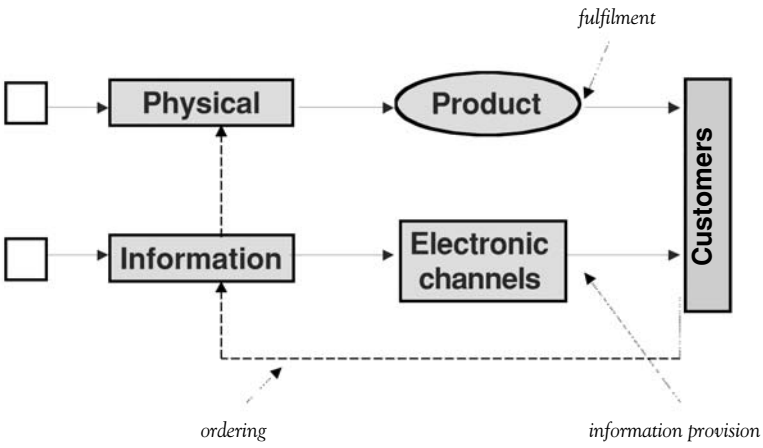
In Amazon.com the business model is not simply the replacement of an expensive shopping process (for the customer) by cheap electronic means, but also an approach of differentiation aimed at turning a direct online bookseller into a virtual focal point.

The approach of Egg in the UK to online financial services is much the same. The UK financial services firm Prudential launched Egg in 1998 to offer products and services in banking, investments, insurance and online shopping. Egg is designed for the Internet generation using the Internet as the primary medium of consumer contact. It has a strong focus on technology by combining the characteristics of a conventional direct business with automated customer relationship management.

In both the cases of Amazon and Egg, the very high fixed costs of online selling are defrayed not just by high volumes but also by the economies of scope that follow from diversification channelled through the online shopping point. Evans and Wurster (2000) in their well-known book about deconstruction and being 'Blown to Bits' make the same point with an example about separating the information-rich part from the commodity part of the value chain.

The *deconstruction model*, in Figure 2, stems from the need to focus the firm's attention on those few, typically knowledge-based activities that underpin long-term competitive advantage. To do this the firm has to redefine its remaining capabilities as activities which can be bought in from 'best-in-class' suppliers. This applies equally to overhead 'services' as much as it does elsewhere in the value chain. In this way the firm becomes less vertically integrated in the conventional way – it has to deny the old saying that 'what it does not own it cannot control'. However, it does maintain its control of

FIGURE 1: The new competitive model.



the value chain and of the advantages accruing to the customer by actively reinforcing the core competences which it has retained (more on this below) and by investing in the management of its outsourcing so as to maintain its bargaining power with its (new) suppliers and partners.

In Figure 3 there is a battle for control in the deconstructed model of the supply chain – a battle for ownership and control of competitive advantage. One mode of operation occurs when integration gives way to *orchestration*. Successful orchestrators possess powerful brands and other core competences that give them competitive advantage by virtue of which they can control the terms of supply (Figure 3). Nike and Hewlett-Packard are examples of this mode. The orchestrators retain sufficient idiosyncratic capital to preserve some degree of vertical integration sufficient to exercise power throughout the supply chain. But control over the supply chain depends on the location of knowledge in the chain.

Those players who focus on a specific value-added step have incentives for scale and scope effects with the possibility of wresting control from the traditional integrated players (Figure 4). Intel and Microsoft did this to IBM because IBM was not able to control the IBM standard for PCs.

Instead Intel's microprocessor technology and Microsoft's software represented the key knowledge assets which then dominated the supply chain. In the extreme case, the integrated firms deconstruct entirely with each value-added step in the supply chain becoming a business in its own right. Competition is then fragmented, products become near-commodities, and economic rents are minimal and transitory.

The *reconstruction model* has two elements. The first applies existing core competences into other value chains to establish new economies of scale there and in doing so creating new economies of scope (Figure 5).

This is familiar in that it replicates well-known processes of related diversification. But it is different in that it represents an attempt to dominate other apparently related supply chains with existing knowledge-based competences. In this process, the nature of scope has changed from product-market relatedness to knowledge (or resource) relatedness. In this resource-based view of the corporate portfolio, competition is as much a competition for competences and for knowledge as its more familiar application to products. Evans and Wurster have dubbed this the 'rewiring' of the firm in which knowledge-based competences have become the

FIGURE 2: The deconstruction model: disintermediating.

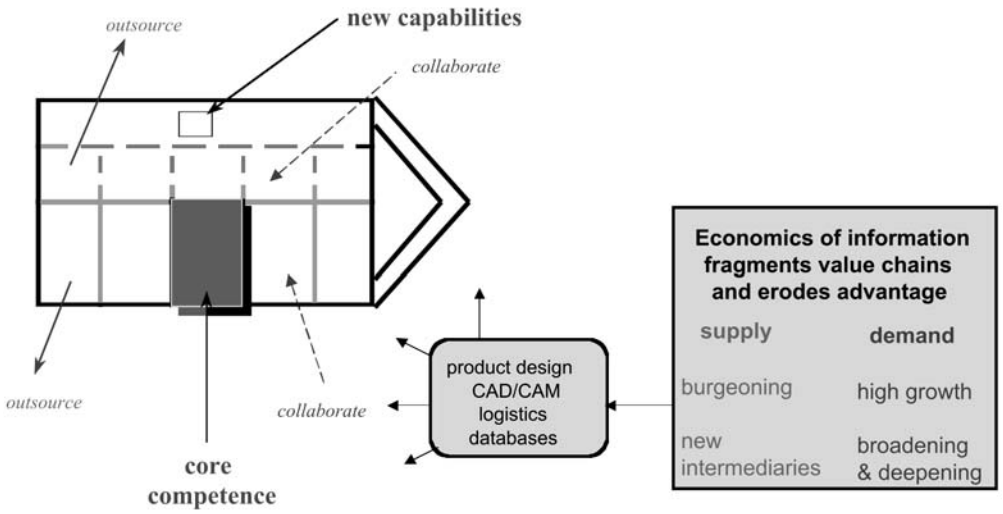


FIGURE 3: The deconstruction model: orchestrating downstream.

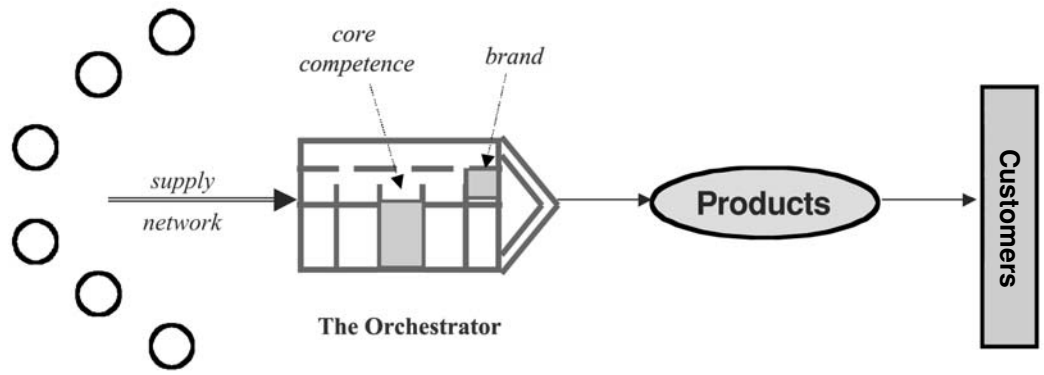
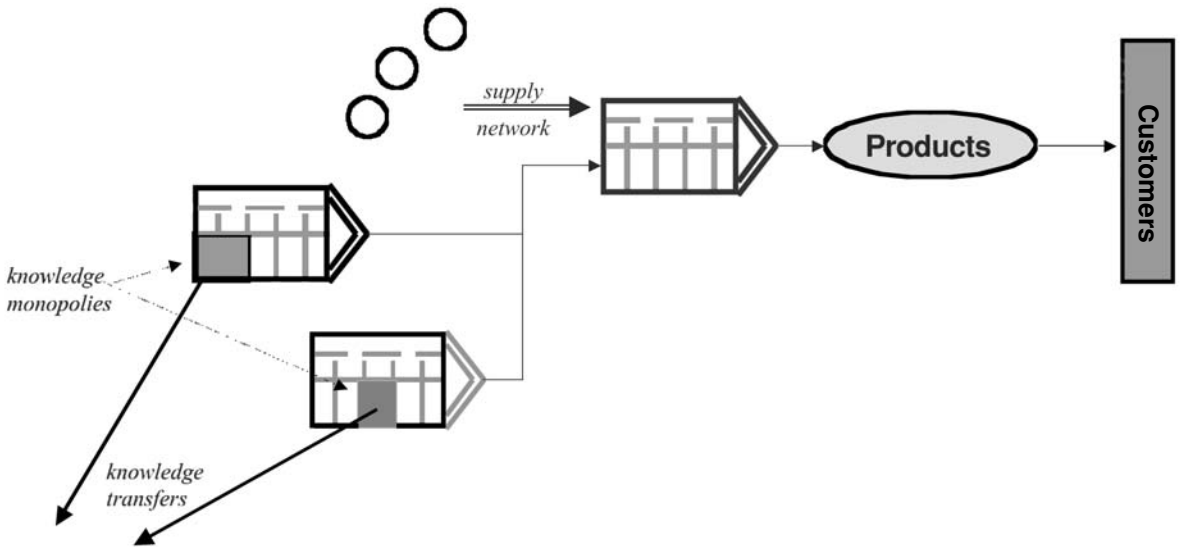


FIGURE 4: The deconstruction model: orchestrating upstream.



controlling element in multiple supply chains through several highly focused (i.e. short) value chains (Figure 6).

The second element is the creation of a new set of corporate level capabilities whose purpose is to identify and manage the set of collaborative relationships that make up the web of partners and strategic suppliers. Indeed the vertical integration metaphor of the value chain gives way to the language of the *value web* (Figure 7). The centre of the web contains the new corporate glue (idiosyncratic knowledge) that maintains the new style of portfolio positions across value chains and across industries and orchestrates strategic linkages so as to retain control over the traditional value chains. The points of leverage for this core competence are the specific knowledge-based assets that are applied across different industries. In this new game compa-

FIGURE 5: The reconstruction model: knowledge transfer.

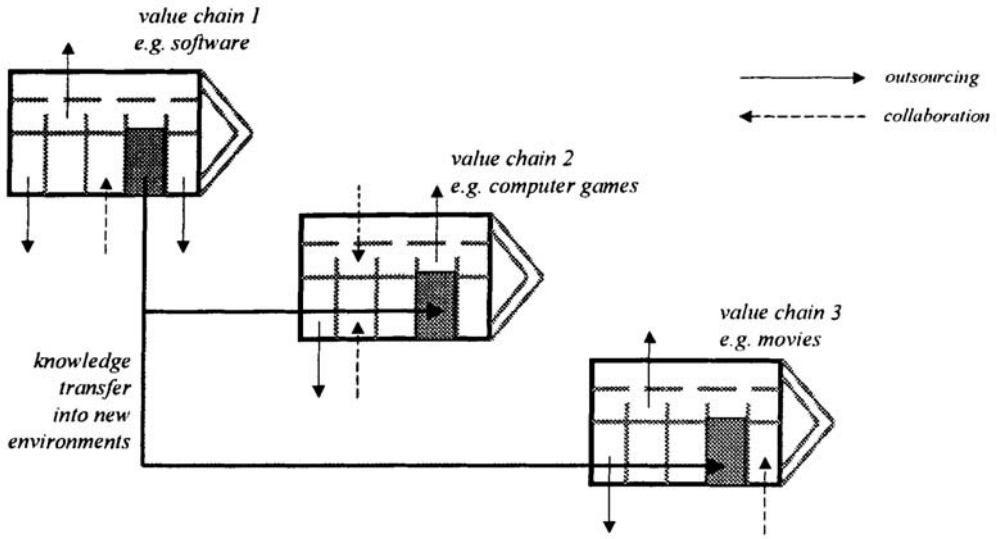
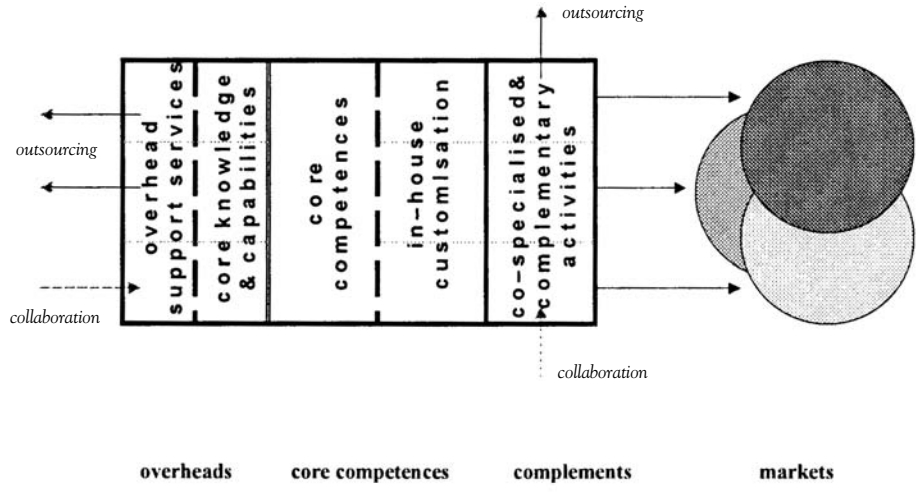


FIGURE 6: The reconstruction model: the composite new value chain.



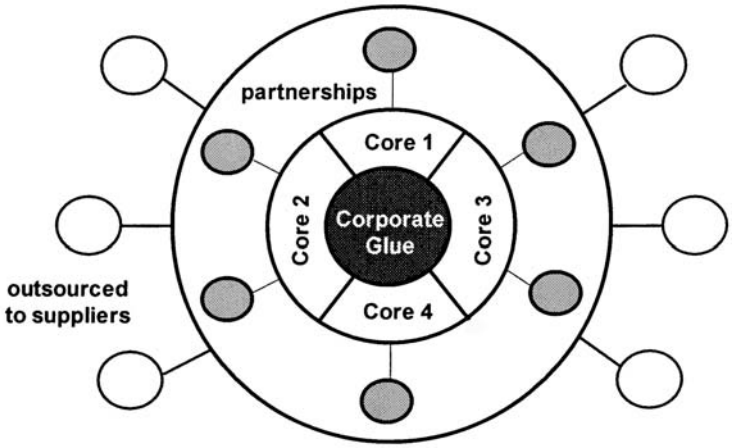
nies can develop a much higher degree of focus in applying their strategies through their knowledge-based assets rather than through traditional product market strategies.

In general, the focus of strategic thinking has shifted from domination of markets and of supply chains through a vertically integrated set of activities to an assertion of leading positions in key intellectual assets coupled with new thinking at the corporate centre about knowledge management, risk assessment, and the management of information and knowledge. The focus has shifted from products and product classes to activities, knowledge and competences.

Positive feedback in network industries

In this section we look at how the old industrial economy was characterised by economies of scale

FIGURE 7: The value web.



and scope, whilst the new information economy is driven by the economics of positive feedback in network industries. We will discuss the nature of network industries, such as railroads, telecommunications, software and hardware networks. We will see how network companies benefit from positive feedback on both the demand side and supply side. On the demand side, the more customers join a network, such as a telecommunications service, the higher the incentive for other customers to join. On the supply side, the larger a network becomes in terms of users and also in size of assets deployed, the easier it is for a company to lower costs and prices. The lower the price introduced by a network company, the more subscribers will join the network and positive feedback kicks in. The result is a positive spiral. The importance of critical mass, competition and standards is discussed in the light of the dynamics of positive feedback.

The Old World: economies of scale and scope

In the Old World of the industrial economy, firms seeking market power could do so by virtue of barriers of entry created by economies of scale. Saloner et al. (2001, p. 138) identify three types of entry barriers: barriers from production or distribution technology, barriers from brand name or reputation, and legal barriers. The first two are essentially cost barriers in that replication of the incumbent's assets is inhibited by the costs of so doing. The third type is an absolute barrier that arises from institutional characteristics.

Diversified firms required an additional force: economies of scope. These are defined as 'the cost savings realised when two different products are produced within the same organisation rather than at separate organisations' (Saloner et al., 2001, p. 364). The products would share a common input such as plant or equipment, obtaining volume discounts on purchases, or applying common expertise or reputation. In such a situation, where competition is monopolistic (or imperfect) producers may attempt to shape customer preferences. They may succeed in modifying the demand function to become downward sloping in the conventional manner and producers can then price according to the nature of their marginal cost curves and the price elasticities in the market. However, demand and supply are mediated through a market mechanism in which product demand is independent of other products and demand is not time dependent. This exaggerates the point, as we will see later when we discuss product complementarity.

There is a class of markets and industries that do not conform with the assumptions of the Old World. These are the network industries. In this discussion the terms 'industry' and 'market' are used as if interchangeable.

The New World of network industries

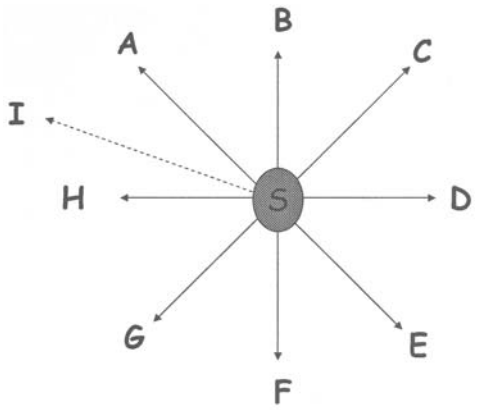
The concept of network can be segmented into real and virtual networks (Shapiro and Varian, 1999). Real networks are found in industries such as telephony and railways where a physical network is

present. Virtual networks are typified by computer and software platforms where the interconnection between users is intangible. The two types of networks are discussed below.

In real networks the interconnection between users is tangible. Examples are cable networks for telephone users and radio transmissions in mobile phones. Electricity grids, telecommunications networks encompassing telephones, fax machines, online services, and the Internet, are typical examples of products or services within real networks. There are one-way networks such as broadcast television where information flows in one direction only. In two-way networks, such as railroads and telephone systems, links are operated in both directions. Any network may be viewed as a set of connections (links) between nodes. A two-way network allows the links to be operated in both directions whereas a one-way network has specific direction. Two-way networks include railroads and telephone systems.

Figure 8 shows a simple star network where A can communicate with B through a switch S. B can also communicate with A by reversing the direction of the link (viz. a telephone call). In Figure 8 we have eight nodes (A through G) linked through a switch S. If this were a two-way-network AB and BA would be distinct products (different telephone calls, different rail journeys). The total number of products in the network would be 56, i.e. $n(n-1)$ where n = the number of nodes. If there were to be a ninth member (the dotted lines to I in Figure 8) this would increase the total number of products to 72 (n is now 9), a total increase of 16 products available from the expanded network. If the value to each user of being in the network is proportional to the number of users, then the value of this network has just increased by 28.5% (16 as a percentage of 56) even though the size of the network has increased by only 12.5% (one added to eight). This is an algebraic characteristic of network economies of scale that *the value rises disproportionately higher than the increase in network size* as long as prices are constant and products are independent. Intuitively we might expect that an increase in network size beyond a certain point has little value. (Using calculus we would expect the first derivative to be positive but the second derivative to be negative. Therefore total value increases but at a decreasing rate.) If this network were a one-way network there would be half the number of products but the value of the network would nevertheless increase at the same rate but achieving only half the value.

FIGURE 8: A two-way network.



In virtual networks the interconnections between users are intangible, but users remain interdependent. Computer systems are typical of virtual networks. For example, Mac users are part of the Mac network, with Apple as the sponsor of the network. Mac users are locked into a network determined by the technology standard of this platform. They can only use software that is compatible to the system and will exchange files with users within the system. Operating systems such as Windows and Unix are other examples of virtual networks. Virtual network dynamics also operate in the entertainment industry for Sony Playstation, Microsoft Xbox and Nintendo's Gamecube networks.

Network size is still important in virtual networks in that a large consumer base makes production viable and usage possible. In addition, the value of a product increases as the number of, or the variety of, the complementary goods or services increases. Indirect network effects in the computer industry are referred to as the hardware–software paradigm. The success of an operating system for personal computers depends on the variety of software applications available in the market. Value may depend more critically on software applications.

Network externalities: the new economic forces

Earlier in the paper we have looked at the traditional economic model for the 'Old World' which

was driven by economies of scale and scope. The 'New World', characterised by information and communications technology is governed by a different dynamic. Network externalities are the new drivers of the network economy. It is important to recognise that economies of scale/scope and network externalities represent the extreme ends of a spectrum of effects, and that the presence of one does not imply the exclusion of the other. Companies may experience the effects of both to varying degrees, with a tendency for network externalities to have more strategic relevance in the new network economy.

The concept of network externalities has attracted the attention of academics and practitioners alike. The extent to which network industries have proliferated in the economy is a recent phenomenon. The effects of network externalities, however, have been recognised for some time with the development of the older network companies such as the railroads and the electricity systems. (In 1804 Trevithick constructed the first practical locomotive in England. In 1882 the Edison Electric Lighting Company completed the first commercial generating station at Holborn Viaduct in London. The first commercial telephone line was installed in Boston, Massachusetts in 1877.)

Network externalities are defined as the increasing utility that a user derives from consumption of a product as the number of other users who consume the same product increases (Katz and Shapiro, 1985). For example, the more people there are in a telephone network the more users can be reached on the network, thereby increasing its usability. Fax machines, broadcast industry services, credit card networks, and computer hardware and software are examples of products exhibiting network externalities.

According to Economides and Flyer (1997):

> The value of nearly every good is influenced by aggregate consumption levels in its market and in the markets for related goods. In many cases, high aggregate consumption in its own market, and in the markets for complementary goods affects positively the value of a good. Traditionally such effects have been called network externalities, since they were first identified in network industries. While such effects are salient in some markets, such as for telephones, fax machines and computer operating systems, for most goods these influences are more subtle and tend to be smaller.

Complementarity in network products

An important element is the notion of complementarity; thus the value of a railway station is derived from the existence of other railway stations on the network. A weaker definition relies also on complementarity between products (or nodes, in network language) but allows the links to be created *by* the customer rather than *for* the customer. Economides and Flyer (1997) have some powerful examples:

> ... the value of a washing machine is affected by the aggregate consumption of washing machines and the consumption level of the particular brand, since this determines the availability of parts, repairmen, detergents, fabric softeners and various other related goods and services. The value of a sporting event is influenced by the aggregate size of the audience, as this enhances the excitement level, analysis, discussion and remembrance of the event. Even a grapefruit is influenced by network externalities, since the variety of accessible complements, such as peeler, slicers, juicers, recipes, nutritional information and specialised spoons, are affected by the aggregate consumption of the fruit.

The essence of this idea is that the demand for a product is influenced by total demand for the product class or by total demand in a complementary product class. Thus demand is conditioned by a consumer externality. Where these consumer externalities are powerful, the feedback effect on demand is such that there is a tendency towards a single network, or platform, or standard. The value for consumers of being on a common standard outweighs any specific differences between alternative standards. We see that the VHS video-recording standard was preferred to a 'technically better' Betamax rival to the extent that the rival standard disappeared. The Wintel standard is greatly preferred to the Apple standard and the rival exists only as a small niche in the market. Where the externality is smaller and the intrinsic difference between standards is relatively larger then we might observe multiple competing and coexisting 'platforms'. (To observe multiple standards defies common sense, hence the term 'platform' which denotes an array of linked complementary products that together are compatible with other products.) An example of a platform can be seen in the automobile industry where a company might develop a core of components and sub-assemblies that can be used to support alternative body styling to create a

product range. Such a platform can coexist with other platforms because the scale efficiencies associated with platforms are modest in relation to market size.

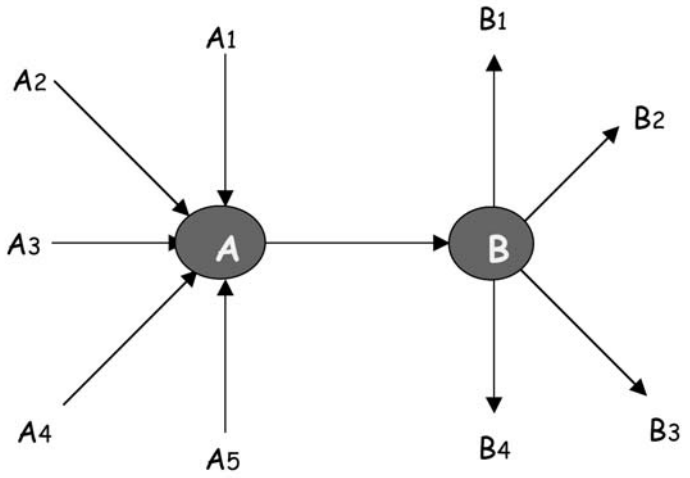
The analysis of complementarity is equivalent to the analysis of a one-way network. Figure 8 can be extended as in Figure 9 to show a typical one-way network. Here we can interpret the A_i as automatic teller machines (ATMs) and the B_j as banks. The network runs only from A to B, i.e. the two switches have only one link. This means that there is compatibility between all ATMs and all banks. This maximises the value of the network but increases the competition between banks for customers through ATMs. (Two complementary components A and B are compatible when they can be combined to produce a composite good. A VHS player is compatible with VHS tapes. Two substitute components A1 and A2 are compatible when each of them can be combined with a complementary good B to produce a composite good. Thus two VHS tapes are compatible, and two VHS players are compatible.) It is this compatibility that makes the complementarity actual and the network operational. For complex products actual complementarity has to be achieved through adherence to specific technical standards. Other complementary products can be visualised in terms of Figure 9. VHS tapes could be the A_i and VHS players could be the B_j. Think also of copier paper and copiers, or printer paper and printers, or car accessories and cars, or local and long distance telephone networks.

Networks were originally analysed on the assumption that each network was owned by a single firm and research concentrated on the efficient use of the network structure and on the appropriate allocation of costs (Economides, 1996; Sharkey, 1993). With the anti-trust cases against AT&T and its later break-up, attention shifted towards economies of scope, the efficiency gains from joint operation of complementary components of networks (Baumol et al., 1982). This led to issues of interconnection and compatibility in parallel with the reduced role of IBM in the 1980s and 1990s in the setting of technical standards in computer hardware and software. As technology has advanced, there have been significant reductions in telecommunications costs and a shift towards fragmented ownership of telecommunications networks. Market structure has shifted from natural monopoly to oligopoly. Similar trends are evident in other IT-intensive industries. Thus, the focus of interest in network economics has shifted from the analysis of natural monopoly towards issues of interconnection, compatibility, interoperability and coordination of quality.

Network externalities and the battle for critical mass

For normal goods, the demand curve slopes downwards. As price decreases, more of the product is demanded. Other elements in the demand function such as income or advertising serve as 'demand shifters' and would elevate the demand to a higher level. Figure 10 illustrates the traditional role of a

FIGURE 9: A one-way network.



demand shifter. Higher levels of consumption are derived from higher incomes (positive income elasticities) or from lower prices (negative price elasticities).

This fundamental relationship is greatly distorted in the presence of network externalities. In the presence of network externalities, we specify that sales rise as accumulated sales (the installed base) rise and we obtain a chicken and egg paradox. Customers may not be interested in purchasing because the installed base is small and/or not expected to grow. Imagine the purchase of complex software without Internet support, help lines and user groups. Alternatively, there may be confident expectations that the installed base will grow substantially and therefore consumers will confidently make purchases. The paradox is that consumers will not buy if the installed base is too low. However, the installed base is too low because customers will not buy. The crux of the paradox lies in the management of expectations (see the second paper in this series). In markets for normal goods, equilibrium is explained in terms of a balance between costs and demand, between marginal costs and marginal utility. In network markets, there is also equilibrium to be struck between actual demand and expectations of total demand.

This gives rise to an economic paradox. Almost the first law of economics is that value comes from scarcity. However, in the New World economy value comes from plenty: the more something is demanded and the more it is expected to be demanded then the more valuable it becomes. Expectations are so important in driving demand that a point exists where the momentum is so overwhelming that success becomes a runaway event and we observe a 'winner takes all' phenomenon.

The 'tipping point' is when the installed base (or size of network) tips expectations sharply towards one player (or one network) and away from its rival. We have experienced this effect when we moved towards Windows as our prevailing computer operating system, rather than OS2. Another example of tipping would be IBM-compatibles versus Apple as shown in Figure 12.

The exception to the winner takes all phenomenon would be a regulated network market with strong interconnections between competing platforms. The mobile telephone industry is a classic example. The standards are harmonised across the network providers, at least by continental region. The platforms are interlinked and the sales curves of the regulated network providers follow the pattern of the overall subscription curve for the industry.

Traditional economic thinking is based on negative feedback systems in which the strong get weaker at the margin and the weak get stronger, thus providing a drive towards a competitive equilibrium. This is captured in economics by the concept of diminishing marginal utility as consumption grows. In the New World of networks,

FIGURE: 10: Demand shifts due to the installed base.

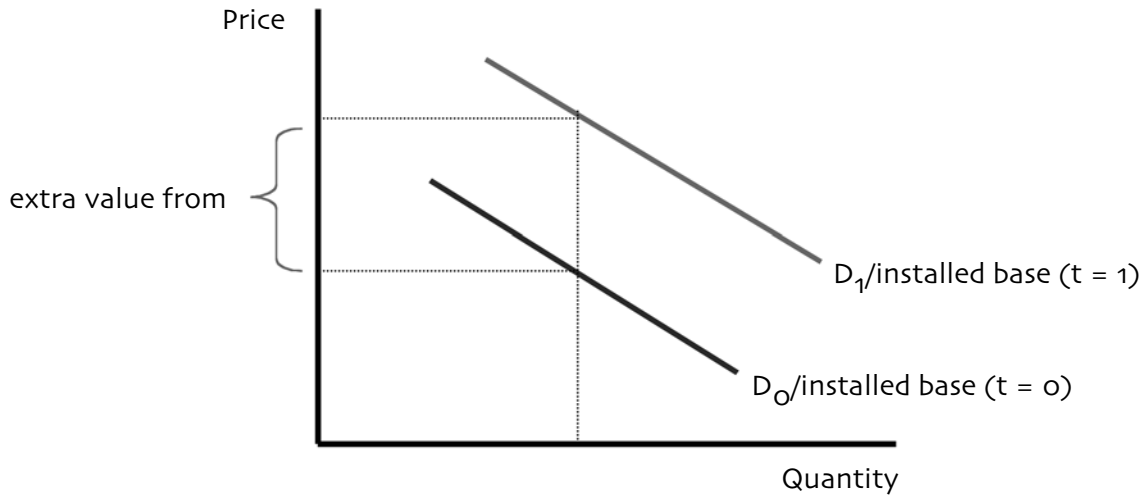


FIGURE 11: Winner takes all.

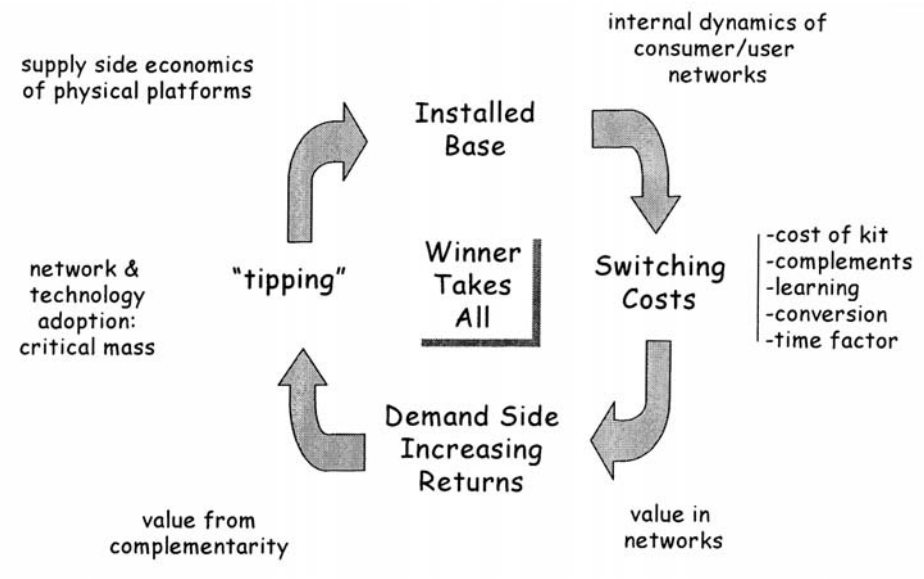
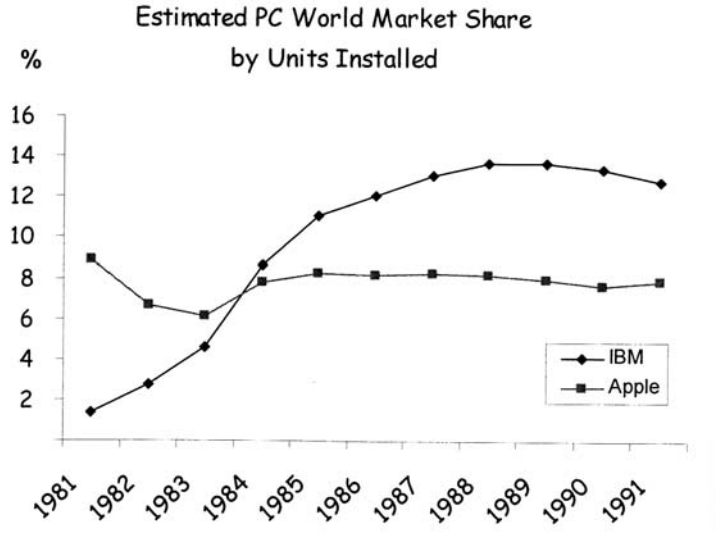


FIGURE 12: Market share of IBM and Apple.



positive feedback rules. In this world the valuation of a product increases the more that others consume the product. Strictly speaking, it arises from the interdependence of consumer decisions whereas diminishing marginal utility dominates when consumer decisions are independent – the normal assumption in economics.

The price–quantity relationship is normally held to be downward sloping, but the demand curve for a network product should be drawn differently

(Figure 13). The value to the consumer of a network product is reflected in the price he is willing to pay – the vertical axis. The principal driver of value is the size of the network, also referred to as the installed base, and is shown on the horizontal axis. Quantity demanded does still have an effect on price but for these products, this is secondary to the network effect.

The initial upward slope of the curve reflects a rising valuation at the margin, as consumers perceive that they gain value by virtue of other consumers having the product. Being on the Wintel standard gives value to new users. However, as the network grows, the extra consumers at the margin are less valuable – i.e. this shape assumes that those users with higher potential valuation of the network will join first. As the network gets very large, further growth has less value for future customers. The intercept on the vertical axis represents the value the network product has as a stand-alone product. Thus a Wintel computer has some stand-alone value, but a telephone has no value on its own and is a pure network good.

There is a notion of an optimal size of a network. This can be seen from the interaction of demand and cost so that as less and less valuable customers join the network there may come a point where the costs of acquiring and servicing new customers begins to exceed the price those customers are willing to pay. This determines the optimal size and has significant implications for competition.

The three configurations shown in Figure 14 indicate the range of possibilities. The first is a pure network good, such as a telephone system, in which the optimal size of network is a very high proportion of the available market. This implies there is little or no room for rival networks. The second is a product with a significant intrinsic value that attracts a modest size group of users. For example, this could be a corporate software package (e.g. enterprise solutions) that attracts dedicated user support from the supplier through the web. Alternative networks could coexist. The third case is one of very high intrinsic demand but extensive consumer interactions (small in size but several in number) provide a substantial total network value. The obvious example is word processing software where the value from standardising on MS Word is very high with the result that alternative standards (such as WordPerfect) are being frozen out of the market even though the intrinsic value of any word processing package is high.

Networks, standards and competition

According to Economides and Flyer (1997):

> Firms that compete in markets where network externalities are present face unique trade-offs regarding the choice of a technical standard.

FIGURE 13: The network demand curve. The idea of optimal size (OS).

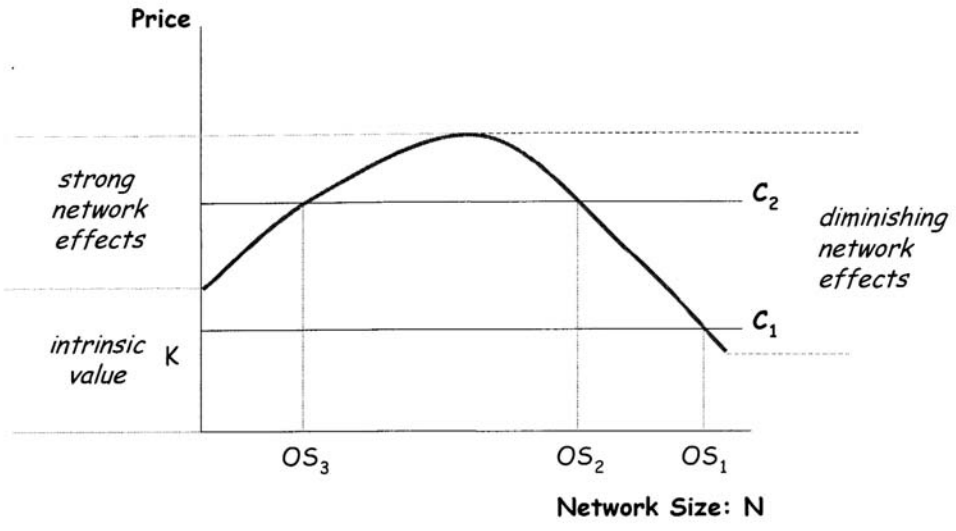
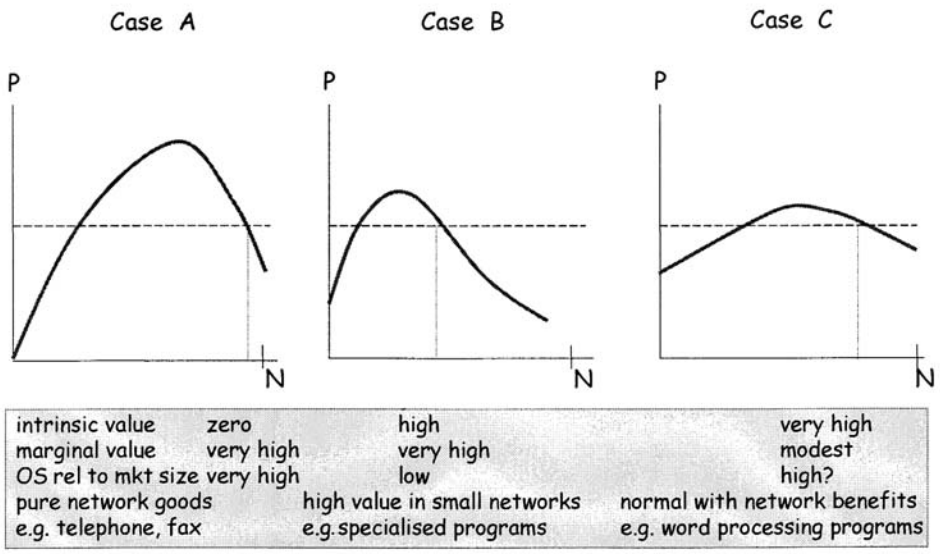


FIGURE 14: Alternative network demand configurations.



Adhering to a leading compatibility standard allows a firm's product to capture the value added by a large network. However, simultaneously the firm loses direct control over the market supply of the good and faces (direct) intra-platform competition. Alternatively, adhering to a unique standard allows the firm to face less or no intra-platform competition, but it sacrifices the added value associated with a large network.

This trade-off is a key strategic decision that depends in part on the control that firms have in making their output compatible with competitors' outputs and complementary products. The ability to conform to a common standard opens the opportunity to make this trade-off. Where standards are proprietary, the decision rests with the owner of the standard. The owner's trade-off is the pay-off associated with developing the existing network and its spillovers versus the introduction of more intra-platform competition. Essentially the trade-off is the same: to adhere to a common standard or to seek uniqueness. This can be expressed as a sequential game: at the outset, one chooses the appropriate technical standard (and, therefore, the network to join), and later one chooses how to compete. Normal markets do not have this choice of network and there are consequences for market structure and competition of the presence of network externalities. The mathematical model in Economides and Flyer (1997) defines networks as coalition structures and analyses the stability of coalitions under different standards regimes and varying levels of network externalities. There are a number of implications for market structure in the presence of network externalities.

First, it is intuitively clear that industry output will be higher when there are network externalities and when standards are open. Firms are free to choose which standard to adopt and are deterred only by the costs of adoption. Second, when standards are incompatible and the owners of standards can exercise proprietary control, then incumbents are more strongly protected against the consequences of new entrants. Moreover, there will usually be considerable asymmetries between firms in terms of outputs, prices, and profits. (Under incompatibility regimes firms are equivalent to platforms and constitute one firm networks.) For pure network goods the asymmetries are particularly marked.

In general, with total incompatibility of standards, market concentration, output inequality and price and profit inequality increase with the extent of the network externality. This is an important result because it explains why network industries are so often dominated by one or two firms. The

mechanism is straightforward. The leading network establishes its critical mass, leaving the second network to establish a critical mass across the remaining untapped market coverage. The third network follows in the same fashion and so on. It follows that there will be a tendency to provide large incentives to organise customers into few platforms so as to maximise the added value from the available networks. Firms will be keen to abandon their own weak standards in favour of the higher value obtainable from a leading network.

There is a third implication. Where there are proprietary standards and strong network effects there is no natural equilibrium in terms of network offerings. There are always incentives for at least one firm to move to a stronger network and the consequences of any one move is to shift the incentives for all other firms. However, equilibrium can be reinforced by the refusal of firms to make their proprietary standards available. Again, the mechanism is straightforward. Under strong externalities, the owner of a standard has a considerable incentive to exploit the standard by itself and to exclude other firms with weaker standards. Conversely, where the externality is weak, the owner will find a stronger incentive to admit other firms to its proprietary standard in order to grow the network through collective effort and thus generate more added value.

In summary, strong network externalities suggest the following conclusions:

1. Larger industry output.
2. Very large asymmetries between firms/platforms.
3. Likelihood of market dominance.
4. Enhancement and protection of proprietary standards.
5. Equilibrium market structures that are the reverse of the world without network externalities.

This suggests some rules that govern the new economy.

1. The information economy depends on connectivity. Without connectivity, consumer interdependence is indirect. Positive feedback gives an economic law of plenty – more gives more.
2. The competition between rival networks/standards can be hard to call in advance. Management of expectations is key and 'tippy markets' are common.
3. Commonly this is a game where the upfront costs are very large and the revenues are substantially delayed and are significantly at risk.
4. As a result, this is a 'winner takes all' world.
5. It is also a world of immense uncertainty where even the range of potential outcomes is not known but also where there is a significant probability that future technological change might undermine an apparently winning position.
6. There is a law of inverse pricing. The best (i.e. the most valuable in the future) products are given away, such as web browsers, in order to create a consumer standard, and sheer volume causes both marginal costs and prices to fall over time as the product becomes more valuable. The cash flow machine is modest (even small) margins multiplied by gigantic volumes to defray massive investments. The machine is volume driven and protected by very large switching costs.
7. Open standards are the key to volume. Protected standards are only viable as small high-priced niche markets.
8. The first strategic choice is what network to join. The second, and a long way behind, is how to compete within the network of choice.

A new set of strategies are emerging to offset the risks and pressures exerted by the rules listed above. This is visible in the setting up of global standards and their ensuing platforms. For example, Group Speciale Mobile, commonly known as GSM, is an association of 600 network operators and suppliers of the mobile phone industry. Their primary objective is to set a common standard for mobile communications in order to create a homogeneous industry where equipment, software and networks can seamlessly talk to each other. Strategies of standardisation are stabilising the markets and charting the course for research and development policies.

Conclusion

This paper establishes two propositions. *First*, the supply of knowledge has a distinctive cost structure that makes large volumes extremely important. As (globalising) markets have grown, so has specialisation in the production of knowledge with consequences for companies in terms of outsourcing of activities and for new, more focused approaches to their own knowledge base and to the acquisition of knowledge from outside. This has led to three new, distinctive business models:

1. The new competitor who uses new, typically electronic channels to challenge existing businesses and to engineer new economies of scope which have the effect of transforming (potentially) the wider competitive landscape.
2. The deconstruction model in which the value chain is slimmed down to those elements in which clear advantage is evident. Control over the entire supply chain, and therefore profitability, depends on the location of the most significant knowledge components in the supply chain.
3. The reconstruction model builds on deconstruction but the knowledge assets are deployed into adjacent and apparently unrelated supply chains with the effect of transferring knowledge and (potentially) controlling these supply chains.

These economic characteristics of knowledge are dependent in large part on the interconnectivity that is characteristic of the technologies of information goods. Interconnectivity allows customers to view, use and link products, giving rise to virtual networks of customers. In these networks, powerful demand-side increasing returns can operate, giving rise to our *second proposition*. Where consumer-based externalities are powerful there are strong pressures towards 'winner takes all' phenomena (e.g. Wintel globally, and Sky TV in the UK). In these circumstances conventional economic laws are challenged. De facto monopoly can emerge: but uncertainty is high and markets may be intrinsically unstable. Successive waves of technology may outmode old monopolies and serve as the basis for new monopolies.

In our second companion paper we argue that there are significant implications for firms arising from both of these propositions. In particular, collaboration between firms becomes almost essential, requiring a complex balance to be struck between collaboration and competition.

References

Arthur WB (1989) Competing technologies, increasing returns, and lock-in by historical events. *Economic Journal* 99: 116–31.

Arthur WB (1990) Positive feedback in the economy. *Scientific American* 262: 92–9.

Baumol W, Panzar J, Willig R (1982) Contestable markets and the theory of industry. *American Economic Review* 72.

DiMaggio PJ, Powell WW (1983) The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. *American Sociological Review* 48(April): 147–60.

Dumont B (2000) Book Review: Carl Shapiro and Hal R. Varian, Information Rules: A Strategic Guide to the Network Economy. *Journal of Network Industries*.

Economides N (1996) The economics of networks. *International Journal of Industrial Organization* 14(2): 675–99.

Economides N, Flyer F (1997) Compatibility and Market Structure for Network Goods. Discussion Paper EC-98-02, Stern School of Business, NYU.

Evans P, Wurster TS (2000) *Blown to Bits*. Boston, MA: Harvard Business School Press.

Hawley A (1968) Human ecology. In: Sills DL (ed) *International Encyclopaedia of the Social Sciences*, pp 328–37. New York: Macmillan.

Katz M, Shapiro C (1985) Network externalities, competition and compatibility. *American Economic Review* 75(3): 424–40.

Kelly K 1994) *Out of Control – The New Biology of Machines*. Addison Wesley.

Quinn JB (2001) *Services and Technology, Revolutionizing Economics, Business and Education*. Darmond College.

Saloner G, Shepard A, Podolny J (2001) *Strategic Management*. New York: Wiley.

Shapiro C, Varian H (1999) *Information Rules: A Strategic Guide to the Network Economy*, p 2. Boston, MA: Harvard Business School Press.

Sharkey NE, Sharkey AJC (1993) Adaptive generalisation and the transfer of knowledge. *Artificial Intelligence Review: Special Issue on Connectionism* 7: 313–28.

Address for correspondence: John McGee, BA (Econ) Manchester) PhD (Stanford), Professor of Strategic Management, and Academic Director, Full Time MBA Programme, Warwick Business School, University of Warwick, Coventry CV4 7AL, United Kingdom
E-mail: J.McGee@WBS.ac.uk