Comp. by: DKandavel Date:20/11/06 Time:12:23:48 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides

# 8 The Internet and network economics

*Nicholas Economides*

## 8.1 Introduction: the Internet and network economics

The Internet is the most important new network of the last part of the 20th century. As a global network of interconnected networks that connect computers, the Internet allows data transfers as well as the provision of a variety of interactive real-time and time-delayed telecommunications services. Originally developed through grants by the US Department of Defense (DOD) to connect disparate computers of its research division and its various contractors, the Internet served the academic community for over a decade before reaching the wide public in the mid-1990s. Internet communication is based on common and public protocols. Presently, hundreds of millions of computers are connected to the Internet.

The Internet exemplifies to the utmost "network effects", that is the fact that the value of a connection to a network increases with the size of the network. The Internet was not the first or last of electronic networks, but it is definitely the largest on earth in terms of hosts (computers) connected to it, content stored in them, and bits of traffic traversing it every day. Many electronic networks pre-dated the Internet, some with very sophisticated protocols compared with the Internet. But the strength of the Internet in terms of size and simplicity of its basic protocols forced pre-existing proprietary networks such as AOL, Prodigy, AT&TMail, and MCIMail to conform to the open standards of the Internet and become part of it. Competition in the provision of services is vigorous on the Internet.[1]

This chapter revives a paper originally published in the *IJIO*, "The Economics of Networks",[2] which was written in 1996 when the Internet was just reaching out to the mass market. A number of applications

[1] For a comprehensive discussion of competition issues on the Internet backbone, see Economides (2005a).
[2] Economides "The Economics of Networks", *International Journal of Industrial Organization*, vol. 14, no. 6, pp. 673–699 (October 1996).

Comp. by: DKandavel Date:20/11/06 Time:12:23:48 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



benefiting from network effects mushroomed after that and were wildly successful "killer applications" on the Internet, including electronic mail, file transfers (Napster), interactive search, and interactive advertising. Telephone calls over the Internet, commonly called "Voice Over Internet Protocol" or VOIP, technically feasible for a number of years, are becoming a significant force in telecommunications and may, in the long run, become a "killer application".[3] VOIP converts voice calls to data and takes advantage of the very low cost of transmitting data over the Internet.[4]

Below we discuss in some detail the consequences of network effects for market structure in network industries, as well as the private incentives of each firm to deviate from common standards and make its product incompatible with those of competitors. In the extreme case, in telecommunications, incompatibility would mean refusal to interconnect. In the history of telecommunications, we have seen significant cases where a dominant firm refused to interconnect with others. From the expiration of its original patent before the turn of the 19th century until it was regulated in the 1930s and interconnection became mandatory, AT&T, leveraging its long-distance bottleneck position, refused to interconnect with independent telephone companies. This led to the absurd but then common situation of a customer having two different phones, each connected to a separate telephone network.

Even when each competitor potentially has its own bottleneck, as in the case of unique subscribers who need to be reached over the network by others, it can be shown that foreclosure of rivals can occur. In a local telecommunications network, if cost-based reciprocal compensation were not the rule, and networks were able to set prices for terminating calls at profit-maximizing levels, Economides, Lopomo, and Woroch (1996a,b) have shown that a large network will try to impose very high termination charges on an opponent's small network so that no

[3] In the United States, the growth of VOIP was significantly helped by the fact that, by the middle of 2004, the main alternative way (other than through the incumbent local exchange carrier) to traverse the "last mile" to reach residential customers (which was through leasing of parts of the incumbent's local network at cost-based prices) was practically eliminated as incumbents were allowed to charge commercial rates from such leases. See Economides (2005b), "Telecommunications Regulation: An Introduction", at http://www.stern.nyu.edu/networks/Economides_Telecommunications_Regulation.pdf. In many other parts of the world, VOIP has grown significantly as an alternative to extremely high long-distance, international, and interconnection rates.

[4] VOIP works best in a broadband connection. A significant limitation to wide adoption of VOIP in the United States is the fact that incumbent local exchange carriers typically do not allow access to a DSL loop unless they also provide voice service on the same line. Thus, most customers are unable to get their first line through VOIP.

Comp. by: DKandavel Date:20/11/06 Time:12:23:48 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



calls terminate from the small network to the large one. Without the possibility of such across-networks calls, a small network will be able to provide only within-network calls and, being small, will be of little value to potential subscribers. As a consequence, large networks are able to offer subscribers more value and the small network is foreclosed. Starting from a regime of a large local incumbent and a small potential entrant, the large incumbent can set up termination access fees so that the entrant is kept out of the market.[5]

In the context of these historical examples of using bottlenecks to leverage market power and foreclose rivals, the Internet stands out as an example that breaks the rules. On the Internet, compatibility is the rule, pricing is independent of distance or direction of origination and does not even depend on the number, duration, and type of transactions – pricing depends only on the total bandwidth capacity bought. This feat is even more profound given the fact that the Internet is the least regulated part of the telecommunications world today. The success of competition on the Internet backbone is based on the public nature of Internet protocols, ease of entry, very fast network expansion, connections by the same Internet service provider (ISP) to multiple backbones (ISP multi-homing), and connections by the same large web site to multiple ISPs (customer multi-homing) that enhance price competition and make it very unlikely that any firm providing Internet backbone connectivity would find it profitable to degrade or sever interconnection with other backbones in an attempt to monopolize the Internet backbone.

On the "last mile" of the Internet, that is in reaching from the backbone to the customer's location, there is much less competition than on the backbone. There are two main avenues of residential broadband access, through "cable modems" that attach the cable TV wire and through digital subscriber lines (DSL) that use the higher frequencies of copper wire lines. Unfortunately, the United States lags behind countries like South Korea in broadband penetration. Wireless alternatives

[5] See Economides, Lopomo, and Woroch (1996a), "Regulatory Pricing Policies to Neutralize Network Dominance", *Industrial and Corporate Change,* vol. 5, no. 4, pp. 1013–1028, pre-publication copy at http://www.stern.nyu.edu/networks/96-14.pdf, and Economides, Lopomo, and Woroch (1996b), "Strategic Commitments and the Principle of Reciprocity in Interconnection Pricing", Discussion Paper EC-96–13, Stern School of Business, at http://www.stern.nyu.edu/networks/96-13.pdf. Note that this is not just a theoretical possibility. Telecom New Zealand (TNZ), operating in an environment of weak antitrust and regulatory intervention (so-called "light-handed regulation"), offered such high termination fees that the first entrant into local telecommunications, Clear, survives only by refusing to pay interconnection fees to TNZ, while the second entrant, BellSouth New Zealand, exited the local telecommunications market.

Comp. by: DKandavel Date:20/11/06 Time:12:23:48 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



exist but presently are very expensive if done through traditional cellular carriers or very limited in geographic coverage if done through municipal WiFi networks. In such a context, our seminal work on the "Economics of Networks", presented below from section 2 to 7, examines the consequences of network effects for market structure and the various ways in which companies may attempt to take advantage of network effects by creating or leveraging bottlenecks.

## 8.2 Classification of networks[6]

Network industries play a crucial role in modern life. The modern economy would be very much diminished without the transportation, communications, information, and railroad networks. This chapter will analyze the major economic features of networks. In the course of the analysis it will become clear that many important non-network industries share many essential economic features with network industries. These non-network industries are characterized by strong complementary relations. Thus, the lessons of networks can be applied to industries where vertical relations play a crucial role; conversely, the economic and legal learning developed in the analysis of vertically related industries can be applied to network industries.

Formally, networks are composed of links that connect nodes. It is inherent in the structure of a network that many components of a network are required for the provision of a typical service. Thus, network components are complementary to each other. Figure 8.1 represents the emerging *information superhighway* network. Clearly, services demanded by consumers are composed of many complementary components. For example, interactive ordering while browsing in a "department store" as it appears in successive video frames requires a number of components: a database engine at the service provider, transmission of signals, decoding through an interface, display on a TV or computer monitor, etc. Clearly, there are close substitutes for each of these components; for example, transmission can be done through a cable TV line, a fixed telephone line, a wireless satellite, public cable network (PCN), etc.; the in-home interface may be a TV-top box or an add-on to a PC, etc. It is likely that the combinations of various components will not result in identical services. Thus, the information superhighway will provide substitutes made of complements – this is a typical feature of networks.

[6] The literature on networks is so extensive that it is futile to attempt to cover it. This contribution discusses only some issues that arise in networks and attempts to point out areas in which further research is necessary.

Comp. by: DKandavel Date:20/11/06 Time:12:23:48 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



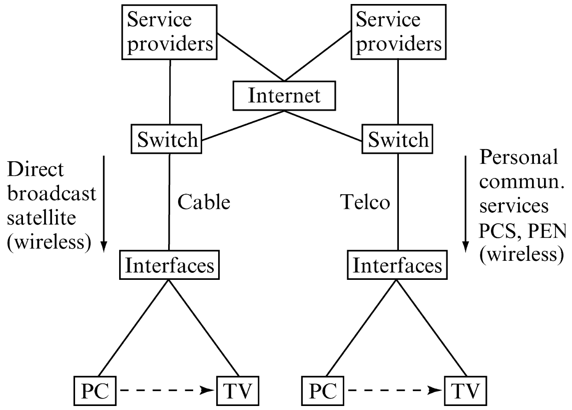


Figure 8.1. An information superhighway.

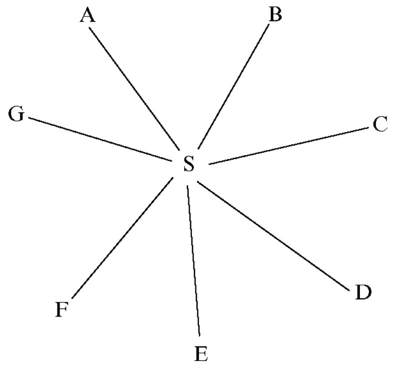


Figure 8.2. A simple star network.

Figure 8.2 shows this feature in a simple star telephone network. A phone call from A to B is composed of AS (access to the switch of customer A), BS (access to the switch of customer B), and switching services at S. Despite the fact that goods AS and BS look very similar and have the same industrial classification, they are *complements* and not substitutes.[7]

Networks where services AB and BA are distinct are named "two-way" networks in Economides and White (1994). Two-way networks

[7] AS and BS can also be components of *substitute* phone calls ASC and BSC.

Comp. by: DKandavel Date:20/11/06 Time:12:23:48 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



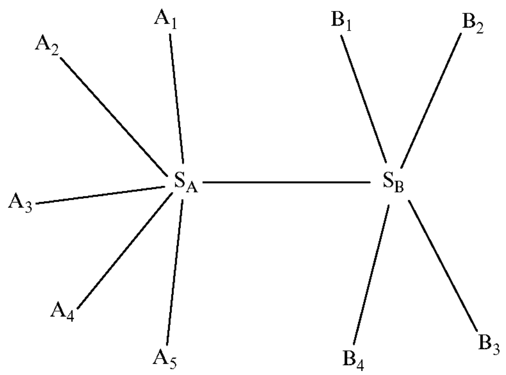


Figure 8.3. A simple local and long-distance network.

include railroad, road, and many telecommunications networks. When one of AB or BA is unfeasible or does not make economic sense, or when there is no sense of direction in the network so that AB and BA are identical, the network is called a one-way network. In a typical one-way network, there are two types of components, and composite goods are formed only by combining a component of each type and customers are often not identified with components but instead demand composite goods. For example, broadcasting and paging are one-way networks.[8]

The classification in network type (one-way or two-way) is not a function of the topological structure of the network. Rather, it depends on the interpretation of the structure to represent a specific service. For example, the network of Figure 8.3 can be interpreted as a two-way telephone network where $S_A$ represents a local switch in city A, $A_i$ represents a customer in city A, and similarly for $S_B$ and $B_j$.[9] In this network, there are two types of local phone calls, $A_iS_AA_k$ and $B_jS_BB_l$, as well as long-distance phone call $A_iS_AS_BB_j$. We can also interpret the network of Figure 8.3 as an automatic teller machine (ATM) network. Then a transaction (say a withdrawal) from bank $B_j$ from ATM $A_i$ is $A_iS_AS_BB_j$. Connections $A_iS_AA_k$ and $B_jS_BB_l$ may be feasible but there is no demand for them.

We have pointed out that the crucial relationship in both one-way and two-way networks is the complementarity between the pieces of the

[8] The 1994 spectrum auction will allow for a large two-way paging network.

[9] In this network, we may identify end-nodes, such as $A_i$ and $B_j$, end-links, such as $A_iS_A$ and $S_BB_j$, the interface or gateway $S_AS_B$, and switches $S_A$ and $S_B$.

Comp. by: DKandavel Date:20/11/06 Time:12:23:49 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



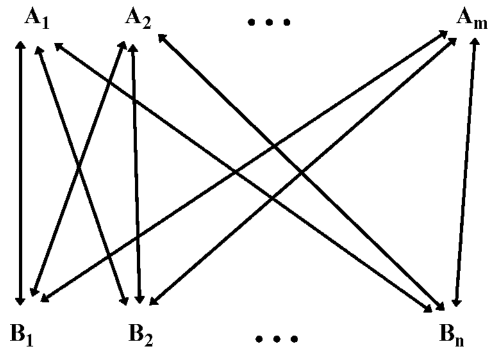


Figure 8.4. A pair of vertically related markets.

network. This crucial economic relationship is also often observed between different classes of goods in non-network industries. In fact, Economides and White (1994) point out that a pair of vertically related industries is formally equivalent to a one-way network. Figure 8.4 can represent two industries of complementary goods A and B, where consumers demand combinations $A_iB_j$. Notice that this formulation is formally identical to our long-distance network of Figure 8.3 in the ATM interpretation.

The discussion so far has been carried out under the assumption of *compatibility*, i.e. that various links and nodes on the network are costlessly combinable to produce demanded goods. We have pointed out that links on a network are potentially complementary, but it is *compatibility that makes complementarity actual*. Some network goods and some vertically related goods are immediately combinable because of their inherent properties. However, for many complex products, actual complementarity can be achieved only through the adherence to specific technical compatibility standards. Thus, many providers of network or vertically related goods have the option of making their products partially or fully incompatible with components produced by other firms. This can be done through the creation of proprietary designs or the outright exclusion or refusal to interconnect with some firms.

Traditionally, networks were analyzed under the assumption that each network was owned by a single firm. Thus, economic research focused on the efficient use of the network structure as well as on the appropriate allocation of costs.[10] In the 1970s, partly prompted by the antitrust suit

[10] See Sharkey (1995) for an excellent survey.

Comp. by: DKandavel Date:20/11/06 Time:12:23:49 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



against AT&T, there was a considerable amount of research on economies of scope, i.e. on the efficiency gains from joint operation of complementary components of networks.[11]

Once one of the most important networks (the AT&T telecommunications network in the US) was broken to pieces, economic research focused in the 1980s and 1990s on issues of interconnection and compatibility. Similar research on issues of compatibility was prompted by the reduced role of IBM in the 1980s and 1990s in the setting of technical standards in computer hardware and software. Significant reductions in costs also contributed and will play a part in the transformation toward fragmented ownership in the telecommunications sector in both the United States and abroad. Costs of transmission have fallen dramatically with the introduction of fiberoptic lines. Switching costs have followed the fast cost decreases of microchips and integrated circuits. These cost reductions have transformed the telecommunications industry from a natural monopoly to an oligopoly. The same cost reductions have made many new services, such as interactive video and interactive games, feasible at low cost. Technological change now allows for joint transmission of digital signals of various communications services. Thus, the monopoly of the last link closest to home is in the process of being eliminated,[12] since both telephone lines and cable lines (and in some cases personal communication services (PCS) and terrestrial satellites) will provide similar services.[13,14]

In a network where complementary as well as substitute links are owned by different firms, the questions of interconnection, compatibility, interoperability, and coordination of quality of services become of

[11] See Baumol, Panzar, and Willig (1982).

[12] It is already eliminated in some parts of the United Kingdom, where cable TV operators offer telephone service at significantly lower prices than British Telecom.

[13] These significant changes in costs and the convergence of communications services open a number of policy questions on pricing, unbundling, deregulation, and possibly mandated segmentation in this sector. It is possible that ownership break-up of local and long-distance lines is no longer necessary to improve competition. For example, European Union policy mandates open competition by 1998 in any part of the telecommunications network, but does not advocate vertical fragmentation of the existing integrated national monopolies; see the Bangemann Report. The reduction in costs and the elimination of natural monopoly in many services may make it possible for this policy to lead the industry to competition.

[14] Another important network, the airline network, faces significant change in Europe. Airlines have not benefitted from significant cost reductions and technological change; the reform is just the abolition by the European Union of the antiquated regime of national airline monopolies, and its replacement by a more competitive environment.

Comp. by: DKandavel Date:20/11/06 Time:12:23:49 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



paramount importance. We will examine these issues in detail in the next few sections. We first focus on a fundamental property of networks, i.e. the fact that they exhibit *network externalities*.

## 8.3 Network externalities

Networks exhibit positive consumption and production externalities. A positive consumption externality (or network externality) signifies the fact that the value of a unit of the good increases with the number of units sold. To economists, this fact seems quite counterintuitive, since they all know that, except for potatoes in Irish famines, market demand slopes downwards. So the earlier statement "the value of a unit of the good increases with the number of units sold" should be interpreted as "the value of a unit of the good increases with the *expected* number of units to be sold". Thus, the demand slopes downward but shifts upward with increases in the number of units expected to be sold.

### *8.3.1 Sources of network externalities*

The key reason for the appearance of network externalities is the complementarity between the components of a network. Depending on the network, the externality may be direct or indirect. When customers are identified with components, the externality is direct. Consider, for example, a typical two-way network, such as the local telephone network of Figure 8.2. In this n-component network, there are $n(n-1)$ potential goods. An additional $(n+1)$th customer provides direct externalities to all other customers in the network by adding 2n potential new goods through the provision of a complementary link say ES to the existing links.[15]

In typical one-way networks, the externality is only indirect. When there are m varieties of component A and n varieties of component B as in Figure 8.4 (and all A-type goods are compatible with all B-type), there are mn potential composite goods. An extra customer yields indirect externalities to other customers, by increasing the demand for components of types A and B and thereby (because of the presence of economies

[15] This property of two-way networks was pointed out in telecommunications networks by Rohlfs (1974) in a very early paper on network externalities. See also Oren and Smith (1981).

Comp. by: DKandavel Date:20/11/06 Time:12:23:49 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



of scale) potentially increasing the number of varieties of each component available in the market.

Financial exchange networks also exhibit indirect network externalities. There are two ways in which these externalities arise: first, in the act of exchanging assets or goods, and second, in the array of vertically related services that compose a financial transaction. These include the services of a broker, of bringing the offer to the floor, matching the offer, etc. The second type is similar to other vertically related markets. The first way in which externalities arise in financial markets is more important.

The act of exchanging goods or assets brings together a trader who is willing to sell with a trader who is willing to buy. The exchange brings together the two complementary goods, "willingness to sell at price p" (the "offer") and "willingness to buy at price p" (the "counteroffer") and creates a composite good, the "exchange transaction". The two original goods were complementary and each had no value without the other. Clearly, the availability of the counteroffer is critical for the exchange to occur. Put in terms commonly used in finance, minimal liquidity is necessary for the transaction to occur.

Financial markets also exhibit positive size externalities in the sense that the increasing size (or thickness) of an exchange market increases the expected utility of all participants. Higher participation of traders on both sides of the market (drawn from the same distribution) decreases the variance of the expected market price and increases the expected utility of risk-averse traders. *Ceteris paribus*, higher liquidity increases traders' utility. Thus, financial exchange markets also exhibit network externalities.[16,17]

[16] For a more detailed discussion of networks in finance see Economides (1993a). Economides and Schwartz (1995a) discuss how to set up *electronic call markets* that bunch transactions and execute them all at once. Call markets have inherently higher liquidity because they take advantage of network externalities in exchange. Thus, transaction costs are lower in call markets. Economides (1994a) and Economides and Heisler (1994) discuss how to increase liquidity in call markets. The survey of institutional investors reported by Economides and Schwartz (1995b) finds that many traders who work in the present continuous market environment would be willing to wait a number of hours for execution of their orders if they could save in transaction costs, including bid-ask spreads. Thus, the time is right for the establishment of call markets in parallel operation with the continuous market.

[17] The increase of utility in expectation due to market thickness was pointed out by Economides and Siow (1988), and earlier and in less formal terms by Garbade and Silber (1976, 1979). The effects are similar to those of search models as in Diamond (1982, 1984).

Comp. by: DKandavel Date:20/11/06 Time:12:23:49 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



### 8.3.2 *The "macro" approach*

There are two approaches and two strands of literature in the analysis of network externalities. The first approach assumes that network externalities exist, and attempts to model their consequences. I call this the "macro" approach. Conceptually this approach is easier, and it has produced strong results. It was the predominant approach during the 1980s. The second approach attempts to find the root cause of the network externalities. I call this the "micro" approach. In industrial organization, it started with the analysis of mix-and-match models and has evolved to the analysis of various structures of vertically related markets. In finance, it started with the analysis of price dispersion models. The "micro" approach is harder, and in many ways more constrained, as it has to rely on the underlying microstructure. However, the "micro" approach has a very significant benefit in defining the market structure. We discuss the "macro" approach first.

*Perfect competition*

As we have noted, network externalities arise out of the complementarity of different network pieces. Thus, they arise naturally in both one- and two-way networks, as well as in vertically related markets. The value of good X increases as more of the complementary good Y is sold, and vice versa. Thus, more of Y is sold as more X is sold. It follows that the value of X increases as more of it is sold. This positive feedback loop seems explosive, and indeed it would be, except for the inherent downward slope of the demand curve. To understand this better, consider a fulfilled expectations formulation of network externalities as in Katz and Shapiro (1985), Economides (1993b, 1996a), and Economides and Himmelberg (1995). Let the willingness to pay for the nth unit of the good when $n^e$ units are expected to be sold be $p(n, n^e)$.[18] This is a decreasing function of its first argument because the demand slopes downward. $p(n, n^e)$ increases in $n^e$; this captures the network externalities effect. At a market equilibrium of the simple single-period world, expectations are fulfilled, $n = n^e$, thus defining the fulfilled expectations demand $p(n, n)$. Figure 8.5 shows the construction of a typical fulfilled expectations demand. Each curve $D_i$, $i = 1, \ldots, 4$, shows the willingness to pay for a varying quantity n, given an expectation of sales $n^e = n_i$. At $n = n_i$, expectations are fulfilled and the point belongs to $p(n, n)$ as $p(n_i, n_i)$. Thus $p(n, n)$ is constructed as a collection of points $p(n_i, n_i)$.

[18] In this formulation n and $n^e$ are normalized so that they represent market shares rather than absolute quantities.

Comp. by: DKandavel Date:20/11/06 Time:12:23:49 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



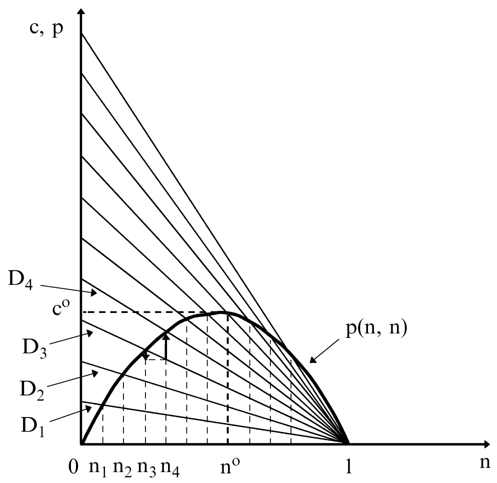


Figure 8.5. Construction of the fulfilled expectations demand.

To avoid explosions and infinite sales, it is reasonable to impose $\lim_{n\to\infty}$ (p(n, n) = 0; it then follows that p(n, n) is decreasing for large n. Economides and Himmelberg (1995) show that the fulfilled expectations demand is increasing for small n if either one of three conditions hold: (i) the utility of every consumer in a network of zero size is zero, or (ii) there are immediate and large external benefits to network expansion for very small networks, or (iii) there is a significant density of high-willingness-to-pay consumers who are just indifferent on joining a network of approximately zero size. The first condition is straightforward and applies directly to all two-way networks. The other two conditions are a bit more subtle, but commonly observed in networks and vertically related industries.

When the fulfilled expectations demand increases for small n, we say that the network exhibits a positive critical mass under perfect competition. This means that, if we imagine a constant marginal cost c decreasing parametrically, the network will start at a positive and significant size $n^o$ (corresponding to marginal cost $c^o$). For each smaller marginal cost, $c < c^o$, there are three network sizes consistent with marginal cost pricing: a zero size network; an unstable network size at the first intersection of the horizontal through c with p(n, n); and the Pareto optimal stable network size at the largest intersection of the horizontal with p(n, n). The multiplicity of equilibria is a direct result of the coordination problem that arises naturally in the typical network externalities model.

Comp. by: DKandavel Date:20/11/06 Time:12:23:50 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



In such a setting, it is natural to assume that the Pareto optimal network size will result.[19]

In the presence of network externalities, it is evident that perfect competition is inefficient: the marginal social benefit of network expansion is larger than the benefit that accrues to a particular firm under perfect competition. Thus, perfect competition will provide a smaller network than is socially optimal, and for some relatively high marginal costs perfect competition will not provide the good while it is socially optimal to provide it.

One interesting question that remains virtually unanswered is how to decentralize the welfare-maximizing solution in the presence of network externalities. Clearly, the welfare-maximizing solution can be implemented through perfect price discrimination, but typically such discrimination is unfeasible. It remains to be seen to what extent mechanisms that allow for non-linear pricing and self-selection by consumers will come close to the first best.

### *Monopoly*

Economides and Himmelberg (1995) show that a monopolist who is unable to price-discriminate will support a smaller network and charge higher prices than perfectly competitive firms. This is despite the fact that the monopolist has influence over the expectations of the consumers, and he recognizes this influence, while no perfectly competitive firm has such influence.[20] Influence over expectations drives the monopolist to higher production, but the monopolist's profit-maximizing tendency toward restricted production is stronger and leads it to lower production levels than perfect competition. Thus, consumers and total surplus will be lower in monopoly than in perfect competition. Therefore, the existence of network externalities does not reverse the standard welfare comparison between monopoly and competition; it follows that the existence of network externalities cannot be claimed as a reason in favor of a monopoly market structure.

### *Oligopoly and monopolistic competition under compatibility*

Cournot oligopolists producing compatible components also have some influence over expectations. A natural way to model the influence of

[19] It is possible to have other shapes of the fulfilled expectations demand. In general, $p(n, n)$ is quasiconcave under weak conditions on the distribution of preferences and the network externality function. Then, if none of the three causes mentioned above is present, the fulfilled expectations demand is downward sloping.

[20] A monopolist unable to influence expectations will clearly produce less than a monopolist able to influence expectations.

Comp. by: DKandavel Date:20/11/06 Time:12:23:50 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



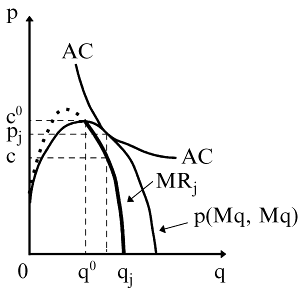


Figure 8.6. Monopolistic competition with network externalities and M compatible goods.

oligopolists on output expectations is to assume that every oligopolist takes the output of all others as given and sets the expectation of consumers of his own output. In this setting, M compatible Cournot oligopolists support a network of a size between monopoly ($M = 1$) and perfect competition ($M = \infty$). The analysis can easily be extended to monopolistic competition among compatible oligopolists if firms face downward-sloping average cost curves as shown in Figure 8.6. Firms produce on the downward-sloping part of the firm-scaled fulfilled expectations demand. At a symmetric equilibrium, firm j's output is determined at the intersection of marginal cost c and marginal revenue $MR_j$. Price is read off the fulfilled expectations firm-scaled inverse demand p(Mq, Mq). At a monopolistically competitive equilibrium, the AC curve is tangent to the fulfilled expectations demand at $q_j$.

*Oligopoly under incompatibility*

One of the most interesting issues in the economics of networks is the interaction of oligopolists producing incompatible goods. A full analysis of such a market, in conjunction with the analysis of compatible oligopolists, will allow us to determine the incentives of individual firms to choose technologies that are compatible or incompatible with others.

Given any set of firms $S = \{1, \ldots, N\}$, we can identify a subset of S that adheres to the same technical "standard" as a coalition. Then the partition of S into subsets defines a coalition structure $C_S = \{C_1, \ldots, C_k\}$. Compatibility by all firms means that there is a single coalition that includes all firms. Total incompatibility, where every firm adheres to its own unique standard, means that $k = N$.

A number of criteria can be used to define the equilibrium coalition structure. A purely non-cooperative concept without side payments requires that, after a firm joins a coalition, it is better off at the resulting

Comp. by: DKandavel Date:20/11/06 Time:12:23:51 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



market equilibrium, just from revenues from its own sales.[21] At a non-cooperative equilibrium with side payments, firms divide the profits of a coalition arbitrarily to induce firms to join a coalition. Yet firms do not cooperate in output decisions. Katz and Shapiro (1985) show that the level of industry output is greater under compatibility than at any equilibrium with some incompatible firm(s). This is not sufficient to characterize the incentives of firms to opt for compatibility.

Intuitively, a firm benefits from a move to compatibility if (i) the marginal externality is strong; (ii) it joins a large coalition; and (iii) it does not thereby increase competition to a significant degree by its action. Yet the coalition benefits from a firm joining its "standard" if (i) the marginal externality is strong; (ii) the firm that joins the coalition is large; (iii) competition does not increase significantly as a result of the firm joining the coalition. Clearly, in both cases, the second and the third criteria may create incentives that are in conflict; this will help define the equilibrium coalition structure.[22]

Katz and Shapiro (1985) show that if the costs of achieving compatibility are lower for all firms than the increase in profits because of compatibility, the industry move toward compatibility is socially beneficial. However, it may be true that the (fixed) cost of achieving compatibility is larger than the increase in profits for some firms, while these costs are lower than the increase in total surplus from compatibility. Then profit-maximizing firms will not achieve industry-wide compatibility while this regime is socially optimal. Further, if a change leads to less than industry-wide compatibility, the private incentives to standardize may be excessive or inadequate. This is because of the output changes that a change of regime has on all firms. Similarly, the incentive of a firm to produce a one-way adapter, that allows it to achieve compatibility without affecting the compatibility of other firms, may be deficient or excessive because the firm ignores the change it creates on other firms' profits and on consumers surplus.

### *Coordination to technical standards with asymmetric technologies*

So far it has been assumed that the cost of standardization was fixed and the same for both firms. If standardization costs are different, firms play a standard coordination game. A $2 \times 2$ version of this game is presented in Figure 8.7. Entries represent profits.

[21] See Economides (1984), and Yi and Shin (1992a,b).
[22] Economides and Flyer (1995) examine the incentives for coalition formation around compatibility standards.

Comp. by: DKandavel Date:20/11/06 Time:12:23:51 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



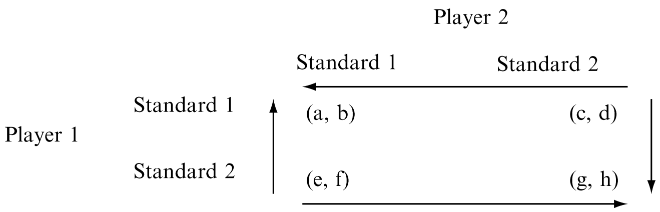


Figure 8.7. Choice between compatibility and incompatibility.

In this game, we will assume that firm i has higher profits when "its" standard i gets adopted, $a > g$, $b < h$. Profits, in case of disagreement, will depend on the particulars of the industry. One standard assumption that captures many industries is that in case of disagreement profits will be lower than those of either standard, $e, c < g$; $d, f < b$. Under these circumstances, the setting of either standard will constitute a non-cooperative equilibrium.[23] There is no guarantee that the highest joint profit standard will be adopted. Since consumer surplus does not appear in the matrix, there is no guarantee of maximization of social welfare at equilibrium. For an analysis with continuous choice of standard specification, see Berg (1988).

### 8.3.3 *The "micro" approach*

The "micro" approach starts with an analysis of the specific micro-structure of a network. After identifying the physical aspects of a network, such as nodes and links, we identify the goods and services that are demanded on the network. We distinguish between the case where only end-to-end services are demanded and the case when there is also demand for some services that do not reach from end to end. The case when only end-to-end services exist is easier and has been dealt with in much more detail in the literature. However, many important networks, such as the railroad and telephone networks, provide both end-to-end and partial-coverage service. We examine this case later.

We start with a simple case where only end-to-end services are demanded. Suppose there are two complementary types of goods, A and B. Suppose that each type of good has a number of brands available, $A_i$, $i = 1, \ldots, m$, $B_j$, $j = 1, \ldots, n$, as in Figure 8.4. Let consumers demand 1:1 combinations $A_iB_j$. We call each of the

[23] Standard 1 is an equilibrium if $a > e$, $b > d$. Similarly, standard 2 is an equilibrium if $g > c$, $h > f$.

Comp. by: DKandavel Date:20/11/06 Time:12:23:51 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



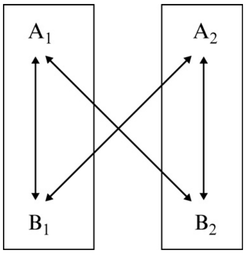


Figure 8.8. Mix-and-match compatibility.

complementary goods $A_i$ or $B_j$ *components*, while the combined good $A_iB_j$ is called a *composite good* or *system*. Potentially all combinations $A_iB_j$, $i = 1, \ldots, m$; $j = 1, \ldots, n$, are possible. Thus complementarity exists in potential. Complementarity is actualized when the components $A_i$ and $B_j$ are combinable and function together without extra cost, i.e. when the components are *compatible*. Often it is an explicit decision of the producers of individual components to make their products compatible with those of other producers. Thus, compatibility is a *strategic* decision and should be analyzed as such.

Modern industrial organization provides a rich collection of environments for the analysis of strategic decisions; because of shortage of time and space, this chapter will discuss the decision on compatibility in only a few environments.

### *Mix and match: compatibility vs. incompatibility*

The mix-and-match literature does not assume *a priori* network externalities; however, it is clear that demand in mix-and-match models exhibits network externalities. The mix-and-match approach was originated by Matutes and Regibeau (1988), and followed by Economides (1988, 1989a, 1991a, 1991b, 1993c), Economides and Salop (1992), Economides and Lehr (1995), Matutes and Regibeau (1989, 1992), and others. To fix ideas, consider the case of Figure 8.4 with $m = 2$, $n = 2$, technologies are known, coordination is costless, price discrimination is not allowed, and there are no cost asymmetries created by any particular compatibility standard. Figure 8.8 shows the case of compatibility. The incentive for compatibility of a vertically integrated firm (producing $A_1$ and $B_1$) depends on the relative sizes of each combination of complementary components. Reciprocal compatibility (i.e. simultaneous compatibility between $A_1$ and $B_2$, as well as between $A_2$ and $B_1$) increases demand (by allowing for the sale of $A_1B_2$ and $A_2B_1$) but also increases competition for the individual components. Therefore, when the hybrid demand is

Comp. by: DKandavel Date:20/11/06 Time:12:23:51 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



large compared with the own-product demand (including the case where the two demands are equal at equal prices), a firm has an incentive to want compatibility.[24] When the demand for hybrids is small, a firm does not want compatibility. Thus, it is possible, with two vertically integrated firms, that one firm wants compatibility (because it has small own-product demand compared with the hybrids demand) while the other one prefers incompatibility (because its own-product demand is large compared with the hybrids demand). Thus, there can be conflict across firms in their incentives for compatibility, even when the technology is well known. The presumption is that opponents will not be able to counteract and correct all incompatibilities introduced by an opponent, and therefore in situations of conflict we expect that incompatibility wins.

These results hold both for zero-one decisions – i.e. compatibility vs. incompatibility – and for decisions of partial (or variable) incompatibility. The intuition of the pro-compatibility result for the zero-one decision in the equal hybrid- and own-demand is simple. Starting from the same level of prices and demand in both the compatibility and incompatibility regimes, consider a price increase in one component that produces the same decrease in demand in both regimes. Under incompatibility, the loss of profits is higher since *systems* sales are lost rather than sales of *one component*. Therefore, profits are more responsive to price under incompatibility; it follows that the residual demand facing firms is more elastic under incompatibility, and therefore firms will choose lower prices in that regime.[25] This is reminiscent of Cournot's (1838) celebrated result that a vertically integrated monopolist faces a more elastic demand and will choose a lower price than the sum of the prices of two vertically disintegrated monopolists.[26]

So far we have assumed that compatibility is reciprocal – i.e. that the same adapter is required to make both $A_1B_2$ and $A_2B_1$ functional. If compatibility is not reciprocal – i.e. if different adapters are required for $A_1B_2$ and $A_2B_1$ – the incentive of firms to achieve compatibility depends on the cross substitution between own-products and hybrids. Roughly, if the substitutability among A-type components is equal to the

[24] Matutes and Regibeau (1988) and Economides (1989a) find that compatibility is always the firm's choice because they assume a locational setting with uniform distribution of consumers in space that results in equal own-product and hybrid demands at equal prices. The exposition here follows the more general framework of Economides (1988, 1991a).

[25] These results also hold when firms can price-discriminate between buyers who buy the pure combination $A_iB_i$ and buyers who buy only one component from firm i. Thus, firms practice mixed bundling. See Matutes and Regibeau (1992) and Economides (1993c).

[26] See Economides (1988) for a discussion of Cournot's result, and Economides and Salop (1992) for an extension of the result to (parallel) vertical integration among two pairs of vertically related firms.

Comp. by: DKandavel Date:20/11/06 Time:12:23:51 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



substitutability among B-type components, the earlier results of the reciprocal setup still hold.[27] Nevertheless, if the degree of substitutability among the As is different than among the Bs, one firm may create an advantage for itself by introducing some incompatibilities. However, it is *never* to the advantage of *both* vertically integrated firms to create incompatibilities.

The issue of compatibility and coordination is much more complicated if there are more than two firms. A number of coalitions can each be formed around a specific technical standard, and standards may allow for partial compatibility, or may be mutually incompatible. Not enough research has been done on this issue. Research in this area is made particularly difficult by the lack of established models of coalition formation in non-cooperative settings. The analysis based on coalition structures is more complicated in the "micro" approach because of the specifics of the ownership structure.

The studies we have referred to this far take the ownership structure as given (i.e. as parallel vertical integration) and proceed to discuss the choice of the degree of compatibility. In many cases, vertical integration is a decision that is more flexible (and less irreversible) than a decision on compatibility. Thus, it makes sense to think of a game structure where the choice of technology (which implies the degree of compatibility) *precedes* the choice of the degree of vertical integration. Economides (1996b) analyzes the choice of asset ownership as a consequence of the choice of technology (and of the implied degree of compatibility). It posits a three-stage game of compatibility choice in the first stage, vertical integration in the second stage, and price choice in the third stage. Incentives for vertical mergers in industries with varying degrees of compatibility are compared. In analyzing the stage of compatibility choice, the influence of the anticipation of decisions on (vertical) industry structure on compatibility decisions is evaluated.

### *Changes in the number of varieties as a result of compatibility decisions*

Economides (1991b) considers the interplay of compatibility and the number of varieties of complementary goods. There are two types of goods, A and B, consumed in 1:1 ratio. There are two brands of good A, $A_1$ and $A_2$, each produced by an independent firm. The number of B-type brands, each also produced by an independent firm, is determined by a free-entry condition, so that industry B is in monopolistic competition.

[27] Economides (1991a), p. 52.

Comp. by: DKandavel Date:20/11/06 Time:12:23:51 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



In a regime of compatibility, each B-type component is immediately compatible with either $A_1$ or $A_2$. In a regime of incompatibility, each brand $B_i$ produces two versions, one compatible with $A_1$ and one compatible with $A_2$. The two cases are shown in Figures 8.10 and 8.11.

Under incompatibility, each B-type firm incurs higher fixed costs; it follows that *ceteris paribus* the number of B-type brands will be smaller under incompatibility. An A-type firm prefers incompatibility or compatibility according to the equilibrium profits it realizes in each regime. These profits, and the decision on compatibility, depend on the specifics of the utility function of consumers, and in particular on the impact of an increase in the number of varieties on utility. If industry demand is not sensitive to increases in the number of varieties of composite goods n (and does not increase much as n increases), then equilibrium profits of an A-type firm decrease in the number of firms; therefore profits of an A-type firm are higher at the smaller number of firms implied by incompatibility, and an A-type firm prefers incompatibility. Conversely, when consumers have a strong preference for variety and demand for composite goods increases significantly in n, equilibrium profits of an A-type firm increase in the number of firms; therefore its profits are higher at the larger number of firms implied by compatibility, and an A-type firm prefers compatibility.

Church and Gandal (1992b) and Chou and Shy (1990a,b,c) also examine the impact of the number of varieties of complementary

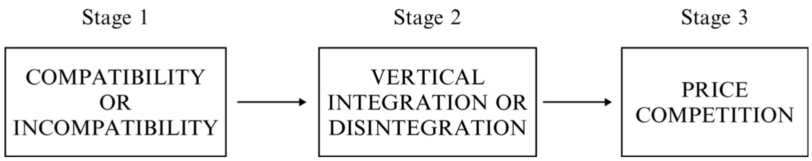


Figure 8.9. Compatibility decisions are less flexible than vertical integration decisions.

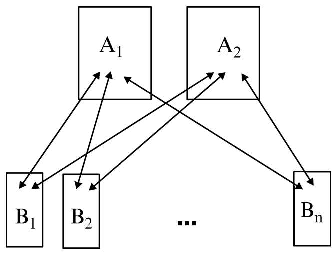


Figure 8.10. Compatibility.

Comp. by: DKandavel Date:20/11/06 Time:12:23:51 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



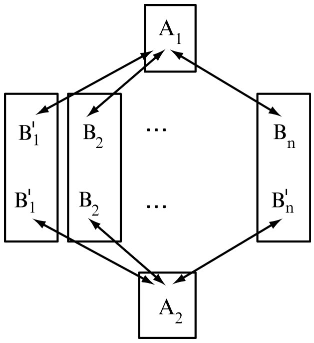


Figure 8.11. Incompatibility.

(B-type) goods on the decisions of consumers to buy one of the A-type goods under conditions of incompatibility.

*Quality coordination in mix and match*

The framework of mix-and-match models applies to both variety and quality features that are combinable additively in the utility function. That is, in the standard mix-and-match model, the utility accruing to a consumer from component $A_i$ is added to the utility from component $B_j$. However, in some networks, including telecommunications,[28] the utility of the composite good $A_iB_j$ is not the sum of the respective qualities. In particular, the quality of voice in a long-distance call is the minimum of the qualities of the component parts of the network, i.e. the local and the long-distance transmission. Thus, significant quality coordination problems arise in a network with fragmented ownership. Economides (1994b) and Economides and Lehr (1995) examine this coordination problem.

Let A and B be components that are combinable in a 1:1 ratio. Suppose that the quality levels of the components are $q_A$ and $q_B$, while the quality level of the composite good is $q_{AB} = \min(q_A, q_B)$. Consumers vary in their willingness to pay for quality improvements as in Gabszewicz and Thisse (1979) and Shaked and Sutton (1982), and firms play a two-stage game of quality choice in the first stage, followed by price choice in the second stage. As mentioned earlier, Cournot (1838) has shown that an integrated monopolist producing both A and B will charge less than

[28] See also Encaoua et al. (1992) for a discussion of the coordination of the timing of different legs of airport transportation.

Comp. by: DKandavel Date:20/11/06 Time:12:23:51 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



two vertically related monopolists, each producing one component only. This is because of the elimination of double marginalization by the integrated monopolist. Economides (1994b) and Economides and Lehr (1995) show that an integrated monopolist also provides a higher quality than the two independent monopolists. In bilateral monopoly, marginal increases in quality have a bigger impact on price. Being able to sell the same quality at a higher price than under integrated monopoly, the bilateral monopolists choose lower quality levels, which are less costly. Despite that, because of double marginalization, prices are higher than in integrated monopoly, a lower portion of the market is served, and firms realize lower profits.[29] Thus, *lack of vertical integration leads to a reduction in quality*. Note that this is not because of lack of coordination between the bilateral monopolists in the choice of quality, since they both choose the same quality level.[30]

In this setting, Economides and Lehr (1995) examine various ownership structures where for at least one of the types of components there is more than one quality level available. Clearly, a situation where all components have the same quality is not viable, since competition would then drive prices to marginal cost. Further, for a "high" quality composite good to be available, both an A- and a B-type good must be of "high" quality. They find that third (and fourth) "low" quality goods have a hard time surviving if they are produced by independent firms. In contrast, in parallel vertical integration (with firm i, $i = 1, 2$, producing $A_i$ and $B_i$), firms prefer not to interconnect – i.e. to produce components that are incompatible with those of the opponent.

## 8.4 Network externalities and industry structure

### *8.4.1 Invitations to enter*

In the presence of strong network externalities, a monopolist exclusive holder of a technology may have an incentive to invite competitors and even subsidize them. The realization of network externalities requires high output. A monopolist may be unable credibly to commit to a high output as long as he is operating by himself. However, if he licenses the technology to a number of firms and invites them to enter and compete with him, market output will be higher; and since the level of market

[29] Consumers also receive lower surplus in comparison to vertically integrated monopoly.
[30] The *reliability* of the network, measured by the percentage of time that the network is in operation, or by the probability of a successful connection, is measured by the product of the respective reliabilities of the components (another non-linear function).

Comp. by: DKandavel Date:20/11/06 Time:12:23:51 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



output depends mainly upon other firms, the commitment to high output is credible.

The invitation to enter and the consequent increase in market output has two effects: a *competitive effect* and a *network effect*. The competitive effect is an expected increase in competition because of the increase in the number of firms. The network effect tends to increase the willingness to pay and the market price because of the high expected sales. Economides (1993b, 1996a) shows that if the network externality is strong enough, the network effect is larger than the competitive effect, and therefore an innovator-monopolist invites competitors and even subsidizes them on the margin to induce them to increase production.

### *8.4.2 Interconnection or foreclosure by a local monopolist?*

Many telecommunications, airline and railroad networks have the structure of Figure 8.12. In a railroad network, there may be direct consumer demand for links AB, BC, as well as AC. This figure can also represent a telephone network with demand for local telephone services (AB) and for long-distance services (ABC); in that case, there is no direct demand for BC, but only the indirect demand arising from long-distance calls ABC. In many cases, one firm has a monopoly of a link that is necessary for a number of services (here AB), and this link is a natural monopoly. This bottleneck link is often called an essential facility. The monopolist can foreclose any firm by denying access to the bottleneck facility. What are his incentives do so?

Economides and Woroch (1992) examine intermodal competition in the context of a simple network pictured in Figure 8.13. S and R are local switches; AS and BR are local services (in different cities); SR and STR are alternative long-distance services. The diagram is simplified by eliminating R without any essential loss. Suppose that an integrated firm offers end-to-end service (ASB), while a second firm offers service of partial coverage only (STB). They find that although the integrated firm has the opportunity to foreclose the opponent, it prefers not to. In fact, the integrated firm is better off by implementing a vertical price squeeze on the opponent and charging a significantly higher price to the opponent for the use of the monopolized link than it "charges" itself.[31] Thus, foreclosure, although feasible, is not optimal for the monopolist.[32]

[31] This result is dependent on the linear structure of the demand system, and may not hold for any demand structure.

[32] Church and Gandal (1992a) find that sometimes firms prefer foreclosure, but their model does not allow for a vertical price squeeze.

Comp. by: DKandavel Date:20/11/06 Time:12:23:51 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



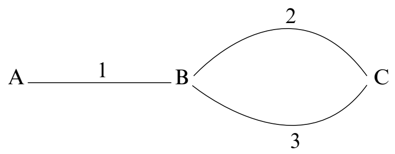


Figure 8.12. *AB* is a bottleneck facility.

Network in extensive and collapsed form

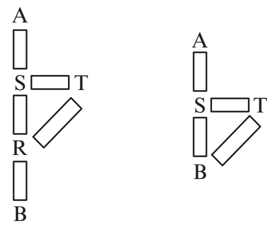


Figure 8.13. Intermodal competition.

Economides and Woroch (1992) also find that vertical disintegration is not desirable for the firm that offers end-to-end service. Once disintegrated, its constituent parts realize lower total profits. This is because, besides appropriating monopoly rents for its AS monopoly, the integrated firm (ASB) was creating a significant restriction of competition in the SB-STB market by its *de facto* price-discriminating strategy. After disintegration, the SB-STB market becomes much more competitive, even if AS price-discriminates between SB and STB. Thus, even if network ASB were to receive the full rent earned by the new owner of SB, its after-divestiture profits would be lower than before divestiture.[33]

Even in simple networks, there may be relations among firms that are neither purely vertical nor purely horizontal. Thus, the conventional wisdom about vertical and horizontal integration fails. Economides and Salop (1992) discuss pricing in various ownership structures in the model of Figure 8.8. They call the ownership structure of this figure, where each firm produces a component of each type, *parallel vertical integration*. They also consider the *independent ownership* structure, where each of the four components is owned by a different firm. In both of these structures, no firm is purely vertically or purely horizontally related

[33] This result is in contrast to Bonanno and Vickers (1988) because of the absence of two-part contracts in Economides and Woroch (1992).

Comp. by: DKandavel Date:20/11/06 Time:12:23:52 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



to another firm. Thus, starting from independent ownership, or starting from parallel vertical integration, a merger to *joint ownership*, where all components are produced by the same firm, can either increase or decrease prices. Thus, simple prescriptions against mergers may easily fail.

In the model of Figure 8.13, Economides and Woroch (1992) consider the case where link ST is owned by a firm that owns a vertically related link (either AS or BT), or is owned by an independent firm. Clearly, the strategic structure of the game remains unaffected when link ST changes hands between two firms that also own a link that is vertically related to ST. Therefore, if ST has a fixed cost, it is a liability to such a firm; each firm would like the opponent to own it. However, if the link is owned by a third party, it has a positive value because of its monopoly position in the chain. Thus, each original owner has an incentive to sell ST to a third party. The direct implication is that the value of links depends on what other links a firm owns. Thus, general prescriptions on the desirability of unbundling of ownership are suspect.

Often parts of the network are regulated, while other parts are not. This is the typical arrangement in telephony in the US, where only local telephone companies are tightly regulated, since their market is traditionally considered a natural monopoly.[34] Baumol and Sidak (1994a,b) propose that to attract efficient entrants in the long-distance market and to discourage inefficient entrants, a local telephone company should charge them an *interconnection (or access) fee* equal to the marginal cost of provision of service plus any opportunity cost that the local telephone company incurs.[35] This is correct under a set of strict assumptions: first, that the end-to-end good is sold originally at the competitive price; second, that the entrant produces the same complementary good (long-distance service) as the incumbent;[36] third, that there are no economies of scale in either one of the complements. Economides and White (1995) discuss how the relaxation of these assumptions leads to different interconnection charges.[37] For example, if competition between an entrant and the incumbent reduces the market power of the incumbent, entry may increase social welfare even when the entrant produces at higher cost than the incumbent.

[34] This is changing for some customers through the existence of competitive access providers, which directly compete with the local telephone company for large customers, and the potential for competition from cable companies.
[35] Kahn and Taylor (1994) have very similar views.
[36] Armstrong and Doyle (1994) relax this assumption.
[37] See also Ergas and Ralph (1994).

Comp. by: DKandavel Date:20/11/06 Time:12:23:52 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



## 8.5 Sequential games

In network markets, and more generally in markets with network externalities, when firms and consumers interact in more than period, *history matters*. Both consumers and firms make production and consumption decisions based on sizes of installed base and on expectations of its increases over time. The same underlying technology and consumer preferences and distribution can lead to different industrial structures depending on the way things start. Thus, strategic advantages, such as first-mover advantages, can have long-run effects.[38]

Network externalities and historical events are particularly important in the speed of adoption of an innovation that creates services on a network. Cabral (1990) discusses the adoption of innovations under perfect competition in the presence of network externalities. His main conclusion is that when network externalities are strong, the equilibrium adoption path may be discontinuous. This is another way of saying that there are two network sizes supported as equilibria at the same time. This may occur at the start of the network, and then it is called positive critical mass by Economides and Himmelberg (1995). It may also occur at other points in the network evolution. In practice, discontinuities in the size of the network over time do not occur since that would imply an infinite size of sales at some points in time. Continuity and smoothness of the network path is restored if instantaneous marginal production costs are increasing. Under this assumption, Economides and Himmelberg (1995) find that the adoption path is much steeper in the presence of externalities. Further, driven by the externality, in early stages the network can expand so quickly as to exhibit increasing retail prices even when marginal costs are falling over time. Their analysis is applied to the fax market in the US and Japan.

The analysis is more complex when we depart from the assumption of perfect competition. Accordingly, this analysis tends to be in the form of simple two-period models. We analyze it with reference to the standard simultaneous choice coordination game above, where we now interpret the first strategy as sticking to the old technology and the second as the adoption of a new one.

Network externalities for both technologies mean that $a > c, e; b > d, f; g > c, e; h > d, f$. If both firms are worse off when they are not coordinated, both the "new technology" (i.e. (N, N)) and the "old technology"

[38] See Arthur (1988, 1989) and David (1985). David argues that the QWERTY keyboard was adopted mainly because it appeared first while the DVORAK keyboard was superior. This is disputed by Liebowitz and Margolis (1990).

Comp. by: DKandavel Date:20/11/06 Time:12:23:52 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



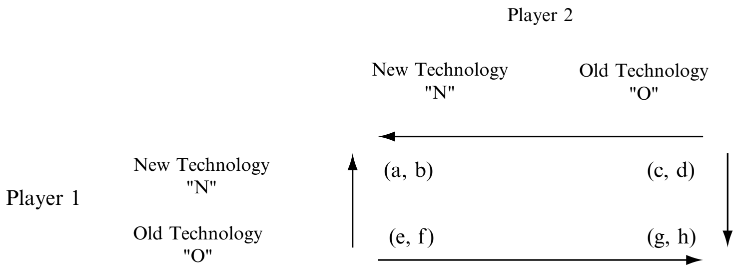


Figure 8.14. Choice between old and new technology.

(i.e. (O, O)) will arise as equilibria. Clearly, one of the equilibria can be inefficient. If the (O, O) equilibrium is inefficient and is adopted, Farrell and Saloner (1985) call the situation *excess inertia*. Similarly, if the (N, N) equilibrium is inefficient and it is adopted, the situation is called *excess momentum*.

Farrell and Saloner (1985) discuss a two-period model where consumers vary in their willingness to pay for the change of the technology, measured by $\theta$. Users can switch in period 1 or 2, and switching is irreversible. Users fall into four categories according to the strategy they pick: (i) they never switch, whatever the behavior of others in the first period; (ii) they switch in period 2 if other users have switched in period 1 – jumping on the bandwagon; (iii) they switch in period 1; (iv) they switch in period 2 even if others have not switched in period 1. The last strategy is dominated by strategy (iii). Consumers of low $\theta$ use strategy (i), consumers of intermediate $\theta$ use strategy (ii), and consumers of high $\theta$ use strategy (iii). Consumers would like to coordinate themselves and switch in the first period (thereby getting the bandwagon rolling) but are unable to do so, thus creating excess inertia.[39] This inertia can be reduced through communication among the consumers, through contracts, through coordination in committees, or through new product sponsorship and special introductory pricing.[40]

In a sequential setting, preannouncement (i.e. announcement of a new product before its introduction) may induce some users to delay their purchase. Also penetration pricing can be important. Katz and Shapiro (1986a) examine the effects of sponsorship (allowing firms to

[39] See Katz and Shapiro (1992) for a different view arguing for excess momentum (which they call *insufficient friction*).

[40] See also Farrell and Saloner (1988) for mechanisms to achieve coordination, and Farrell and Saloner (1985) for a discussion of network product sponsorship.

Comp. by: DKandavel Date:20/11/06 Time:12:23:52 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



price differently than at marginal cost). Katz and Shapiro (1986b) examine the effects of uncertainty in product adoption and introduction.

Nevertheless, there is much more work to be done on multi-period and on continuous-time dynamic games with network externalities. The issues of foreclosure and predation have not been sufficiently discussed in the context of network externalities. More generally, much more work is required on multi-period dynamic games in this context, especially for durable goods.

## 8.6 Markets for adapters and add-ons

Not enough research has been done on the economics of adapters and interfaces. One strand of the mix-and-match literature assumes that compatibilities introduced by one firm cannot be corrected by the other, so that adapters are unfeasible. Economides (1991a) assumes that adapters are provided by a competitive industry at cost, but firms' decisions determine the extent of incompatibility, and therefore the cost of the adapters. Farrell and Saloner (1992) assume that converters make the technologies only partially compatible, in the sense that hybrid goods that utilize incompatible components as well as an adapter give lower utility than a system composed of fully compatible components. In this framework, the availability of converters can *reduce* social welfare, since, in the presence of converters, some consumers would buy the converter and the "inferior" technology rather than the "best" technology, although the "best" technology gives more externalities.

## 8.7 Concluding remarks

In this chapter, we have noted some of the interesting issues that arise in networks and vertically related industries, especially in the presence of a fragmented ownership structure. Evidently, many open questions remain. One of the most important issues still largely unresolved is the joint determination of an equilibrium market structure (including the degree of vertical integration) together with the degree of compatibility across firms. The extent of standardization in markets with more than two participants and the structure of "standards" coalitions also remain open questions. Markets for adapters and add-ons have not been sufficiently analyzed. An analysis of market structure in multi-period dynamic games with network externalities is also unavailable. Further, issues of predation and foreclosure in networks have not been fully analyzed yet. On a more fundamental level, there is no good prediction yet of the "break points" that define the complementary components in

Comp. by: DKandavel Date:20/11/06 Time:12:23:52 Stage:1st Revises
File Path://spiina1001z/womat/production/PRODENV/0000000009/0000000102/
0000000005/0000048928.3D Proof by: QC by: Author: Economides



a modular design structure. Even if these break points are known, little analysis has been done of competition in a multi-layered structure of vertically related components. Nevertheless, it is exactly this kind of modelling that is needed for an analysis and evaluation of the potential structures of the "information superhighway".