13 Public Policy in Network Industries

Nicholas Economides

Network industries are a large, significant, and often fast-growing part of the world economy. A key network industry is telecommunications, providing voice and data services, including the Internet and the World Wide Web. Another key network industry is computer software and hardware. These two sectors, telecommunications and computers, have been the engines of fast growth of the world economy. In the news and entertainment sector, network industries include broadcasting and cable television, which in recent years are reaching into traditional telecommunications services. In transportation, networks include airlines, railroads, roads, and shipping, and the delivery services that "live" on these, such as the postal service and its competitors. In the financial sector, networks include traditional financial exchanges for bonds, equities, and derivatives, clearing houses, B2B and B2C exchanges, credit and debit card networks, as well as automated transactions banking networks, such as ATM networks.

Besides traditional network industries, many of the features of networks apply to *virtual networks*. A virtual network is a collection of compatible goods that share a common technical platform. For example, all VHS video players make up a virtual network. Similarly all computers running Windows can be thought of as a virtual network. Compatible computer software and hardware make up a network, and so do computer operating systems and compatible applications. More generally, networks are composed of complementary components, so they also encompass wholesale and retail networks, as well as information networks and servers such as telephone yellow pages, Yahoo, MSN, and Google.

In recent years high-technology industries have been playing an even more central role in the US and world economy, exhibiting very fast growth and high valuations of their equity. Many of the high-technology industries are based on networks (e.g., the telecommunications network and the Internet). Other high-technology industries, such as the computer software and hardware industries, exhibit properties that are typically observed in networks, as will be explained in detail. So, to understand the "new economy," we need to understand the economics of networks. Another key infrastructure industry, electricity production and distribution, is also a network industry.

Adding to the importance of networks from a public policy point of view is the fact that network industries often provide necessities. Monopolization in such a setting can have significant social and political implications.

There may be a number of anticompetitive concerns in a network industry. The focus of this chapter are the following questions: Since network industries have special common features, are there special competition policy issues arising out of key features of network industries? If yes, what is the framework of the public policies that can be pursued to address these issues?

13.1 The Logic of Competition Law

The logic of competition and antitrust law in the United States and the European Union is to guard against restrictions and impediments to competition that are not likely to be naturally corrected by competitive forces. For this chapter, I will posit that the maximization of efficiency (allocative, productive, and dynamic) is the desired outcome of competition and antitrust law, and that typically competition is the means of achieving efficiency.

As an alternative to antitrust and competition law, economic regulation has been established in three exceptional case: (1) for those markets where it is clear that competition cannot be achieved by market forces, (2) for markets where deviation from efficiency is deemed socially desirable, and (3) for markets where the social and private benefits are clearly different. In each of these cases it is clear that a market without intervention will not result in the desired outcome. In the first case, this is true by the definition of the category. In the second case, markets may lead to efficiency, but society prefers a different outcome, so intervention is necessary to achieve this. In the third case, maximization of social surplus does not coincide with maximization of the sum of profits and consumers' surplus because of "externalities."

Some key network industries are regulated at least in part or in some aspects. Telecommunications has significant regulation at both the federal and state level. Railroads, electricity, and air and ground transportation are also heavily regulated. Financial exchanges are under "light" regulation and to a significant extent under self-regulation. In contrast, B2B exchanges, credit card, and banking networks, as well as computers and their virtual networks, are almost completely unregulated.

A full discussion of the merits and problems with regulation of each of these network industries is impossible in the context of this brief chapter. Instead, I will outline the parameters that would necessitate regulation or deregulation based on the broad features of network markets. In future work, I will examine the full application of these principles in all network industries. I expect to observe that the principles of economic regulation are not applied equally to all industries, and in a number of cases, the present regulatory regime is based on historical reasons (political, social, and technological) and cannot be justified based on the application of the economic principles to the present technology.

13.2 Special Features of Markets with Network Effects

13.2.1 Sources of Network Effects and the Reversal of the Law of Demand

Many network industries exhibit increasing returns to scale in production: unit cost (average cost) decreases with as production increases. Often incremental cost is negligible, for example, in software. But these are also features of nonnetwork industries and are *not* the defining feature of network industries. Thus increasing returns to scale in production or negligible incremental production cost are not the defining feature of the competition policy issues that are rooted in the existence of networks.

Networks are composed of *complementary* nodes and links. The crucial defining feature of networks is the complementarity among the various nodes and links. A service delivered over a network requires the use of two or more network components. Thus, network components are complementary to each other.

Figure 13.1 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, and so forth. There are substitutes for each of these components. For example, transmission can be done through a cable TV line, a fixed telephone line, a wireless satellite, or PCS; the in-home interface may be a TV-top box or an add-on to a PC. It is likely that the combinations of various components will result in substitute but not identical services.

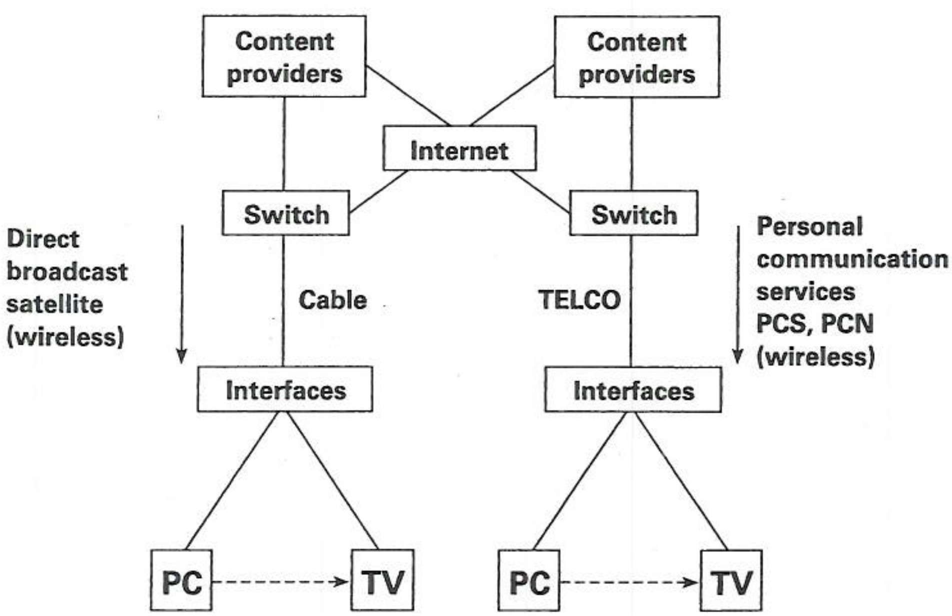

Figure 13.1
The Information superhighway

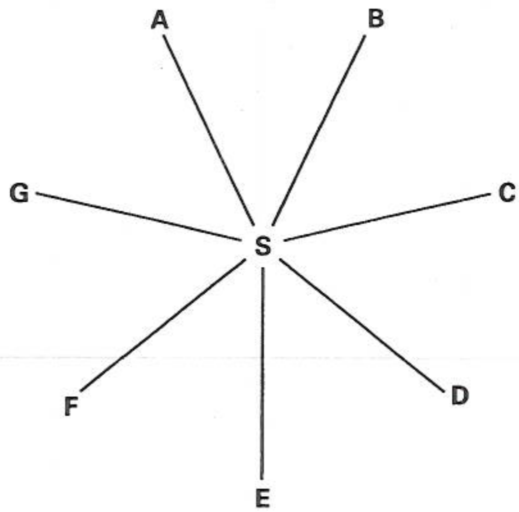

Figure 13.2
A star network

Thus the information superhighway will provide substitutes made of complements, and this is a typical feature of networks.

A common and defining feature of network industries is that they exhibit increasing returns to scale in consumption, commonly called network effects. The existence of network externalities is the key reason for the importance, growth, and profitability of network industries and the "new economy." A market exhibits network effects (or network externalities) when the value to a buyer of an extra unit is higher as more units are sold, everything else being equal.[1]

Network effects arise because of complementarities. In a traditional network, network externalities arise because a typical subscriber can reach more subscribers in a larger network. Figure 13.2 depicts a traditional telecommunications network where customers $A, B, \ldots, G$ are connected to a switch at S. Although goods "access to the switch," $AS, BS, \ldots, GS$ have the same industrial classification, and traditional economics would classify them as substitutes, they are used as *complements*. In particular, when customer A makes a phone call to customer B, he uses *both* AS and BS.

Networks where services AB and BA are distinct are called "two-way" networks. Two-way networks 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, then the network is called a one-way network. In a typical one-way network there are two types of components. 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, radio and TV broadcasting and early paging networks are one-way networks.

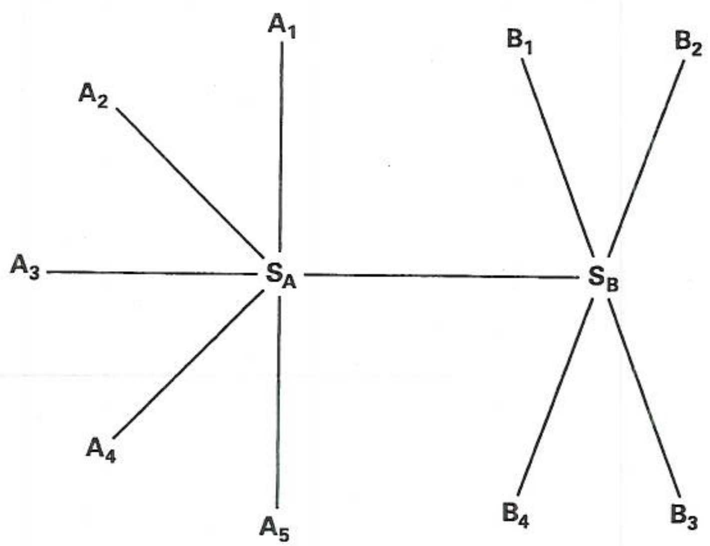

Figure 13.3
A long-distance network (alternatively, an ATM network)

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 13.3 can be interpreted as a two-way telephone network where *SA* represents a local switch in city *A*, A_1, represents a customer in city *A*, and similarly for *SB* and B_j. We may identify end nodes, such as A_i and B_i, 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. In this network there are two types of local phone calls $A_iS_AA_k$ and $B_jS_BB_R$, as well as long-distance phone calls $A_iS_AS_BB_j$. We can also interpret the network of figure 13.3 as an Automatic Teller Machine 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 $B_jS_BB_R$ may be feasible but there is no demand for them.

A *virtual network* can be thought of as a collection of compatible goods that share a common technical platform. For example, all VHS video players make up a virtual network. Similarly all computers running Windows XP can be thought of as a virtual network. More generally, a virtual network can be thought of a combination of two collections of two types of goods $\{A_1, \ldots, A_m\}$ and $\{B_1, \ldots, B_n\}$ such that (1) each of the *A*-type good is a substitute to any other *A*-type good, (2) each of the *B*-type good is a substitute to any other *B*-type good, and (3) each of the *A*-type good is a complement to any *A*-type good. Virtual networks are one-way networks. Examples of virtual networks are computer hardware and software, computer operating systems and software applications, cameras and compatible film, and razors and compatible blades. There are many more virtual networks than traditional networks.

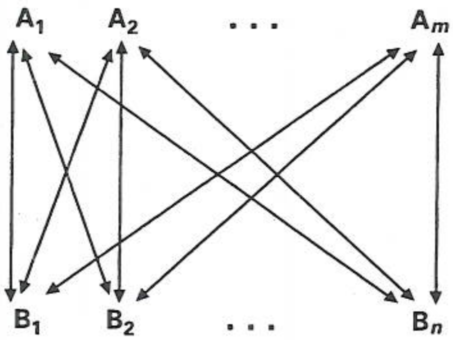

Figure 13.4
A virtual network of complementary goods

In a virtual network, externalities arise because larger sales of components of type A induce larger availability of complementary components $B_1, \ldots, B_n$, thereby increasing the value of components of type A. As figure 13.4 shows, the increased value of component A results in further positive feedback. Despite the cycle of positive feedbacks, it is typically expected that the value of component A does not explode to infinity because the additional positive feedback is expected to decrease with increases in the size of the network.[2]

The crucial relationship in both one-way and two-way networks is the complementarity between the pieces of the network. This crucial economic relationship is also often observed between different classes of goods in nonnetwork industries. Figure 13.4 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 the long-distance network of figure 13.3 in the ATM interpretation.

In traditional nonnetwork industries the willingness to pay for the last unit of a good decreases with the number of units sold. This is called *the law of demand*, and it is traditionally considered to hold for almost all goods.[3] However, the existence of network effects implies that as more units are sold, the willingness to pay for the last unit is higher. This means that for network goods, the fundamental law of demand is violated, so some portions of the demand curve can slope upward. In other words, for some portions of the demand curve, as sales expand, people are willing to pay more for the last unit.

The law of demand is still correct if one disregards the effects of the expansion of sales on complementary goods. But, as increased sales of a network good imply an expansion in the sales of complementary goods, the value of the last unit increases. Combining the traditional downward-slopping effect with the positive effect due to network expansion can result in a demand curve that has an upward-slopping part.

The key reason for the appearance of network externalities is the complementarity among network components. Depending on the network, the network effect may be direct or indirect. When customers are identified with components, the network effect is direct. Consider a typical two-way network, such as the local telephone network of figure 13.2.

In this n nodes two-way 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 (e.g., ES) to the existing links.[4]

In typical one-way networks, the network effect is only indirect. When there are m varieties of component A and n varieties of component B as in figure 13.4 (and all A-type goods are compatible with all of B-type), there are $m \times n$ potential composite goods. An extra customer yields indirect externalities to other customers, by increasing the demand for components of types A and B. In the presence of economies of scale in production, the increase in demand may potentially increase the number of varieties of each component that are available in the market.

Exchange networks or financial networks (e.g., the NYSE and NASDAQ, commodities, futures, and options exchanges as well as business-to-business "B2B" exchanges), also exhibit indirect network externalities. There are two ways in which these externalities arise. First, externalities arise in the act of exchanging assets or goods. Second, externalities may arise in the array of vertically related services that compose a financial transaction. These include the services of a broker, bringing the offer to the floor, and matching the offer. The second type of externalities are 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 this creates a composite good, the "exchange transaction." The two original goods are complementary, and each has no value without the other one. 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 and business-to-business exchanges 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.[5] Thus financial exchange markets also exhibit network externalities.[6]

As I noted earlier, network externalities arise out of the complementarity of different network pieces. 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.

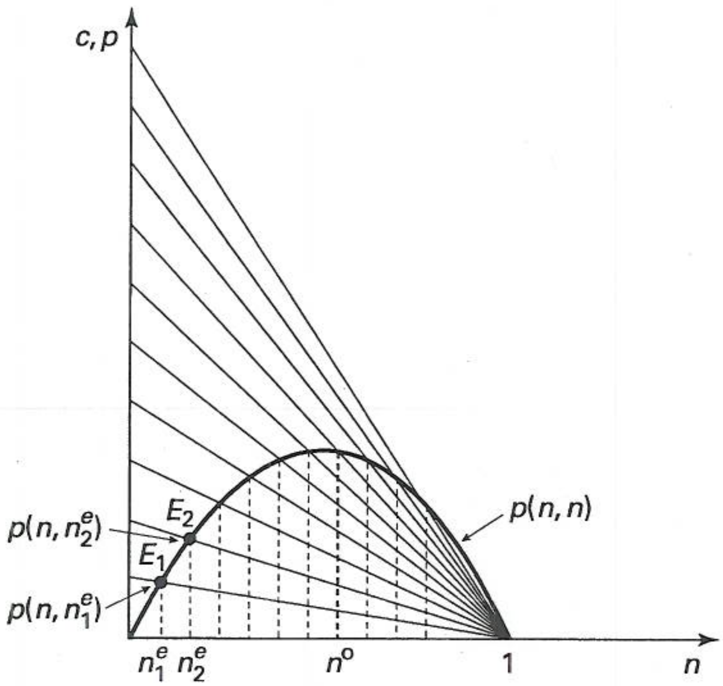

Figure 13.5
Fulfilled expectations demand and critical mass

To understand this better, consider a fulfilled expectations formulation of network externalities. Let the willingness to pay for the nth unit of the good when n^e units are expected to be sold be $p(n, n^e)$. In this formulation, n and n^e are normalized so that they represent market coverage, ranging from 0 to 1, rather than absolute quantities. Willingness to pay $p(n, n^e)$ is a decreasing function of its first argument because the demand slopes downward. As $p(n, n^e)$ increases in n^e, this captures the network externalities effect, meaning the good is more valuable when the expected sales n^e are higher. At a market equilibrium of the simple single-period world, expectations are fulfilled, $n = n^e$, and this defines the fulfilled expectations demand $p(n, n)$.

Figure 13.5 shows the construction of a typical fulfilled expectations demand in a network industry. Each willingness-to-pay curve, $p(n, n_i^e)$, $i = 1, 2, \ldots$, shows the willingness to pay for a varying quantity n, given an expectation of sales $n^e = n_i^e$. At $n = n_i^e$, expectations are fulfilled and the point belongs to $p(n, n)$ as $p(n_i^e, n_i^e)$. Thus $p(n, n)$ is constructed as a collection of points $p(n_i^e, n_i^e)$. It is reasonable to impose the condition $\lim_{n \to 1} p(n, n) = 0$. This means that as the market is more and more covered, eventually the network reaches consumers who are willing to pay very little for the good even if they are able to reap very large network externalities. It follows that $p(n, n)$ is decreasing for large n. In figure 13.5 the fulfilled expectations demand at quantity zero is $p(0, 0) = 0$. This means that consumers think that the good has negligible value when its sales (and network effect) are zero. Although this is true for many network goods, some network goods have positive inherent value even at zero sales and no network effects. If the good has an inherent value k, $p(0,0) = k$, the fulfilled expectations demand curve in figure 13.5 starts at $(0,k)$.

Economides and Himmelberg (1995), taking off from Rohlfs (1974), have shown that the fulfilled expectations demand is increasing for small n if *either one* of three conditions holds:[7]

1. The utility of every consumer in a network of zero size is zero.
2. There are immediate and large external benefits to network expansion for very small networks.
3. There are a significant number 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, such as the telecommunications and fax networks where the good has no value unless there is another user to connect to. The other two conditions are more subtle but are commonly observed in networks and vertically related industries. The second condition holds for networks where the addition of even few users increases the value of the network significantly. A good example of this is a newsgroup on an obscure subject, where the addition of very few users starts a discussion and increases its value significantly. The third condition is most common in software markets. A software application has value to a user even if no one else uses it. The addition of an extra user has a network benefit to other users (because they can share files or find trained workers in the specifics of the application), but this benefit is small. However, when large numbers of users are added, the network benefit can be significant.

13.2.2 Critical Mass

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 as technology improves, the network will start at a positive and significant size n^0 (corresponding to marginal cost c^0). For each smaller marginal cost, $c < c^0$, 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 line through c with $p(n,n)$, and the Pareto optimal stable network size at the largest intersection of the horizontal line 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. In such a setting it is natural to assume that the Pareto optimal network size will result.

The existence of an upward-sloping part of the demand curve and the multiplicity of equilibria even under perfect competition also allows for a network to start with a small size and then expand significantly. For example, take the case where marginal cost is at $c < c^0$ and a new invention creates a new product with significant network effects. Then it is possible that the industry starts at the left intersection of the horizontal at c with $p(n, n)$ as expectations are originally low, and later on advances suddenly and quickly to the right intersection of the horizontal at c with $p(n, n)$. Thus the multiplicity of equilibria in network industries can lead to sudden significant expansions of network size.

13.2.3 Features of Markets with Network Effects

Ability to Charge Prices on Both Sides of a Network There are a number of fundamental properties of network industries that arise out of the existence of network effects.

A firm can make money from either side of the network or from both. For example, a telecommunications services provider can charge subscribers when they originate calls or when they receive calls or for both.[8] When a network consists of software clients and servers, both provided by the same firm, the firm can use the prices of the client and server software to maximize the network effect and its profits. For example, it can distribute the client software at marginal cost (free) and make all its profits from the server software. In a similar vein, Adobe distributes the "Acrobat Reader" free while it makes its profits from the "Acrobat Distiller" product, which allows the creation of files that can be read by the Acrobat Reader. The availability of prices on both sides of the network allows for complex pricing strategies and, depending on the dynamics and market shares on the two sides of the market, can be used strategically to enhance and leverage a firm's strong strategic position on one side of the network. Of course, this is not confined to high-technology or software industries; it applies wherever complementary components are present.[9]

Externalities May or May Not Be Internalized In network industries often the additional subscriber/user is not rewarded for the benefit that he/she brings to others by subscribing. Hence typically there are "externalities," meaning benefits not fully intermediated by the market. However, firms can use price discrimination to provide favorable terms to large users to maximize their network effect contribution to the market. For example, a large customer in a financial market can be given a very low price to be compensated for the positive network effect it brings to the market.[10]

Fast Network Expansion Generally, the pace of market penetration (network expansion) is much faster in network industries than in nonnetwork industries. In the earlier discussion on critical mass, we saw that in a one-period model, as unit cost decreases, the network starts with significant market coverage. In the presence of frictions and not perfectly elastic supply, the network expansion is not instantaneous from 0 to n^0 but rather is a rapid expansion following an S-shaped curve, as seen in figure 13.6. Notice how the market share expansion compares for a new good (diffusion) in the presence (delta = 1) and absence (delta = 0) of network effects as a function of time. The self-reinforcing nature of network effects leads to a much faster expansion when network effects are present.[11]

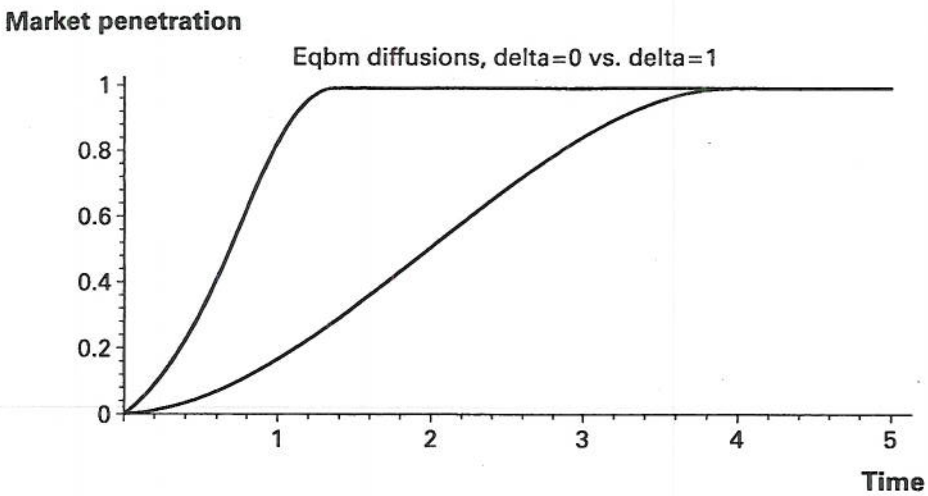

Figure 13.6
Diffusion of an innovation with and without network effects

Efficiency: Perfect Competition Is Inefficient 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 provides a smaller network than is socially optimal, and for some relatively high marginal costs, perfect competition does not provide the good while it is socially optimal to provide it.[12]

Since perfect competition is inefficient, state subsidization of network industries is beneficial to society. The Internet is a very successful network that was subsidized by the US government for many years. However, the subsidized Internet was aimed at promoting interaction among military research projects. During the period of its subsidization, almost no one imagined that the Internet would become a ubiquitous commercial network. Yet the foundation of the Internet on publicly and freely available standards has facilitated its expansion and provided a guarantee that no firm can dominate it.

Standards Wars So far we have assumed *compatibility*, namely that various links and nodes on the network are costlessly combinable to produce demanded goods. Two complementary components A and B are compatible when they can be combined to produce a composite good or service. For example, we say that a VHS-format video player is compatible with a VHS-format tape. Two substitute components A_1 and A_2 are compatible when each of them can be combined with a complementary good B to produce a composite good or service. For example, two VHS tapes are compatible. Similarly we say that two VHS video players are compatible. Also we say that two software products are compatible (more precisely *two-way* compatible) when they each can read and write files in a common format. So compatibility may be *one-way* when the files of format B_1 of software A_1 can be read by software A_2, but the files format B_2 of software A_2 cannot be read by software

A_1. Moreover compatibility may be only partial in the sense that software A_1 is able to read files of format B_2 but unable to write files in that format.

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. As we will see, it is not always in the best interests of a firm to allow full compatibility of its products with those of its competitors. The extent to which a firm is compatible with the products of other firms is an important strategic decision for a firm (as will be discussed in detail further on).

A key strategic decision for a firm is the extent to which it will be compatible with other firms. As I noted earlier, a network good has higher value because of the existence of network effects. Different firms conforming to the same technical standard can create a larger network effect while still competing with each other in other dimensions (e.g., quality and price). But even the decision to conform to the same technical standard is a strategic one. A firm can choose to be compatible with a rival and thereby create a larger network effect and share it with the rival. A firm could alternatively choose to be incompatible with the rival but keep all the network effects it creates to itself. Which way the decision will go depends on a number of factors. First, in some network industries, such as telecommunications, interconnection and compatibility at the level of voice and low-capacity data transmission are mandated by law. Second, the decision will depend on the expertise that a firm has on a particular standard (and therefore on the costs that it would incur to conform to it). Third, the choice on compatibility will depend on the relative benefit of keeping all the network effects to itself by choosing incompatibility versus receiving half of the larger network benefits by choosing compatibility. Fourth, the choice on compatibility depends on the ability of a firm to sustain a dominant position in an ensuing standards war if incompatibility is chosen. Finally, the compatibility choice depends on the ability of firms to leverage any monopoly power that they manage to attain in a regime of incompatibility to new markets.

Standards may be defined by the government (e.g., at the beginning of the Internet), a world engineering body (e.g., in the FAX), an industrywide committee, or just sponsored by one or more firms. Even when industrywide committees are available, firms have been known to introduce and sponsor their own standards.[13] The discussion here is on the *incentives* of firms to choose to be compatible with others. We first examine the simple case where standardization costs are different and firms play a coordination game. A 2×2 possible version of this game is presented below. Entries represent profits as shown in figure 13.7, where there is full compatibility at both noncooperative equilibria. The arrows

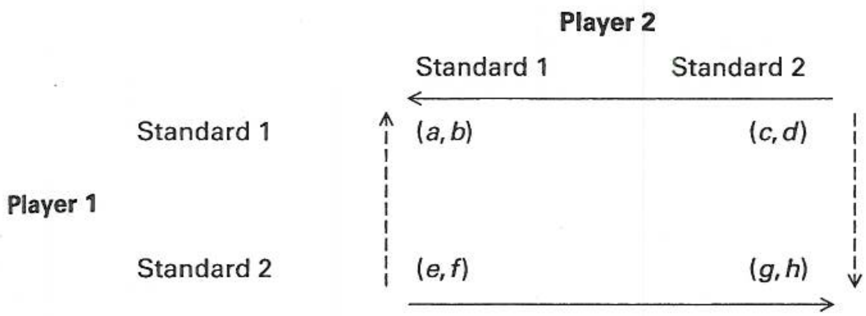

Figure 13.7
Standards war leading to incompatibility

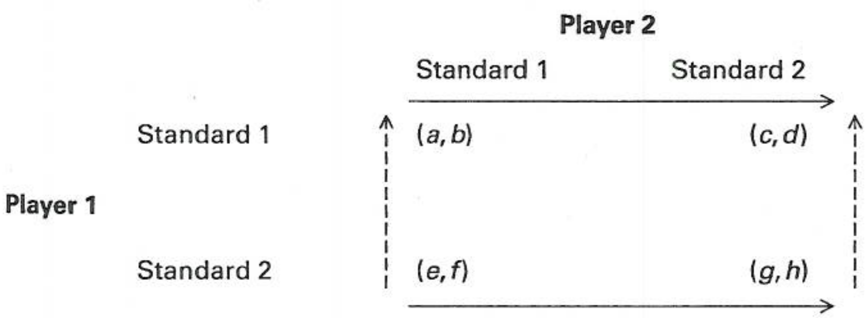

Figure 13.8
Standards war leading to incompatibility

signify the direction of individually incentives, assuming that the opponent will not change his/her strategy.

In the game of figure 13.7, standard 1 is a noncooperative equilibrium if $a > e$, $b > d$. Similarly standard 2 is an equilibrium if $g > c$, $h > f$. In this game assume that a firm has higher profits when "its" standard i get 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 cases of disagreement profits are lower than those of argument on either standard, $e, c < g$; $d, f < b$. Under disagreement, the setting of either standard will constitute a noncooperative equilibrium. There is no guarantee that the highest joint profit standard will be adopted. Since the consumers' surplus does not appear in the payoff matrix, there is no guarantee of maximization of social welfare at the noncooperative equilibrium.

In the same setup we could have each side preferring its own sponsored standard no matter what the opponent does; that is, each side has a dominant strategy and chooses its own technical standard. This game results in an in incompatibility equilibrium.

To understand the relative benefits of the compatibility decision, we need to examine the industry structures that would arise under either choice. When all firms are compatible, one expects equality to the extent that it is the rule in nonnetwork industries. However, in industries exhibiting strong network effects, in a regime of total incompatibility (where each firm has its own incompatible standard), we expect to observe extreme inequality in market shares and profits. This is commonly observed in the computer software and hardware industries and in most of the new markets created by the Internet. Sometimes such extreme inequality is commonly explained in industry circles by attribution to history. Stories abound on who or which company "was at the right place at the right time" and therefore now leads the pack. Traditional economic theory cannot easily explain such extreme inequality and may also resort to "managerial," "entrepreneurship," or "historical" explanations that are brought over in economics only when all else fails. As a last resort, if all else fails to explain a market phenomenon, economists tend to dismiss what they cannot explain as an "aberration" or a temporary phenomenon that will certainly disappear in the long-run equilibrium! Such explanations are deficient not only because they may be incorrect but also because they tend to treat situations as isolated events and therefore lose all potential predictive power that is derived from correct modeling of economic phenomena.

There is a simple explanation of market structure in network industries without resorting to managerial, entrepreneurship, or historical explanations. The explanation is based on two fundamental features that network industries have and other industries lack: the existence of network externalities and the crucial role of technical compatibility in making the network externalities function.

Firms can make a strategic choice on if they are going to be compatible with others, and sometimes on if they will allow others to be compatible with them. The ability of a firm to exclude other firms from sharing a technical standard depends on the property rights that a firm has. For example, a firm may have a copyright or a patent on the technical platform or design, and can therefore exclude others from using it.

Compatibility with competitors brings higher network externality benefits ("network effect") and therefore is desirable. At the same time compatibility makes product X a closer substitute to competing products ("competition effect"), and it is therefore undesirable. In making a choice on compatibility, a firm has to balance these opposing incentives. In a network industry the traditional decisions of output and price take special importance, since higher output can increase the network externalities benefits that a firm can reap.

Inequality in market shares and profitability is a natural consequence of incompatibility. Under incompatibility, network externalities act as a *quality feature* that differentiates the products. Firms want to differentiate their products because they want to avoid intense competition.

In making the choice between compatibility and incompatibility, firms take into account the intensity of the network externality. The more intense the network externality, the stronger is the incentive for a firm to break away and be incompatible from substitutes. It follows that in industries with very intense network externalities, firms are more likely to choose incompatibility. As we will see in detail below, incompatibility implies inequality.

Inequality is accentuated by output expansion and an increase of the network. Moreover a firm of higher output has a higher perceived quality and is therefore able to quote a higher price. Thus the inequality in profits is even more acute than the inequality of outputs.

Inequality of Market Shares and Profits Markets with strong network effects where firms can choose their own technical standards are "winner-take-most" markets. In these markets there is extreme market share and profits inequality. The market share of the largest firm can easily be a multiple of the market share of the second largest, the second largest firm's market share can be a multiple of the market share of the third, and so on. This geometric sequence of market shares implies that even for a small number of firms *n*, the *n*th firm's market share is tiny. In equilibrium there is extreme market share and profits inequality.[14]

The reason for the inequality is straightforward. A firm with a large market share has higher sales of complementary goods and therefore its good is more valuable to consumers. This feeds back resulting in even higher sales. Conversely, a firm with small market share has lower sales of complementary goods, and the feedback results in even lower sales. However, the low-sales firm is not necessarily driven out of business because that would require too low a price by the high-sales firm. In the absence of fixed costs, an infinite number of firms can survive, but there is tremendous inequality in market shares, prices, and profits among them. Good examples of this market structure are the PC operating systems market and many software applications markets.

To understand the extent of market share, price, and profits inequality in network industries, we provide results from Economides and Flyer (1998). As a benchmark they assume that all firms produce identical products, except for whatever quality is added to them by network externalities. Also they assume that no firm has any technical advantage in production over any other with respect to any particular platform and that there are no production costs. We consider here only the extreme case of "pure network goods" where there is no value to the good in the absence of network externalities. The summary of the equilibria under total incompatibility (which can be enforced when firms have proprietary standards) is in tables 13.1 and 13.2. The ith firm sells quantity q_i at price p_i, and firms are ordered in decreasing quantity so that $q_1 > q_2 > q_3$, meaning firm 1 has the largest sales, firm 2 is the second largest, and so on. The maximum potential sales is normalized to equal 1. At equilibrium not all consumers buy the good, that is, total sales are $\sum_{j=i}^{I} q_j < 1$.

The market equilibria exhibit extreme inequality. The ratio of outputs of consecutive firms is over 2.6. Ratios of prices of consecutive firms are higher than 7. The ratio of profits of consecutive firms is about 20. This means that a firm that has about 38 percent of the sales of the immediately larger firm, can charge only 15 percent of the price of the next larger firm, and receives only 5 percent of the profits of the immediately larger firm. Entry after the third firm has practically no influence on the output, prices, and profits of the top three firms as well as the consumers' and producers' surplus. From the fourth one on, firms are so small that their entry hardly influences the market.

Table 13.1
Quantities, market coverage, and prices under incompatibility

Number of firms I	Sales of largest firm q_1	Sales of second firm q_2	Sales of third firm q_3	Market coverage $\sum_{j=I}^{I} q_j$	Price of largest firm p_1	Price of second firm p_2	Price of third firm p_3	Price of smallest firm p_I
1	0.6666			0.6666	0.222222			2.222e-1
2	0.6357	0.2428		0.8785	0.172604	0.0294		2.948e-2
3	0.6340	0.2326	0.0888	0.9555	0.170007	0.0231	0.0035	3.508e-3
4	0.6339	0.2320	0.0851	0.9837	0.169881	0.0227	0.0030	4.533e-4
5	0.6339	0.2320	0.0849	0.9940	0.169873	0.0227	0.0030	7.086e-5
6	0.6339	0.2320	0.0849	0.9999	0.169873	0.0227	0.0030	9.88e-11
7	0.6339	0.2320	0.0849	0.9999	0.169873	0.0227	0.0030	≈ 0

Note: Even with no fixed costs and an *infinite* number of firms, the Herfindahl-Hirschman index is HHI = 0.464, which corresponds to between two and three firms of equal size.

Table 13.2
Profits, consumers' and total surplus under incompatibility

Number of firms I	Π_1	Π_2	Π_3	Profits of last firm Π_I	Total industry profits $\sum_{j=i}^{I} \Pi_j$	Consumers' surplus (CS)	Total surplus (TS)
1	0.1481			0.1481	0.1481	0.148197	0.29629651
2	0.1097	7.159e-3		7.159e-3	0.1168	0.173219	0.29001881
3	0.1077	5.377e-3	3.508e-4	3.508e-4	0.1135	0.175288	0.28878819
4	0.1077	5.285e-3	3.096e-4	1.474e-5	0.1132	0.175483	0.28868321
5	0.1077	5.281e-3	2.592e-4	8.44e-7	0.1132	0.175478	0.28867817
6	0.1077	5.281e-3	2.589e-4	1.18e-14	0.1132	0.175478	0.28867799
7	0.1077	5.281e-3	2.589e-4	0	0.1132	0.175478	0.28867799

Monopoly May Maximize Total Surplus In industries with significant network externalities, under conditions of incompatibility between competing platforms, monopoly may maximize social surplus. This is because, when strong network effects are present, a very large market share of one platform creates significant network benefits for this platform, which contribute to large consumers' and producers' surpluses. It is possible to have situations where a breakup of a monopoly into two competing firms of incompatible standards *reduces* rather than increases social surplus because network externalities benefits are reduced. This is because de facto standardization is valuable, even if done by a monopolist.

In the Economides-Flyer model, although consumers' surplus is increasing in the number of active firms, total surplus is decreasing in the number of firms. That is, the more firms in the market, the lower is total welfare. This remarkable result comes from the fact that when there are fewer firms in the market there is more coordination and the network effects are larger. As the number of firms decreases, the positive network effects increase more than the deadweight loss so that total surplus is maximized in a monopoly! Total surplus is highest while consumers' surplus is lowest in a monopoly. This poses an interesting dilemma for antitrust authorities. Should they intervene or not? In nonnetwork industries typically both consumers' and total surplus are lowest in a monopoly. In this network model, maximizing consumer's surplus would imply minimizing total surplus.

Compared to the market equilibrium under compatibility, the incompatibility equilibrium is deficient along many dimensions. Consumers' and total surplus are higher under compatibility, the profits of all except the highest production firm are higher under incompatibility, and prices are lower under compatibility except possibly in a duopoly.

No Anticompetitive Acts Are *Necessary* to Create Market Inequality Because inequality is natural in the market structure of network industries, there should be no presumption that anticompetitive actions are responsible for the creation of market share inequality or very high profitability of a top firm. Thus no anticompetitive acts are *necessary* to create this inequality. The "but for" benchmark against which anticompetitive actions in network industries are to be judged should not be "perfect competition" but an environment of significant inequality and profits.

In Network Industries, Free Entry Does Not Lead to Perfect Competition The existence of network effects imply that in network industries, free entry does not lead to perfect competition. In a market with strong network effects, once few firms are in operation, the addition of new competitors, even under conditions of free entry, does not change the market structure in any significant way. Although eliminating barriers to entry can encourage competition, the resulting competition may not significantly affect market structure. This implies that in markets with strong network effects, antitrust authorities may not be able to significantly affect market *structure* by eliminating barriers to entry. See the earlier example where the addition of the fifth firm hardly changes the output of the first four firms.

The remarkable property of the incompatibility equilibrium is the extreme inequality in market shares and profits that is sustained under conditions of free entry. Antitrust and competition law have placed a tremendous amount of hope on the ability of free entry to spur competition, reduce prices, and ultimately eliminate profits. In network industries, free entry brings into the industry an infinity of firms, but it fails miserably to reduce inequality in market shares, prices, and profits. Entry does not eliminate the profits of the high-production firms. And it is worth noting that at the equilibrium of this market there is no anticompetitive behavior. Firms do not reach their high output and market domination by exclusion, coercion, tying, erecting barriers to entry, or any other anticompetitive behavior. The extreme inequality is a natural feature of the market equilibrium.

At the long-run equilibrium of this model with free entry, an infinity of firms have entered, yet the equilibrium is far from competitive. No anticompetitive activity has led firms to this equilibrium. Traditional antitrust intervention cannot accomplish much because the conditions that such intervention seeks to establish already exist in this market.

Can there be an improvement over the market incompatibility equilibrium? Yes, a switch to the compatibility equilibrium has higher consumers' and total surpluses for any number of firms. Is it within the scope of competition law to impose such a change? It depends. Firms might have a legally protected intellectual property right that arises from their creation of the design of the platform. Only if anticompetitive behavior was involved, can the antitrust authorities clearly intervene.

Imposing a "Competitive" Market Structure Is Likely to Be Counterproductive An implication of network effects is that antitrust interventions may be futile. Because "winner takes most" is the natural equilibrium in these markets, attempting to superimpose a different market structure (e.g., one in which all firms have approximately equal market shares) may be both futile and counterproductive.

Nature of Competition Is Different in Network Industries Strong network effects imply that competition *for the market* takes precedence over competition *in the market*. The fact that the natural equilibrium market structure in network industries is winner take most with very significant market inequality does not imply that competition is weak. Competition on which firm will create the predominant (top) platform and reap most of the benefits is often intense. In network industries there is typically an intense race to be the dominant firm. In network industries we often observe Schumpeterian races for market dominance.

A good recent example of Schumpeterian competition is the competition among dot-coms in 1999 and 2000. As explained earlier, economic models imply a high valuation of the dominant firm compared to other firms in the same network industry. The same perception prevailed on Wall Street. During that period, dot-com firms advertised very intensely and subsidized consumers so as to be able to achieve the coveted dominant position in the market. The easy availability of capital for dot-coms at the time facilitated this behavior as firms "burned" almost all the cash they had in their attempts to get the top market share. Many of the dot-coms failed because demand for their services was much lower than predicted or because of flaws in their business models. However, all the successful dot-coms, such as eBay, Amazon, Yahoo, and later Google, also followed this strategy. Generally, in network industries, the costs of entry may be higher, but the rewards of success may be higher compared to nonnetwork industries.

Path Dependence The presence of network effects gives special importance to *path dependence*. Path dependence is the dependence of a system or network on past decisions of producers and consumers. For example, the price at which a VHS player can be sold today is path dependent because it depends on the number of VHS players sold earlier (the installed base of VHS players). The existence of an installed base of consumers favors an incumbent. However, competitors with significant product advantages or a better pricing strategy can overcome the advantage of an installed base.

For example, in the market for video players, VHS overcame Beta after six years of higher installed base by Beta. This was an implication of

1. Sony's mistake in disregarding network externalities and not licensing the Beta format.
2. JVC's widespread licensing of VHS.
3. The fact that one low-end, low-priced VHS player can contribute as much to the network effect as a high-end, high-priced Beta player.

In the Beta/VHS case, it is clear that Sony mistakenly ignored the network effects that arose from the availability of rental tapes of pre-recorded movies. The main function of video recorders was originally thought to be "time delay" in watching material recorded from the TV. The pre-recorded market emerged later, first as a market where movies were sold and later as a movies rental market. The emergence of markets for "movies for sale" and "movies for rent," which had to be recorded in a particular format, created a significant complementary good for Beta and VHS players. The significant cost of physical distribution of tapes throughout the country and the costs of carrying a significant inventory of titles made the choice of what movies to bring and in what format crucially dependent on present and forecasted demand. This forecast was highly correlated with the present and forecast installed base of video players in each format. Thus, although network effects and path dependence played a crucial role in determining the fate of Beta, the outcome was far from predetermined. Early, more aggressive licensing of the Beta format by Sony or the early promotion of low-end Beta players could have reversed the demise of the Beta format.

An often cited example on path dependence is the prevalence of the QWERTY keyboard despite claims of more efficient function by the alternative Dvorak keyboard. For many business applications, and for antitrust purposes, the QWERTY example is not crucial because there was no significant strategic business interest in the success of either design. There is also a factual dispute on whether the Dvorak keyboard was significantly more efficient than the QWERTY one.[15]

13.3 Competition Policy Issues in Network Industries

13.3.1 One-sided Bottlenecks

Interconnection issues in telecommunications, railroads, airline, and other transportation networks are very common. Often one company controls exclusively a part of the network, which is required by others to provide services. We call this network part "a bottleneck." Generally, bottlenecks can be divided into two categories: one-sided and two-sided. A one-sided bottleneck is monopolized by a firm, and this firm does not require the use of a different bottleneck. Such a bottleneck is shown as link AB in figure 13.9. An example of

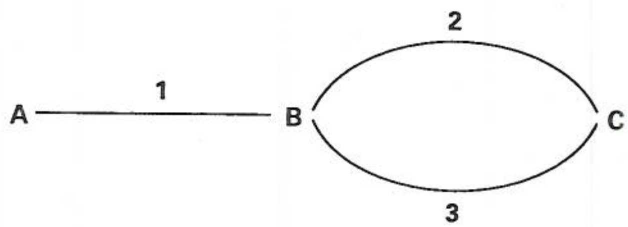

Figure 13.9
One-sided bottleneck

such a bottleneck is the connection of local service telecommunications subscribers to a switch. This is typically called "the last mile," and often called "the local loop." After the 1984 breakup of AT&T, the local loop has been monopolized by the local exchange carrier, typically a Regional Bell Operating Company (RBOC) or GTE. Excluding cellular phones, the local loop is a required input in the production of long-distance services, and typically long-distance companies do not have a comparable local loop. Similarly such a one-way bottleneck can arise when a firm monopolizes a railroad track such as *AB*. In telecommunications, the local exchange bottleneck has traditionally resulted in high prices for use of the bottleneck to originate ("access origination") or terminate calls ("access termination").

The potential anticompetitive consequences of a one-sided bottleneck are obvious, and have been understood since the early days of the telecommunications network when AT&T enjoyed a monopoly in long distance (here *AB*) but faced competition in local markets. In the context of figure 13.9, the early AT&T was in possession of links 1 (long distance) and 2 (local) but did not allow an independent firm that possessed link 3 to interconnect at *B* and provide part of the long-distance service *CBA*. For over two decades in the beginning of the twentieth century, AT&T refused to interconnect independent local telecommunications companies to its long-distance network unless they became part of the Bell System, which essentially meant unless they were acquired.[16]

The early AT&T foreclosure of independents through a refusal to interconnect shows the importance of complementarities in networks and the way that companies can leverage dominance in one market to create dominance in a market for complementary goods, especially when the complementary good requires the monopolized input to provide a final service. In this case AT&T monopolized long distance and was able to leverage its position in long distance (through the refusal to interconnect with independent locals) and gain a dominant position in local telecommunications markets throughout the country.

The continued foreclosure of the independents by AT&T and its "refusal to deal" with them caused regulation to be established at the state and federal levels in the 1930s. The 1934 Federal Communications Act ("1934 Act") imposed mandatory interconnection in an attempt to stop the foreclosure of independents and stabilized the market share of local lines held by AT&T. However, at that point AT&T's market share of local lines had increased from about 50 percent in 1914 to close to the 89 percent that AT&T had in 1981, prior to the 1982 agreement with the Department of Justice to be broken up.

A major revision of the 1934 Act, the Telecommunications Act of 1996 ("1996 Act"), mandates interconnection of all public switched telecommunications networks at any technically feasible point. The 1996 Act and similar EU regulations attempt to solve the problem of the monopolization of the key parts of local telecommunications network. They impose unbundling of the network and forced leasing to entrants of some of the monopolized parts of the network, including the local loop. The goal is to make "mix-and-match" entry strategies feasible for local voice telephone service as well as broadband Internet access through digital subscriber lines (DSL) that utilizes high-frequency transmission though copper local loops. Thus the 1996 Act and EU regulations mandate access prices for unbundled parts of the network (unbundled network elements, or UNEs) at cost-based prices. The FCC and state PUCs accepted the view that lease prices should be based on *forward-looking* costs of a network that could be built today rather than on historical, accounting, or embedded costs (as favored by RBOCs). In setting prices for unbundled network elements, the FCC and state public utility commissions (PUCs) also rejected the relevance of prices based on *private* opportunity cost, such as the efficient components pricing rule (ECPR). Such rules derive prices for components from the monopoly prices of end-to-end services. Thus the ECPR and its varieties would guarantee the monopolist's profits despite market structure changes in the markets for components that are used to create final services.[17] To prevent anticompetitive actions in telecommunications, the 1996 Act also imposes a number of rules, such as number portability, mandatory resale of services, transparency, and nondiscrimination. A full discussion of these rules can be found at Economides (1999). Still the 1996 Act missed opportunities to define technical standards and require technical compatibility of telecommunications equipment.

Unfortunately, legal maneuvers by the incumbent local exchange carriers and high prices for the unbundled network elements considerably delayed very significant entry in local telecommunications markets. Despite the delayed entry, Economides, Seim, and Viard (2007), analyzing local telecommunications entry in New York State, show that entry resulted in substantial benefits to consumers. Local telecommunications entry in residential markets through leasing of the incumbent's network by major carriers such as AT&T and MCI has practically stopped after the Court of Appeals in Washington, DC, struck down the FCC rules implementing leasing, and the new FCC rules did not enforce leasing "at cost plus reasonable profit" prices.[18]

13.3.2 Two-sided Bottlenecks

In a two-sided bottleneck, each of two firms is a monopolist, each with a different bottleneck, and each firm requires the other's bottleneck to produce its output. Take the example

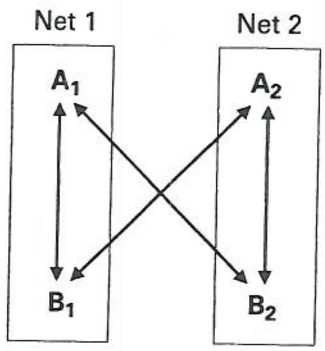

Figure 13.10
Two-sided bottleneck

of two local telephone companies. Customers subscribe only to one local telephone company but may require the other company's network to complete calls. This case is represented in figure 13.10. Calls originate at A_1 or A_2 and terminate at B_1 or B_2. In the context of this example, each of firms 1 and 2 buys *access termination* from the other. If each firm $i = 1, 2$, sells both services A_1B_2 and B_1A_2, then each firms buys both *access origination* and access termination from the other. Termination charges at B_1 or B_2 for calls from the opponent network can be used to disadvantage and foreclose the opponent network.

Many issues of traditional bottlenecks have been dealt by regulation in the United States and the European Union. In monopolized one-way bottlenecks, such as access origination and termination used in the creation of long-distance calls, access prices are regulated and there has been a tendency to decrease the regulated prices. However, prices are still high. In the two-way bottleneck of access used in the creation of local calls by competing local exchange carriers, the Telecommunications Act of 1996 imposes cost-based reciprocal fees and allows the possibility of "bill and keep," meaning zero prices. If cost-based reciprocal compensation were not the rule, and firms were able to set termination at profit-maximizing levels, Economides, Lopomo, and Woroch (1996, 2007) show that a large network will try to impose very high termination charges on an opponent's small network so that no calls terminate from the small network to the large one. Without the possibility of such across-networks calls, a small network will only be able to provide within-network calls and, being small, will be of little value to potential subscribers. As a consequence large networks would be able to provide more value to subscribers, and the small network would be foreclosed. Starting from a regime of a large local incumbent and a small potential entrant, the large incumbent could set up termination access fees to keep the entrant out of the market.[19]

In summary, in the absence of specific regulatory rules, two-sided bottlenecks can lead to foreclosure of competitors, even when each firm requires use of the bottleneck of the other to complete calls or provide a service.[20]

13.3.3 Market Power Creation Specific to Networks: The Importance of Technical Standards

The example of early AT&T's refusal to deal (interconnect) with independents (and with interconnected networks of independents) can also arise in milder terms when a firm X that has a significant position in its industry insists that firms that provide complementary products Y do not also provide them to X's competitors. For example, in the mid-1980s Nintendo refused to allow third-party games (software) to play on its game console (hardware) unless the software manufacturers agreed not to write a similar game for competing game systems for two years. Faced with this condition imposed by Nintendo, software developers had to make a choice to either write a game for Nintendo or for the competing platforms of Atari and Sega. Clearly, this restriction reduced the potential revenue of a game developer who would like, for a small additional cost, to port its game to the alternative systems. But also, more important, the restriction forced developers to predict which game system would have higher sales, and create software just for this system. Thus Nintendo used its dominance of the game market at that point in time to coerce developers to write software just for its platform, and thereby increased the value of the Nintendo virtual network (of hardware and software). Nintendo abandoned this requirement under antitrust challenge.

Because of the extreme inequality of market shares, prices, and profits in a network industry, restriction of the installed base of a firm in a network industry can be very detrimental. It can push a firm to a lower rank with significantly lower profits or, in extreme cases, push a firm to such a low market share that it has to close down because it cannot recover its fixed costs.

Another example from the computing industry illustrates a situation of market power creation specific to networks. Suppose that firm A chooses to make its product A incompatible with the products of other firms that perform similar functions, and it also subsidizes firms that produce complementary goods B to its product A.[21] Alternatively, we may assume that firm A subsidizes its own division that sells complementary goods B. The results are as follows:

1. The value of firm A's product increases.
2. The entry hurdle of firm A's rivals increases.
3. There is possible creation of market power.

Firm A's defense will be that its actions are pro-competitive, since their primary cause is the enhancement of the value of product A. From the point of view of A's competitors, the actions of A look very much like anticompetitive behavior since the abundance of complementary goods B for product A puts them at a competitive disadvantage.

Note that the existence of incompatibility is a *necessary* condition for possible creation of market power. Moreover the key to increasing social welfare is to move to compatibility. That is, assuming that innovation and product availability will not be reduced, the best of all worlds is to have public standards and full compatibility. However, it is very difficult for US antitrust authorities to intervene and/or define standards.

Besides the use of technical standards, firms can use bundling and other pricing strategies as well as non–price discrimination strategies to leverage market power across markets.

13.3.4 Vertical Integration and Vertical Control Issues in Network Industries

In networks, as in other settings, there are potentially anticompetitive issues arising from the possibility of vertical integration and the behavior of vertically integrated firms. These may include the bundling of components through vertical integration, contract, or manipulation of technical standards so that an entrant must enter both components markets even if it desires to enter only one of the markets. Often firms have expertise or a technical advantage in only one component and desire to enter only in the market for that component. An incumbent can strategically alter the market environment through acquisition or contract so that the entrant can only be successful if it enters more than one market. This increases the financial hurdle for an entrant, and it also forces it to sell components where it does not have expertise. Thus it makes it more likely that entry will not occur.

A vertically integrated firm can also use discrimination in price charged to a subsidiary compared to the price charged to a downstream competitor, or discrimination in quality provided to subsidiary compared to quality provided to a downstream competitor, that is, raising rivals' costs. These issues are discussed in more detail in Economides (1996b).

Firms in network industries can also use a variety of way to manipulate technical standards in joint ventures to achieve market power. The issue of market power then arises in "aftermarkets," where consumers are "locked in" to a service that arises out of commitments of a durable nature. For example, in an important case that reached the Supreme Court, Kodak refused to supply parts to independent firms that serviced Kodak photocopiers. Although one could argue that there was significant competition in the market for new photocopiers and Kodak did not have a dominant position in that market, once customers had bought a Kodak photocopier, they were "locked in," and faced significant costs to buy a new photocopier of a different brand. So Kodak's actions could be anticompetitive in the "aftermarket" for repair services of consumers who have already bought Kodak photocopiers. A similar case of anticompetitive actions can be made in aftermarkets where consumers are locked in by having made an investment in a durable good that is incompatible with other comparable durable goods or are locked in other ways. For example, consumers without number portability in wireless cellular and PCS markets may be "locked in" to the service of a particular provider or network. Similarly consumers can be "locked in" to the e-mail service of an Internet service provider (ISP), since there is no portability of e-mail addresses and many ISPs do not offer forwarding of incoming e-mail messages to another ISP. The problem becomes even more complex when the conditions that establish dominance in the aftermarket are set in the shrink-wrap contract that the consumer agrees to when he bys the original product. For example, what if car manufacturers imposed the condition that by buying the car, the consumer agrees to only buy replacement parts from this manufacturer? Or more generally in a network industry, what if a hardware or software platform seller imposed conditions on what software (or complementary goods) will it allow consumers to use with its platform. These are important issues that need to be examined.

13.3.5 B2B, B2C, and Other Exchanges Issues

The world of business to business (B2B) and business to consumers (B2C) exchanges lacks the regulation of traditional financial and commodity exchanges. Many proposed B2B exchanges are run by the firms that also are trading. For example, Enron was proud of the fact that it was participating as a trading party in B2B exchanges that it organized and ran. Such a situation would be strictly prohibited in traditional financial and commodity exchanges because of the possibility that the organizer of the exchange would take advantage of the information created in the trading process to fashion privately beneficial trades. In another example, Covisint, an exchange for automobile parts organized by automobile manufacturers, was accused of acting to consolidate the monopsony power of car manufacturers. In general, B2B exchanges can provide substantial benefits by consolidating trades, increasing market liquidity, improving standardization, and reducing search costs.[22] But B2B exchanges also have the potential of creating significant antitrust issues.

13.3.6 Dynamic Efficiency Issues

The world of networks and dynamic effects brings to the forefront the fact that behavior that exhibits static efficiency may lack dynamic inter-temporal efficiency. The possibility exists of a lock-in to a technology and a path that, when decisions are taken in every period, looks optimal given past decisions but is suboptimal if earlier investment decisions had been delayed and all the decisions were taken at once. In a world with network effects a "lock-in" to an inferior technology can easily occur as firms (and countries) find it more desirable to invest in the technology in which they have already invested. This can occur under perfect competition, but the problem can easily become much more important under oligopoly, as firms race to become dominant, given the importance of dominance in a network industry. With rapid technological change, firms in an oligopoly race for dominance can easily get stuck investing heavily in the currently best technology and unable to invest sufficiently in the next technology, thereby placing themselves off the optimal investment path.

13.3.7 Innovation Issues

An important antitrust issue is the speed of innovation in a network industry as affected by strategic decisions of firms and potentially anticompetitive actions. The effects of actions on innovation are important because innovation affects the welfare of future consumers, and this should be taken into consideration in an antitrust decisions. The difficulty in dealing with innovation issues in antitrust arises from the fact that the efficiency and intensity of innovation in monopoly compared to perfect competition and oligopoly are open questions in economics. Thus it is very hard to make general statements on innovation in an antitrust context.

13.3.8 Criteria to Be Used for Antitrust Intervention in Network Industries

When an antitrust intervention is considered in a network industry, a number of considerations that arise out of the nature of network industries have to be taken account. These are explained in detail in earlier sections, primarily in section 13.2. First, the benchmark of the "but for" world that should be considered is a network industries equilibrium with significant inequality, rather than a perfectly competitive equilibrium. Second, competitors' harm is not a sufficient reason for intervention. The right question is, Were consumers (past, present, future) harmed by specific actions? Third, uncertainty must be taken into account, and caution must be used in guessing how a high-technology industry would have evolved but for the anticompetitive action(s). Fourth, it is possible for monopoly to maximize total surplus. Fifth, it is not possible to sustain a long-term equilibrium with equal market shares, and a short-term equilibrium with equal market shares generally has low total surplus. Sixth, path dependence and the value of installed base are limited by Schumpeterian competition, and upheavals are not uncommon in network industries. Seventh, especially in software industries, the extent and functionality of products is flexible.[23] The definition of the market of a potential antitrust violation may be much harder.

13.3.9 Criteria to Be Used for Remedies

When a remedies phase is reached, a liability finding has already been made. The objective of remedies is to stop practices that are found to be illegal, prevent the recurrence of such practices, and restore any recurring threat posed by such practices.

Any intervention by antitrust authorities creates a disruption in the workings of markets. The objective of the remedial relief is to accomplish the objectives mentioned in the previous paragraph without damaging efficient production and competition in the market. The potential damage that antitrust intervention can produce is larger when it is applied to an industry with fast technological change, where leaps to new and more efficient technologies are expected while the specific nature of the future winning technology is unknown. Often it is difficult to predict future winning technologies, and therefore it is hard to fashion an antitrust remedy with an accurate prediction of its effect on industry structure and competition a few years down the road. Of course, this uncertainty is multiplied when the remedy creates a significant intervention in the industry. Therefore, lacking the knowledge of the effects of their actions, it is in the public interest that antitrust authorities and courts avoid extensive intervention in industries with fast technological change. It is best to intervene only to the extent that (1) intervention reverses the effects of actions for which liability was established and (2) the effects of the intervention are predictable.

In markets with network effects, as I have explained in detail above, the existence of network effects has crucial implications on market structure and the ability of antitrust authorities to affect it. In markets with strong network effects, even in the absence of anticompetitive acts, the existence of network effects in markets results in significant inequalities in market shares and profits. The resulting equilibrium market structure can be called a *natural oligopoly* where very few firms dominate the market. The structural features of natural oligopoly for a software market cannot be altered by antitrust intervention without significant losses for society. The very nature of markets with network effects implies that the ability of antitrust authorities to alter market structure in such industries is limited, as discussed above.

As an alternative to antitrust and competition law, sectoral economic regulation can and has been established in three exceptional case: (1) for markets where it is clear that competition cannot be achieved by market forces, (2) for markets where deviation from efficiency is deemed socially desirable, and (3) for markets where the social and private benefits are clearly different. In each of these cases it should be clear that a market without intervention will not result in the desired outcome. I will leave case 2 aside, since a discussion of it will lead us to a detailed discussion of specific industries. The requirements for case 3 are typically met in many network industries, since expansion of the network creates network effects that are typically not fully internalized by markets. However, it would be foolish to advocate regulation as the standard solution in network industries because of the existence of network effects. Often a much smaller intervention, such as subsidization of the network to help network effects, will be enough.

In case 1, where it is clear that competition and its benefits cannot be achieved by market forces, regulation may be a solution. The significant advantage of industry-specific regulation is that it can be tailored to the specifics of the industry, and specific rules on pricing and availability of particular products and services. Regulators, such as the FCC, also have staffs that can provide impartial technical advice that is unavailable to a court.

However, regulation has a number of drawbacks. First, it is best suited for industries with well-defined and slow-changing products and services. With stable product definitions, rules can be devised and specific pricing can be implemented if necessary. Second, as a corollary to the first observation, regulation is not well suited in industries with rapid technological change and frequently changing product definitions. Moreover, in an industry with fast technical progress, regulation can be used by the regulated companies to keep prices relatively high, as exemplified by telecommunications regulation. Third, often regulators are very close to the interests of the regulated parties rather than to the interests of the public. Fourth, experience has shown that often regulators are not well informed about key variables as well as changes in the industry. Fifth, regulators at both the state and federal levels are under pressure and influence by both the executive and the legislative part of government, so they cannot be as impartial as a court. Sixth, there is a tendency for regulators to expand their reach into related and new markets. For example, the California Public Service Commission has recently asserted its authority over the Internet.[24] These drawbacks can create significant surplus loss due to regulation.

In summary, regulation should be used sparingly in industries with stable products if it is clear that antitrust has failed. It should be kept in mind that regulation can cause a significant surplus loss as well.

13.3.10 Recent Antitrust Cases in Network Industries

The Microsoft Antitrust Cases Over the last few years the Federal Trade Commission and the Department of Justice of the United States have investigated Microsoft on various antitrust allegations. The 1991 to 1993 and 1993 to 1994 investigations by the Federal Trade Commission (FTC) ended with no lawsuits. The 1994 investigation[25] by the United States Department of Justice (DOJ) was terminated with a consent decree in 1995.[26] The key provisions of the 1995 consent decree were as follows:

1. Microsoft agreed to end "per-processor" contracts with computer manufacturers (original equipment manufacturers, or OEMs), but it was allowed to use unrestricted quantity discounts.
2. "Microsoft shall not enter into any License Agreement in which the terms of that agreement are expressly or impliedly conditioned upon the licensing of any other Covered Product, Operating System Software product or other product (provided, however, that this provision in and of itself shall not be construed to prohibit Microsoft from developing integrated products); or the OEM not licensing, purchasing, using or distributing any non-Microsoft product."[27]

Thus the 1995 consent decree imposed two restrictions, one horizontal, and one vertical. The horizontal restriction stops Microsoft from charging an effectively zero price for the last 5 or 10 percent of the consumers sold to an OEM. By setting a lump-sum fee (based on the total number of computers manufactured by an OEM), Microsoft effectively set the price of the last 10 percent or more to zero. For example, suppose that an OEM produces 1 million units and Microsoft sets the per unit price of Windows at $40. In a per-processor contract Microsoft can offer to the OEM the option to install Windows in all 1 million units at a price of $30 million. This is equivalent to 750,000 units sold at $40 and 250,000 units sold at zero price. Thus this particular per-processor contract essentially sets the price of the last 25 percent of the production of the OEM at zero. If the OEM was installing Windows to approximately 75 percent of its production, the per-processor contract gives strong incentives to expand this percentage, since competing operating systems are sold at a positive price. The per-processor offer was an aggressive pricing contract by Microsoft that helped it gain and retain market share in competition with DR DOS and IBM's

OS-2. The 1995 consent decree stopped the zero marginal price offer but allowed quantity discounts to OEMs. Since this was not quantified and since software has practically zero marginal cost, there is no clear guidance from this case on the extent of quantity discount that would be considered illegal.

The vertical restriction of the 1995 consent decree prohibits product bundling created by contract but allows Microsoft to keep expanding the number and type of functions of its products, including Windows. In short, in the 1995 consent decree contractual bundling was disallowed, but technological bundling was explicitly allowed. As we will see, this issue is of crucial importance in the next Microsoft case. The 1995 consent decree was a crucial win for Microsoft because it allowed it to expand the functionality of Windows.

Over time a number of new functions were incorporated in Windows, including memory management, disk compression, disk defragmentation, Web browsing, and many others. On October 20, 1997, DOJ alleged that Microsoft violated the 1995 consent decree by bundling Internet Explorer (IE) with the Windows operating systems, and requiring computer manufacturers to distribute IE with Windows 95. DOJ petitioned the district court to find Microsoft in civil contempt. On December 11, 1997, Judge Thomas Penfield Jackson issued a preliminary injunction barring the bundling of IE with Windows.[28] On May 12, 1998, the court of appeals for the DC Circuit voided the 1997 preliminary injunction. On June 23, 1998, the court of appeals ruled that the 1995 consent decree did not apply to Windows 98, which was shipped with an integrated IE as part of the operating system and an IE icon on the PC desktop, arguing that "courts are ill equipped to evaluate the benefits of high-tech product design."[29]

During the week following the court of appeals stay of Judge Jackson's preliminary injunction that barred the bundling of IE with Windows because of the alleged violation of the 1995 consent decree, DOJ filed a major antitrust suit against Microsoft. In this action (DOJ Complaint 98-12320), filed on May 18, 1998, DOJ was joined by the attorneys general of 20 states and the District of Columbia. The court of appeals in its June 23, 1998, decision affirmed that Microsoft's practice of bundling IE with Windows was legal under the terms of the 1995 consent decree. To overcome this interpretation of the law as far as the integration of the browser is concerned, DOJ argued that Microsoft's bundling of IE with Windows and its attempt to eliminate Netscape as a competitor in the browser market was much more than adding functionality to Windows and marginalizing a series of add-on software manufacturers. DOJ alleged (and the district court concurred) that Microsoft added browser functionality to Windows and marginalized Netscape because Netscape posed a potential competitive threat to the Windows operating system. This distinctive threat posed by Netscape was a crucial part of the DOJ allegations. DOJ alleged that applications could be written to be executed "on top" of Netscape; since Netscape could be run on a number of operating systems, DOJ alleged that Netscape could erode the market power of Windows. In DOJ's logic, Microsoft gave away IE and integrated it in Windows so that Netscape would not become a platform that would compete with Windows. Thus DOJ alleged that Microsoft's free distribution of IE, its bundling with Windows, and all its attempts to win the browser wars were *defensive* moves by Microsoft to protect its Windows monopoly, which DOJ and Microsoft agreed was originally created legally.[30]

Formally, the allegations were as follows:

1. Microsoft illegally monopolized the market for operating systems (OSs) for personal computers (PCs) under Section 2 of the Sherman Antitrust Act.
2. Microsoft had anticompetitive contractual arrangements with various vendors of related goods, such as with computer manufacturers (OEMs) and Internet service providers (ISPs), and had taken other actions to preserve and enhance its monopoly; these contractual arrangements and other actions were illegal under Section 2 of the Sherman Antitrust Act.
3. Microsoft illegally attempted to monopolize the market for Internet browsers (but failed to succeed), an act that is illegal under Section 2 of the Sherman Antitrust Act.
4. Microsoft bundled anticompetitively its Internet browser, IE, the Microsoft Internet browser, with its Windows operating systems; this is illegal under Section 1 of the Sherman Antitrust Act.

The Microsoft trial took place at an accelerated schedule at the US District Court for the District of Columbia from October 19, 1998, to June 24, 1999.[31] Microsoft's CEO Bill Gates was not called as a witness, but his video-taped deposition was extensively used during the trial. Judge Jackson pre-announced that he would announce his "findings of fact" *before* his "conclusions of law." This was widely interpreted as implying that the judge was trying to give an opportunity to the sides to reach a compromise and resolve the case through a consent decree.

On November 5, 1999, Judge Jackson issued his "findings of fact," siding very strongly with the plaintiffs. In December 1999, Richard Posner, a prominent antitrust scholar and the Chief Judge of the Court of Appeals for the Seventh Circuit, agreed to serve as mediator for settlement discussions.[32] On April 1, 2000, settlement talks broke down after some states reportedly disagreed with the proposed agreement.[33] On April 3, 2000, Judge Jackson issued his "conclusions of law" finding for the plaintiffs on almost all points. In particular, Judge Jackson found Microsoft liable for monopolization and anticompetitive tying of IE with Windows but found that Microsoft's exclusive contracts did not make it liable for preventing Netscape from being distributed. On June 7, 2000, after an extremely short hearing,[34] Judge Jackson issued his remedies decision, splitting Microsoft into two companies, one for operating systems and one for applications and everything else, and imposing severe business conduct restrictions.

There a number of fundamental mistakes with Judge Jackson's decision on Microsoft. First, the judge did not consider the natural equilibrium in network industries. Instead, implicitly it was assumed that in the absence of anticompetitive actions, the natural equilibrium would be perfect competition. However, there is ample theoretical and empirical evidence discussed earlier in this chapter that markets with strong network effects, such as the OS market, are "winner-take-most" markets with significant market share and profits inequality as well as high concentration. Thus from the high profits and high concentration in these markets there does not necessarily follow the presumption of anticompetitive behavior. That is, the market share inequality and high profits for top firms are not necessarily the effects of anticompetitive actions but rather part of the natural equilibrium in these industries. Second, imposing a breakup will create an egalitarian market structure, since the natural equilibrium is "winner take most."

Third, there was the issue of exercising monopoly power in the OS market. In antitrust it is generally understood that a firm has monopoly power when it has the sustained ability to control price or exclude competitors. The existence of significant barriers to entry and the very high market share of Microsoft in the operating systems market gave indications that Microsoft had monopoly power. But there was also a very strong indication to the contrary. Microsoft priced its operating system to original equipment manufacturers (OEMs) at an average price of $40 to $60, a ridiculously low price compared to the static monopoly price.[35] Microsoft's economic witness showed that the static monopoly price was about $1,800, a large multiple of Microsoft's actual price.[36] At first glance, it seems that Microsoft could not possibly have monopoly power in OSs when its OS price is about 3 percent of the monopoly price.

Why was the price of Windows low compared to the monopoly price? I have argued elsewhere that all other explanations fail except that the price was low because of potential competition.[37] Briefly, Microsoft might have set the price of Windows low to hook consumers. But the low prices prevailed even when Microsoft had 90 percent market share. When would Microsoft increase the price? An alternative explanation might be that price was low because of competition from the installed base of Windows. However, (1) it is very difficult to uninstall Windows, (2) consumers buy much better new PCs faster than traditional obsolescence rates imply, and (3) Windows' price is small compared to the PC plus Windows bundle. Thus this explanation is also unlikely. Yet another explanation might be to reduce pirating. However, the price of other Microsoft products, notably MS-Office, was high. This is not consistent with the piracy explanation. Finally, another proposed explanation of the low Windows price is because it allows for higher prices of complementary goods. But since MS does not monopolize *all* the complementary goods markets, it would be optimal to charge the monopoly price on Windows. The only remaining explanation is the low price was due to the existence of actual and potential competition. Since actual competition was very limited, the remaining explanation is that potential competition constrained the price of Windows.

The Appeals Court in Washington, DC, ruled on Microsoft's appeal on June 28, 2001. The Court vacated Microsoft's breakup and other remedies imposed by the District Court. Microsoft was found liable of monopolization of the operating systems market for PCs. Microsoft was found not liable of bundling. Microsoft was found not liable of attempting to monopolize the Internet browser market. Judge Thomas Penfield Jackson of the District Court was taken off the case for improper behavior.[38] The case was remanded to the District Court for remedies determination for the monopolization charge. The Appeals Court instructed the District Court to examine the bundling of IE and Windows (if plaintiffs bring it up) under "a rule of reason" where the consumer benefits of bundling are balanced against the damage of anticompetitive actions. In face of the Appeals Court decision, DOJ decided not to pursue the bundling issue and announced that it would not seek the breakup of Microsoft.

On November 6, 2001, the United States—New York, Illinois, North Carolina, Kentucky, Michigan, Louisiana, Wisconsin, Maryland, and Ohio—and Microsoft proposed a settlement in the major antitrust case. California, Connecticut, Iowa, Massachusetts, Minnesota, West Virginia, Florida, Kansas, Utah, and the District of Columbia pursued the suit further to a full remedies trial (started March 11, 2002) in front of US District Judge Colleen Kollar-Kotelly. These states proposed making the source code of Windows and IE public, "freezing Windows" so that additional functionality would be sold as an additional good, making all APIs public, and other severe remedies.

On November 12, 2002, Judge Colleen Kollar-Kotelly imposed the final judgment that had only small differences from the original proposed settlement. The settlement terms are as follows:

- Provisions seen as favorable to Microsoft

1. No breakup
2. Microsoft can expand functions of Windows
3. No general restrictions on bundling
4. No wide disclosure of source code

- Provisions seen as favorable to the plaintiffs

1. Broad scope of definition of middleware products (browser, e-mail clients, media players, instant messaging software, etc.)
2. Requirement to partially disclose middleware interfaces. Microsoft is required to provide software developers with the interfaces used by Microsoft's middleware to interoperate with the operating system.
3. Requirement to partially disclose server protocols. The settlement imposes interoperability between Windows and non-Microsoft servers of the same level as between Windows and Microsoft servers.
4. Freedom to install middleware software. Computer manufacturers and consumers will be free to substitute competing middleware software on Microsoft's operating system.
5. Ban on retaliation. Microsoft will be prohibited from retaliating against computer manufacturers or software developers for supporting or developing certain competing software.
6. Uniform pricing of Windows for same volume sale. Microsoft will be required to license its operating system to key computer manufacturers on uniform terms for five years. Microsoft will be allowed to provide quantity discounts.

7. Ban on exclusive agreements; contract restrictions. Microsoft is prohibited from entering into agreements requiring the exclusive support or development of certain Microsoft software.
8. Compliance and enforcement. A panel of three independent, on-site, full-time computer experts will help enforce the settlement.[39]

In my opinion, this was a fair settlement. It seems that DOJ got a bit more than what it would have gotten in a full remedies trial. It is unlikely that the dissenting states will, in the end, be able to get anything substantially different from this settlement.[40]

The EU Competition Commission under Mario Monti started an important proceeding against Microsoft in August 2000. It alleged that Microsoft abused its dominant position in operating systems for PCs by leveraging this power (1) in server software and (2) in media players. In the server market, the European Union alleged that Microsoft was using its market power in PC clients to disadvantage the non-Microsoft server (in particular, Sun servers) by not disclosing sufficient information that would allow full interoperability between Microsoft clients and servers and non-Microsoft servers. In the media player market, the European Union alleged that Microsoft was using its dominant position in PC operating systems and bundling the Windows Media Player (WMP) with Windows to disadvantage RealAudio and other media players. The statement of objections was finalized in August 2003. The European Commission found Microsoft liable on both counts in March 2004 and imposed a $497.2 million fine. Microsoft was also required (1) to fully disclose the interface that would allow non-Microsoft servers to achieve full interoperability with Windows PCs and servers and (2) to offer to PC manufacturers a version of its Windows without WMP.[41] Microsoft appealed but did not win a stay. The EU decision was affirmed by the Court of First Instance in September 2007.

I have very serious concerns about the usefulness of the requirement imposed by the European Union to create a version of Windows without WMP. Competing media players such as RealAudio and Quicken are easily downloadable and available for free. Thus the cost to consumers of any of these players is just the five minutes or so it takes to download them. Even if the European Union is perfectly correct on liability, the remedy it imposed is way out of proportion and, in the end, may reduce rather than increase consumers' welfare. Since the version of Windows without WMP was sold at the same price as the one with the media player—the European Union had not imposed any restrictions on pricing—practically no computer manufacturer bought Windows without WMP. On the issue of interoperability, it is hard to offer an opinion without full knowledge of facts that are not publicly available. However, it is clear that even with full disclosure of the interface there may be advantages to components (e.g., clients and servers) produced by the same company.

Local Telecommunications, Trinko, and the Sacrifice Principle The *Trinko* case is the most important recent Section 2 case in the intersection of antitrust and regulation. The Law

Offices of Curtis V. Trinko bought local telecommunications services from AT&T. AT&T was providing these services by combining leased parts of the Verizon local telecommunications network (unbundled network elements, or UNEs) and adding retail services of its own, such as billing and marketing. Under the rules of the Telecommunications Act of 1996, incumbent local exchange carriers, including Verizon, were obligated to lease parts of their local telecommunications network to any firm at "cost plus a reasonable profit" prices that could combine them at will, add retailing services, and sell local telecommunication service as a rival to the incumbent.

The 1996 Act was a brave but failed attempt to introduce competition in all telecommunications markets.[42] Congress understood that it was uneconomic for firms to enter in local telecommunications by replicating the infrastructure of the incumbents.[43] Thus it set up two additional possibilities for entrants (besides entering with their own facilities): (1) to enter by leasing parts of the incumbents' local telecommunications network (leasing unbundled network elements (UNEs)[44] and (2) to enter by buying in wholesale the incumbents services and reselling them. The most important avenue to entry was leasing UNEs, combining them with the entrant's retailing services (e.g., marketing and billing) and selling to final consumers. To facilitate entry, the 1996 Act set the price for UNEs at "cost plus reasonable profit."[45] The 1996 Act additionally mandated that unbundled network elements be sold at "rates, terms, and conditions that are just, reasonable, and nondiscriminatory."[46] To facilitate entry, the 1996 Act also imposed the requirement on an incumbent to allow for physical collocation of equipment at its premises,[47] and on all companies the duty to provide number portability, so that consumers can keep their phone numbers if they change local service provider.[48]

In *Verizon Communications Inc. v. Law Offices of Curtis V. Trinko*,[49] Trinko sued Verizon for raising the costs of its retail rival AT&T (which had entered the market as a competitive local exchange carrier, or CLEC) and otherwise disadvantaging AT&T through anticompetitive conduct (including discrimination in fulfilling customer transfer orders to entrants) under Section 2 of the Sherman Act.[50] The Supreme Court held that Trinko's complaint failed to state a claim under Section 2 of the Sherman Act, and dismissed the complaint.

In arguing for dismissal, the Supreme Court noted that the markets for leasing parts of the local telecommunications network were created by the 1996 Act and did not exist voluntarily earlier. The Court somehow believed that Verizon's refusal to deal and its related raising rivals' costs practices were justified because infrastructure leasing prices were based on cost plus a reasonable profit: "Verizon's reluctance to interconnect at the cost-based rate of compensation available under Section 251(c)(3) tells us nothing about dreams of monopoly."[51] But Verizon was a monopolist in the network infrastructure and in the network services markets.[52] In my opinion, reluctance to sell leases at above average cost prices is a clear indication that the monopolist in the networks infrastructure market is attempting through this action to prevent entry in the network services market, which requires access to the networks infrastructure market. In my opinion, the fact that Verizon was obligated to lease local telecommunications infrastructure at cost plus a reasonable profit and did not write such leases at any price earlier does not imply that Verizon's refusal to deal and raising rivals' costs strategies does not create antitrust liability.

Additionally the Court missed the vertical leveraging issue in *Trinko*. Verizon provides two local telecommunications services: (1) network infrastructure services (NET services), which it provides to itself and to competitors in local telecommunications, and (2) retail services. End-users consume a composite service comprised of NET services and retail services. Competitors to Verizon in retail local telecommunications buy only NET services, adding their own retailing services for sale to end-users. At the point of the initial implementation of the 1996 Act, Verizon had a monopoly in both NET services and retail services. The conduct of Verizon in *Trinko* can be seen as the result of Verizon leveraging its monopoly in NET services to preserve its monopoly in retail services. This was recognized by the Second Circuit, which noted that Trinko could have "a monopoly leveraging claim" based on the fact that "the defendant '(1) possessed monopoly power in one market; (2) used that power to gain a competitive advantage ... in another distinct market; and (3) caused injury by such anticompetitive conduct.'"[53] In contrast, the Supreme Court dismissed the vertical issue, using a fallacious circular argument in footnote four of its decision, stating, "*In any event, leveraging presupposes anticompetitive conduct, which in this case could only be the refusal-to-deal claim we have rejected.*"[54] In my opinion, the key anticompetitive conduct was the leveraging from NET services to retailing services and the Court missed that.[55]

It is hard to offer a definitive opinion on the implications of *Trinko*. Some believe that its importance is confined to regulated industries. Others believe that it has significantly weakened enforcement of Sherman Section 2. This opinion is based their view on *Trinko* Courts statement: "*Aspen Skiing* is at or near the outer boundary of Section 2 liability."[56] The facts in *Aspen Skiing* were as follows: Aspen Skiing Co. controlled three out of four ski slopes in Aspen, Colorado, with the fourth slope controlled by Aspen Highlands. Aspen Skiing and Aspen Highlands offered a joint ticket that allowed the buyer to ski on all four slopes with revenue shared according to use. Aspen Skiing discontinued the joint ticket in 1978–79 and refused to sell its tickets to Aspen Highlands even at full retail price, to prevent Aspen Highlands from bundling them with its own tickets and recreating the joint ticket that had formerly been available. The Supreme Court ruled that Aspen Skiing's refusal to deal was anticompetitive.[57]

Despite the Court's statement in *Trinko*, one would expect *Trinko* to fall within the outer boundary set by *Aspen Skiing*. Because Verizon's price was set by regulation at cost plus a reasonable profit, it is reasonable to infer that Verizon's price to cost margin was lower than in the duopoly of *Aspen Skiing*. Thus, all else being equal, one would expect Verizon more likely to refuse to sell than Aspen Skiing Co. From the point of view of the firm committing the anticompetitive act, the incentive seems stronger for Verizon than for

Aspen Skiing. Therefore, if the Supreme Court deems the refusal to deal by a duopolist in *Aspen* anticompetitive, it should find the refusal to deal by the monopolist in *Trinko* even more so.

Being forced by regulation to sell below the monopoly price, and unable to discriminate in price by regulatory restraints, the monopolist in Trinko has an incentive to raise the costs of its rivals. If regulation were not present, price discrimination and monopoly pricing would have likely made raising rivals' costs strategies suboptimal form the monopolist's point of view and would not have been used. In the regulatory environment of the 1996 Act, raising rivals' costs is a natural response of a monopolist to the restraints of regulation.[58] Raising rivals' costs strategies reduce competition and social welfare associated from the existence of a free market.[59] In applying the Aspen Skiing standard, the Court erred in not considering the significant difference in incentives of the potentially liable party between the unregulated environment in *Aspen Skiing* and the regulated environment of *Trinko*.

In *Trinko* the Supreme Court did not state a rule under which specific conduct will be found to be "willful monopolization." In its brief in *Trinko* the government proposed such a standard based on the "sacrifice principle."[60] I define the sacrifice principle as follows: *a defendant is liable of anticompetitive behavior if its conduct "involves a sacrifice of short-term profits or goodwill that makes sense only insofar as it helps the defendant maintain or obtain monopoly power."*[61] This definition only partially coincides with the definition of the same principle in the government's brief in *Trinko*. Specifically, the government's brief allows *all* behavior that does *not* involve sacrifice of short-term profits to be characterized as not "exclusionary" and not "predatory."[62] I disagree. *Conduct can be exclusionary even without a sacrifice of short-term profits.* But when such a sacrifice is observed, it shows directly that this conduct is anticompetitive.

In my opinion, the vertical leveraging in *Trinko* passes the "sacrifice test." The behavior of Verizon to raise the costs of rivals in local telecommunications services entailed a sacrifice of profits from potential leasing of UNEs. This sacrifice would not have occurred if Verizon were not trying to protect its monopoly in the retail market for local telecommunications services. Thus, in applying the sacrifice principle, Verizon's actions are found to be anticompetitive. If Verizon did not have a retailing division, it would have no incentive to foreclose or disadvantage independent retailing firms. Instead, if its strategy were not to preserve its monopoly position in retailing, Verizon would have had every incentive to sell its NET services to all at prices above cost, as mandated by the 1996 Act. Since Verizon sells its NET services to its retailing division at cost while any NET services price sold to third parties includes a reasonable profit (according to the 1996 Act's rules), raising rivals' costs actions that disadvantage third party retailing service firms and result in smaller sales of NET services to these firms clearly imposes a sacrifice of profits for Verizon. Therefore the "sacrifice" principle can be applied in the *Trinko* case in the same way that the Supreme Court articulated it in *Aspen Skiing* to conclude that Verizon's raising rivals costs actions resulted in a short-term sacrifice of profits and therefore would not have been taken except to preserve its monopoly.

Market Power in Broadband Internet Access and Net Neutrality The Internet is 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. Internet communication is based on common and public protocols. Hundreds of millions of computers are presently connected to the Internet. The vast majority of computers owned by individuals or businesses connect to the Internet through commercial Internet service providers (ISPs).[63] Users connect to the Internet either by dialing their ISP, connecting through cable modems, residential DSL, or through corporate networks. Typically routers and switches owned by the ISP send the caller's packets to a local point of presence (POP) of the Internet. Dial-up, cable modem, and DSL access POPs as well as corporate networks dedicated access circuits connect to high-speed hubs. High-speed circuits, leased from or owned by telephone companies, connect the high-speed hubs forming an Internet backbone network (IBN).

The Internet is based on three basic separate levels of functions of the network:

1. The hardware/electronics level of the physical network
2. The (logical) network level where basic communication and interoperability are established
3. The applications/services level

Thus the Internet separates the network interoperability level from the applications/services level. Unlike earlier centralized digital electronic communications networks, such as CompuServe, AT&T Mail, Prodigy, and early AOL, the Internet allows a large variety of applications and services to be run "at the edge" of the network and not centrally.

An example of complex pricing discussed in section 13.2.3 is the present attempt of AT&T, Verizon, Deutsche Telecom, and other broadband Internet access providers to implement a complex price discrimination scheme on the Internet. A number of different services are provided on the Internet, including e-mail, browsing (using Internet Explorer, Firefox, Opera, etc.), peer-to-peer services, Internet telephony (voice over Internet protocol, or VOIP), among many others. A number of different functions/applications run on top of the Internet browser, including information services (Google, Yahoo, MSN), display of images, and transmission of video. Since the advent of Mosaic in 1993 and Netscape in 1994, the text-based Internet was enhanced to allow for images and video to be transmitted on it in digital form.

On the Internet, users pay ISPs for access to the whole Internet. Similarly ISPs pay backbones for access to the whole Internet.[64] ISPs pay per month for a pipe of a certain bandwidth, presumably according to their expected use.[65] When digital content, for example, is downloaded by consumer *A* from provider *B*, both sides; that is, both *A* and *B* pay.

A pays to his ISP through his monthly subscription, and *B* pays similarly. In turn ISPs pay to their respective backbones through their monthly subscription. The present regime on the Internet does not distinguish in terms of price (or in any other way) among bits or information packets depending on the services that these bits and packets are used for. This regime, called "net neutrality," has prevailed on the Internet since its inception. Presently a bit or information packet used for VOIP, for e-mail, for an image, or for a video is priced equally as a part of the large number of packets that correspond to the subscription services of the originating and terminating ISP.

After the acquisition of AT&T by SBC[66] and of MCI by Verizon, taking advantage of a change in regulatory rules by the Federal Communications Commission, AT&T and Verizon now advocate price discrimination based on which application and on which provider the bits they transport came from. AT&T and Verizon would like to abolish the regime of net neutrality and substitute for it a complex pricing schedule where, besides the basic service for transmission of bits, there will be additional charges by the Internet access operator applied to the originating party (e.g., Google, Yahoo, or MSN) even when the application provider is not directly connected to AT&T or Verizon, that is, even when Google's Internet service provider (ISP) is not AT&T or Verizon.[67]

The proposal is to impose price discrimination on the provider side of the market and not on the subscriber; that is, it is a version of two-sided pricing. This is uniquely possible to firms operating within a network structure. Besides traditional networks, such two-sided pricing is also possible by intermediaries in exchange networks (e.g., the exchanges themselves). There is presently considerable debate on the legality as well as the efficiency properties of the implementation of such complex rules by broadband Internet access firms mainly because of the very considerable market power of such firms.

Residential retail customers may well have difficulty changing ISPs. In the United States at least 95 percent of households have only one or two broadband Internet access choices, a digital subscriber line (DSL), or through cable TV, and their resellers, and many households are facing a monopoly of either cable or DSL. There are also switching costs to residential customers when changing equipment. Finally, residential customers are much more affected by contracts that bundle broadband Internet access with other services such as telecommunications and cable television.

As discussed earlier, the Internet under net neutrality separated the network layer from the applications/services layer. Net neutrality has allowed firms to innovate "at the edge of the network" without seeking approval from network operator(s). The decentralization of the Internet based on net neutrality facilitated innovation resulting in big successes such as Google, MSN, Yahoo, and Skype. Net neutrality also increased competition among the applications and services "at the edge of the network," which did not need to own a network to compete. Additionally the existence of network effects on the Internet implies that efficient prices to users on both sides (consumers and applications) should be lower than in a market without network effects. We see instead an attempt to increase prices, which will reduce network effects and innovation.

Abolition of net neutrality raises both horizontal and vertical antitrust issues. Starting with horizontal issues, last mile carriers (who are in duopoly or monopoly) may reduce capacity of "plain" broadband Internet access service and/or degrade it so that they can establish a "premium" service for which they will charge additionally content/applications provider. Coordinated reduction of capacity in "plain" service is reminiscent of cartel behavior. In general, the coordinated introduction of price discrimination schemes may reduce output. There is a general theorem in economics that price discrimination that reduces output reduces total surplus.[68] Therefore introduction of coordinated price discrimination may have anticompetitive consequences.

There is a also a variety of potentially anticompetitive vertical effects. For example, a carrier may favor its own content or application over that of an independent. VOIP provided over broadband Internet competes with traditional circuit-switched service provided by AT&T and Verizon and can be subject to discrimination. Additionally both AT&T and Verizon are gearing to distribute video, so they may favor their video service over that of others. But the anticompetitive concerns are hardly limited to products and services currently provided by the firms with market power in the access market. The carriers can also leverage market power in broadband access to the content or applications markets through contractual relationships. For example, a carrier can contract with a Internet search engine to put it in "premium" service while searches using other search engines have considerable delays using "plain" service.

The question posed before the US Congress is whether it should intervene immediately by imposing nondiscrimination restrictions or wait instead for antitrust suits. In my opinion, it is better to impose the nondiscrimination restrictions by law for a number of good reasons:

1. Suits take time, and much damage can be done before they are resolved. The legal system is slow, and lawsuits do not get resolved in "Internet time."
2. There are a variety of antitrust concerns, whereas typically each suit will deal with one issue. Thus delays can be compounded.
3. The Internet is a key essential network for growth of the US economy. The United States is already lagging behind fourteen less developed and lower income countries in Internet market penetration, as seen in figure 13.11.
4. Increasing prices through two-sided pricing will not increase network traffic nor grow the network!
5. Even if in the end there are no antitrust violations connected with the abolition of net neutrality (which I think is very unlikely), the abolition of net neutrality is likely to have significant negative consequences for innovation on the Internet, and therefore it is in the public interest to prevent it by law.

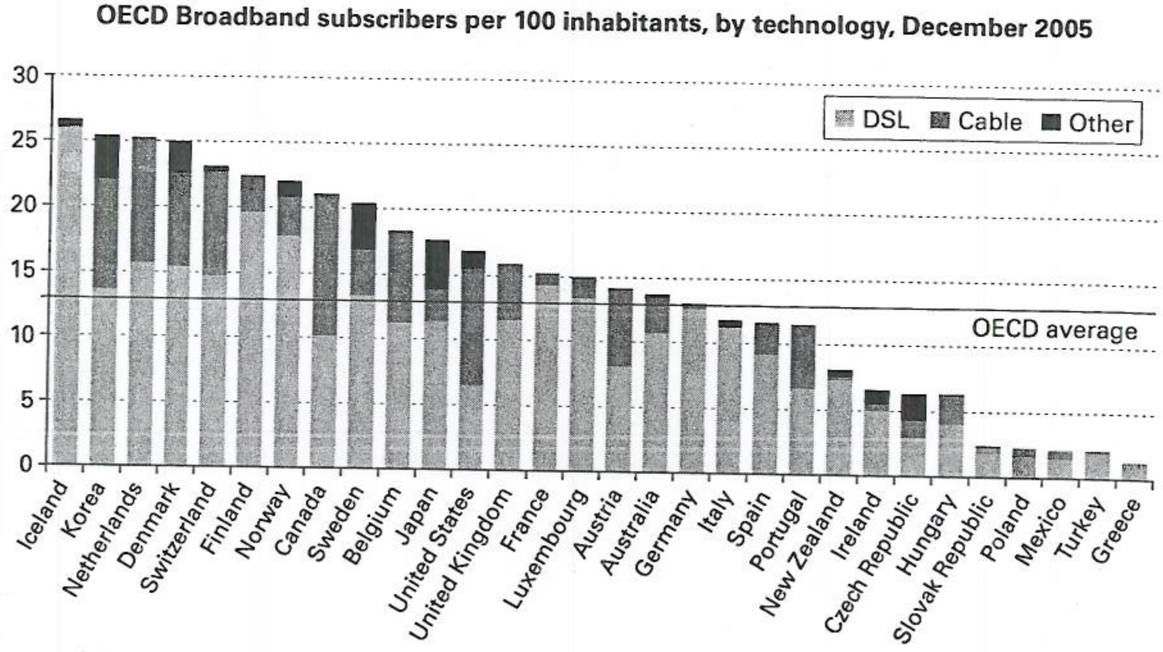

Figure 13.11
Internet broadband penetration among OECD countries (source: OECD)

Competitiveness of the Internet Backbone and the MCI Mergers When WorldCom proposed to buy MCI, the European Union objected,[69] with concern that the combined Internet backbones of MCI and WorldCom would dominate the European market with anticompetitive consequences. As a result of EU objections, MCI agreed to spin off its backbone, MCIi, which was bought by Cable and Wireless. The subsequent proposed merger of WorldCom and Sprint faced similar objections by the European Union and in the United States by the DOJ, as well as objections by DOJ in the long-distance and unswitched private lines markets, so that was shelved as well. The market shares at the time and contemporary projections are summarized in table 13.3.[70]

In papers filed in support of the merger of SBC and AT&T as well as of the merger of Verizon with MCI, there was mention of two recent traffic studies by RHK. These studies showing traffic for 2004, summarized in table 13.4, show a dramatic change in the ranking of the networks, with AT&T now being first and MCI being fourth. They also show that now a much bigger share of traffic (over 40 percent) is carried by smaller networks. These latest traffic studies show that the concern of the European Union and of the DOJ in the United States that the Internet backbone market would tilt to monopoly has proved to be overstated.

GTE, which had made a smaller bid for MCI than WorldCom, supported by Cremer et al. (1998, 2000), raised a number of anticompetitive concerns for networks that obey the following assumptions:

Table 13.3
Market shares of national Internet backbones

Company	1997	1999	2001 (projected in 1999)	2003 (projected in 1999)
MCI WorldCom	43%	38%	35%	32%
GTE-BBN	13%	15%	16%	17%
AT&T	12%	11%	14%	19%
Sprint	12%	9%	8%	7%
Cable and wireless	9%	6%	6%	6%
All other	11%	21%	22%	19%
Total	100%	100%	100%	100%

Source: *Hearing on the MCI WorldCom-Sprint Merger before the Senate Committee on the Judiciary*, Exhibit 3 (November 4, 1999). Testimony of Tod A. Jacobs, Senior Telecommunications Analyst, Sanford C. Bernstein & Co., Inc. Bernstein Research, *MCI WorldCom* (March 1999) at p. 51.

Table 13.4
Carrier traffic in petabytes per month in 2004

	Traffic				Market share among all networks
Company	1Q2004	2Q2004	3Q2004	4Q2004	4Q2004
A (AT&T)	37.19	38.66	44.54	52.33	12.58%
B	36.48	36.50	41.41	51.31	12.33%
C	34.11	35.60	36.75	45.89	11.03%
D (MCI)	24.71	25.81	26.86	30.87	7.42%
E	18.04	18.89	21.08	25.46	6.12%
F	16.33	17.78	17.47	19.33	4.65%
G	16.67	15.04	14.93	15.19	3.65%
Total traffic top 7 networks	183.53	188.28	203.04	240.38	57.78%
Total traffic all networks	313	313	353	416	100%

Source: Data from *RHK Traffic Analysis—Methodology and Results*, May 2005.
Note: The identities of all networks are not provided, but it is likely that B, C, E, and F are level 3; Quest, Sprint, and SBC are in unknown order.

1. Consumers do *not* demand universal connectivity.
2. There is an installed base of clients (ISPs) of Internet backbone networks who cannot migrate to other providers.

Using these assumptions, Cremer et al. (1998, 2000) argue that (1) a large IBN has an incentive to introduce incompatibilities and to degrade interconnection with one rival but not with all rivals, (2) even small differences in network size will lead to a spiral of ever-increasing dominance by a larger IBN, since dominance is defined by size, (3) large IBN will refuse to cooperate with small networks, and (4) in the case where switching costs are low, large IBNs will still be able to dominate small networks.

These results are based on assumptions that do not presently hold on the Internet backbone. First, customers on the Internet demand *universal connectivity*. Users of the Internet do not know in advance what Internet site they may want to contact or to whom they might want to send e-mail. Thus Internet users demand from their ISPs and expect to receive universal connectivity. This is the same expectation that users of telephones, mail, and fax machines have: that they can connect to any other user of the network without concern about compatibility, location, or, in the case of telephone or fax, any concern about the manufacturer of the appliance, the type of connection (wireline or wireless) or the owners of the networks over which the connection is made. Because of the users' demand for universal connectivity, ISPs providing services to end-users or to Web sites must make arrangements with other networks so that they can exchange traffic with *any* Internet customer.

Second, there are no "captive customers" on the Internet backbone. The reasons for this are as follows:

1. ISPs can easily and with low cost migrate all or part of their transport traffic to other network providers.
2. Many ISPs already purchase transport from more than one backbone to guard against network failures and for competitive reasons (ISP "multihoming").
3. Many large Web sites/providers use more than one ISP for their sites ("customer multihoming").
4. Competitive pressure from their customers makes ISPs agile and likely to respond quickly to changes in conditions in the backbone market.

Competitive conditions imply that significant price increases, or raising rivals' costs or degrading interconnection, are unlikely to be profitable on the Internet backbone. If the large Internet backbone connectivity provider's strategy were to impose equal increases in transport costs on all customers, the response of other backbone providers and ISPs will be to reduce the traffic for which they buy transit from the large IBN, and to instead reroute traffic and purchase more transit from each other. Thus, in response to a price increase by the large Internet backbone, other IBNs and ISPs reduce the traffic for which they buy transit from the large IBN down to the minimum level necessary to reach IBNs that are *exclusively* connected to the large IBN. All other IBNs and ISPs exchange all other traffic with each other, bypassing the large IBN. For these and other reasons it is unlikely that large IBNs at the time of the MCI mergers will find it profitable to limit connectivity to other backbones.[71]

13.4 Concluding Remarks

The discussion of this chapter is a start of an in-depth study of public policy in network industries. The legal system does not yet have a framework for analysis of competition policy issues in network industries. This was to a large extent exemplified in *United States v. Microsoft*.

The Microsoft case has certainly been the most important antitrust case of the "new economy" this far. Unfortunately, its legal battle was fought to a large extent without the use of the economics tools discussed above, which are at the foundation of the new economy and were key to the business success of Microsoft. There are several explanations for this omission. First, often legal cases are created and filed before an economist is found who will create the appropriate economic model to support the case. Second, the economic theory of networks is so incomplete and unsettled that there is no commonly accepted body of knowledge on market structure with network effects. Based on such theory one could have evaluated deviations toward anticompetitive behavior. Third, the legal system has tremendous inertia to new ideas and models. Fourth, the legal system is ill-equipped to deal with complex technical matters. Fifth, because of all these facts lawyers on both sides find it easier to fight the issues on well-trotted ground even if the problems are really of a different nature. It is as if there is a dispute among two parties in the middle of a heavily forested area, but the lawyers of both parties fight it as if the dispute happened on the open plains. They know the way that disputes on the plains are resolved, whereas the law of dispute resolution in forests has yet to be established.

The bottom line is that as things stand today, there are many areas of antitrust law where there is significant uncertainty on how to apply the law in network industries and when to judge various practices as legal or illegal. Guidance to firms on avoidance of anticompetitive actions is further complicated by the divergence of competition policy standards between the European Union and the United States.

With further academic analysis of antitrust issues, I hope that when the next major new economy antitrust case appears, there will be a deeper understanding and application of the economics of networks and of the way that antitrust law should be applied to network industries.

Notes

1. I will typically use the term "network effects" and reserve the term "network externalities" for those cases where a market fails to fully intermediate network effects.

2. Some authors consider two-sided networks a separate category than markets with network effects. *See* Armstrong (2005), Evans (2003), Jullien (2004), and Rochet and Tirole (2003, 2004). However, the two-sided setup is a standard setup that has been discussed in the economics of networks literature since the middle 1980s, and there is no crucial distinguishing factor to separate the analysis of two-sided networks from the traditional analysis of virtual networks.

3. More precisely, the law of demand is true for *normal* goods, that is, for goods for which an increase in income leads to a higher quantity of sales. If increases in income drove sales sharply down, the possibility of a *Giffen good* arises where sales increase as prices increase. Giffen goods are truly exceptional and rarely observed.

4. Besides the positive network effects described in detail in this chapter, there is the possibility of negative effects, such as congestion in transportation networks, or interference in radio, broadcast TV, and wireless telecommunications networks.

5. See Economides and Siow (1988).

6. The existence of network externalities in exchange markets is one of the reasons behind the proposal of Economides-Schwartz to bunch orders and execute them at once in pre-defined times, thereby creating a *call market* concurrent with the continuous NYSE market; see Economides and Schwartz (1995a). A call market has higher liquidity than continuous markets and can provide anonymity to large orders. A survey of equity traders has established that most of traders are willing to wait for execution of their orders in a call market; see Economides and Schwartz (1995b).

7. Suppose that consumers expect a network of size n^e, $0 \leq n^e \leq 1$. Let network externalities be summarized by function $h(n^e) = k + f(ne)$, with $f(0) = 0$, $f'(\cdot) > 0$, $f''(\cdot) \leq 0$. Consumer of type y has utility $u(y, n^e) = yh(n^e)$ from one unit of the good, and let the distribution of consumer types be $G(y)$. Then the marginal consumer y^* is defined by $u(y^*, n^e) = p$. Therefore $y^* = p/h(n^e)$ for $0 < p/h(n^e) < 1$. It follows that demand is $n = 1 - G(y^*) \Rightarrow p(n, n^e) = h(n^e)G^{-1}(1-n)$. Imposing fulfilled expectation, $n = n^e$, we derive the fulfilled expectations demand $p(n, n) = h(n)G^{-1}(1-n)$. Now we check for upward-slopping demand at $n = 0$. In general, the slope of the fulfilled expectations demand is $dp(n, n)/dn = -h(n)/G' + h'(n)G^{-1}(1-n)$. The slope at $n = 0$ is positive if and only if $\lim_{n\to 0} dp(n, n)/dn = \lim_{n\to 0} h'(n) - \lim_{n\to 0} h/G'(1) = \lim_{n\to 0} f'(n) - k/G'(1) > 0$. Therefore the fulfilled expectations demand slopes upward at $n = 0$ in the three cases:

- When $k = 0$ (pure network good)
- When $\lim_{n\to 0} f'(n)$ large (very strong marginal network externalities)
- When $G'(1)$ large (density of high-valuation types is sufficiently large)

8. The pricing schemes used vary considerably depending on the telecommunications service. Traditionally, in fixed networks, as in most places in the United States, local calls are free with a local connection that requires a fixed monthly fee. Long-distance subscribers were charged only for outgoing calls. In the last three decades 800, 866, 877, 888–prefix "toll-free" services allow for no charge to the calling party but impose a fee to the receiving party, while 900–prefix services allow the receiving party to charge a positive price to the initiator. In wireless cellular and PCS telecommunications, subscribers in the United States pay for both incoming and outgoing calls, whereas in most of the rest of the world, wireless subscribers pay only for outgoing calls. On the Internet, typically retail subscribers pay a flat fee regardless of the amount of time they use the service, the number of bits exchanged, or whether the calls are incoming or outgoing, and regardless of the destination. Similarly Internet service providers buy backbone connectivity at rates that depend just on the size of the pipe they utilize and regardless of utilization or whether the information packets are incoming or outgoing, and regardless of the destination.

9. Low-technology examples are razors and blades and cameras and film.

10. It is anecdotally known that Cantor Fitzgerald, which has a 70 percent market share in the secondary market for US government thirty-year bonds, offered to Salomon (the largest "primary dealer" and trader of US bonds) prices equal to 1/10 to 1/5 of those charged to small traders. This is consistent with profit maximization by Cantor Fitzgerald because of the liquidity (network effect) brought to the market by Salomon, which is by far the largest buyer ("primary dealer") in the auctions of US government bonds.

11. For a detailed discussion, see Economides and Himmelberg (1995).

12. This is easy to prove mathematically in the model of Economides and Himmelberg (1995) discussed earlier. Let $B(n, n) = \int_0^n p(q, n)\, dq$ be the gross benefit function at sales level n. Then the incremental benefit of network expansion is $B(n, n)/dn = p(n, n) + \int_0^n \partial p(q, n)/\partial n\, dq > p(n, n)$, since $\partial p(q, n)/\partial n > 0$ represents the positive network effect of expansion of the network on the willingness to pay of other network participants. Under perfect

competition, consumers face a price equal to marginal cost. Since the incremental benefit of network expansion exceeds price, perfect competition is inefficient.

13. See Stango (2004) for a recent survey of the economics of standards setting. There are a number of recent empirical studies on standard competition and competition with competing with incompatible standards. Dranove and Gandal (2001) discuss the effects of the "preannouncement" of the DIVX standard on adoption of DVDs. Also see the earlier analyses of Farrell and Saloner (1985, 1986), the survey of Besen and Saloner (1994), and Lerner and Tirole (2004). Additionally often firms pre-announce new products to make consumers delay their purchasing decisions and therefore make it less likely that consumers will buy products from competitors. Often, products are pre-announced and never appear in the market, although the pre-announcement had the effect discussed above. Such products that never appear have been called "vaporware." For a more detailed discussion, see Choi et al. (2005).

14. If the distribution of the willingness to pay is distributed away from 0, an industry with network effects exhibits the finiteness property (Shaked and Sutton 1983), with a finite maximum number of active firms despite all realizing strictly positive profits.

15. See David (1985) and Liebowitz and Margolis (1990).

16. AT&T claimed that the main reason for its refusal to interconnect was the low technical standards of the independents, as well as incompatibilities, that would jeopardize AT&T's network after interconnection. While there is some truth to those claims, it is unlikely that they applied to all independents. Moreover, once acquired by AT&T, independents were interconnected with AT&T's network, after some modifications. This shows that the refusal to interconnect was mainly a strategic and commercial decision rather than a technical one.

17. See Economides and White (1995) and Economides (2003).

18. AT&T, the largest long-distance company, announced in the summer of 2004 that it stopped marketing both local and long-distance service to residential customers, and MCI followed the same strategy without a formal announcement. In 2005 SBC announced that it would buy AT&T, and Verizon announced that it would buy MCI. Both mergers cleared antitrust and regulatory review in March 2007. In filings to the US Department of Justice and the FCC, both AT&T and MCI claimed that they had abandoned the residential local and long-distance markets. For a detailed discussion, see Economides (2005a, b).

19. 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, whereas the second entrant, BellSouth New Zealand, exited the local telecommunications market.

20. Also see Armstrong (1998), Cambini and Valletti and (2005), Carter and Wright (2003), Dessein (2004), Gans and King (2001), and Laffont et al. (1998).

21. For example, one can think of A as a computer operating system (OS), and B as an application. OS manufacturers can and do embed software routines that are useful to application software developers, since they reduce the cost of writing applications.

22. See Economides and Siow (1988) for a discussion of the benefits of B2B and other exchanges.

23. This can help an incumbent because it can expand the functionality of its product, but can also help its rivals as they may incorporate functionalities of the incumbent's product into theirs.

24. The California Public Utilities Commission investigated the WorldCom-Sprint proposed merger in the context of the "transfer of licenses" to the merged entity. However, after the merger proposal was withdrawn by the parties, the CPUC continued to investigate the participation of these two firms on the Internet, and asserted its regulatory authority over the part of the Internet that is physically located in California.

25. DOJ sued Microsoft on July 15, 1994, under 15 USC §2 of the Sherman Act, alleging that Microsoft had entered into licensing agreements with OEMs that prevented other operating system vendors from gaining widespread distribution of their products.

26. The Microsoft court entered the consent decree as its Final Judgment on April 21, 1995.

27. Final Judgment, Civil Action No. 94-1564.

28. The Microsoft court also referred the issue to a special master, Professor Lawrence Lessig of Stanford.

29. The court of appeals further noted that "the limited competence of courts to evaluate high-tech product designs and the high cost of error should make them wary of second-guessing the claimed benefits of a particular design decision"; *see* 147 F.3d at 950 n. 13.

30. For a detailed discussion of *United States v. Microsoft*, see Economides (2001a, 2001b, 2002) and http://www.stern.nyu.edu/networks/.

31. Only twelve witnesses testified for each side.

32. As mediator, Judge Posner was *not* acting in his judicial capacity.

33. See generally, *New York Times*, April 2, 2000.

34. The hearing was on May 24, 2000. It started at 10:15 am, ended around 3:30 pm, and included a two-hour lunch break.

35. It is likely that the marginal price for the last unit sold to the same OEM was extremely low, since the 1995 consent decree allowed Microsoft to have quantity discounts but barred it from zero marginal cost pricing.

36. The derivation of the monopoly price for Windows follows. Let p_H be the price of the PC hardware (everything except Windows), and let p_W be the price of Windows. Assume that Windows is installed on all PCs, meaning Microsoft has 100 percent market share. Since hardware and software are combined in a ratio of 1:1, the demand of a PC with Windows is $D(p_H + p_W)$. Profits of Microsoft from Windows sales are

$$\Pi_W = p_W D(p_H + p_W) - F_W,$$

where F_W is the fixed cost of developing Windows, and the marginal cost is negligible. Maximizing Π_W implies marginal revenue equals marginal cost:

$$\frac{D(p_H + p_W) + p_W\, dD}{dp_W} = 0 \Leftrightarrow 1 + \left(\frac{p_W}{p_H + p_W}\right)\left(\frac{p_H + p_W}{D}\right)\left(\frac{dD}{dp_W}\right) = 0 \Leftrightarrow \frac{p_W}{p_H + p_W} = \frac{1}{|\varepsilon|},$$

or equivalently, the monopoly price of Windows is

$$p_W = \frac{p_H}{|\varepsilon| - 1},$$

where $|\varepsilon| = -[(p_H + p_W)/D]/[dD/dp_W]$ is the market elasticity of demand for PCs with Windows. Say that the average price of PC hardware is \$1,800 and the elasticity is $|\varepsilon| = 2$. Then the monopoly price of Windows is $p_W = \$1{,}800$. Even if the elasticity were much higher $|\varepsilon| = 3$, and the average price of PC hardware much lower at \$1,200, the monopoly price would be \$600, which is ten to twelve times the price charged by Microsoft to OEMs. An elasticity of $|\varepsilon| = 31$ is required to get a Windows monopoly price of \$60. It is extremely unlikely that that the market for PCs exhibits such large market elasticities.

37. See Economides (2001a, b).

38. During the trial Judge Jackson gave extensive interviews on his views of the case. In one of them, he compared the leadership of Microsoft (Bill Gates, Steve Ballmer) to a drug-dealing gang that he had convicted in a recent case.

39. One panel member is selected by Microsoft, one by the Justice Department, and one by both. The panel has full access to all of Microsoft's books, records, systems, and personnel, including source code, and helps resolve disputes about Microsoft's compliance with the disclosure provisions of the settlement.

40. See my detailed position at my amicus brief, Economides, Nicholas (2002), *http://www.usdoj.gov/atr/cases/ms_tuncom/major/mtc-00022465.htm.*

41. See "EU Commission Concludes Microsoft Investigation, Imposes Conduct Remedies and a Fine," *at http://www.eurunion.org/news/press/2004/20040045.htm.*

42. For an extensive discussion of the 1996 Act, see Economides (2005a, 2003).

43. As the Federal Communications Commission stated:

Because an incumbent LEC currently serves virtually all subscribers in its local serving area, an incumbent LEC has little economic incentive to assist new entrants in their efforts to secure a greater share of that market. An incumbent LEC also has the ability to act on its incentive to discourage entry and robust competition by not interconnecting its network with the new entrant's network or by insisting on supracompetitive prices or other unreasonable conditions for terminating calls from the entrant's customers to the incumbent LEC's subscribers.

Congress addressed these problems in the 1996 Act by mandating that the most significant economic impediments to efficient entry into the monopolized local market must be removed. The incumbent LECs have economies of density, connectivity, and scale; traditionally, these have been viewed as creating a natural monopoly. As we pointed out in our NPRM, the local competition provisions of the Act require that these economies be shared with entrants. (Federal Communications Commission, First Report and Order, FCC No. 96-325, at ¶ 10–11 [August 1, 1996])

44. 47 USC §§251(c)(3), (c)(6). The Federal Communications Commission defined the key UNEs as the "local loop," local switching, and local transport. See Federal Communications Commission, First Report and Order, FCC No. 96-325, at §51.319 (August 1, 1996).

45. The 1996 Act at 47 USC §252(d)(1) orders that pricing of interconnection or unbundled network elements:

(A) shall be
(i) based on the cost (determined without reference to a rate-of-return or other rate-based proceeding) of providing the interconnection or network element (whichever is applicable), and
(ii) nondiscriminatory, and
(B) may include a reasonable profit.

46. 47 USC §251(c)(3).

47. 47 USC §251(c)(6).

48. 47 USC §251(b)(2).

49. 540 US 398 (2004).

50. Ibid. at 402–405. Trinko originally sued NYNEX, which was later bought by Bell Atlantic. Bell Atlantic merged with GTE to created Verizon.

51. *Trinko*, 540 US at 409.

52. *See* Economides, Seim, and Viard (2007).

53. *Trinko*, 305 F.3d at 108.

54. *Trinko*, 540 US at 415 n. 4.

55. For a full discussion of the *Trinko* case, see Economides (2005b, 2007).

56. *Trinko*, 540 US at 409.

57. *Aspen Skiing*, 472 US at 593–608.

58. *See* Economides (1998).

59. See generally, Salop and Scheffman (1983) and Krattenmaker and Salop (1986).

60. Brief of Amici Curiae United States and the Federal Trade Commission in support of *Trinko*, 540 US 398 (2004) (No. 02-682) at 16.

61. As the government's brief notes, the sacrifice principle has been used in various versions in *Aspen Skiing*, 472 US at 608, 610–611 (conduct that "sacrifice[s] short-run benefits," such as immediate income and consumer goodwill, undertaken because it "reduc[es] competition over the long run"); *General Industries Corp. v. Hartz Mountain Corp.*, 810 F.2d 795, 803 (8th Cir. 1987) (conduct anticompetitive if "its 'anticipated benefits were dependent upon its tendency to discipline or eliminate competition and thereby enhance the firm's long term ability to reap the benefits of monopoly power.'"); *Stearns Airport Equipment Co. v. FMC Corp.*, 170 F.3d 518, 523–524 n. 3 (5th Cir. 1999) (conduct exclusionary if it harms the monopolist but is justified because it causes rivals more harm); *Advanced Health-Care Services v. Radford Community Hospital*, 910 F.2d 139, 148 (4th Cir. 1990) ("making a short-term sacrifice" that "harm[s] consumers and competition" to further "exclusive, anti-competitive objectives"). Ibid.

62. "Conduct is not exclusionary or predatory unless it would make no economic sense for the defendant but for its tendency to eliminate or lessen competition." Brief of Amici Curiae United States and the Federal Trade Commission at 15, *Verizon Communications Inc. v. Law Offices of Curtis V. Trinko*, 540 US 398 (2004) (No. 02-682). For a more detailed discussion of the *Trinko* case, see Economides (2005b, 2007).

63. Educational institutions and government departments are also connected to the Internet but do not offer commercial ISP services.

64. This service is called "transit." See Economides (2005c).

65. See Economides (2005c, fig. 2).

66. SBC changed its name to AT&T after it acquired AT&T.

67. The proposed Internet model without net neutrality would be closer to the traditional pre–Internet telecommunications model where customers pay per service. See Economides (2005b).

68. This is contingent on all markets are served under uniform pricing, which holds here because we are starting with all markets served under net neutrality.

69. See European Union Commission (1998).

70. See Economides (2005c).

71. For a more detailed discussion, see Economides (2005c).

References

Armstrong, M. 1998. Network interconnection in telecommunications. *Economic Journal* 108: 545–64.

Armstrong, M. 2005. Competition in two-sided markets. *RAND Journal of Economics* 37: 668–91.

Advanced Health-Care Services v. Radford Community Hospital. 910 F.2d 139, 148 (4th Cir. 1990).

Aspen Skiing Co. v. Aspen Highlands Skiing Corp. 472 US 585, 593–608 (1985).

Besen, S., and J. Farrell. 1994. Choosing how to compete: Strategies and tactics in standardization. *Journal of Economic Perspectives* 8: 117–31.

Brief of Amici Curiae United States and the Federal Trade Commission. Verizon Communications Inc. v. Law Offices of Curtis V. Trinko, 540 US 398 (2004).

Carter, M., and J. Wright. 2003. Asymmetric network interconnection. *Review of Industrial Organization* 22: 27–46.

Choi, J. P., E. G. Kristiansen, and J. Nahm. 2005. An economic analysis of product pre-announcements. *CESifo Economic Studies* 51(2–3): 299–319.

Chou, C.-F., and O. Shy. 1990. Network effects without network externalities. *International Journal of Industrial Organization* 8: 259–70.

Church, J., and N. Gandal. 1993. Complementary network externalities and technology adoption. *International Journal of Industrial Organization* 11: 239–60.

Cremer, J., P. Rey, and J. Tirole. 1998. The degradation of quality and the domination of the Internet. *Submission of GTE to the European Union for the merger of MCI with WorldCom*, app. 5.

Cremer, J., P. Rey, and J. Tirole. 2000. Connectivity in the commercial Internet. *Journal of Industrial Economics* 48: 433–72.

David, P. A. 1986. Understanding the economics of QWERTY. In W. N. Parker, ed., *Economic History and the Modern Economist*. Oxford: Basil Blackwell.

Dessein, W. 2003. Network competition in nonlinear pricing. *RAND Journal of Economics* 34(4): 593–611.

Dranove, D., and N. Gandal. 2003. The DVD vs. DIVX standard war: Empirical evidence of preannouncement effects. *Journal of Economics and Management Strategy* 12: 363–86.

Economides, N. 1996. The economics of networks. *International Journal of Industrial Organization* 16(4): 675–99. Prepublication copy available at ⟨http://www.stern.nyu.edu/networks/Economics_of_Networks.pdf⟩.

Economides, N. 1998. The incentive for non-price discrimination by an input monopolist. *International Journal of Industrial Organization* 16: 271–84. Prepublication copy available at ⟨http://www.stern.nyu.edu/networks/1136.pdf⟩.

Economides, N. 1999. The Telecommunications Act of 1996 and its impact. *Japan and the World Economy* 11: 455–83. Prepublication copy available at ⟨http://www.stern.nyu.edu/networks/98-08.pdf⟩.

Economides, N. 2001a. The Microsoft antitrust case. *Journal of Industry, Competition and Trade: From Theory to Policy*. Prepublication copy available at ⟨http://www.stern.nyu.edu/networks/Microsoft_Antitrust.final.pdf⟩.

Economides, N. 2001b. The Microsoft antitrust case: Rejoinder. *Journal of Industry, Competition and Trade: From Theory to Policy* 1(1): 71–79. Prepublication copy available at ⟨http://www.stern.nyu.edu/networks/Microsoft_Antitrust.Rejoinder.pdf⟩.

Economides, N. 2002. Amicus brief of Nicholas S. Economides on the revised proposed final judgment in the Microsoft case. Submitted to the United States Department of Justice. Available at ⟨http://www.usdoj.gov/atr/cases/ms_tuncom/major/mtc-00022465.htm⟩.

Economides, N. 2003a. US telecommunications today. In C. V. Brown and H. Topi, eds., *IS Management Handbook*. Boca Raton, FL: Auerbach Publications, pp. 191–212. Prepublication electronic copy available at ⟨http://www.stern.nyu.edu/networks/Economides_US_Telecommunications_Today_October_2002.pdf⟩.

Economides, N. 2003b. The tragic inefficiency of M-ECPR. In A. Shampine, ed., *Down to the Wire: Studies in the Diffusion and Regulation of Telecommunications Technologies*. New York: Nova Science Publishers, pp. 142–54. Prepublication copy available at ⟨http://www.stern.nyu.edu/networks/tragic.pdf⟩.

Economides, N. 2005a. Telecommunications regulation: An introduction. In R. R. Nelson, ed., *The Limits and Complexity of Organizations*. New York: Russell Sage Foundation Press, pp. 48–76. Prepublication copy available at ⟨http://www.stern.nyu.edu/networks/Telecommunications_Regulation.pdf⟩.

Economides, N. 2005b. Vertical leverage and the sacrifice principle: Why the Supreme Court got Trinko wrong. *New York University Annual Survey of American Law* 63(3): 379–413. Prepublication copy available at ⟨http://www.stern.nyu.edu/networks/Trinko.pdf⟩.

Economides, N. 2005c. The economics of the Internet backbone. In I. Vogelsang, ed., *Handbook of Telecommunications*. Amsterdam: Elsevier. Prepublication copy available at ⟨http://www.stern.nyu.edu/networks/Economides_ECONOMICS_OF_THE_INTERNET_BACKBONE.pdf⟩.

Economides, N. 2007. Hit and miss: Leverage, sacrifice, and refusal to deal and the supreme court decision in *Trinko*. Forthcoming in *Vanderbilt Journal of Entertainment and Technology Law*. Prepublication available copy at ⟨http://www.stern.nyu.edu/networks/Hit_and_Miss.pdf⟩.

Economides, N. 2008. Net neutrality, non-discrimination, and digital distribution of content through the Internet. Forthcoming in *I/S: A Journal of Law and Policy for the Information Society*. Prepublication copy available at ⟨http://www.stern.nyu.edu/networks/Economides_Net_Neutrality.pdf⟩.

Economides, N., and F. Flyer. 1998. Compatibility and market structure for network goods. Discussion paper EC-98-02. Stern School of Business, NYU. Available at ⟨http://www.stern.nyu.edu/networks/98-02.pdf⟩.

Economides, N., and C. Himmelberg. 1995. Critical mass and network evolution in telecommunications. In G. Brock, ed., *Toward a Competitive Telecommunications Industry: Selected Papers from the 1994 Telecommunications Policy Research Conference*. Prepublication copy available at ⟨http://www.stern.nyu.edu/networks/tprc.pdf⟩.

Economides, N., G. Lopomo, and G. Woroch. 1996. Regulatory pricing policies to neutralize network dominance. *Industrial and Corporate Change* 5(4): 1013–28. Prepublication copy available at ⟨http://www.stern.nyu.edu/networks/96-14.pdf⟩.

Economides, N., G. Lopomo, and G. Woroch. 2007. Strategic commitments and the principle of reciprocity in interconnection pricing. In G. Madden, ed., *The Economics of Digital Markets*. Cheltenham, UK: Edward Elgaz. Available at ⟨http://www.stern.nyu.edu/networks/reciprocity.pdf⟩.

Economides, N., and R. Schwartz. 1995a. Electronic call market trading. *Journal of Portfolio Management* 21(3): 10–18. Prepublication copy available at ⟨http://www.stern.nyu.edu/networks/Electronic_Call_Market_Trading.pdf⟩.

Economides, N., and R. Schwartz. 1995b. Equity trading practices and market structure: Assessing asset managers' demand for immediacy. *Financial Markets, Institutions and Instruments* 4(4): 1–46. Prepublication copy available at ⟨http://www.stern.nyu.edu/networks/Equity_Trading_Practices.pdf⟩.

Economides, N., K. Seim, and B. Viard. 2007. Quantifying the benefits of entry into local phone service. Available at ⟨http://www.stern.nyu.edu/networks/Local_Telecommunications.pdf⟩.

Economides, N., and A. Siow. 1988. The division of markets is limited by the extent of liquidity. *American Economic Review* 78(1): 108–21. Available at ⟨http://www.stern.nyu.edu/networks/aer88.pdf⟩.

Economides, N., and J. Tåg 2007. Net neutrality on the Internet: A two-sided market analysis. Available at ⟨http://www.stern.nyu.edu/networks/Economides_Tag_Net_Neutrality.pdf⟩.

Economides, N., and L. J. White. 1994. Networks and compatibility: Implications for antitrust. *European Economic Review* 38: 651–62.

Economides, N., and L. J. White. 1995. Access and interconnection pricing: How efficient is the efficient component pricing rule? *Antitrust Bulletin* 40(3): 557–79. Prepublication copy available at ⟨http://www.stern.nyu.edu/networks/Access_and_interconnection_pricing.pdf⟩.

Economides, N., and L. J. White. 1998. The inefficiency of the ECPR yet again: A reply to Larson. *Antitrust Bulletin* 43(2): 429–44. Prepublication copy available at ⟨http://www.stern.nyu.edu/networks/The_Inefficiency_of_the_ECPR.pdf⟩.

European Union Commission. 1998. *Statement of Objections to the MCI WorldCom Merger.*

Evans, D. 2003. The antitrust economics of multi-sided platform markets. *Yale Journal on Regulation* 20(2): 324–81.

Farrell, J., and P. Klemperer. 2003. Coordination and lock-in: Competition with switching costs and network effects. In R. Schmalensee and R. Willig, eds., *Handbook of Industrial Organization*, vol. 3. Amsterdam: North Holland.

Farrell, J., and G. Saloner. 1985. Standardization, compatibility and innovation. *RAND Journal of Economics* 16: 70–83.

Farrell, J., and G. Saloner. 1986. Installed base and compatibility: Innovation, product preannoucements, and predation. *American Economic Review* 76: 940–55.

Federal Communications Commission. 1996. *First Report and Order*, FCC No. 96-325 (August 1, 1996).

Gabel, D., and D. F. Weiman. 1994. Historical perspectives on interconnection between competing local operating companies: The United States, 1894–1914. In D. Gabel and D. F. Weiman, eds., *Opening Networks to Competition: The Regulation and Pricing of Access.* Norwell, MA: Kluwer Academic, pp. 75–105.

Gans, J., and S. King. 2001. Using "bill and keep" interconnect arrangements to soften network competition. *Economic Letters* 71(3): 413–20.

General Industries Corp. v. Hartz Mountain Corp. 810 F.2d 795, 803 (8th Cir. 1987).

Katz, M. L., and C. Shapiro. 1992. Product introduction with network externalities. *Journal of Industrial Economics* 40: 55–83.

Katz, M. L., and C. Shapiro. 1994. Systems competition and network effects. *Journal of Economic Perspectives* 8: 93–115.

Katz, M. L., and C. Shapiro. 1986. Technology adoption in the presence of network externalities. *Journal of Political Economy* 94: 822–41.

Krattenmaker, T. G., and S. C. Salop. 1986. Anticompetitive exclusion: Raising rivals' costs to achieve power over price. *Yale Law Journal* 96: 209.

Laffont, J.-J., P. Rey, and J. Tirole. 1998. Network competition. II: Price discrimination. *RAND Journal of Economics* 29(1): 38–56.

Lerner, J., and J. Tirole. 2006. A model of forum shopping, with special reference to standard setting organizations. *American Economic Review* 96: 1091–103.

Liebowitz, S. J., and S. E. Margolis. 1990. The fable of the keys. *Journal of Law and Economics* 33: 1–25.

Matutes, C., and P. Regibeau. 1988. Mix and match: Product compatibility without network externalities. *RAND Journal of Economics* 19: 221–34.

Posner, R. A. 2000. Antitrust in the new economy. *Tech Law Journal.* Available at ⟨http://www.techlawjournal.com/atr/20000914posner.asp⟩.

Rochet, J.-C., and J. Tirole. 2003. Platform competition in two-sided markets. *Journal of the European Economic Association* 1(4): 990–1029.

Rochet, J.-C., and J. Tirole. 2004. Two-sided markets: An overview. Mimeo. University of Toulonse.

Rohlfs, J. 1974. A theory of interdependent demand for a communications service. *Bell Journal of Economics* 5(1): 16–37.

Rohlfs, J. 2003. *Bandwagon Effects in High Technology Industries.* Cambridge: MIT Press.

Salop, S. C., and D. T. Scheffman. 1983. Raising rivals' cost. American Economic Review 73: 267.

Shaked, A., and J. Sutton. 1983. Natural oligopolies. *Econometrica* 51: 1469–84.

Stango, V. 2004. The economics of standards wars. *Review of Network Economics* 3(1): 1–19.

Stearns Airport Equipment Co. v. FMC Corp. 170 F.3d 518, 523–524 n. 3 (5th Cir. 1999).

Valletti, T. M., and C. Cambini. 2005. Investments and network competition. *RAND Journal of Economics* 36: 446–67.

Verizon Communications Inc. v. Law Offices of Curtis V. Trinko. 540 US 398 (2004).