Introduction

The purpose of this book is to examine the implications of network theory for public policy toward communications. Network theory provides a highly useful and powerful tool for modeling and understanding communications networks. Network theory opens the "black box" of networks to reveal the complexities of network architecture. Rather than working with the reduced form of the standard economic cost function, network theory pays attention to the nodes and links that form networks. The architecture of networks affects transmission capacity and the quality of service. The structure of networks also is important for understanding the boundaries of private networks provided by individual firms and the interconnections between networks. By applying network theory to communications, we develop a general Coasian theory of networks.

Using network theory, we develop a comprehensive framework for analyzing public policy toward communications. The discussion identifies the shift in regulatory policy from utility regulation to mandated access. Access to networks refers to the market transactions that connect a firm's network to its customers, suppliers, competitors, and partners. Our discussion presents a classification scheme for analyzing access to networks. We apply network theory based on the mathematics of graph theory tools to characterize the structure of networks. We examine how regulatory access mandates can distort access price. In competitive markets, optimization by firms and access transactions lead to efficient network boundaries and interconnections. Regulatory access mandates potentially reduce economic efficiency by changing the market-equilibrium structure and boundaries of communications networks.

A. The Changing Network Economy

Economic growth in the modern economy has accompanied a dramatic upsurge in the importance of communications networks. Consumers and businesses have

become linked together with ever-faster connections transmitting a growing amount of data for electronic commerce, organizational management, and social interaction. Communications networks have joined more traditional network industries in energy and transportation in playing a pivotal role in the functioning of national economies. The information and communications technology (ICT) industries form a growing share of the global economy (U.S. Department of Commerce 2003). International trade has gone online, with worldwide electronic transmission of information, technology, services, entertainment, and money. Telecommunications companies have played a starring role in a significant number of megamergers that have transformed the business environment and also served as a driving force behind the spectacular rise and equally spectacular fall of the NASDAQ index. Perhaps most dramatically, the failure of WorldCom produced the largest bankruptcy in U.S. history.

In addition, scientific advances have rendered the technological environment increasingly dynamic, with different types of communications now available through an ever-increasing array of transmission technologies. These developments have made different communications media increasingly interchangeable and have turned different technological platforms that had previously constituted universes unto themselves into competitors. Not only has the emergence of platform competition provided consumers and firms with a wider variety of ways to access network services; it has also begun to put pressure on the traditional regulatory distinction among voice, video, and data communications, under which each type of service was governed by a separate regulatory regime (Yoo 2002). It is also forcing policy makers to abandon their traditional approach of framing regulation in largely static terms and to begin focusing on issues of dynamic efficiency, with the accompanying emphasis on providing incentives for investing in the deployment of new technologies.

These developments have heightened the importance of understanding how networks function and how regulation affects their behavior. Not only does government policy play a key role in shaping returns and investment incentives; a growing number of commentators have suggested that regulation has also played a decisive role in precipitating much of the turmoil that has wracked the industry of late, having shaped both the recent wave of mergers (Chen 1999) and the WorldCom bankruptcy (*Wall Street Journal* 2002). The direct link between regulation and industry performance makes understanding the economic implications of current regulatory policy all the more imperative.

B. A Fundamental Shift in Regulatory Policy

The need for a more sophisticated understanding of networks has been made all the more critical by a fundamental shift in the basic approach to regulating networks. For over a century, policy makers charged with regulating networks relied almost

exclusively on the set of tools associated with *rate regulation*, in which regulatory authorities used tariffs to dictate the prices that network owners could charge end customers for network services. Because rate regulation targeted final goods that represented the output of the entire system, this approach did not require much of a theory of how networks are configured or how the various network components interact with one another.

In recent years, however, regulators have begun to turn to a new approach, known as *access regulation*. Unlike rate regulation, where regulators focus on the prices that network owners charge for the services of the entire network, access regulation dictates the terms under which network owners must allow customers, partners, other network firms, and even competitors to use portions of their networks. In short, rather than following the approach dictated by rate regulation and controlling the terms under which *consumers* purchase access to *outputs*, access regulation instead controls the terms under which *competitors* may purchase access to *inputs*. This shift in regulatory approach is exemplified by the Telecommunications Act of 1996, which attempted to introduce competition into local telephone service by potentially compelling incumbent local telephone companies to provide competitors with unbundled access to every element of their networks. Access requirements are also being implemented with respect to cable television systems, networks of utility poles, and broadband technologies. Some scholars have even suggested that the shift to access regulation represents a paradigm shift in the approach to regulating network industries (Kearney and Merrill 1998). As one commentator aptly acknowledges, we do indeed live in "the Age of Access" (Rifkin 2000).

The shift to access regulation has changed the primary unit of regulatory analysis from the outputs of the entire network to the services provided by individual network elements. In so doing, it has created the need for a more comprehensive understanding of how network components interact within the context of a complex system, as well as some basis for determining the impact of access regulation on optimal network design. Absent some greater insight into these considerations, regulatory authorities will be hard pressed to shape policy in ways that are both coherent and constructive.

We demonstrate that the complexity of networks implies the need for additional regulatory forbearance. The law of unintended consequences applies with a vengeance to network access regulation. Small regulatory changes that affect network utilization and interconnection can significantly affect network performance and capacity. This raises the bar for network access regulation in comparison with traditional utility regulation.

Put differently, there must be a substantial market failure in telecommunications to justify regulatory intervention through mandatory network access. In addition, regulators must show that their intervention will address the alleged market failure. To satisfy this test, regulators must identify the potential consequences of their intervention. This requires application of the theory of networks to understand the possible effects of access regulation.

C. The Limits of Existing Scholarship

Unfortunately, the existing economic and legal commentary on networks provides few insights into network architecture and design. The literature has been hampered by the absence of a terminology that identifies networks' essential components and captures the manner in which they interact with one another. Even more importantly, the recent upsurge in attention from economists has focused primarily on the phenomenon known as *network economic effects*, which occurs when a network's value is largely a function of the number of other users connected to it.[1] The problem with this approach is that it focuses on only one aspect of networks: their size. As a result, it is unable to provide insights into the relative benefits of different network architectures or the interrelationship among the various network components.

Legal scholarship and the substantive decisions of regulatory authorities have proven similarly unhelpful. They have reflexively adhered to the cost-based approaches associated with traditional rate regulation and have based access prices on the costs of particular network elements, with the primary policy issue centering on whether these calculations should be based on historical or replacement cost. Cost-based approaches violate one of the central precepts of economics by focusing solely on the supply side without providing a way to take demand-side considerations into account (Yoo 2003a). Even more problematically, by considering each component as if it existed in isolation, this approach fails to capture one of the central characteristics of networks, which is how the aggregation of individual network components into an integrated system causes them to interact with one another in complex ways. In the words of one noted network theorist:

> First, real networks represent populations of individual components that are actually *doing something* – generating power, sending data, or even making decisions. Although the structure of the relationships between a network's components is interesting, it is *important* principally because it affects either their individual behavior or the behavior of the system as a whole. Second, networks are dynamic objects not just because things happen in networked systems, but because the networks themselves are evolving and changing in time, driven by the activities or decisions of those very components. In the connected age, therefore, *what happens and how it happens depend on the network*. And the network in turn depends on what has happened previously.
>
> (Watts 2003, p. 28)

As with all complex systems, the costs and benefits associated with any one component cannot be assessed without an appreciation for the function it plays

1 The literature on network externalities is vast. For the seminal articles, see Rohlfs (1974); Farrell and Saloner (1985); and Katz and Shapiro (1985). For a survey appearing in the legal literature, see Lemley and McGowan (1998).

within the system as a whole. Each network is designed so that each element interacts with other elements in ways that analysis of individual elements cannot adequately take into account. Thus, the impact of compelling access to a particular network component can only be understood if one has a theoretical model of the interrelationship of the network's various constituent parts.

D. Network Law and Economics

This book is designed to overcome these conceptual limitations and to place the issues surrounding network economics and policy on a sounder economic and legal foundation. We begin by demystifying networks as an economic phenomenon by offering a more precise definition of what constitutes a network. The terminology we establish not only provides a technical background for our study; it also provides important insights into network design.

In addition, we offer a more general theory of networks that makes it possible to analyze how networks function as integrated systems based on the critical engineering and management processes employed for the design and operation of networks. Our theory applies and extends the insights provided by the mathematics associated with *graph theory*, which has served as the foundation for a substantial scientific literature on the science of networks that has been largely overlooked by economists, legal academics, and policymakers. The rapidly developing science of networks is chronicled in a number of popular works (Watts 1999, 2003; Barabási 2002; Buchanan 2003).

Economists have begun to take into account the complexity of networks. The cost model of Gasmi, Kennett, et al. (2002) reflects the architectural structure of telecommunications networks. Cost allocation in networks is examined by Bird (1976), Sharkey (1995), and Henriet and Moulin (1996). Networks have also been studied in markets for electric power, including the application of Kirchoff's laws; see for example Bohn et al. (1984) and Hogan (1992). For an introduction to the economics of networks, see Economides (1996). We also consider some principles of network design from engineering and operations research.

Our analysis shows how demand interacts with the cost, capacity, geography, and directional flow of the various network components to create a theory of optimal network architecture. Then we present an economic analysis of pricing and efficient choice of network capacity. Applying the insights of graph theory to the economic and legal issues surrounding access regulation captures one of the essential qualities of networks, which is that the properties of the overall network cannot be understood solely in terms of the individual components, any more than the behavior of organisms can be understood solely in terms of individual cells. By providing a basis for analyzing how networks function as integrated systems, our theory makes possible an appreciation for how the entire network can be more

than merely the sum of its constituent parts. The model we advance also provides insights into the complex interrelationships among network components that cause them to interact in ways that are often surprising and unpredictable.[2]

We augment the insights provided by graph theory with the economics of transaction costs, pioneered and developed by Ronald Coase (1937). This framework allows us to analyze two other elements critical to network design and management: (i) a method for understanding *market transactions* for network services and (ii) a method for understanding the *organizational governance* of networks. Transaction cost economics provides a method of integrating these two components in the analysis of network firms. Network firms own and operate networks to provide a wide variety of transmission and transportation services. The organizational structure and boundaries of network firms are determined by the relative costs of market transactions and organizational governance.

Coase explained that firms choose their activities by comparing the costs of engaging in market transactions with the costs of internal resource allocation. For Coase, firms arise as a means of economizing on market transaction costs by internalizing those transactions within the organization. The transaction-cost approach to the economic theory of the firm emphasizes the need for vertical integration as a means of avoiding contracting costs.

Network firms optimize access and network design by trading off the costs of market transactions and organizational governance. Compelled access regulations interfere with these decisions. The result will be departures from the efficient degree of vertical integration and institutional organization in network industries. Over time, there may be dynamic inefficiencies affecting investment in network capacity and research and development of network-related technology.

Economic analysis indicates that efficiency would best be promoted if access to those network services were based on market value. Reliance on market-based pricing mechanisms not only tends to allocate goods to their best use; it also provides the proper signals to parties who consider investing in network technologies. Rather than basing access prices on the market value of the network services provided, regulators continue to employ access-pricing methodologies that focus on the cost of the inputs used to establish the physical network. Such an approach might be appropriate in an ideal, frictionless world in which all welfare-enhancing transactions clear instantly. Over time, competition tends to drive the market prices of outputs and the costs of production together, so that ideally the purchase cost

2 One classic example of this phenomenon is the widespread failure of the power grid across much of the western United States in August 1996. The failure of a single transmission line in western Oregon interacted with a handful of other seemingly similarly minor discrepancies in the system to plunge 7.5 million people into darkness within a matter of minutes (Watts 2003, pp. 21–23). A similar problem occurred in August 2003, when the failure of three high-voltage transmission lines in Ohio cut off power to 50 million people in the northeastern and midwestern United States and parts of Ontario, Canada (see Chapter 5).

of inputs would represent a good approximation to the earning potential and thus the market value of those inputs.

In practice, however, both purchase cost and market value are moving targets. Improvements in production technology, innovations in goods and services, shifts in consumer demand, entry and exit by producers, and changes in factor prices represent exogenous shocks that temporarily prevent the market from reaching general equilibrium, during which time the market value of inputs deviates from their initial cost. The greater the rate of change of technology and other forces, the greater this disparity is likely to be. Given the unpredictability of such changes, the deviations from market value caused by basing access prices on the cost of the inputs used to create the network will tend to lead to gluts or shortages and will eventually induce entrants to over- or underinvest in certain types of network capacity. Furthermore, basing access prices on input costs ignores the fact that the whole is typically greater than the sum of its parts. So long as a firm is efficient and creative, the value of the services it provides is likely to exceed the cost of the inputs it uses.

E. Distinguishing among Different Types of Access to Networks

Applying graph theory and transaction costs to networks provides a basis for a more specific definition of access. *Access* in our framework refers to economic transactions between network firms and their customers, partners, and other network firms. Access is the critical economic mechanism for allocating network services and organizing network industries. Our framework makes it possible to distinguish among five different types of access, each with its own distinct economic and legal implications. (1) Retail access refers to transactions between network firms and retail customers. (2) Wholesale access denotes transactions between network firms and resellers. (3) Interconnection access refers to transactions between network firms for origination, termination, and transit of transmissions. (4) Platform access is the set of transactions between network firms and suppliers of complementary services, such as programming on cable networks or Internet content providers. (5) Unbundled access includes transactions between network firms for leasing network elements such as lines and switches.

Distinguishing among different types of access in this manner provides a clearer understanding of how compelling access can have unintended effects. For example, the manner in which regulators have implemented access regimes has resulted in major distortions in market transactions for access. In addition, compelled access potentially affects the design and planning of networks, resulting in inefficient usage of existing networks and inefficient investment in network facilities over the long term. The combination of these regulatory effects further alters the tradeoff between market transactions and the internalization of particular business activities within the boundaries of the network.

F. Toward a Market-Based Approach to Access Prices

In addition, our analysis offers a powerful critique of regulators' current tendency to implement access regimes by basing access rates on cost. The only plausible justification for basing regulated prices on costs incurred was that the absence of external markets caused by the lack of technological substitutes made it impossible to base rates on market prices. By stimulating direct facilities-based competition, the emergence of platform competition and the shift to access regulation have made market-based pricing both feasible and desirable. We refute arguments advanced by other scholars suggesting that network industries are somehow prone to unique forms of market failure that justify adhering to cost-based pricing. Economies of scale and scope, sunk costs, and network economic effects do not generally cause market prices to deviate from levels that promote efficiency and do not change the basic analysis.

Finally, we examine the constitutional implications of the emergence of access regulation by evaluating the limits that the Takings Clause places on such regulation. Because rate regulation simply restricts the terms and conditions under which parties may contract for finished goods and services, it represents the type of adjustment of economic burdens traditionally subject to the more permissive analysis applied to nonpossessory takings. Access regulation, in contrast, often requires network owners to permit third parties to place equipment on their property. When that is the case, access regulation necessarily falls within the Supreme Court's physical takings jurisprudence, which mandates that the government reimburse property owners for the market value of their property without regard to the economic impact of the regulation or whether the regulation in question furthers important public interests. Therefore, when physical collocation is involved, just compensation for compelled access exactly corresponds to economically efficient prices for compelled access.

Established principles of economics and constitutional law thus require that regulators adopt methodologies that allow voluntary access transactions, or as a second-best alternative, adopt methods that base access rates on market prices. Because access transactions are central to the organization of network industries and the organizational structure of firms in those industries, competition and efficiency require unfettered access transactions.

To summarize, we argue that the fundamental shift in regulatory approach toward compelled access demands an equally fundamental shift in the approach to setting prices. If there is to be public intervention in private access transactions, economic principles still dictate that access prices should be based on the market value of the incremental network services provided by the relevant input. In addition, to the extent that compelled access to a telecommunications network requires that competitors be permitted to place equipment on the network owner's property, access requirements constitute physical takings for which market-based compensation must be paid. Although the unavailability of market-based determinants

once justified basing prices on some measure of cost, the shift in regulatory policy (especially when combined with the emergence of direct, facilities-based competition made possible by technological convergence) has caused the justifications for refusing to set rates on the basis of market prices to fall away.

Although the absence of comparable transactions in external markets historically led regulatory authorities to eschew market-based pricing in favor of cost-based pricing, platform competition and the shift to access regulation have drained this justification of its vitality. The emergence of direct facilities-based competition from alternative telecommunications networks has created market-based benchmarks that can serve as independent bases for setting rates. Contrary to the suggestion of some commentators, distinctive economic features of networks such as sunk costs, economies of scale and scope, and network economic effects do not alter this core conclusion.

Our public policy analysis focuses primarily on telecommunications networks. Although we suspect that the economic and constitutional concepts that we discuss may have implications for other network industries, such as energy and transportation, developing those ideas would require detailed consideration of the technological and regulatory conditions of those industries and would exceed the scope of this book. We therefore withhold any firm policy conclusions with respect to energy and transportation networks, although we draw on examples from the energy and transportation industries to help illustrate the analysis of networks. We believe, however, that the definitions of basic network concepts advanced here have general applicability to all network industries.

PART I

THE ECONOMICS OF NETWORKS