13

THE INTERNET AND THE NETWORK NEUTRALITY DEBATE

IN JUST A few decades, the Internet has grown to become an integral part of our society, economy, and daily lives. In addition to the amazing growth in electronic commerce and innovations unimaginable only two decades ago, the Internet has radically increased entrepreneurship, political discourse, the production and consumption of media, social network formation, and community building, among many other things. The Internet has transformed and continues to transform various information- and communications-dependent systems—our economic, cultural, political, and other social systems. Though such a strong claim of transformation may seem hyperbolic, one need only stop and think for a few moments about how the Internet, like other major infrastructures, affects different aspects of our lives. Consider how the Internet provides and shapes opportunities of individuals, firms, households, and other organizations to interact with each other and participate in various social systems. The scale and scope of possible and actual social interactions alone is staggering. Federal Communications Chairman Julius Genachowski stated the following during a speech he made on December 1, 2010:

> Millions of us depend on the Internet every day: at home, at work, in school—and everywhere in between. The high-speed networks we call broadband are transforming health care, education, and energy usage for the better. It's hard to imagine life today without the Internet—any more than we can imagine life without running water or electricity.

> The Internet has been an unprecedented platform for speech and democratic engagement, and a place where the American spirit of innovation has flourished. We've seen new media tools like Twitter and YouTube used by democratic movements around the world. . . . Internet companies have begun as small start-ups, some of them famously in dorm rooms and garages with little more than a computer and access to the open Internet. Many have become large businesses, providing high-paying, high-tech jobs in communities across our country. It's the American dream at work.[1]

It is easy to take such routine benefits for granted, as is commonplace for infrastructure. In a sense, this reveals just how much the Internet has seeped into our lives and subtly affected our perceptions of the world and expectations for what we can do. But we should not take these social benefits and opportunities for granted. The infrastructure, its management, and the various infrastructure-dependent systems are not fixed, predetermined, or a given to be ignored. It is important to understand and appreciate how the Internet generates social value, how Internet infrastructure supports widespread user participation in an incredibly diverse range of productive activities, and how managing the Internet infrastructure as a commons sustains a spillover-rich environment.

* * * *

There are many policy debates in the Internet context involving infrastructure. For example, debates about the domain name system, spectrum allocation, standards and protocols, broadband build-out, and peer production and distribution systems involve concerns about managing/governing infrastructural resources. Nonetheless, this chapter focuses on the particularly contentious "network neutrality" debate.[2] At the heart of this debate is whether the Internet infrastructure will continue to be managed as a commons. Ultimately, the outcome of this debate may very well determine whether the Internet continues to operate as a mixed infrastructure that supports widespread user production of commercial, public, and social goods, or whether it evolves into a commercial infrastructure optimized for the production and delivery of commercial outputs.

The network neutrality debate has been going on for over a decade and will likely persist for another. Some may believe that the debate is winding down because the Federal Communication Commission has promulgated "Open Internet" rules, and others may claim that the debate has been and, if it persists, will continue to be a sideshow because there are many other, more important developments afoot. I do not share these sentiments. The FCC's rule is likely only the first step on a long path, for at least three reasons:

---

[1] Genachowski 1–2 (2010).

[2] This chapter builds on and borrows from a series of published articles, including Frischmann & van Schewick (2007); Frischmann & Lemley (2007); Frischmann (2005a). In particular, it owes a lot to the article with Barbara van Schewick, our extensive discussions on this topic, and her recent book. Van Schewick (2010).

first, it will be challenged in the courts for years; second, it will be implemented at the agency level and will evolve case by case in regulatory proceedings and, consequently, litigation in courts for years; third, it is substantively incomplete, as I discuss below.[3]

A. Internet Infrastructure and Commons Management through End-to-End Design

The Internet consists of many infrastructure resources. Scholars have delineated two macro-level infrastructure resources. The *physical infrastructure* consists of a wide variety of physical networks interconnected with each other, while the *logical infrastructure* consists of the standards and protocols that facilitate seamless transmission of data across different types of physical networks.[4] Both the physical and the logical infrastructure act as essential inputs into end-user production of an incredibly wide variety of outputs, typically described in terms of applications and content. In contrast with the upstream-downstream/input-output terminology used elsewhere in this book to describe the functional relationships between infrastructure and infrastructure-dependent activities and the goods or services that flow from such activities, Internet scholars use layered models of the Internet to distinguish complementary resources as layers based on the functions each layer performs.[5] The number of layers in particular models varies.[6]

For our purposes, the five-layered model illustrated in table 13.1 is sufficient. The physical and logical infrastructures are the foundational layers on which the Internet environment has been built. I refer to the physical and logical infrastructure together as Internet infrastructure and to the applications, content, and social relationships as downstream outputs. In past work, I did not include the social layer,[7] but I do so here to reinforce the point that social goods, including social networks and social capital, are an incredibly important and socially valuable output from Internet use.

---

[3] Regarding the claim that it is a sideshow, I cannot offer a counterargument without getting into the details of other important developments and a comparative analysis. Such a move is unnecessary. While I disagree with the claim, it is perhaps more appropriate to simply emphasize that whatever the candidates for more important developments might be (e.g., interconnection, peering arrangements, spectrum allocation, the domain name system), they are highly likely to implicate the same set of underlying demand-side issues.

[4] Benkler (2000b).

[5] Farrell & Weiser 90–91 (2004); Werbach 57–64 (2002); Sicker & Mindel (2002).

[6] The basic Open Systems Interconnection Model separates Internet infrastructure into the following seven layers: physical; data; network; transport; session, presentation, application. The application layer is the layer with which the user interacts. OSI Model, OSI Model, http://www.osimodel.org/ (last visited Mar. 10, 2011).

[7] I discussed social goods but not a social layer. Frischmann (2005a). On a social layer, see Zittrain 67 (2008) (describing a "'social layer,' where new behaviors and interactions among people are enabled by the technologies underneath"); Madison (2006) (social software); Crawford 699 n. 12 (2005) (social protocol layer).

TABLE 13.1 Five-Layer Model of the Internet

Layer	Description	Examples
Social	Relations and social ties among users	Social networks, affiliations, groups
Content	Information/data conveyed to end-users	E-mail communication, music, web page
Applications	Programs and functions used by end-users	E-mail program, media player, web browser
Logical Infrastructure	Standards and protocols that facilitate transmission of data across physical networks	TCP/IP, domain name system
Physical Infrastructure	Physical hardware that comprises interconnected networks	Telecommunications, cable and satellite networks, routers and servers, backbone networks

The current Internet infrastructure evolved with the so-called "end-to-end" design principle as its central tenet.[8] To preserve its "robustness and evolvability" and to allow applications to be easily layered on top of it, the broad version of this design principle recommends that the lower layers of the network be as general as possible, while all application-specific functionality should be concentrated in higher layers at end hosts.[9] End-to-end design is implemented in the logical infrastructure through the Internet Protocol (IP), which provides a general technology- and application-independent interface to the lower layers of the network.[10]

As a consequence of this design, the lower layers of the network were blind as to the identity of the use (often referred to as "application-blind"); this prevented infrastructure providers from distinguishing between and affecting execution of the applications and content of network communications.[11] In other words, it precludes discrimination based on the identity of the use.

[8] There are two versions of the end-to-end arguments. VAN SCHEWICK 96–105 (2010). When I refer to "end-to-end," I am referring to the broad version.

[9] VAN SCHEWICK 96–105 (2010); Isenberg 24–31 (1998).

[10] VAN SCHEWICK 116–23 (2010) (describing how the Internet Protocol implements the two versions of end-to-end arguments). *Internet Protocol: DARPA Internet Program Protocol Specification*, IETF RFC 791 (September 1981), http://www.ietf.org/rfc/rfc0791.txt? number=791 (formally describing IP). Notably, the logical infrastructure—protocols, standards, and so on—are managed as commons.

[11] VAN SCHEWICK 101–3 (2010); Lemley & Lessig 931 (2001).

User identity also was largely obscured. Data packets routed through the various networks that comprise the physical infrastructure utilize IP addresses and do not directly reveal user identity.[12] Of course, access networks know the identities of their own customers,[13] but they did not know the identity of everyone else with whom their customers interact online. Networks could discriminate on the basis of IP address itself—for example, by blocking traffic originating from a particular IP address, perhaps because an end node (destination) is under persistent attack by another end node (source) sending spam or malicious viruses.[14] Such blocking is end-to-end compliant when practiced at the "edges" (end hosts), for example, when end-users choose to block unwanted content using their browser or block spam using their e-mail program or when an e-mail service provider blocks spam at the mail server. In many situations, access networks are vertically integrated and also act as edge providers (end-systems), for example where access networks host content, run server-based applications, and provide related services for their customers. A mail server is an end-point, whether owned and managed by an access network such as Comcast or by a customer of the access network. As noted, blocking traffic at the mail server is a routine method for handling spam and other forms of unwanted content that is end-to-end compliant.[15] On the other hand, if an access network blocked or otherwise discriminated in the treatment of traffic not destined for its mail servers or comparable end-points—for example, by blocking peer-to-peer traffic destined for a customer's home computer—then the access network would violate end-to-end principles.

---

[12] "Every end-user's computer that is connected to the Internet is assigned a unique Internet Protocol number ('IP address'), such as 123.456.78.90, that identifies its location (i.e., a particular computer-to-network connection) and serves as the routing address for e-mail, pictures, requests to view a web page, and [any] other data sent across the Internet from other end-users. This IP address routing system is essential to the basic functionality of the Internet, in a similar fashion as mailing addresses and telephone numbers are essential to the functionality of the postal service and telecommunications system." Register.com v. Verio, 356 F.3d 393 (2d Cir. 2004). See *Glossary of Internet Terminology*, Internet Corporation for Assigned Names and Numbers, http://www.icann.org/en/general/glossary.htm (last modified Aug. 13, 2010) ("Internet Protocol Address is the numerical address by which a location in the Internet is identified. Computers on the Internet use IP addresses to route traffic and establish connections among themselves . . .").

[13] Access networks are the subset of infrastructure providers that connect end-users. Some infrastructure providers, such as backbone networks, may interconnect networks but not connect end-users.

[14] Still, user identity was obscure; it was very difficult, if not impossible, to block effectively based on the identity of the user because it is not easy to tie addresses to specific users. For example, in residential networks, IP addresses often are assigned dynamically and shared over time as people log off and on; several users may share an IP address through a network address translator; a server may host a wide variety of different users; and so on.

[15] Comcast filters inbound and outbound spam at the server level and blacklists IP addresses that send excessive amounts of spam. See *What Is Comcast Doing about Spam*, Comcast, http://customer.comcast.com/Pages/FAQViewer.aspx?seoid=What-is-Comcast-doing-about-spam (last visited Feb. 26, 2011). Comcast does not apply its spam filter to other e-mail applications, such as Gmail. This is acceptable because discrimination effectively occurs at the application layer rather than the infrastructure layers. There may be reasons to be concerned about this practice, but this chapter will not address nondiscrimination rules at the higher layers.

End-to-end design sustains an infrastructure commons by insulating end-users from market-driven restrictions on access and use of the infrastructure.[16] If infrastructure providers follow end-to-end principles strictly, they cannot distinguish between end-uses, base access decisions or pricing on how packets may be used, or optimize the infrastructure for a particular class of end-uses. For the most part, infrastructure providers are ignorant of the identity of the end-users and end-uses.[17] At the same time, end-users and end-uses are ignorant of the various networks that transport data packets (with the exception of access networks).[18] In a sense, shared ignorance is built into the infrastructure through widespread compliance with the end-to-end design principle.[19]

Functionally, the end-to-end principle acts as a limitation on the property rights of network owners, in much the same way that common carriage operates as a limitation on the rights of common carriers, fair use operates as a limitation on the rights of copyright owners, and environmental regulations operate as a limit on the rights of various property owners. But the end-to-end principle is not law, and because circumstances have changed, network owners may decide to abandon it.[20]

There is considerable pressure for change. First, there is pressure to replace the existing "dumb," open architecture with an "intelligent," restrictive architecture capable of differentiating and discriminating among end-uses and end-users. Second, there is pressure for network owners to internalize externalities more fully and appropriate the value of the Internet. This pressure comes from many sources, including the Internet's evolution from narrowband to broadband,[21] the rapid increase in users,[22] demand for latency-sensitive and jitter-sensitive applications such as massively multiplayer online role-playing games (MMORGs) and IP telephony,[23] demand for security measures and spam-regulation

---

[16] Lessig 46 (2001b); Frischmann 1007–22 (2005a); Frischmann & Lemley 294–96 (2007).

[17] Ohm 1450–51 (2009) (referring to this as a structural constraint right in privacy, because providers did not have the capacity or technology to determine identity).

[18] Frischmann 27 (2001).

[19] *Id.*

[20] Some have, at least in some contexts. *Comcast Network Management Practices Order*, 23 FCC Rcd 13028, at ¶ 9 (2008), http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-08-183A1.pdf. As with many voluntary commons arrangements, defection by community members may be an attractive private strategy, especially if detection is difficult and punishment or enforcement mechanisms are weak.

[21] *Broadband Access Grows 29 Percent, While Narrowband Declines, According to Nielsen//NetRatings*, Nielsen//Netratings (Jan. 15, 2003); Kerner (2008); Smith 6 (2010) (showing that broadband use overtook narrowband use at the end of 2004, and in May 2010, broadband had been adopted by 66 percent of American adults (increasingly), and narrowband had been adopted by 5 percent of adults (decreasingly)).

[22] Between 2000 and 2010, the total number of global Internet users grew by 444.8 percent. *Top 20 Countries with the Highest Number of Internet Users*, Internet World Stats (last modified Aug. 21, 2010).

[23] Regarding MMORGs, World of Warcraft, for example, has over 12 million subscribers as of Oct. 7, 2010. Press release, Blizzard Entertainment, World of Warcraft Subscriber Base Reaches 12 Million Worldwide (Oct. 17, 2010), http://us.blizzard.com/en-us/company/press/pressreleases.html?101007. World of Warcraft has a 58 percent share in the MMORG market as of 2009. Crossley (2009). Regarding IP telephony, Skype has shown immense

measures implemented at the "core" of the Internet (rather than at the ends), and demand for increased returns on infrastructure investments.[24]

In response to these pressures, technology has become available that enables network owners to "look into" the packets traveling across their networks to determine the application or web page they belong to and affect the transport of packets based on this information.[25] This technology violates the broad end-to-end principle, but again, because end-to-end is just a design principle, there is nothing that forces technology to comply with it.[26] Many other methods for getting around end-to-end design and identifying users or uses have emerged. For example, access providers routinely monitor traffic, relying on pattern recognition and other techniques to identify uses—for example, to identify packets used in e-mail, IP telephony, or peer-to-peer applications.[27] In addition, the Federal Communications Commission (FCC) removed most of the regulations that governed the behavior of providers of broadband networks in the past by classifying the provision of broadband Internet access services over cable or DSL as an "information service."[28]

These developments have given rise to the "network neutrality" debate.[29] The central issues are whether and if so how government regulation should disable the ability of network providers to discriminate among uses or users of the Internet.

---

growth since its inception. By August 10, 2010, the service had 560 million registered users, though only 124 million of them used it on a monthly basis. Waters 21 (2010).

[24] Blumenthal & Clark 71 (2001). Zittrain (2008).

[25] Anderson (2007); Cherry 61 (2005); Cisco Systems, Inc., *Network-Based Application Recognition*, http://www.cisco.com/en/US/products/ps6616/products_ios_protocol_group_home.html (last visited Mar. 4, 2011).

[26] Van Schewick (2010).

[27] These methods include Packet Sniffers/Analyzers (see R. Kayne, *What Is a Packet Sniffer*, WiseGeek, http://www.wisegeek.com/what-is-a-packet-sniffer.htm (last modified Feb. 24, 2011)) and Deep Packet Capture (see Solera Networks, *Network Forensics*, http://www.soleranetworks.com/network-forensics/our-network-forensics-technology/capture (last visited Feb. 26, 2011)).

[28] Appropriate Framework for Broadband Access to the Internet over Wireline Facilities, Report and Order and Notice of Proposed Rulemaking, 20 FCC Rcd. 14853 (2005) (hereinafter Wireline Broadband Access Order). Before, the FCC's decision to classify the provision of broadband Internet access services over cable modems as an "information service" had been upheld by the Supreme Court (see Nat'l Cable & Telecomm. Ass'n v. Brand X Internet Servs., 545 U.S. 967 (2005), aff'g Inquiry Concerning High-Speed Access to the Internet Over Cable and Other Facilities, Internet Over Cable Declaratory Ruling, Appropriate Regulatory Treatment for Broadband Access to the Internet Over Cable Facilities, Declaratory Ruling and Notice of Proposed Rulemaking, 17 FCC Rcd. 4798 (2002) (hereinafter Cable Modem Declaratory Ruling and NPRM)).

[29] Though the debate has roots in an earlier debate about whether to impose open-access requirements on owners of physical networks, our focus is on the network neutrality debate, as reflected in the recent FCC Report and Order. FCC Report and Order, *Preserving the Open Internet; Broadband Industry Practices*, GN Docket No. 09-191, WC Docket No. 07-52, FCC 10-201, http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-10-201A1.pdf (released Dec. 23, 2010).

B. The Network Neutrality Debate

In this section, I first explain why network neutrality is not exactly about neutrality. Because bias of one form or another is inevitable, the real issue is (or ought to be) which type of bias to tolerate. Second, I explain how the debate often is framed in narrow economic terms, and how this framing myopically ignores the most important issues and distorts the debate. Third, I explain how framing the debate in terms of innovation is also too narrow. Innovation is extremely important and a critical source of much social value; however, it is but one of many user activities that generates public and social goods. In short, the current framing of the debate risks missing the forest for the trees. The next section explains how a demand-side approach helps to reframe the debate.

1. NETWORK "NEUTRALITY"

In an influential pair of articles, Tim Wu framed the network neutrality debate by asking whether and how the Internet should be made neutral.[30] Wu and Lawrence Lessig submitted an ex parte letter to the FCC explaining their view that the Commission should embrace network neutrality.[31]

How does the end-to-end design principle relate to network neutrality? Although often conflated, network neutrality is not equivalent to retaining the end-to-end architecture of the Internet.[32] On one hand, the blindness of the network is only one consequence of applying end-to-end principles; end-to-end design is much broader than network neutrality. On the other hand, network neutrality does not necessarily require end-to-end compatible protocols, such as the Internet Protocol. There are alternative means for implementing a nondiscrimination rule.

Implementing a commons via end-to-end network design might appear "neutral" to applications, while shifting to an "intelligent" network design capable of allocating access to the infrastructure based on the identity of the uses (users) appears "non-neutral." Yet end-to-end design is not really neutral. Like commons management in general, it effectively precludes prioritization based on the demands of users or uses. Such prioritization is referred to as differentiated quality of service (QoS)[33] because such an allocation regime would provide quality-of-service guarantees to particular users and uses. End-to-end design effectively disfavors those users or uses that demand such QoS guarantees.

---

[30] Wu (2003); Wu (2004). Wu is credited with coining the term "network neutrality."

[31] Wu & Lessig 3 n. 3 (2003).

[32] Van Schewick 72–73 (2010).

[33] The Internet currently provides best effort data delivery, a simple form of QoS. *Id.* at 7–8. There are different types of QoS, some of which are "more consistent" with end-to-end than others. Lessig 47 (2001b).

One might wonder why an application user might demand guaranteed QoS. The answer is that the Internet infrastructure is congestible and the consequential costs of congestion vary among users and uses. *If and only if there is congestion*, some applications feel the effects more than others, and thus the costs of congestion may be greater for some users than others. Put another way, *user demand for QoS* depends completely on the existence of congestion; users have no need for QoS on an uncongested network.[34] Thus, the shape of user demand will depend on the degree, distribution, and persistence of congestion and the demand for various user activities. In sharp contrast, *provider demand for QoS* does not depend on congestion, because the possibility of extracting greater surplus from consumers via price discrimination makes QoS attractive, even when there is no congestion.

User demand for QoS is a common phenomenon on congestible infrastructure, and it is not unique to or necessarily more severe on the Internet. To take a non-Internet example, consider how congestion might impact drivers on the highway differently—compare the person going to the hospital or an important time-sensitive business meeting with the person going to a friend's house or out to eat. Delays associated with congestion impact these drivers differently. Yet, as discussed in chapter 9, guaranteed QoS/priority is not for sale on most roads.[35] On the Internet, certain applications feel the effects of congestion more intensely than others, and, knowing this, certain application-providers and users might prefer to purchase guaranteed prioritization through QoS. For example, latency- or jitter-sensitive applications, such as World of Warcraft and Skype, might suffer in quality because congestion causes delays in the routing of packets, variations in such delays, or packet loss, and these congestion-related effects can lead customers to value their experience less. (By contrast, a user checking e-mail, surfing the Internet, or using AOL Instant Messenger is less inconvenienced by congestion.) Latency- or jitter-sensitive applications would fare better during congestion on an infrastructure that prioritized access and use according to the identity of the application. Thus, despite the appeal to neutrality, critics of network neutrality are quick to point out (correctly) that

[34] One exception is that some users may want to purchase priority for strategic reasons, for example, to crowd out others and/or obtain a competitive advantage.

[35] Of course, in many commercial contexts, priority arrangements are available. Consider, for example, shelf space at the supermarket. Some economists place great weight on the efficiencies of such arrangements in those contexts and attack restrictions on the freedom to do the same in the Internet infrastructure context. But it is important to be clear that those contexts do not involve basic infrastructure at all. Shelf space at the grocery store is rivalrous, and it solely involves the distribution of private goods to consumers (rather than the user-generated public or social goods). Accordingly, for shelf space, the three economic principles discussed in chapter 3 apply, and it seems reasonable to expect that priority arrangements, like many other vertical arrangements among private market actors, would be efficient; the main concern is whether there are significant anti-competitive effects. There is much more at stake in the Internet context. As the next section explains, it would be a major mistake to conflate the Internet infrastructure context with commercial contexts such as shelf space at the grocery store.

nondiscrimination rules are not neutral because they effectively discriminate against users and uses that would benefit under a regime that permits discrimination.

Yet such applications would fare as well on an uncongested infrastructure that did not provide prioritization. The critics assume a congested infrastructure and are most concerned with efficiently allocating scarce infrastructure capacity to its highest valued use, measured, not surprisingly, by users' willingness to pay. The appropriateness of the assumption and the magnitude of congestion and related effects remain quite contentious.[36]

End-to-end design favors one set of applications; shifting to differentiated QoS favors another set. This point escapes critics of network neutrality and remains unappreciated in the debate. But it is worth emphasis: Just as the current end-to-end design favors certain applications at the expense of time-sensitive applications (assuming, for purposes of argument, a congested infrastructure), shifting to a fine-grained QoS regime also would exhibit a systematic bias for particular applications—specifically, for commercial applications that generate observable and appropriable returns. The bias would not be technologically determined, as in the case of end-to-end design; rather, it would be determined by the conventional operation of the market mechanism. Given the ability to discriminate among end-users and end-uses on a packet-by-packet basis and the inability to perfectly price discriminate, infrastructure suppliers rationally may prioritize access and use of the infrastructure via imperfect price or quality discrimination and/or optimize infrastructure design in favor of output markets that generate the highest levels of appropriable returns (producer surplus), at the expense of output markets that generate a larger aggregate surplus (direct consumer surplus, producer surplus, and external surplus).

The bottom line is that one way or another, some bias in the system is inevitable. When confronted with this argument, critics shift gears and suggest that no such bias would occur because all users would obtain access and use. Such a move depends on an assumption either that discrimination will increase output dramatically or that congestion no longer matters. Neither assumption makes sense. I explain these points in further detail below. To overemphasize neutrality or non-neutrality merely distracts from the more important issue, which is how best to manage the infrastructure, given predictable biases.

2. THE ROLE OF ANTITRUST AND REGULATORY ECONOMICS

Network neutrality involves complex interdisciplinary issues, and economics plays a critical role in framing the debate. The central issue is whether government regulation should

[36] The congestion-related effects are *not* a consequence of interaction effects or incompatibilities among uses. See chapter 7. (Perhaps the easiest way to recall how such effects arise is to imagine the fighting buffaloes and goats hypo.)

disable the ability of network providers to discriminate among uses or users. This issue triggers familiar economic considerations, discussed extensively in earlier chapters.

Unfortunately, here as in other infrastructure contexts, the case for regulation imposing nondiscrimination rules tends to be evaluated exclusively in terms of antitrust and regulatory economics.[37] As discussed in chapter 5, the antitrust and regulatory economics traditions largely disfavor government intervention into private infrastructure markets for a variety of reasons, most of which permeate the network neutrality debate. Recall that antitrust and regulatory economics adopt the premises that intervention is needed only when markets are not competitive and that even when markets are not competitive, intervention is justified only in very narrow circumstances where demonstrable harm to consumers in the relevant markets can be shown and not outweighed by efficiency gains.

The focus on competition and demonstrable harm to consumers is completely misguided. It distorts the debate dramatically and distracts participants from the more important, fundamental question, which is what type of Internet environment our society demands. Before I attempt to reframe the debate around this question, I highlight three related ways in which the debate is distorted: (1) the supply-chain view of the Internet, (2) the false supplier/consumer dichotomy, and (3) the competition red herring.

a. The Supply-Chain View of the Internet

First, the antitrust and regulatory economics framework views the Internet as a mere supply chain: network access → applications and content → consumers. More "sophisticated" versions of the supply chain view incorporate so-called "two-sided" markets, where economists put networks in between application/content providers and consumers (essentially, applications and content → network access → consumers) and consider how networks will be able to efficiently mediate transactions between each side of the market.[38] But either of these frameworks leaves the partial-equilibrium blinders on and

---

[37] Barbara van Schewick suggested to me that I go too far in this section because proponents of network neutrality regulation have incorporated arguments from outside of antitrust and regulatory economics and fought against the distortions I note in the text. While I support these efforts and have participated in some of them, for example by submitting a comment in the FCC's Open Internet proceeding, I am not convinced that the network neutrality debate has fully internalized the arguments or broken out of the antitrust and regulatory economics framework. I would love to be persuaded otherwise. I do appreciate that in its Report and Order, the FCC expressed similar concerns and an awareness of the broader set of issues that are not well accounted for by the antitrust and regulatory economics framework. Regardless, the point of this section remains relevant and important because the antitrust and regulatory economics framework still has considerable influence.

[38] Weisman & Kulick 87 (2010) (considering auctions, credit cards, dating bars, newspapers, video game consoles, and the Yellow Pages as examples; defining a two-sided market as any market that brings together two distinct user groups that both benefit from the presence of the other); Yoo 96–97 (2010) ("[B]roadcast television . . . brings together two groups: viewers and advertisers."); Ratliff & Rubinfeld 659–60 n. 34 (2010) ("The sale of

reduces an incredibly complex, open system to an inappropriately simple closed system for purposes of analysis, evaluation, and policy making.[39]

The simple supply-chain view assumes complete markets and ignores incomplete and missing markets despite their prevalence in the actual Internet environment. As discussed below, the incomplete and missing markets often involve user-generated public and social goods. The slightly more sophisticated, two-sided market view may incorporate some incomplete markets, but only where the relevant externalities are within the closed system of markets being considered—that is, when the external effect are effects on one side or the other of the mediated market. This is convenient because it allows modelers to examine whether the intermediary will set prices efficiently, but it still assumes away externalities associated with the production, sharing, use, and reuse of public and social goods that are felt outside the closed system. It also fails to appreciate the complexity associated with the fact that the Internet involves many, interdependent many-sided markets and nonmarkets.

Let me make the point using a series of simple illustrations in which CM = complete market; IM = incomplete market; MM = missing market; and NI = network intermediary. (Alternatively, we could illustrate the point by using private, public, social, and network goods rather than markets and network intermediaries.) First, consider the Internet reduced to a two-sided market/closed system:[40]

(1) $CM_1 \longleftrightarrow \mathbf{NI} \longleftrightarrow CM_2$

(2) $IM_1 \longleftrightarrow \mathbf{NI} \longleftrightarrow IM_2$

In (2), incompleteness in the markets constitutes externalities flowing from one side to the other (from IM_1 to IM_2, or vice versa) and capable of being internalized if the intermediary sets differential prices to account for the externalities.

Second, consider a series of slight transformations that move incrementally toward reality but drastically away from the closed-system view in (1) and (2). Consider what happens when we make the following transformations.

(3) $IM_1, CM_1 \longleftrightarrow NI \longleftrightarrow CM_2, IM_2, MM_1$

In (3), the missing market constitutes externalities that flow outside the system (or off-network) and thus do not flow to any of the other markets within the system.

---

advertising to businesses and the display of advertisements to consumers take place in a *two-sided market* at the hub of which sits the content publisher . . . ").

[39] On the problems with these blinders, see Lipsey & Lancaster (1956); Lunney (2008).

[40] I describe this as a supply chain and thus multiple markets with the intermediary being a participant in the various markets, for example, as a service provider. The two-sided market literature sometimes focuses narrowly on a single market with two or more sides.

(4) MM1, IM1, CM1 ←→ **NI** ←→ CM2, IM2, MM2

(5) MM1, MM2, IM1, IM2, CM1, CM2 ←→ **NI** ←→ CM3, CM4, IM3, IM4, MM3, MM4

(6) MM1, IM1, CM1 ←→ **NI1** ←→ CM3, **NI2** ←→ CM2, IM2, MM2

In (6), a second network intermediary connects various user populations.

(7) And so on.

Although we are still not close to reality, it should be clear how the models in (1) and (2) distort dramatically by omission.

Keep in mind that the externalities need not flow to the other side of the two-sided or multisided market; the externalities may flow off the network altogether; the network intermediary is incapable of accounting for, much less efficiently pricing, various types of externalities. The NI certainly has incentives to set prices *as if* the Internet fit model (1) or (2), and it is precisely these distorted (and distorting) incentives that jeopardize the Internet that exists in reality.

b. The False Supplier/Consumer Dichotomy

Second, the framework leads to misconceptions of the actors involved. Specifically, in accordance with the supply-chain view of the Internet, the focus is how network neutrality would affect (a) network providers, (b) application and content providers, and (c) consumers.[41] This framing conflates too much, by creating a false distinction between application and content providers and other end-users and understating the role of consumers as producers. Users produce a wide range of private, public, and social goods, including various applications and content. Google is an end-user, like you, me, or anyone else. We are different in some ways—for example, in the amount of traffic or revenue we generate—but not necessarily in a way that matters (or should matter) to the network neutrality debate. I, you, or anyone else could be the next Google or perhaps the Google of our own "space" on the Internet. Who knows where the next killer app, idea, or YouTube video will come from?

The key point is that viewing the Internet as a means for distributing content, applications, and services to consumers biases the debate in a way that misses the forest for the trees. (As the next section discusses, it reduces the Internet to a commercial infrastructure, although the Internet is a mixed infrastructure.) The social value produced by users as well as the basic capabilities/opportunities for users to be productive are

[41] Many commentators recognize the flaws in this point of view. See, e.g., Mehra (2011); Benkler (2006b). In its Report and Order, the FCC seems to as well. FCC R&O paragraphs 2, 20.

marginalized if not outright ignored. The Internet is much more than a low-cost delivery system for application or content providers or a two-sided market with networks acting as intermediaries between providers and consumers. While the supply-chain view of the relationships among different actors and the corresponding categorization of actors yield a familiar conceptual map, tractable models, and (perhaps) measurable data points, such comforts do not excuse the distortions.

c. The Competition Red Herring

Finally, the debate has been fixated on competition or the lack thereof.[42] As Jonathan Nuechterlein explains:

> most advocacy for net neutrality regulation argues that there is inadequate competition in the market for broadband Internet access and that the government should step in to prevent abuses of the resulting market power. If each American consumer had a choice of ten broadband Internet access providers, *there would be no credible basis for such intervention, because competition would ensure each provider's responsiveness to consumer choice*.[43]

I suspect that many (though not all) proponents of network neutrality would probably concede that regulation would not be necessary if Internet access markets were competitive. The debate has revolved around whether network owners have market power and whether discrimination among data packets causes anticompetitive effects or demonstrable harm to consumers.[44] Proponents of network neutrality regulation claim that network owners have the market power, capability, and incentives to engage in harmful discrimination; opponents claim that the relevant markets are competitive and that any discrimination practiced by network owners is presumably beneficial. To be fair, participants on both sides of the debate appeal to other considerations.[45] For example, an important strand of the debate focuses on innovation, which I discuss below. Another important strand of the debate focuses on the question of what form of public regulation might be appropriate. Does antitrust provide a sufficient check on discrimination, or should the FCC create additional sector-specific regulations?[46] If more regulations are

---

[42] Barbara Van Schewick criticized me for saying this once at a conference. As I mentioned, she told me that the debate has moved on. For her rejection of the idea that competition would make network neutrality regulation unnecessary, see VAN SCHEWICK 255–64 (2010).

[43] Nuechterlein 19, 34 (2009) (emphasis added).

[44] The FCC rejected the idea that only anticompetitive discrimination should be prohibited, FCC R&O paragraphs 77–78, and instead prohibited unreasonable discrimination. I discuss this briefly below.

[45] A survey of the voluminous literature or comments filed in the FCC Proceeding reveals an extremely wide range of arguments.

[46] See, e.g., Nuechterlein (2009).

needed, what form should the regulations take (e.g., a rule versus a standard)? Nonetheless, these strands are heavily influenced by the more basic competition policy framework and its premises. For example, the one "intervention" accepted by both sides of the debate—transparency—suffers from this myopia and its attendant problems. Transparency is perceived as an adequate solution because it would enable consumers to be effective market participants, capable of disciplining access providers by making a fuss about harmful discrimination and switching providers.[47] There are at least three problems with this "solution." First, it presumes that consumers have the time, inclination, and capability to process and react to the information made available. Second, it presumes competition. Third, and perhaps most important, it presumes complete markets in which consumer demand effectively manifests societal demand. Simply put, transparency is important, but transparency alone would not be enough to preserve the social benefits of an open Internet.[48]

Framed as a "lack of competition" problem, the debate ignores demand-side issues by assuming that private demand reflected in markets fully reflects social demand (as the quote above from Nuechterlein suggests). As a result, it fails to appreciate that the social value of the Internet greatly exceeds its market value and that relying on unfettered market allocation of infrastructure access and use runs the risk of demand-side market failures of the sort described in chapters 4 and 5.

Competition alone does not alleviate the demand-side concerns discussed at length in chapter 5.[49] Competition does not ensure an efficient allocation of resources. It does not assure us an Internet environment that maximizes social welfare. Competition does not address these interests for the same reasons that antitrust law is orthogonal to environmental law—antitrust law does not address market failures associated with externalities, whether environmental pollution (negative externalities) or the production, sharing, and productive reuse of public and social goods (positive externalities). Indeed, it is well established in economics that competitive markets overproduce pollution and underproduce public and social goods. Moreover, conventional economic solutions to the underproduction of public goods, such as directing subsidies to public goods producers, do not work well in this context because of the incredible variety of producers and of public goods, and governments' predictable failure in choosing how to direct subsidies in

[47] There are other reasons transparency is desirable. FCC R&O paragraphs 53–60.

[48] The FCC also concluded that transparency alone is not sufficient. See *id.* paragraph 61.

[49] First, it does not alleviate concerns about undersupply and underuse of infrastructure, and undersupply of infrastructure-dependent public and social goods. Second, it does not alleviate concerns about dynamic shifts in the nature of infrastructure resources—that infrastructure development may be skewed in socially undesirable directions. This could happen, for example, if private infrastructure owners prematurely optimize infrastructure for uses that they expect will maximize their *private* returns and, in doing so, choose a path that forecloses production of various public or social goods that would yield greater net *social* returns.

this context. Further, the conventional economic solutions do not work well for social goods that are jointly produced in a more distributed fashion.[50]

Even if we assume robust competition,[51] the case for network neutrality remains strong.[52] As the next section discusses in more detail, the Internet infrastructure is a mixed infrastructure that when managed as a commons supports a spillover-rich environment. A tremendous amount of the social value derives from activities by and among users associated with producing, sharing, and reusing ideas and other intellectual and social goods. Spillover effects associated with speech, information flows, social interactions, and other related activities are rampant and arguably should define the relevant policy space.

This does not mean that market power and anticompetitive effects do not matter or do not provide justification for network neutrality regulation, nor does it mean that recognizing the existence and importance of spillovers associated with user-generated public and social goods provides easy, determinate answers. The network neutrality debate is and must be complicated; it should not be reduced to a competition policy framework. It needs to grapple with the demand-side issues, including the social value associated with spillovers from a wide variety of user-generated public and social goods, and the role of commons management in sustaining the social option value of the Internet.

3. INNOVATION

For many, the network neutrality debate is fundamentally about innovation policy.[53] There are very good reasons to focus on innovation. First, innovation itself is crucial because it fuels technological advancement, economic growth, human development, and social progress. Second, the Internet infrastructure has undoubtedly facilitated an incredible amount of innovation, contributed significantly to economic growth, and distributed the capability to innovate widely among Internet users.

---

[50] See chapters 3, 5. On social production, see BENKLER (2006b).

[51] Admittedly, this is a heroic assumption. Still, I will not address the debate over the existence and robustness of competition. For a discussion, see VAN SCHEWICK 251–64 (2010).

[52] I thus strongly disagree with Nuechterlein 42 (2009):

> Proposals for net neutrality rules could have merit only if (i) the broadband Internet access market is inadequately competitive and will remain so indefinitely; (ii) such market concentration will give incumbent broadband providers both the incentive and the ability to discriminate against specific applications providers; (iii) such discrimination would harm *consumers* and not just particular *providers*; and (iv) any such consumer harm would exceed the costs of regulatory intervention. In short, the net neutrality debate, properly conceived, is fundamentally about core antitrust concepts: about market power, market failures, market definition, and the costs and benefits of government intervention in a rapidly evolving, high-technology market.

[53] For an excellent discussion of the relationship between innovation and network neutrality, see VAN SCHEWICK 270–73, 289–92 (2010).

Proponents of network neutrality emphasize, among other things, how an open Internet has supported an incredibly productive innovation system; opponents emphasize, among other things, how government regulation threatens to limit innovation by networks and foreclose opportunities to develop new technologies and services that might thrive in an environment that permitted prioritization and optimization of networks for specialized uses. As Wu has noted, both sides "idoliz[e] innovation."[54] He characterizes the debate in terms of a battle over competing views of what best promotes innovation: an open system that sustains decentralized innovation and prevents centralized control (whether by governments or private infrastructure owners), or an unregulated system free of government interference and subject only to market forces.[55]

Lessig, a major proponent of sustaining the end-to-end design, focuses extensively on the notion of an *innovation commons* and the idea that experimenting and tinkering with—and creating without inhibition—new applications and content are critical productive activities facilitated by the end-to-end architecture of the Internet.[56] In an earlier article, I suggested that Lessig should not focus exclusively on innovation because the term generally referred to commercial innovation;[57] he replied that our disagreement was semantic and made clear that he viewed innovation much more broadly than commercial innovation.[58] I appreciate the broader perspective on innovation and agree that many noncommercial entities innovate and that there are many examples of socially valuable noncommercial innovations. Still, I think that a focus on innovation is not enough.

Innovation is an integral part of the debate, but it ought not be the linchpin on which the end-to-end architecture of the Internet hangs. Innovation—commercial and otherwise—is a vitally important activity furthered by end-to-end architecture. The open Internet greatly enhances users' innovative capability, the capability to experiment, tinker, and put new ideas into practice without the approval of networks acting as gatekeepers.[59] Many users exercise that capability and thereby generate substantial social value. According to Christiaan Hogendorn, "[t]he most dramatic source of spillovers on the Internet is innovation."[60] As one of the critical drivers of economic growth, innovation is an important economic activity to encourage.

Yet innovation entails some degree of novelty, and many of the most valuable Internet applications are not (or are no longer) new, improved, or very innovative—for example, e-mail or instant messaging.[61] While end-to-end architecture undoubtedly promotes

---

[54] Wu 80–84 (2004).

[55] *Id.* at 80–84; Wu 152–54 (2003); Wu & Lessig 5–7 (2003).

[56] Lessig (2001b); Lessig (2005).

[57] Frischmann (2005a).

[58] Lessig (2005).

[59] Lessig (2004b). See also Kelty (2008) (making similar points with respect to open-source software).

[60] Hogendorn 14 (2010).

[61] Of course, there may be plenty of room for further innovation in these communications technologies.

innovation, bringing us valuable new applications and opportunities,[62] it also promotes the continued use of socially valuable, but perhaps less innovative, applications. To put it another way, we should acknowledge the limits of innovation as a concept because the capability to innovate is but one of the important Internet-enabled user capabilities sustained by a nondiscriminatory infrastructure.

C. Reframing the Debate

This section first discusses the nature of the Internet as mixed infrastructure and the significant role of users in generating social value. It then considers whether commons management is an attractive public strategy.

1. THE INTERNET AS MIXED INFRASTRUCTURE

The Internet satisfies all three demand-side criteria for infrastructure. The Internet infrastructure is a partially (non)rival resource; it is consumed both nonrivalrously and rivalrously, depending on available capacity. The physical infrastructure and certain components of the logical infrastructure, such as domain name space, are partially (non)rival in the sense that they are congestible but not necessarily congested; the risk of congestion depends on the amount of capacity, number of users, and other contextual factors.[63] As the next section discusses, congestion can be managed in a nondiscriminatory manner that sustains the commons and leverages nonrivalry.

The benefits of the Internet are generated at the ends. Like a road system, a telecommunications network, an ocean, and basic research, the Internet is socially valuable primarily because of the wide variety of productive activities it facilitates. End-users generate value and realize benefits through their activities, which involve running applications on their computers; generating, consuming, and using content; and creating and engaging in various social, economic, or other relations with other users. End-users create demand for Internet infrastructure through their demand for applications, content, and relations. Keep in mind that *activities on the Internet* always involve *interactions* among *end-users*; that the interactions may be commercial, educational, social, political, and so on; and that end-users may be individuals, corporations, government actors, or other entities.

Currently the Internet is a mixed commercial, public, and social infrastructure. The Internet is perhaps the clearest example of an infrastructure resource that enables the production of a wide variety of private, public, and social goods. The supply-chain view

[62] Van Schewick (2010).

[63] Moreover, just like other networked infrastructure systems (e.g., transportation), there are many subsystems and components that can be congested at different times. See chapter 9.

of the Internet captures the commercial nature of the Internet: Applications and content providers use the Internet as a means of engaging in a host of commercial transactions with passive consumers. Many Internet-enabled commercial transactions involve the distribution of applications and/or content for personal consumption under (technical and legal) conditions that strictly limit sharing or reuse; thus, although the applications/content are public goods in the technical sense, the transaction effectively concerns passive consumption of a private good.[64] The Internet also acts as an input for information dissemination and exchange for commercial advertising and marketing and facilitates business transactions and information gathering for product development, consumer demand assessment, and operations management.[65] Many Internet-enabled commercial transactions involve the exchange of private goods offline. In these cases, the Internet may reduce the transaction costs or increase the scale of an already existing offline market. Without doubt, the value of the Internet as commercial infrastructure is immense. Here are a few rough estimates:

- According to John Quelch, in the United States the Internet directly and indirectly has led to 3.05 million jobs. The direct monetary value (including advertising services, retail transactions, and payments to Internet service providers [ISPs]) is estimated at $175 billion. Around 190 million Americans spend an average of sixty-eight hours per month online, which means that, by a conservative estimate, the time value of the Internet is around $680 billion per month.[66]
- According to the US Census Bureau, total US commerce in 2008 (the latest year reported on) was about $22 trillion; about $3.7 trillion was in the form of e-commerce, mostly over the Internet, and about 92 percent of e-commerce was business-to-business.[67]
- Globally, the OECD estimates that of the top 250 ICT firms in terms of revenue, Internet firms accounted for $18.3 billion in revenue in 2000, growing to $56 billion in revenue in 2006, with employment growing from 47,539 to 93,380 over that time period.[68]

---

[64] Frischmann 1015 n. 383 (2005a).

[65] Litan & Rivlin 4–5, 19–38 (2001). These processes are often tailored to channeling end-users toward purchasing and consuming commercial products. Balkin 14 (2004).

[66] Quelch (2009).

[67] Bradner (2011), citing U.S. Census Bureau, Measuring the Electronic Economy, 2008 E-Commerce multisector "E-Stats" Report (released May 27, 2010).

[68] *Id.* at 51 (citations removed). According to the Boston Consulting Group, in 2009 the Internet contributed £100, or 7.2 percent of GDP, to the UK economy. Kalapesi et al. 5, 33 (2010). This contribution included online consumption (including cost of access), investment (in telecom companies, hardware/software by corporations), government spending on information and communications technology, and net exports.

Yet the value of the Internet as public and social infrastructure dwarfs its value as commercial infrastructure. I recognize that it is extremely difficult to substantiate such a claim with empirical data that purports to measure value. As many previous chapters explained, it is precisely this difficulty that leads us to take the social value for granted. The public and social aspects of the Internet infrastructure are dramatically undervalued in the current debate. Bringing these aspects of the Internet into focus strengthens the case for managing the Internet infrastructure as a commons.

Consider what makes the Internet valuable to *society*. The Internet's value to society is tied to the range of capabilities it provides for individuals, firms, households, and other organizations to interact with each other and to participate in various activities and social systems. It is very difficult to estimate the full social value of the Internet, in large part because of the wide variety of user activities and interactions that generate public and social goods. Despite this difficulty, we know that the Internet is transforming our society.[69] The transformation is similar to transformations experienced in the past with other infrastructure, but things are changing in a more rapid, widespread, and dramatic fashion.

The Internet is integral to the lives, affairs, and relationships of individuals, companies, universities, organizations, and governments worldwide. It is having significant effects on fundamental social processes and resource systems that generate value for society. Commerce, community, culture, education, government, health, politics, and science are all information- and communications-intensive systems that the Internet is transforming. The transformation is taking place at the ends, where people are empowered to participate and are engaged in socially valuable, productive activities. As Jack Balkin observed, the "digital revolution makes possible widespread cultural participation and interaction that previously could not have existed on the same scale."[70]

The Internet opens the door widely for users[71] to become involved in many different productive activities. Users actively engage in innovation and creation; create new markets and disrupt old ones; create and inhabit new virtual worlds; engage in commercial and noncommercial exchange; speak about anything and everything; maintain family connections and friendships; debate, comment, and engage in political and nonpolitical discourse; organize clubs and protests; meet new people; search, research, learn, and educate; and build and sustain communities. (This paragraph alone could be expanded to a book-length discussion.)

---

[69] National Broadband Plan–Executive Summary, http://www.broadband.gov/plan/executive-summary/; President's Info. Tech. Advisory Comm. 11–20 (1999).

[70] Balkin (2004).

[71] Lest one forget, "users" refers to individuals, corporations, government actors, or other entities using the Internet infrastructure to interact, whether Microsoft or the reader.

These are the types of productive activities that generate substantial social value—value that evades observation or consideration within conventional economic transactions. When engaged in these activities, end-users are not passively consuming content delivered to them or producing content solely for controlled distribution on a pay-to-consume basis. Instead, end-users are interacting with each other to build, develop, produce, and distribute public and social goods. Public participation in such activities not only benefits the participants directly (as indicated by their decision to participate), but also results in external benefits that accrue to society as a whole, *both online and offline.* These benefits are not fully captured, or necessarily even appreciated, by the participants.

Active participation in these activities by some portion of society benefits even those who do not participate. In other words, the social benefits of Internet-based innovation, creativity, cultural production, education, political discourse, and so on are not confined to the Internet; they often spill over. For example, when bloggers engage in a heated discussion about the merits of proposed legislation or the Iraq war, citizens who never use the Internet may benefit because others have deliberated. With respect to weblogs, in particular, political scientists, journalists, economists, and lawyers, among others, have begun to appreciate and more carefully study the dynamic relationships between this new medium of communication and traditional, offline modes of communication and social interaction (whether economic, political, social, or otherwise).

Consider the fact that a significant portion of the content traveling on the Internet is noncommercial, speech-oriented information—whether personal e-mails and web pages, blog or Twitter postings, instant messaging, or government documentation—and the economic fact that such information is a pure public good generally available for both consumption and productive use by recipients.[72] The productive use and reuse of such information creates benefits for the user, the downstream recipients, and even (as discussed above) people offline. These benefits are positive externalities that are not fully appropriated or even appreciated by the initial output producer. The magnitude of the external effects may be quite small on average. Diffusion of small-scale positive externalities, however, can lead to a significant social surplus when the externality-producing activity is widespread, as it is on the Internet.[73] The "killer app phenomenon" is also

[72] As is generally the case with speech-related activities, there may be external costs to consider as well (for example, hate speech); the same is true of innovation (for example, malicious software applications).

[73] Widespread, interactive participation in the creation, molding, distribution, and preservation of culture, in its many different forms and contexts, is an ideal worth pursuing from an economic perspective because of the aggregate social welfare gains that accrue to society when its members are actively and productively engaged. This seems to reflect, in economic terms, the basic idea underlying Jack Balkin's democratic culture theory. Balkin (2004); Balkin (2009) This view also complements the arguments, persuasively made by Yochai Benkler, concerning the social value of diversity in both the types and sources of content. Benkler (2001a).

prevalent on the Internet. The history of the Internet is riddled with examples of killer apps emerging from the "backwaters."[74] I discuss a few below.

To illustrate both the small-scale and killer app phenomena, consider YouTube, a video-sharing platform made possible by the Internet. Most videos on YouTube are intended for a small audience, but many probably reach a slightly larger audience. For example, suppose I post a video of my kids singing and dancing. I might send a link to family and friends, expecting the video to be viewed by twenty-five people. Perhaps one hundred people actually visit, watch, and enjoy. This would generate small-scale spillovers to seventy-five extra people. While YouTube may capture some of the benefits through advertising revenue, the benefits are irrelevant to my own decision to post the video. The benefits are incidental and seem like small potatoes, but they add up considerably when millions of people participate in this activity.[75] As you may know, every once in a while (actually, each week), user-generated and posted videos attract millions of viewers and becomes a cultural phenomenon.[76] For example, in May 2007, Howard Davies-Carr posted a video of his two boys, Harry and Charlie, titled *Charlie Bit My Finger—Again!* Apparently, the video "was intended for just one person: the boys' godfather, who lives in the United States."[77] Unexpectedly, the video went "viral" and has now been viewed over 385 million times. That video generated enormous spillovers, social value above and beyond what the poster possibly could have anticipated or is capable of appropriating.[78] Neither the government nor the market would have selected the video ex ante and funded the producer. The bottom line is that an open video-sharing platform

[74] Zittrain (2008). Everything from e-mail (Hotmail), web browsers (Firefox), search engines (Google), online auction sites (eBay), matchmaking sites (Match.com), and Instant Messenger (AOL IM) constitute "killer apps" and significantly influence how our society interacts and collaborates online. Dasgupta (2002).

[75] YouTube, like many other sites, also permits users to post comments and interact with each other, and this also may enhance existing social relationships or even create new social relationships. The new relationships may not always be beneficial or the social ties may be incredibly weak; still, the opportunity to socialize is made available and easy, and users actively participate. Related to this is the growth of comment fields in blogs and on news sites, allowing an easy way for the average user to create content, even if it is one sentence.

[76] There are plenty of examples. YouTube Charts, http://www.youtube.com/charts/videos_views?t=a. You can browse videos according to various criteria, such as (a) the most viewed, most discussed, or most favorite on a (b) daily, weekly, monthly, or all-time basis, and (c) by category. It would be interesting to study how these rankings change over time, and to know, for example, whether the ratio of user-generated to commercially produced content has changed.

[77] Chittenden (2009).

[78] YouTube captures some additional advertising revenue but that does not undermine the point about substantial spillovers. The advertising revenue cannot possibly fully internalize the externalities. Nor does the fact that the poster also gains financially undermine the point about substantial spillovers. Apparently, "Howard Davies-Carr . . . managed to parlay [the] video of one of his sons biting the other's finger into a lucrative advertising partnership with YouTube. . . .With that kind of attention, spoofs and the ad revenue followed. Davies-Carr won't say exactly how much his family has earned from the video, but the added income made it more possible to afford a new house, he says." Ransom (2010). Both YouTube and the poster capture some of the surplus, but a substantial amount is not appropriated by either.

provides users with a basic capability and keeps social options open; when users choose to exercise the capability, they presumably do so because it sufficiently satisfies their own self-interest, and at the same time they incidentally generate spillovers because they have shared a public good.[79]

These types of spillover-producing activities are *not* limited in any way to YouTube. In fact, there are countless examples of public-good-sharing platforms enabled by the Internet that generate social value in this fashion. Some are proprietary platforms managed by a particular firm; many are not. Some are organized around a particular type of media; others are organized around a particular topic; others are organized around a particular community; others are simply unorganized. Like YouTube, many platforms facilitate more than just sharing. For example, SlashDot.org, Reddit.com, and Digg.com allow users to contribute links to articles, pictures, and a variety of other content, and once the material is submitted, other users vote the contributions up or down, and can comment on each contribution independently of the linked website. User-generated rating systems are a common feature for many commercial and noncommercial websites (including Amazon, Newegg, Barnes & Noble, and CNN, to name just a few).[80]

In fact, there are many infrastructural platforms for producing and sharing public goods. Wikipedia, for example, provides an incredible example of an infrastructural platform for creating and sharing cultural content. It is a free online encyclopedia that rivals proprietary encyclopedias such as Encyclopedia Britannica in quality[81] and vastly exceeds its rivals in public accessibility. The Wiki platform is a collaborative authorship tool that facilitates an "open, peer-production model."[82] According to its own "About Wikipedia" entry, "Wikipedia is written collaboratively by largely anonymous Internet volunteers who write without pay. Anyone with Internet access can write and make changes to Wikipedia articles (except in certain cases where editing is restricted to prevent disruption or vandalism). Users can contribute anonymously, under a pseudonym, or with their real identity, if they choose."[83] On a daily basis, 91,000 active users contribute to

[79] YouTube is more or less open in the sense that it does not discriminate among users and provides an open platform. But it is worth noting that, unlike some other platforms (e.g., Wikipedia), YouTube limits the ways in which users that view videos can share and reuse the videos.

[80] Benkler 76–80 (2006b) (describing peer production of relevance and accreditation services at Slashdot).

[81] Giles 900 (2005) (claiming "the difference in accuracy was not particularly great: the average science entry in Wikipedia contained around four inaccuracies; Britannica, about three."). But see Encyclopedia Britannica (2006) (claiming Britannica is more accurate than Wikipedia); Poe (2006), (claiming it is a widely accepted view that Wikipedia is comparable to Britannica, regardless of the debate).

[82] Benkler 71 (2006b).

[83] Wikipedia: About, Wikipedia, http://en.wikipedia.org/wiki/Wikipedia:About (last visited Feb. 26, 2011). "Wikipedia's greatest strengths, weaknesses, and differences all arise because it is open to anyone, it has a large contributor base, and its articles are written by consensus, according to editorial guidelines and policies." *Id.* Wikipedia is managed as a commons, but it does have a "Protection Policy" that allows administrators to "protect" a page, for example by restricting editing privileges for a limited or indefinite period. There are different

the website.[84] Those users have created more than 17 million articles that attract 78 million visitors per month. Use of the site has increased over the past years. Between February 2007 and May 2010, Wikipedia use increased by seventeen percentage points (among all adults in the United States), while the number of total Internet users has only increased by eight percentage points.[85] Users receive no monetary compensation; their contributions are voluntary. It is quite clear that the actions of the active users generate substantial spillovers.

For many, this is puzzling. Why would people spend so much time and effort editing encyclopedia entries for free? How does such a large-scale open and collaborative enterprise remain stable? As Jonathan Zittrain suggested, Wikipedia is "the canonical bee that flies despite scientists' skepticism that the aerodynamics add up."[86] I do not intend to solve this puzzle; there is likely a complex mixture of motivations that lead active users to be active. Recall the discussion in chapter 12 of various motivations for people to be creative, to express themselves, or to otherwise invest their own resources in being intellectually productive. It is a mistake to assume that active users are engaging in pure altruism, though it surely is one motivating factor.[87] I suspect many active users simply enjoy the activity, and the costs of participating in a meaningful way are dramatically reduced by the platform and supporting social norms and organizational structures.

Consider also Twitter, a real-time messaging service that allows users to post 140-character messages ("tweets") to the world. In September 2010, Twitter had 175 million registered users, and there were 95 million tweets posted each day. According to Pew Research Center's Internet & American Life Project, "Eight percent of the American adults who use the [I]nternet are Twitter users."[88] Twitter users post content on a variety of topics, including updates on their "personal life, activities or interests" (72 percent); "work life, activities or interests" (62 percent); links to news stories (55 percent); "humorous or philosophical observations about life in general" (54 percent); and photos (40 percent), videos (28 percent), and location information (24 percent), among other things. The platform effectively combines aspects of text messaging and blogging with community-building features, and it also interoperates with other social media, including Facebook. While it may be tempting to dismiss the value of tweets as short bursts of nothingness, the messages are valuable communications and may generate spillovers on a small scale as people broadcast their musings, observations, and feelings to the world.

---

levels of protection that correspond to different types of disruption or vandalism. See Wikipedia: Protection policy, http://en.wikipedia.org/wiki/Wikipedia:Protection_policy (last visited Feb. 26, 2011). There are dispute resolution procedures as well. See Hoffman & Mehra (2010).

[84] Wikipedia: About, Wikipedia, http://en.wikipedia.org/wiki/Wikipedia:About (last visited Feb. 26, 2011).

[85] Zickuhr & Rainie (2011).

[86] Zittrain 148 (2008).

[87] Cf. Hoffman & Mehra (2010) (implying that altruism is the relevant motivation).

[88] Smith & Rainie (2010).

Tweets have also contributed in significant ways to political and cultural discourse. As the Open Internet Coalition noted in its submission to the FCC:

> When opposition protests broke out in Iran following the presidential election, the Iranian government attempted to block cell phones and text messaging and deny access to many social networking sites to prevent the spread of speech and discontent. Quickly, however, Twitter became the medium of choice with protesters tweeting minute-by-minute updates, allowing the world to know what was happening in Tehran and giving the protesters a voice when their government did not want them to be heard. Tweets became so vital to the coverage in Iran that Twitter delayed scheduled site maintenance in order for the political organizing in Iran to continue with minimal disruption. Iranians used all forms of new media to organize themselves and their message: Facebook was used to organize rallies, YouTube was used to distribute to videos of protests, and Google Maps was used to track where government tanks were located.[89]

Similarly, more recent uprisings in Egypt can be, at least in part, attributed to online social networking facilitated by Twitter and Facebook. These resources provided an incredible ability to disseminate information quickly and widely; protesters used the communications technologies to rally supporters, organize protests, and take down the government.[90] When the government sought to take down the sites, the people took to the streets. Even after the government seized control of the Internet in Egypt, citizens had ways to tweet about what was going on, either by using a proxy server[91] or by calling a phone number and using a "speak-to-tweet" system.[92] The opposite side of the story, which is rarely mentioned, is that the government also had access to the sites and used them to its own advantage: Activists were rounded up by Egyptian forces based on user information gleaned from Twitter and Facebook.[93]

Of the many social networking platforms on the web, Facebook is the world's largest.[94] Facebook has more than 500 million active users, 50 percent of whom visit the site on a daily basis. Facebook users create and share a tremendous amount of independent content.

---

[89] Comments of Open Internet Coalition, In the Matter of Preserving the Open Internet, GN Docket No. 09-191, WC Docket No. 07-52 (Jan. 14, 2010) (citing Stone & Cohen (2009)).

[90] *Egypt Internet Users Report Major Network Interruptions*, TVNZ (New Zealand) (Jan. 28, 2011), http://tvnz.co.nz/technology-news/egypt-internet-users-report-major-network-disruptions-4008917.

[91] *Id.*

[92] *Egypt Protesters Use Voice Tweets*, BBC (Feb. 1, 2011), http://www.bbc.co.uk/news/technology-12332850.

[93] Gallagher (2011).

[94] MySpace and LinkedIn are two popular social networking sites. The Open Directory Project lists 215 social networking communities. See http://www.dmoz.org/Computers/Internet/On_the_Web/Online_Communities/Social_Networking/.

Facebook Statistics touts the following: There are over 900 million objects users interact with, including groups, events, and user pages; the average user creates ninety pieces of content each month; and more than 30 billion pieces of content (links, news, blog posts, photo albums, etc.) are shared each *month*.[95] In addition to the creation and sharing of various public goods, including speech and cultural content of all sorts, Facebook enables social interactions, the development of old and new relationships, and the strengthening of social ties (even ties that are relatively weak). As a result of these social capabilities, it enables collective action and coordination through social networking that would be incredibly difficult, and perhaps impossible in some cases, without the platform. The spillover effects offline are immense. Again, the wedge between private market value (value captured in market transactions) and social value is substantial.

* * *

I conclude this section with a brief discussion of the basic capabilities that the Internet provides to users and how the availability and exercise of such capabilities may be affecting people.[96]

[The Internet] offers a wide range of opportunities for individuals to participate productively in political, intellectual, and cultural activities. Yochai Benkler describes many ways that peer production occurs through the use of various Internet-enabled communications technologies, including simple e-mail or blog software.[97] These general-purpose, content-neutral, and easy-to-use technologies facilitate participation in wide-ranging discussions in various communities. From the perspective of liberal autonomy, the increased range of meaningful opportunities—the increased choice—is normatively attractive in itself. In addition, society may benefit from actual participation in these activities and the products and/or changes to the cultural environment that such participation can yield.[98]

The Internet facilitates many different forms of and forums for communication that are open in terms of content and users. E-mail, chat rooms, blogs, and web pages are some of the open communications technologies that have greatly enhanced the communication capacities of individuals and groups (on a one-to-one, one-to-many, and many-to-many basis). Digital cameras, video recorders, editing utilities, and resources such as Wikipedia and Second Life are a few of the technologies and platforms that significantly

[95] Facebook Statistics, FACEBOOK, http://www.facebook.com/press/info.php?statistics (last visited Feb. 26, 2011). In addition, the 200 million people who use Facebook on their mobile phones are twice as active as non–mobile phone users. *Id*. An update: On October 18, 2011, the Facebook Statistics page indicated that the number of active users had grown to 800 million.

[96] The discussion draws from Frischmann (2007a), which reviews BENKLER (2006b).

[97] BENKLER (2006b).

[98] Infrastructures often provide or enhance basic human capabilities, and when people choose to exercise those capabilities to satisfy their own autonomous needs, society often benefits from spillovers. In future work I will explore the relationships between infrastructure, capabilities, and spillovers.

increase the capacities of individuals to produce digital content that can be shared and collaboratively (re)produced online. Not surprisingly, a significant reduction in costs leads to a significant increase in the quantity of speakers, listeners, and content producers, and thus in speech and content. While the quality of speech and content varies considerably, and one might even regard some barriers to entry in communications as socially desirable, on the whole the societal benefits of this incredible expansion in communication capacities seem to substantially outweigh the harms.

Moreover, as many have observed and discussed, these technologies and their complementary cousins—social software—facilitate more than communications between speakers and listeners or the sharing of content; they enable users to develop meaningful associations with others, such that groups, communities, and social networks may thus coalesce. Users actively participate in meaningful social activities that frankly may be oversimplified when discussed solely in terms of either "speech" or "cultural production." At least in some contexts, the formation of social networks around speech-cultural exchange and intellectual pooling may be the more interesting and important phenomenon.

In addition, a qualitative change that is now under way may eclipse the quantitative change in participation. This qualitative change relates to the liberation reflected in an expansion in the choices we experience—our increased autonomy—but it involves liberation in a somewhat different sense. As Benkler explains:

> The qualitative change is represented in the experience of being a potential speaker, as opposed to simply a listener and voter. It relates to the self-perception of individuals in society and the culture of participation they can adopt. The easy possibility of communicating effectively into the public sphere allows individuals to reorient themselves from passive readers and listeners to potential speakers and participants in a conversation. The way we listen to what we hear changes because of this; as does, perhaps most fundamentally, the way we observe and process daily events in our lives. We no longer need to take these as merely private observations, but as potential subjects for public communication.

The key to this qualitative change is that *people* may change for the better with their experiences. Recall our discussion in chapter 12 of the dynamic and interdependent relationships we have with the cultural environment, how we make and shape it, and how it makes and shapes us. Active, productive users may become more aware, conscious of their (potential) roles as listeners, voters, and speakers, but also as consumers and producers, as political, cultural, and social beings, and as members of communities. They may learn to be productive—or learn to want to be productive, if such desire is not simply latent. This very awareness that one can play different roles and that the environment is not fixed or fully determined by others is encouraging. It encourages participation and the development of facilitative social practices, and perhaps, over time, the adoption of a participatory culture.

Coupled with the empowerment and encouragement for individuals is social empowerment. This enables users to develop and sustain old, existing, and new social relationships on a scale and scope impossible without the Internet. The ability to interconnect with such a vast number of people through a host of different networking technologies and organizational tools enables coordination, group formation, and collective action—capabilities not nearly as easily exercised in pre-Internet days.

The Internet infrastructure serves as a foundational infrastructure in a manner analogous to the natural and cultural environment. As the examples discussed show, the Internet infrastructure supports many mixed infrastructure in the higher layers. There are many Internet applications that provide basic communications and socialization capabilities—ranging from e-mail, chat rooms, instant messaging, and the World Wide Web to peer-to-peer technologies, Twitter, massive multiplayer online role-playing games, and social networking platforms (to name just a few). Many of the applications are appropriately celebrated as innovations with significant economic value, but these applications generate substantial social value above and beyond the private value reflected and captured in market transactions. There are network effects, various types of spillovers among users, and substantial spillovers offline. Though important, it is not enough to celebrate these applications as innovations. In a sense, it is the nature of the innovation that seems to matter most. What is innovative about these applications is their mixed infrastructural nature and the expansion of human capabilities to generate and act on ideas, to communicate, to socialize, to participate, and to be productive and engaged rather than passive.[99]

This section could go on endlessly in many different directions describing the wide variety of user-generated public and social goods. The preceding paragraphs contain a laundry list of areas to explore in future research, including (1) the various systems that the Internet is transforming, (2) different types of higher-layer infrastructure (e.g., infrastructural applications, platforms, and organizational systems), (3) different modes of production,[100] and (4) the opportunities and basic capabilities that the Internet provides to users. Each direction leads to similar observations—the Internet generates substantial

[99] Not everyone chooses to exercise the capabilities made available; many Internet users are passive users. This observation does not undermine the point. A recent FCC study indicated that 52 percent of Internet users have submitted a review for a product, 45 percent have uploaded their own content, 23 percent have posted to their own blog or to a group blog, 22 percent have taken a class online, and 14 percent have played a complicated online video game such as a massive multiplayer online role-playing game. Horrigan 16 (2010).

[100] Many of the examples discussed show how the Internet supports infrastructure in the higher layers that enable different modes of production such as commons-based peer production, an important form of nonhierarchical, distributed production by users; collaborative production across organizations (e.g., firms and/or universities); hybrid production that mixes market and nonmarket or proprietary and commons-based production. Participants in these different emerging modes of production capture some of the benefits but often also generate significant spillovers. Benkler (2006b); Frischmann 1113 (2007a).

social value not captured or reflected fully in markets. The Internet is a spillover-rich environment because of the basic user capabilities it provides and the incredibly wide variety of user activities that generate and share public and social goods.

Finally, I acknowledge that I am very optimistic about how the capabilities that the Internet provides, how people choose to exercise those capabilities, the nature of and outputs from the various activities users engage in, and the various ways in which the Internet is transforming us and our environment. Many people do not share my optimism. There is undoubtedly plenty of bad behavior, harmful speech, and malicious computer viruses, among other things. While this is ultimately an empirical question, I am optimistic for two reasons: First, based on my own experience and an anecdotal survey of various Internet activities, I believe the benefits far outweigh the harms; second, I believe that much, though by no means all, of the harmful conduct can be addressed in a manner compliant with commons management principles, as I discuss below.

2. COMMONS MANAGEMENT

The demand-side case for managing the Internet infrastructure as a commons remains quite strong. To begin with, it is worth recognizing that the universal adoption of end-to-end design protocols reflects a strong private commitment to commons management. Private networks voluntarily chose end-to-end design, interoperability, and interconnection with other networks. It has always been legally and technologically permissible for a private network owner to opt out by ceasing to offer Internet access or transport services and instead maintaining a private network. Networks could have chosen to reject the Transmission Control Protocol/Internet Protocol (TCP/IP).[101] Most successful networks did not.

Recall the discussion in chapter 5 of the business case for commons management as private strategy. I noted five reasons why firms choose to adopt a commons management strategy, such as adoption of and commitment to end-to-end design: (1) consumers generally dislike discrimination; (2) an open, uniform platform may be easier and cheaper to manage than a closed, discriminatory one; (3) a commons management strategy may facilitate joint production or cooperation with competitors more generally, for example, to interconnect and create a more robust meta-network; (4) a commons management strategy may support or encourage value-creating activities by users; and (5) a commons

[101] *Transmission Control Protocol: DARPA Internet Program Protocol Specification*, IETF RFC 793 (Sept. 1981), http://www.ietf.org/rfc/rfc0793.txt?number=793 (formally describing TCP); *Internet Protocol: DARPA Internet Program Protocol Specification*, IETF RFC 791 (Sept. 1981), http://www.ietf.org/rfc/rfc0791.txt?number=791 (formally describing IP).

management strategy maintains flexibility in the face of genuine uncertainty.[102] All of these reasons applied, and likely still apply, to private decisions about how to manage the physical networks that jointly comprise the Internet infrastructure.

An astounding number of different networks, both privately and publicly owned, recognized the value of commons management. The collective strategy (commitment, or even social arrangement) sustained the option value of individual networks and the meta-network and avoided premature optimization for any particular use or subset of uses. Persistent uncertainty about the future turned out to be a boon, as anything was possible and no one had the power to control, coordinate, or manage the path of progress.

The network neutrality debate concerns public regulation requiring private networks to comply with commons management, and thus concerns commons management as public strategy. The affirmative case for commons management effectively mirrors the discussion in chapter 5. The Internet infrastructure is a mixed infrastructure, and as such, it faces the two types of demand-driven problems discussed throughout this book: First, it faces concerns about undersupply and underuse of infrastructure to produce infrastructure-dependent public and social goods, which leads to underproduction of those goods. Second, it faces concerns that infrastructure development may be skewed in socially undesirable directions. For example, if private infrastructure owners prematurely optimize infrastructure for uses that they expect will maximize their private returns, and in doing so choose a path that forecloses production of various public or social goods that would yield greater net social returns, the social option value of the Internet is reduced. This latter concern may involve dynamic shifts in the nature of the Internet infrastructure, such as optimizing networks in a manner that shifts from mixed infrastructure toward commercial infrastructure.

Commons management serves the basic functions described in chapter 5. It creates a blunt means for supporting user production of public and social goods. The conventional economic answer to concerns about undersupply of such goods is direct subsidies, but it should be rather obvious why such a solution would be impossible to implement in the Internet context, given the wide variety of activities and the multitude of users. Infrastructure markets subject to commons management have supported such activities and users without the need for government involvement in picking activities or users worthy of subsidy. As chapter 5 explained, commons management effectively creates cross-subsidies among uses/users and eliminates the need to rely on either the market or the government to pick winners by prioritizing or ranking uses (or users). In the Internet context, I would take the cross-subsidy point a bit further. Open infrastructure appears to be what supports the user freedoms and capabilities described in the last section.

[102] This includes uncertainty as to how the infrastructure may evolve, what will be technologically feasible, what unforeseen uses may emerge, what people will want, how much people will be willing to pay, what complementary goods and services may arise in the future, and so on.

Micromanagement of user behavior by networks, whether manifested in price discrimination based on who you are or what you are doing or in preferential treatment or ranking of activities, cuts directly against such freedoms and their exercise.

Critically, the same is true at the level of infrastructural applications. Many of the mixed infrastructure applications discussed in the previous section survive as such and thrive through user participation because they exist within an open environment and are managed as commons. From a dynamic perspective, commons management serves the basic function of maintaining flexibility and the generic nature of the resource. It precludes optimization, and in doing so it sustains the infrastructure's social option value.

Many have argued that such regulation is unnecessary because private networks will not defect from the collective strategy of committing to and complying with the end-to-end principle because they recognize—and if they don't, consumers will discipline and force them to recognize—the private value of commons management.[103] I tend to agree that commons management remains an attractive private strategy, mainly because there is considerable uncertainty about future market value—that is, what the future sources of market returns will be. High market uncertainty provides a strong argument for commons management as a private strategy to maintain the option value of the infrastructure. That said, I am much less confident that many private networks agree, that we can count on the disciplining effect of markets, or that networks will not defect whenever feasible.[104] As we discovered in chapter 5, private infrastructure owners have a number of reasons for choosing to reject a commons management strategy, such as opportunities to price discriminate, vertically integrate and optimize for a subset of downstream markets, and control future progress. Similarly, firms may seek to manage user activities more closely—for example, to manage congestion or other types of interactions among users. Finally, firms may be strongly biased in their estimation of the future market value to favor services that they currently offer or expect to offer, sponsor, or otherwise control, and to disfavor those that they do not. Put another way, firms may not recognize that they are operating in a highly uncertain market environment because they are reasonably certain about the potential market value of known or expected services and significantly discount the potential value of uncertain prospects. Nonetheless, it is not necessary to demonstrate that commons management is no longer an attractive private strategy (or that firms no longer believe it to be attractive). There is too much at stake to bet on private strategy coinciding with public strategy.

There is a very high degree of social value uncertainty. It is impossible to predict with any degree of confidence who or what will be the sources of social value in the future.

[103] Lee (2008).

[104] Even if firms pledge their commitment to commons management and agree to refrain from discrimination or prioritization, it can be incredibly difficult to detect defections from that commitment. Sashkin 306 (2006); Marsden 12 (2008).

Accordingly, there is no reason to defer to private firms in this context. First, there is no reason to believe that firms are better informed, capable of maximizing social value, or likely to resist the pressure to discriminate, prioritize, or optimize the infrastructure based on foreseeable and appropriable private returns. Second, there is no reason to trust that markets will correct misallocations. Such thinking may be what motivates calls for transparency as a solution, to enable consumers to decide and potentially discipline missteps by networks. But, among other things, this line of thought presumes that private demand fully reflects social demand; it reduces users to passive consumers of commercial infrastructure; and it ignores the social opportunity to leverage the nonrival nature of the infrastructure to enable and encourage user activities that generate public and social goods.

The Internet, like much of the other mixed infrastructure discussed in this book, has been and should continue to be managed as a commons. Doing so has generated immeasurable social benefits. To be clear, I do not claim to be able to quantify these benefits so that they can be put on a scale and compared with the benefits of prioritization; rather, I maintain that the case for shifting away from commons management and toward prioritization simply has not been made. Putting aside rhetorical arguments of the sort discussed in chapter 8, the primary arguments in favor of permitting discrimination and prioritization are related to managing congestion, incentives to invest in infrastructure supply, and managing harmful traffic. While superficially appealing, I do not find these arguments persuasive, as I discuss below. The following section proposes a specific nondiscrimination rule and then addresses these arguments.

C. A Proposed Nondiscrimination Rule and Various Complications

A targeted nondiscrimination rule is an appropriate intervention because it would both preclude differentially allocating and prioritizing access and use of the Internet on the basis of expected private returns and limit infrastructure evolution or optimization on that basis. After briefly introducing the recent FCC Order for preserving the Open Internet, I discuss my proposed nondiscrimination rule, which precludes discriminating based on the identity of the user or use.

On September 23, 2011, the FCC published its Open Internet Order in the Federal Register, to go into effect November 20, 2011. The FCC adopted three basic rules to preserve a free and open Internet:

1. Transparency. Fixed and mobile broadband providers must disclose the network management practices, performance characteristics, and terms and conditions of their broadband services.
2. No blocking. Fixed broadband providers may not block lawful content, applications, services, or nonharmful devices; mobile broadband providers may not

block lawful websites, or block applications that compete with their voice or video telephony services.

3. No unreasonable discrimination. Fixed broadband providers may not unreasonably discriminate in transmitting lawful network traffic.[105]

The transparency and no blocking rules are important and not very controversial.[106] The rule prohibiting unreasonable discrimination and thus legitimizing reasonable discrimination is more contentious. Some believe the FCC went too far, others not far enough. The FCC expressly rejected the idea that the nondiscrimination rule should only preclude anticompetitive discrimination proven to harm consumers, and thus envisioned a somewhat broader standard. What is (un)reasonable remains to be seen, however. The FCC Order suggested some guidelines, for example, indicating that pay-for-priority is frowned upon, while use-agnostic discrimination may be reasonable. The FCC also emphasized that the rules would be "applied with the complementary principle of reasonable network management."[107] Thus, with respect to discriminatory practices, such as prioritization, there are two separate reasonableness inquiries. While I agree with much of the FCC's reasoning in the Order about the need to sustain an open Internet, I do not believe the FCC went far enough with its nondiscrimination rule. As I discuss below, I would drop "(un)reasonableness" altogether.

For the remainder of this chapter, I do not further examine the details of the FCC rule. That rule is the first step along what promises to be a long path. Although it is intended to provide a stable and predictable regulatory framework, litigation, regulatory implementation, and even intervention by Congress promise to complicate matters and keep the network neutrality debate alive for the foreseeable future. Even if the FCC rule survives legal challenges and legislative overrides, the rule is substantively incomplete. Accordingly, instead of examining the FCC rule, I propose an alternative formulation that corresponds to the basic definition of commons management employed throughout this book.

Congress or the FCC should consider prohibiting broadband Internet providers from discriminating based on the *identity* of the user or use in the handling of packets. "User" may be defined as sender or receiver; "use" may be defined as application or content type; "handling" may be defined as all transport and related services associated with delivery of packets. This simple formulation of a nondiscrimination rule may seem overly strong in that it appears to rule out a significant range of activities that some might

---

[105] Preserving the Open Internet, 76 Fed. Reg. 59,192, 59,192 (Sept. 23, 2011) (to be codified at 47 C.F.R. pts. 0 and 8), available at http://www.gpo.gov/fdsys/pkg/FR-2011-09-23/pdf/2011-24259.pdf.

[106] Implementation involves controversy, but most agree on the basic principles. I focus on the nondiscrimination rule.

[107] *Id.* at 2.

label "reasonable discrimination" or "reasonable network management." Of course, it depends on what those labels apply to.

The proposed rule primarily precludes certain fine-grained forms of price discrimination, quality discrimination, and prioritization. It does not prohibit other forms of price discrimination that are not based on user/use identity, such as typical second-degree price discrimination, and it does not prohibit more efficient methods for managing congestion, such as traditional usage-sensitive pricing or congestion pricing. After discussing how the rule would impact pricing schemes, I explain why disabling application-based prioritization is not as horrible as opponents suggest.

With respect to pricing, the proposed rule precludes many forms of value-based price discrimination. It does not preclude various forms of cost-based differential pricing, including variable load pricing, congestion pricing, and usage-sensitive pricing. The reason is that these and other forms of cost-based differential pricing do not discriminate on the basis of the identity of users or their specific activities (uses). Instead, they discriminate based on quantity of infrastructure use, capacity utilized, and the marginal cost of such use, taking into account contextual details such as timing and available system capacity; none of this includes identity characteristics. Put simply, to implement most forms of cost-based differential pricing, an infrastructure owner need not know *who is doing what*. Rather, the focus is on *when* and *how much*. In sharp contrast, price discrimination relies directly on identity characteristics that aim to best approximate individual users' subjective valuation of infrastructure use—*who and what are essential to the discrimination or prioritization scheme*.

The proposed rule does not preclude all forms of value-based price discrimination; it precludes discriminating on the basis of the identity of the infrastructure user or use. Second-degree price discrimination, for example, may or may not run afoul of this nondiscrimination rule, depending on how the scheme operates. Second-degree price discrimination may present all infrastructure users with the same price schedule for the same basic service with price variations based on the quantity of use / capacity consumed; users decide what to choose from the menu based on their anticipated demand. Third-degree price discrimination, by contrast, tends to categorize consumers based on their identity, specifically based on identity characteristics that serve as effective proxies for consumers' subjective valuations. The categories are not based on cost. *Who is doing what* is central to the scheme. Third-degree price discrimination generally conflicts with the proposed rule. So do finer-grained identity-based price discrimination schemes.

Value-based price discrimination is at the core of the rule because it is estimated appropriable value that drives private allocation, prioritization, and optimization decisions and potentially leads the infrastructure to evolve in a manner that forecloses production of various public and social goods and thereby reduces social welfare. The proposed nondiscrimination rule precludes differentially allocating, pricing, and prioritizing Internet access and use on the basis of appropriable value. As examined extensively in previous

chapters, this rule (1) can be a more effective—albeit blunt—means for supporting the production of public and social goods than targeted subsidies, and (2) has important dynamic implications. It maintains flexibility and the generic nature of the Internet. The proposed nondiscrimination rule effectively functions as a social option, which makes economic sense because of persistent and systematic uncertainty about the future sources of both market and social value.[108] It precludes premature optimization by infrastructure owners; preserves the "evolvability" of the Internet[109] by supporting experimentation by users; increases the range of potential value-creating activities; leaves room for unforeseen innovations, markets, and value-creating activities to emerge; and facilitates learning over time.

Disabling identity-based price discrimination is not necessarily costless, although it should not be assumed to be costly either. It *may* involve a trade-off. The two principal potential advantages of price discrimination are (1) increased output and thus reduced deadweight losses when compared with uniform pricing, and (2) increased profits for infrastructure providers that may improve incentives to invest in the supply, maintenance, and improvement of infrastructure. I discussed these issues extensively in previous chapters. Neither of these potential advantages appears to be nearly as significant as claimed by many in the network neutrality debate.

Before proceeding, it is helpful to explain that price discrimination in this context typically would occur in tandem with quality discrimination in the sense that networks would prioritize, shape, or otherwise manage traffic in a discriminatory manner, perhaps according to priced quality-of-service guarantees, but not necessarily in such an explicit fashion. Therefore, I do not discuss identity-based price discrimination schemes divorced from prioritization during congestion (which are pure rent-extraction schemes aimed at capturing a greater portion of the surplus derived from various Internet activities).[110] Few, if any, participants in the network neutrality debate argue in favor of such schemes; opponents of network neutrality regulation almost always argue that prioritization is needed to deal efficiently with congestion.[111]

First, the claim that output would increase if network providers can price discriminate and prioritize traffic is misleading. The case for prioritization depends on an assumption of congestion. In the absence of congestion, there is no need to prioritize traffic, other

---

[108] See ch. 5. Gaynor (2003); Gaynor & Bradner (2007); van Schewick (2010).

[109] See *id.*

[110] For an analysis of these schemes, see van Schewick 273–78 (2010).

[111] Comments of Time Warner Cable Inc., In the Matter of Preserving the Open Internet, GN Docket No. 09-191, WC Docket No. 07-52 (Received Jan. 14, 2010), at 66–67; Comments of Cablevision Systems Corp., In the Matter of Preserving the Open Internet, GN Docket No. 09-191, WC Docket No. 07-52 (Received Jan. 14, 2010), at 9–10; Reply Comments of Comcast Corp., In the Matter of Preserving the Open Internet, GN Docket No. 09-191, WC Docket No. 07-52 (Received Apr. 26, 2010), at 35–36.

than plain rent extraction. If the network is congested and infrastructure capacity is scarce, output cannot be increased without making congestion worse or imposing congestion costs on others. If A and B are two use(r)s vying for access and use of a congested resource, prioritization of use(r) A requires deprioritization of use(r) B. The argument for prioritization is really an argument to allocate scarce capacity to uses that would suffer greater consequences from congestion—whether attributed to delay or jitter or packet loss—and thus would realize greater benefits from prioritization. It is not really an argument for increasing output measured by the number of users, uses, or even packets delivered during periods of congestion. Prioritization does not eliminate or minimize congestion. Users A and B *may* adjust their consumption patterns based on how congestion costs are distributed (shared equally or pushed to B because B is less sensitive to timing) and such adjustments *may or may not* lead to increased output over the long run, depending on various factors, but *only if there is sufficient off-peak capacity within an acceptable time frame*. While it is easy to assume the existence and persistence of uncongested periods during which deprioritized users get an opportunity to act, such an assumption is heroically optimistic and plainly unwarranted; as many have pointed out, prioritization creates perverse incentives for network providers to sustain congestion and underinvest in capacity expansion because such a strategy can be more profitable.[112]

Second, the oft-stated claim that increased revenues from price discrimination are necessary to investment incentives is speculative, self-interested, and doubtful. Major providers, such as Comcast, AT&T, and Verizon, have invested considerable sums without any assurance that they would be capable of prioritizing traffic to obtain increased revenues. They would love to capture more of the surplus, but there is no reason to think they are entitled to it, and it is hardly proven that they need it to justify infrastructure investment.[113] As chapter 8 examined extensively, claims that price discrimination is necessary to support incentives to invest are quite easy to state but much more difficult to support theoretically or empirically. Such claims hold only for a special class of cases where (1) the average total cost curve fails to intersect the private demand curve, and (2) price discrimination yields benefits in excess of total costs (i.e., the area under the demand curve exceeds total costs). Even when these conditions are met, there is not necessarily a strong argument in favor of price discrimination. The strength of the argument depends on a number of additional considerations, including (1) a comparative analysis of alternative institutional solutions that might solve the incentive problem in a more efficient manner

[112] See, e.g., Economides 8 (2010); Choi & Kim 29–30 (2008).

[113] See, e.g., FCC Order paragraph 40. I addressed this argument elsewhere. Frischmann & van Schewick (2007); Frischmann & Lemley 297 n. 147 (2007) (arguing there are other ways to incentivize capacity upgrades, such as direct subsidization, tax incentives, and government provision of infrastructure). It should be noted that the US government has recently helped broadband providers expand their networks by providing a $7 billion subsidy. Hinsell (2009).

than price discrimination;[114] (2) the magnitude of the net social welfare gains; (3) the option value of waiting and allowing demand to rise, supply costs to fall, and/or economies of scale to kick in; and (4) the social costs of imperfect price discrimination, including the impacts on the productive activities of users and the generation of spillovers.

Thus, even if arguments that price discrimination is necessary to support incentives to invest were supported with empirical evidence (which they are not), this would only prompt a comparative analysis of various institutions for addressing the supply-side problem. There is no reason to believe that prioritization or discrimination based on the identity of the user or use is the most efficient solution. There are plenty of alternatives, including direct subsidization of infrastructure expansion, tax incentives to support infrastructure expansion, funding research aimed at lowering the cost of infrastructure expansion or entry into the infrastructure market, cooperative research and development projects, joint ventures, government provisioning, and so on.

In addition to arguments about alleged benefits of price discrimination—increased output and incentives to invest—opponents of network neutrality suggest that identity-based discrimination or prioritization may be an efficient and thus reasonable way to manage traffic on networks. In particular, two types of traffic management issues arise: (1) managing congestion and (2) managing unlawful, hazardous, or otherwise harmful traffic.

1. MANAGING CONGESTION

End-to-end design works reasonably well because all packets in fact are the same from a delivery cost and congestion cost perspective. Congestion is associated with the buildup of packets in a queue and the delay associated with processing the queue.[115] With regard to queuing, contributions to delay, and thus congestion, one packet is no different than another.

The TCP/IP, an end-to-end control system for managing congestion, operates at the edges of the network.[116] It allows the source and destination (as well as ISPs) to monitor and communicate about queuing delays. When the delay increases, it allows the

[114] For example, government subsidy can reduce the average total costs and make the investment profitable without relying on price discrimination. Government subsidies also can be directed at research and development in technologies that may reduce the costs of infrastructure supply and facilitate competitive entry.

[115] For an excellent, accessible explanation, see CRTC Regulatory Policy 2009-657, http://www.cippic.ca/uploads/File/Attachment_B_pt_1_-_Reed_Report.pdf (testimony of Dr. David Reed).

[116] See *id.* See also Speta 245 (2002), which provides an accessible description:

> TCP provides the overall "message management" functions that enable computer-to-computer communication. In simplified form, TCP breaks a computer message into the appropriate packet sizes, numbers the packets, creates a check sum so that the receiving computer can check the integrity of the message, provides the information for the reassembly of the message into the proper sequence by the receiving computer, and orders the re-sending of any lost or damaged packet. . . .

end-hosts to reduce the rate at which data is sent until the delay reduces to near zero. There are various technical means for communicating about and responding to congestion; some of these involve managing and prioritizing traffic. But the existence of such means should not lead one to think that packets are different from a delivery-cost or congestion-cost perspective. They are not.

Though packets are the same from a cost perspective, all packets are not the same from a user valuation perspective. Users derive value from the higher-layer uses of delivered packets, and different uses are more or less sensitive to the consequences of congestion. As discussed above, prioritization is primarily about allocating scarcity, typically based on willingness to pay coupled with discriminatory pricing. The economic case for discrimination/prioritization depends on the existence and persistence of congestion. While eliminating congestion completely is not feasible (or even desirable),[117] there are good reasons to direct policy in that direction, or at least toward minimizing congestion, rather than sustaining it to support business models dependent on prioritization. Keep in mind that prioritization as a form of congestion management is offered as a potential justification for dismantling the commons management regime. Simply put, there are more attractive nondiscriminatory means for combating congestion that are compliant with commons management.

Congestion on the Internet should be managed primarily through expanding capacity and implementing usage-sensitive or congestion pricing, rather than accepting prioritization and encouraging persistent congestion. Investment in capacity expansion coupled with pricing sensitive to congestion and/or usage would go a long way toward resolving many of the perceived congestion problems on the Internet.[118] The proposed

. . . IP wraps a header around each packet created through the TCP. This IP header contains the information necessary to route the packets properly from the sending computer to the receiving computer. Thus, the IP header provides the Internet address of the sending computer and the Internet address of the destination computer, as well as the other pieces of information necessary for each network and network gateway to properly handle the packet.

[117] Cornes & Sandler (1996).

[118] For a detailed discussion, see chapter 7; Frischmann & van Schewick 392–409 (2007). Yoo argues that use restrictions may be a more efficient method for managing congestion than usage-based pricing and that this possibility is a reason to refrain from network neutrality regulation; he suggests that networks should be free to choose among different approaches to managing congestion. Yoo (2006). As discussed in chapter 7, use restrictions may be required where incompatibilities arise among uses, but that is not a problem on the Internet. Rather, the argument for using use restrictions is that network providers may use them as cost-effective proxies in the face of allegedly high transaction costs associated with implementing usage-based pricing. See *id.* The argument is based on faulty assumptions about the costs of metering usage. As Barbara van Schewick and I explain, the costs of metering traffic and implementing usage-based pricing and/or imperfect congestion pricing are not inordinately high; rather, the technology exists and has been available and in practice for years. Frischmann & van Schewick 392–409 (2007). (Since 2007, the technology has no doubt improved and the costs fallen.) We also explain why congestion pricing is imperfect and why this is not surprising or especially troubling. As discussed in chapter 7, even somewhat crude forms of congestion pricing, such as time of day pricing, can have significant efficiency benefits as users internalize some of the external costs associated with the

rule is attractive precisely because it would push providers to focus on capacity expansion and congestion- or usage-sensitive pricing rather than prioritization. Moreover, it would continue to push for innovative solution at the ends. As Lemley and I noted, "Even with a dumb architecture, innovators have figured out how to provide certain degrees of quality of service at the periphery of the network and how to make certain latency-sensitive applications, such as IP telephony, work."[119]

Proponents of prioritization as a necessary or cost-effective tool for managing congestion have failed to make their case, to the extent that they have even tried.[120] There are effective nondiscriminatory means for managing congestion on the Internet, and prioritization is an incredibly shortsighted approach that, at best, aims to mitigate the private costs of congestion for a subset of users. Abandoning commons management and the social values it sustains is much too high a price for society to pay.

2. MANAGING UNLAWFUL, HAZARDOUS, OR OTHERWISE HARMFUL TRAFFIC

Many argue that networks should not only be permitted but should be encouraged to actively police Internet traffic, by monitoring, managing, and regulating unlawful, hazardous, and otherwise harmful traffic.[121] I will not fully address this complex argument here. It involves a number of complications that are beyond the scope of this chapter. My bottom line is that care must be taken when we consider whether to rely on private owners of infrastructure to act as regulators of users' speech and behavior, and innovation at the ends has proven remarkable at addressing problems of all sorts. In many situations, there are nondiscriminatory means for addressing such problems. Finally, as discussed, access networks often operate as end-systems, particularly where access networks run server-based applications and provide related services for customers. Blocking or filtering traffic at the mail server is a routine method for dealing with spam that is end-to-end compliant. This does not conflict with the proposed rule.

---

timing of their activities. *Id*. We explain various advantages of imperfect congestion pricing over use restrictions as proxies. For example, in contrast with use restrictions, congestion pricing does not tax users in off-peak periods. Of course, congestion will not disappear altogether, and congestion pricing (much like quality of service assurances) likely cannot be implemented system-wide because of the many different networks and resources involved.

[119] Frischmann & Lemley 295 n. 143 (2007).

[120] Comments filed in the FCC proceeding and the FCC Order itself provide remarkably thin theoretical or empirical support. See Preserving the Open Internet, 76 Fed. Reg. 59,192, 59, 209-10 (Sept. 23, 2011); Cablevision reply comments, pp. 9–10, nn.21–24; Time Warner Cable reply comments p. 86.

[121] Comments of the Motion Picture Association of America Inc., In the Matter of a National Broadband Plan for Our Future, GN Docket No. 09-51, at 31–32 (stating that the National Broadband Plan should include measures to encourage ISPs to deter unlawful online conduct); Comments of Recording Industry Association of American, In the Matter of Preserving the Open Internet, GN Docket No. 09-191, WC Docket No. 07-52, at 12–13.

The proposed rule could admit narrow exceptions for categorical discrimination against traffic that is demonstrated to be harmful to the network itself. Such an exception is incorporated into the FCC rule as a form of reasonable network management. The FCC Order provides only limited discussion of this issue, however, noting that "spam, botnets, and DDoS attacks" are harmful to networks and also including "unwanted" traffic in the group.[122] More is required. The burden should be on networks seeking shelter in the exception to demonstrate that the traffic in question actually causes harm to network resources or poses a substantial risk of such harm *and* that other end-to-end compliant means are ineffective.[123] It remains unclear the extent to which "spam, botnets, and DDoS attacks" are actually harmful to network resources, *above and beyond congestion*. Usage or congestion-based pricing might reduce the alleged harm considerably by imposing costs on the originators of such traffic, and to the extent that originators are "zombie computers" (i.e., end-user computers that may be infected with a program that sends spam or DDoS attacks without the knowledge of the end-user), congestion pricing would create useful signals to the owners. For each example, viable solutions might be implemented at the ends.

Further, narrow exceptions based on demonstrable harm to users, particularly when users themselves raise a red flag, might be acceptable if shown to be necessary. That said, an exception based on harm to users because of the content itself (e.g., spam) may not be needed. It remains unproven (at least, to my knowledge) that such issues cannot be dealt with effectively at the ends. A slightly different type of harm to users concerns security and the integrity of end-hosts. Zittrain argues that sustaining "generativity" on the Internet may require some security functions to be executed by networks because end-users may not be capable, and the incredibly large number of unsecured end-hosts poses a substantial risk of a "catastrophic security attack."[124] The jury is still out on this issue, and there may be innovative solutions developed and implemented by end-users, including firms or other organizations acting on behalf of and in the interest of individual end-users. Nonetheless, again, narrow exceptions based on security concerns might be warranted. Let me emphasize two points regarding such narrow exceptions: First, the end-to-end design principle does not rule out making the network more secure by bringing certain security functions within the "core," as Barbara van Schewick explains in considerable detail.[125] Second, commons management tolerates such exceptions, as the various examples discussed in this book demonstrate.

---

[122] Preserving the Open Internet, 76 Fed. Reg. 59,192, 59,209 n.102 (Sept. 23, 2011).

[123] To this end, the Order also comes up short, simply stating that "a broadband provider should be prepared to provide a substantive explanation for concluding that the particular traffic is harmful to the network . . ." Preserving the Open Internet, 76 Fed. Reg. at 59,209.

[124] Zittrain 165 (2008).

[125] Van Schewick 366–68 (2010).

D. Conclusion

My objective in this chapter has not been to make a dispositive case for network neutrality regulation. My objective has been to demonstrate how the infrastructure analysis, with its focus on demand-side issues and the function of commons management, reframes the debate, weights the scale in favor of sustaining end-to-end architecture and an open infrastructure, points toward a particular rule, and encourages a comparative analysis of various solutions to congestion and supply-side problems. I acknowledge that there are competing considerations and interests to balance, and I acknowledge that quantifying the weight on the scale is difficult, if not impossible. Nonetheless, I maintain that the weight is substantial. The social value attributable to a mixed Internet infrastructure is immense even if immeasurable. The basic capabilities the infrastructure provides, the public and social goods produced by users, and the transformations occurring on and off the meta-network are all indicative of such value.