8-29-2008

Network Neutrality, Consumers, and Innovation

Christopher S. Yoo
University of Pennsylvania Carey Law School

Network Neutrality, Consumers, and Innovation

Christopher S. Yoo*

Abstract

In this Article, Professor Christopher Yoo directly engages claims that mandating network neutrality is essential to protect consumers and to promote innovation on the Internet. It begins by analyzing the forces that are placing pressure on the basic network architecture to evolve, such as the emergence of Internet video and peer-to-peer architectures and the increasing heterogeneity in business relationships and transmission technologies. It then draws on the insights of demand-side price discrimination (such as Ramsey pricing) and the two-sided markets, as well as the economics of product differentiation and congestion, to show how deviating from network neutrality can benefit consumers, a conclusion bolstered by the empirical literature showing that vertical restraints tend to increase rather than reduce consumer welfare. In fact, limiting network providers' ability to vary the prices charged to content and applications providers may actually force consumers to bear a greater proportion of the costs to upgrade the network. Restricting network providers' ability to experiment with different protocols may also reduce innovation by foreclosing applications and content that depend on a different network architecture and by dampening the price signals needed to stimulate investment in new applications and content. In the process, Professor Yoo draws on the distinction between generalizing and exemplifying theory to address some of the arguments advanced by his critics. While the exemplifying theories on which these critics rely are useful for rebutting calls for broad, categorical, ex ante rules, their restrictive nature leaves them ill suited to serve as the foundation for broad, categorical ex ante mandates pointing in the other direction. Thus, in the absence of some empirical showing that the factual preconditions of any particular exemplifying theory have been satisfied, the existence of exemplifying theories pointing in both directions actually supports an ex post, case-by-case approach that allows network providers to experiment with different pricing regimes unless and until a concrete harm to competition can be shown.

* Professor of Law and Communication and Founding Director of the Center for Technology, Innovation, and Competition, University of Pennsylvania. Portions of this Article were incorporated into my testimony at the Public Hearing on Broadband Network Management Practices held before the Federal Communications Commission on February 25, 2008; the Hearing on Net Neutrality and Free Speech on the Internet held before the Task Force on Competition Policy and Antitrust Laws of the House Judiciary Committee on March 11, 2008; and the Hearing on H.R. 5353: The Internet Freedom Preservation Act of 2008 held before the Subcommittee on Telecommunications and the Internet of the House Commerce Committee on May 6, 2008.

INTRODUCTION

Network neutrality remains one of the most controversial issues in internet policy. It represented the most hotly contested issue during the 2006 congressional debates over comprehensive telecommunications reform legislation.[1] It played a starring role in the clearance of the series of megamergers that has recently transformed the telecommunications industry.[2] It

---

1 See Christopher S. Yoo, *Network Neutrality and the Economics of Congestion*, 94 Georgetown L J 1847, 1858-60 (2006) (reviewing the 2006 congressional debate over network neutrality).

2 See AT&T Inc and BellSouth Corp Application for Transfer of Control, Memorandum Opinion and Order, 22 FCC Rec 5662, 5724-27 ¶¶ 116-120 & n 339, 5738-39 ¶¶ 151-153 (2007) (finding a lack of evidence that network owners are likely to block, degrade, or otherwise discriminate against internet content, services, applications, or service providers); Applications for Consent to the Assignment and/or Transfer of Control of Licenses, Adelphia Communications Corporation, Assignors, to Time Warner Cable Inc, Assignees, et al, Memorandum Opinion and Order, 21 FCC Rec 8203, 8296-99 ¶¶ 217-223 (2006) (same); Verizon Communications, Inc and MCI, Inc Applications for Approval of Transfer of Control, Memorandum Opinion and Order, 20 FCC Rec 18433, 18507-09 ¶¶ 139-143 (2005) (same); SBC Communications, Inc and AT&T Corp Applications for Approval of Transfer of Control, Memorandum Opinion and Order, 20 FCC Rec 18290, 18366-68 ¶¶ 140-144 (2005) (same).

has been the subject of hearings and a report issued by the Federal Trade Commission ("FTC").[3] And it is the subject of pending legislation and hearings before the current Congress,[4] as well as a pending Notice of Inquiry[5] and complaint before the Federal Communications Commission (FCC).[6]

The multiplicity of positions taken by various advocates makes it hard to define network neutrality with any precision.[7] Perhaps the easiest definition is the one offered in an op-ed authored by Lawrence Lessig and Robert McChesney, who state that "[n]et neutrality means simply that all like internet content must be treated alike and move at the same speed over the network."[8] Some network neutrality proponents oppose "consumer tiering," in which network providers charge end users higher amounts for more bandwidth or faster internet service.[9]

---

[3] See Federal Trade Commission, Staff Report on Broadband Connectivity Competition Policy 10, 11 (June 2007), available at <http://www.ftc.gov/reports/broadband/v070000report.pdf> (last visited Mar 6, 2008) (recommending "that policy makers proceed with caution in evaluating proposals to enact regulation in the area of broadband Internet access" and that "[p]olicy makers … carefully consider the potentially adverse and unintended effects of regulation in the area of broadband Internet access before enacting any such regulation" and noting that the Commission was "unaware of any significant market failure or demonstrated consumer harm from conduct by broadband providers").

[4] See HR 5994, 10th Cong, 2d Sess (May 8, 2008), in 154 Cong Rec H 3402 (May 8, 2008); HR 5353, 110th Cong, 2d Sess (Feb 12, 2008), in 154 Cong Rec H 869 (Feb 12, 2008); S 215, 110th Cong, 1st Sess (Jan 9, 2007), in 153 Cong Rec S 287 (Jan 9, 2007); The Future of the Internet: Hearing Before the Senate Committee on Commerce, Science & Transportation, 110th Cong (2006); The Internet Freedom Preservation Act of 2008: Hearing on H.R. 5353 Before the Subcommittee on Telecommunications and the Internet, House Committee on Energy and Commerce, 110th Cong (2008); Net Neutrality and Free Speech on the Internet: Hearing Before the Task Force on Competition Policy and Antitrust Laws, House Committee on the Judiciary, 110th Cong (2008).

[5] Broadband Industry Practices, Notice of Inquiry, 22 FCC Rec 7894 (2007).

[6] Comment Sought on Petition for Declaratory Ruling Regarding Internet Management Policies, Public Notice, 23 FCC Rec 340 (2008) (seeking comment on complaint that degradation of peer-to-peer traffic violates the FCC's 2005 Internet Policy Statement and does not constitute reasonable traffic management).

[7] Eli Noam has identified no fewer than seven versions of network neutrality. See Eli Noam, *A Third Way for Net Neutrality*, Fin Times-FT.com (Aug 29, 2006), available at <http://www.ft.com/cms/s/2/acf14410-3776-11db-bc01-0000779e2340.html> (last visited Mar 6, 2008).

[8] Lawrence Lessig and Robert W. McChesney, *No Tolls on the Internet*, Wash Post A23 (June 8, 2006).

[9] See, for example, Marvin Ammori, *Time Warner Goes Back to the Future* (Jan 17 2008), available at <http://www.savetheinternet.com/blog/2008/01/25/back-to-the-future-time-warner-broadband-plan-recalls-aols-walled-garden/> (last visited Apr 19, 2008) (arguing that charging customers for using more bandwidth raises network neutrality issues); Fred von Lohmann, *Time Warner Puts a Meter on the Internet*, Electronic Frontier Foundation (Jan 22, 2008), available at <http://www.eff.org/deeplinks/2008/01/time-warners-puts-meter-internet> (arguing that consumer tiering suffers from "serious potential drawbacks"); Save the Internet Blog, *Time Warner Metered Pricing: Not the Solution* (Jan 17, 2008), available at <http://www.savetheinternet.com/blog/2008/01/17/

Others restrict their objections to "access tiering," which Lessig defines as "any policy by network owners to condition content or service providers' right to provide content or service to the network upon the payment of some fee" in addition to basic internet access fees.[10] Or as Lessig and McChesney more colorfully put it in their op-ed, network providers should not be allowed "to sell access to the express lane to deep-pocketed corporations and relegate everyone else to the digital equivalent of a winding dirt road."[11]

Network neutrality proponents advocate categorical, ex ante prohibitions on discrimination against particular content or applications.[12] These positions are primarily based on two rationales: First, network neutrality proponents argue that permitting network providers to institute such practices would harm consumers by preventing them from freely accessing whatever content and applications they may choose, or attaching to the network whatever equipment they may wish.[13] Second, they argue that such practices would harm innovation in

---

time-warner%e2%80%99s-metered-pricing-not-the-solution/> (quoting network neutrality advocate Ben Scott as arguing that metered pricing may chill innovation); Catherine Holahan, *Time Warner's Pricing Paradox: Proposed Changes in the Cable Provider's Fees for Web Could Crimp Demand for Download Services and Hurt Net Innovation*, BusinessWeek (Jan 18, 2008), available at <http://www.businessweek.com/technology/content/jan2008/tc20080118_598544.htm> (summarizing network neutrality proponents' arguments against metered pricing).

10 Net Neutrality: Hearing Before the Senate Committee on Commerce, Science & Transportation, 109th Cong 2 & n 2, 8-10 (2006) (statement of Prof. Lawrence Lessig), available at <http://commerce.senate.gov/pdf/lessig-020706.pdf> (last visited Mar 6, 2008).

11 Lessig and McChesney, *No Tolls*, Wash Post at A23 (cited in note 8).

12 See, for example, Brett Frischmann and Barbara van Schewick, *Net Neutrality and the Economics of the Information Superhighway: A Reply to Professor Yoo*, 47 Jurimetrics J at 383, 387-88 (2007) (noting that network neutrality proponents "contend that the threat of discrimination will reduce unaffiliated application and content developers' incentives to innovate"); Lessig testimony 8-9 (cited in note 10) (arguing that access tiering represents a threat to innovation in internet applications and content); Network Neutrality: Competition, Innovation and Nondiscriminatory Access, Hearing Before the House Committee on the Judiciary, 109th Cong 4-5 (2006), available at <http://www.judiciary.house.gov/media/pdfs/wu042506.pdf> (arguing that deviations from network neutrality represent a threat to innovation on the internet).

13 See, for example, Common Cause, *Keep the Internet Free and Open!*, available at <http://www.commoncause.org/site/pp.asp?c=dkLNK1MQIwG&b=1234951> (last visited Apr 10, 2008) (arguing that network providers can restricting consumers' ability to access content and applications); John Windhausen, Jr., *Good Fences Make Bad Broadband: Preserving an Open Internet Through Net Neutrality: A Public Knowledge White Paper* (Feb 6, 2006), available at <http://static.publicknowledge.org/pdf/pk-net-neutrality-whitep-20060206.pdf> (last visited Mar 6, 2008) (arguing that broadband network providers can restrict consumers' ability to attach equipment, access websites, or run applications as they see fit).

content, applications, and equipment.[14] They support their arguments by identifying circumstances under which deviations from network neutrality can hurt consumers.[15]

This Article more directly engages claims that mandating network neutrality is essential to protecting consumers and innovation. My analysis shows that the types of prioritized service and access tiering that network neutrality proponents would forbid may actually benefit consumers and promote innovation. It draws on two sources of insight that have often been overlooked in the network neutrality debate. The first is the academic literature on the economics of regulation. In particular, I expand upon my prior work emphasizing supply-side considerations, such as the economics of congestion and product differentiation, to discuss the implications of demand-side considerations, such as Ramsey pricing, and two-sided markets. Interestingly, these analyses suggest that prohibiting network providers from charging different prices to different content and application providers can harm consumers by forcing them to pay a larger proportion of the fixed cost of upgrading the network. In addition, mandating nondiscriminatory access threatens to favor content and applications optimized for the network as it exists today over content and applications that depend on a different network architecture. Indeed, preventing network providers from prioritizing certain content or applications over others may reduce innovation by making it more difficult for those innovations that depend on guaranteed quality of service from emerging.

---

14 See, for example, Lawrence Lessig, *The Future of Ideas* 156, 162, 168, 171, 175 (Vintage 2001) ("A closed network creates an externality on innovation generally. It increases the cost of innovation by increasing the range of actors that must license any new innovation."); Tim Wu, *The Broadband Debate, A User's Guide*, 3 J on Telecommun & High Tech L 69, 72-74, 85-88 (2004) (arguing that network neutrality is necessary to preserve an "innovation commons").

15 See, for example, Frischmann and van Schewick, 47 Jurimetrics J at 412-16 (cited in note 12); Barbara van Schewick, *Towards an Economic Framework for Network Neutrality Regulation*, 5 J on Telecommun & High Tech L 329, 342-52 (2007).

The second source is the empirical evidence about the likely impact of attempting to regulate the terms and conditions of internet access. In this regard, my analysis is informed in part by the lessons from previous attempts to regulate access to communications. It also takes into account the FCC's determination on five separate occasions over the past two and a half years that there was insufficient evidence of degradation and blocking to justify regulatory intervention,[16] a conclusion echoed by the OECD,[17] Justice Department,[18] and FTC,[19] as well as the fact that the ongoing Notice of Inquiry on Broadband Industry Practices[20] has identified only a handful of isolated instances. My position is also informed by the growing empirical literature showing how coordination of content and conduit through vertical integration or contractual exclusivity generally benefits consumers[21] as well as the empirical literature showing how mandating access has deterred investments in new broadband networks.[22]

I also show how the arguments advanced by some network neutrality proponents confuse the role of what economist Franklin Fisher has called "exemplifying theory" and "generalizing theory" in analyzing public policy.[23] Generalizing theory relies on fairly general assumptions to establish broad propositions that apply under a wide range of circumstances. Exemplifying theory, in contrast, employs specialized assumptions to show what can happen under particular

---

[16] See Appropriate Framework for Broadband Access to the Internet over Wireline Facilities, Report and Order and Notice of Proposed Rulemaking, 20 FCC Rec 14853, 14904 ¶ 96 (2005); sources cited in note 2.

[17] OECD Report, Internet Traffic Prioritisation: An Overview 5 (Apr. 6, 2007), available at <http://www.oecd.org/dataoecd//43/63/38405781.pdf> (last visited May 9, 2008).

[18] Ex parte Filing of the Department of Justice, WC Docket No. 07-52 (Sept. 6, 2007), available at <http://www.usdoj.gov/atr/public/comments/225767.pdf> (last visited May 9, 2008).

[19] See Federal Trade Commission, Staff Report on Broadband Connectivity Competition Policy (cited in note 3).

[20] See Broadband Industry Practices, Notice of Inquiry, 22 FCC Rec 7894 (2007).

[21] See Daniel F. Spulber and Christopher S. Yoo, *Mandating Access to Telecom and the Internet: The Hidden Side of* Trinko, 107 Colum L Rev 1822, 1846-47 (2007) (discussing surveys of the empirical literature showing that vertical restraints tend to enhanced economic welfare and benefit consumers).

[22] See Christopher S. Yoo, *Beyond Network Neutrality*, 19 Harv J L & Tech 1, 52 & n 199 (2005) (collecting empirical studies concluding that mandated sharing deterred investment in new broadband networks).

[23] Franklin M. Fisher, *Games Economists Play: A Noncooperative View*, 20 RAND J Econ 113, 117 (1989).

circumstances.[24] The specificity of exemplifying theory can play an important role in isolating the effect of particular economic considerations or in serving as possibility theorems demonstrating the potential existence of particular phenomena. As such, exemplifying theory is very helpful in rebutting calls for categorical rules.

For example, my own work has traced how the Chicago School was able to use exemplifying theory to build a powerful case against treating vertical restraints on trade as per se illegal.[25] Subsequent attempts by Chicago School theorists to expand these theories into a basis for establishing a categorical rule in the other direction under which vertical restraints would be per se legal[26] prompted a series of influential post-Chicago analyses showing the existence of circumstances under which monopoly leveraging is both profitable and inefficient, again effectively rebutting calls for broad categorical rules.[27]

Another example of exemplifying theory is showing how different institutional arrangements can lower transaction costs. Transaction cost theories are often criticized for being all too easy to state, yet all but impossible to verify or falsify empirically.[28] Although there is a

---

24 Id at 117-18.

25 Christopher S. Yoo, *Vertical Integration and Media Regulation in the New Economy*, 19 Yale J Reg 171, 187-200 (2002). One could argue with considerable force that the Chicago School critique represents a form of generalizing theory. That said, the Chicago School theory of vertical exclusion acknowledged the existence of a number of exceptions (including variable proportions and evasion of rate regulation) in which vertical restraints may be both profitable and anticompetitive. Resolving whether the incorporation of these exceptions represents a sufficient departure from generalizing theory to render this critique exemplifying theory is not essential for the argument advanced here. At a minimum, the theories advanced by the Chicago School were sufficiently exemplifying to rebut the then-current doctrine treating many vertical restraints as illegal per se.

26 See Robert Bork, *The Rule of Reason and the Per Se Concept: Price Fixing and Market Division*, 75 Yale L J 373, 397 (1966) ("The thesis advanced here is that every vertical arrangement should be lawful."); Richard A. Posner, *The Next Step in the Antitrust Treatment of Restricted Distribution: Per Se Legality*, 48 U Chi L Rev 6, 22-25 (1981) ("I now think that it would be best to declare that purely vertical restraints on intrabrand competition … are legal per se."); see also Frank H. Easterbrook, *Vertical Arrangements and the Rule of Reason*, 53 Antitrust L J 135, 135 (1984) ("No practice a manufacturer uses to distribute its products should be a subject of serious antitrust attention.").

27 Yoo, 19 Yale J Reg at 202-03 (cited in 25) (reviewing the post-Chicago literature rebutting Chicago School calls for per se legality of all vertical restraints).

28 See, for example, Robert A. Pollak, *A Transaction Cost Approach to Families and Households*, 23 J Econ Lit 581, 584, n 9 (1985) ("Critics of the transaction cost approach often object that its difficult or impossible to test,

burgeoning empirical literature on transaction costs,[29] until more general patterns emerge, transaction cost analyses are unlikely to yield the broad policy inferences that characterize generalizing theory. They nonetheless remain useful in providing counterexamples that can rebut claims that a particular practice inevitably has a particular economic effect.

At the same time, the stylized nature of the assumptions on which exemplifying theories tend to be based limit them to identifying what *can* happen and prevent them from providing any insight into the likelihood that the effects they identify will actually come to pass.[30] Absent empirical support, exemplifying theory cannot provide the broad policy inferences needed to support ex ante categorical prohibitions.[31] In other words, the mere fact that a particular practice may be harmful under certain circumstances does not justify banning that practice categorically. Thus, anyone advocating broad, ex ante prohibitions of the type advocated by network neutrality proponents bears the burden of adducing empirical evidence showing that the conduct they would like to prohibit tends to harm consumers in the vast majority of cases.[32] Failing that, proponents must at least offer a generalizing theory indicating that the harm is sufficiently likely

---

refute, or falsify, claiming that it explains everything and, therefore, explains nothing."); Stanley Fischer, *Long-Term Contracting, Sticky Prices, and Monetary Policy: Comment*, 3 J Monetary Econ 317, 322 n 5 (1977) ("Transaction costs have a well-deserved bad name as a theoretical device … [in part] because there is a suspicion that almost anything can be rationalized by invoking suitably specified transaction costs.").

29 For early surveys of the empirical literature on transaction costs, see Howard A. Shelanski and Peter G. Klein, *Empirical Research in Transaction Cost Economics: A Review and Assessment*, 11 J L Econ & Org 334 (1995); and Aric Rindfleisch and Jan B. Heide, *Transaction Cost Analysis: Past, Present, and Future Applications*, 61 J Marketing 30 (1997). For a more recent survey, see Jeffrey T. Macher and Barak D. Richman, *Transaction Cost Economics: An Assessment of Empirical Research in the Social Sciences*, 10 Bus & Pol 1 (2008).

30 Fisher, 20 RAND J Econ at 118 (cited in note 23) ("Exemplifying theory does not tell us what must happen. Rather it tells us what can happen.").

31 See, for example, Alan J. Meese, *Tying Meets the New Institutional Economics: Farewell to the Chimera of Forcing*, 146 U Pa L Rev 1, 89 (1997) ("Per se rules cannot be established by exemplifying theories.").

32 See, for example, *Leegin Creative Leather Products, Inc v PSKS, Inc*, 127 S Ct 2705, 2712 (2007) ("[T]he *per se* rule is appropriate only after courts have had considerable experience with the type of restraint at issue and only if courts can predict with confidence that it would be invalidated in all or almost all instances under the rule of reason.").

to justify broad-scale remediation and that the proposed regulatory solution addresses that harm without simultaneously proscribing behavior that is potentially beneficial to consumers.[33]

The distinction between generalizing and exemplifying theory simultaneously helps frame the arguments I have advanced in the network neutrality debate and illustrate the shortcomings of the arguments advanced by my critics. My previous work has focused on rebutting calls for imposing categorical, ex ante rules mandating network neutrality by analyzing how deviating from network neutrality can yield consumer benefits by mitigating the sources of market failure that plague the telecommunications industry[34] and by enhancing network providers' ability to manage the mounting problems of congestion.[35]

At the same time, I have never advanced the claim that deviations from network neutrality are always beneficial. Indeed, the exemplifying theories I have offered are by their very nature analytically incapable of supporting such a claim. It is for this reason that I have consistently rejected categorical approaches pushing in either direction in favor of a case-by-case approach that requires a clear showing of harm to competition, or consumers, before imposing

---

33 See, for example, *Northern Pacific Railway Co v United States*, 356 US 1, 5 (1958) (holding that per se illegality should be limited to practices that exhibit such a "pernicious effect on competition and lack of any redeeming virtue" that nothing would be lost if it were presumed to be illegal "without elaborate inquiry as to the precise harm they have caused or the business excuse for their use").

34 See Yoo, 19 Harv J L & Tech at 18-53 (cited in note 22) (describing how allowing network owners to diversify their networks can allow multiple networks to survive despite the supply-side and demand-side scale economies that tend to drive markets for telecommunications services toward natural monopoly); Christopher S. Yoo, *Would Mandating Network Neutrality Help or Hurt Broadband Competition?: A Comment on the End-to-End Debate*, 3 J on Telecommun & High Tech L 23, 54-65 (2004) (showing how preventing network owners from varying their services forces them to complete solely on price and network size, which reinforces the benefits enjoyed by the largest players and thus can become the source of, rather than the solution to, market failure).

35 See Yoo, 94 Georgetown L J at 1863-900 (cited in note 1) (analyzing how deviations from the current regime of network pricing can reduce congestion to more efficient levels).

liability.[36] My work has explicitly pointed to the FCC's prompt action in the *Madison River* case[37] as suggestive of the type of regime I have in mind.[38]

My work is more fairly characterized as an attempt to strike a middle ground that protects consumers while also giving the broadband industry the flexibility it needs to experiment with new ways to meet the increasingly varied and intense demands that consumers are placing on the network. The moderateness of my proposal is underscored by the confusion that has arisen over how to characterize my position. Some scholars have focused on the fact that I favor some level of intervention and called me a proponent of network neutrality regulation.[39] Others have focused on the fact that I oppose ex ante, categorical intervention and characterized my position as deregulationist.[40]

At the same time, the distinction between generalizing and exemplifying theory underscores key weaknesses in the arguments advanced by network neutrality proponents to date. It is insufficient for network neutrality proponents to offer theoretical counter-examples of instances in which mandating network neutrality might be beneficial. The burden remains on those advocating network neutrality not just to offer exemplifying theory, but rather to offer the type of generalizing theory and empirical support necessary to support the type of sweeping, ex

---

36 See Yoo, 19 Harv J L & Tech at 7-8, 24, 75 (cited in note 22) (arguing in favor of a case-by-case approach); Yoo, 94 Georgetown L J at 1854-55, 1900, 1908 (cited in note 1) (same); Yoo, 3 J on Telecommun & High Tech L at 44-47, 58-59 (cited in note 34) (same).

37 Madison River Commc'ns, LLC, Order, 20 FCC Rec 4295 (2005).

38 See Yoo, 19 Harv J L & Tech at 67 (cited in note 22) (pointing to Madison River as an example of a case-by-case approach to network neutrality); Yoo, 94 Georgetown L J at 1855, 1900 (cited in note 1) (same).

39 See, for example, Mark A. Jamison and Janice A. Hauge, *Getting What You Pay for: Analyzing the Net Neutrality Debate* 1 (Aug 16, 2007), available at <http://ssrn.com/abstract=1081690>; Douglas A. Hass, Comment, *The Never-Was-Neutral Net and Why Informed End Users Can End the Net Neutrality Debates*, 22 Berkeley Tech L J 1565, 1569, 1593 (2007).

40 See, for example, Scott Jordan, *A Layered Network Approach to Net Neutrality*, 1 Int'l J Comm 427, 429 (2007), available at <http://ijoc.org/ojs/index.php/ijoc/article/view/168/88> (referring to me as a "[d]eregulationist" who "believe[s] that ISPs are in the best position to determine the most beneficial evolution of the Internet"); Frischmann and van Schewick, 47 Jurimetrics J at 390, 397 (cited in note 12) (claiming that I am arguing in favor of "leav[ing] it to network owners to decide how best to manage congestion on their networks, and rest assured that they will do what is sensible from a social perspective").

ante, categorical prohibitions that they have in mind. In the absence of such empirical support, the impact on consumers is ultimately ambiguous. In the face of theoretical ambiguity and in the absence of any evidence of harm to consumers, there is no justification for prohibiting any particular practices ex ante. The more appropriate course would be to adopt a regulatory that permits experimentation with different practices, but stands ready to intervene should evidence of such consumer harm emerge. Moreover, the empirical literature suggests that vertical integration or exclusivity arrangements between content and conduit are more likely to benefit consumers than harm them.

I. THE FORCES DRIVING THE NETWORK NEUTRALITY DEBATE

A number of technological developments and considerations have added new dimensions to the debate over network neutrality. The emergence of internet video is raising the prospect of a dramatic increase in the growth rate of internet traffic. The growing importance of peer-to-peer technologies also raises significant policy implications. Network providers are also interconnecting in ways and entering into business relationships that are increasingly diverse. Lastly, the controversy surrounding Comcast's treatment of BitTorrent traffic is forcing policymakers to confront variations in the ways that congestion impacts different transmission technologies.

A. The Emergence of Internet Video

Although some industry leaders have occasionally offered somewhat hyperbolic statements about the rate at which internet traffic is expanding,[41] in recent years, internet traffic appears to have settled into a pattern of rapid, but reasonably stable growth. After growing at a rate of 100 percent per year from the early 1990s until about 2002 (not including the ten-fold increase in traffic between 1995 and 1996), internet growth has stabilized at an annual rate of roughly 50-60 percent.[42]

At the same time, reports have begun to appear predicting that the widescale deployment of internet video technologies will cause traffic growth to approach pre-2002 levels.[43] Some estimate that YouTube traffic already constitutes 10 percent of all internet traffic.[44] Other video-based technologies, such as internet distribution of movies (currently being deployed by Netflix), graphics-intensive online games (such as World of Warcraft) and virtual worlds (such as Second Life), and internet protocol television ("IPTV") (currently being deployed by AT&T) are emerging as well.[45] The ongoing transition of high definition television is likely to cause demand to increase still further.[46] Thus, some industry observers predict that video traffic will

---

[41] Univ of Minn Digital Tech Ctr, Minnesota Internet Traffic Studies (Aug 30, 2007), available at <http://www.dtc.umn.edu/mints/home.html> (collecting claims that internet traffic is growing at rates of between 100% and 500% each year).

[42] Univ of Minn Digital Tech Ctr, Internet Growth Trends and Moore's Law (Aug 30, 2007), available at <http://www.dtc.umn.edu/mints/igrowth.html>; see also Cisco Systems, Global IP Traffic Forecast and Methodology, 2006-2011, at 1 (White Paper Jan 14, 2008), available at <http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/net_implementation_white_paper0900aecd806a81aa.pdf> (predicting that internet traffic will grow at a 46% annual rate between 2007 and 2011).

[43] Bret Swanson and George Gilder, *Estimating the Exaflood: The Impact of Video and Rich Media on the Internet* 22 (Jan 2008), available at <http://www.discovery.org/scripts/viewDB/filesDB-download.php?command=download&id=1475> (estimating that internet traffic growth might grow at an annual rate of 90% until 2015).

[44] See Ellacoya Networks, Inc., Press Release, Ellacoya Data Shows Web Traffic Overtakes Peer-to-Peer (P2P) as Largest Percentage of Bandwidth on the Network (June 18, 2007), available at <http://www.ellacoya.com/news/pdf/2007/NXTcommEllacoyamediaalert.pdf>.

[45] See Swanson and Gilder, *Estimating the Exaflood*, at 12-14 (cited in note 43).

[46] See Bret Swanson, *The Coming Exaflood*, Wall St J A11 (Feb 20, 2007).

constitute over 80 percent of all internet traffic by 2010.[47] An oft-cited study by Nemertes Research predicts that video will cause the rate of traffic growth to accelerate beyond the current pace of investment in new network capacity. This means that traffic growth will exhaust the usable network capacity by 2010 unless the world increases its rate of capital expenditures on upgrading the network infrastructure by over $100 billion.[48] Even network neutrality proponents, such as Google and EDUCAUSE, have warned that the internet will struggle to accommodate consumers' increasing demands for bandwidth.[49]

To date, there are no signs that this so-called "exaflood"[50] has begun to materialize. The conflicting reports about the possible acceleration in the rate of internet traffic pose a dilemma for network providers, who must begin plans to expand capacity well in advance of any increase in demand. If they follow the higher estimates, they may end up investing tens of billions of dollars in unnecessary network capacity. Following such an approach would slow national broadband deployment in higher-cost areas by taking up scarce capital and by increasing the number of customers needed for broadband service to break even in any particular area. If they follow the lower estimates, they risk seeing congestion cause their networks to slow to a crawl.

---

47 See William B. Norton, *Video Internet: The Next Wave of Massive Disruption to the U.S. Peering Ecosystem* (v0.91) at 2 (Sept 29, 2006), available at <http://www-tc.pbs.org/cringely/pulpit/media/internetVideo0.91.pdf>.

48 See Nemertes Research, *The Internet Singularity, Delayed: Why Limits in Internet Capacity Will Stifle Innovation on the Web* 31, 45 (Fall 2007), available at <http://www.nemertes.com/system/files/internet+Singularity+Delayed+Fall+2007.pdf>.

49 See *Internet Not Designed for TV, Google Warns*, PC Mag (Feb 8, 2007) (quoting Google head of TV technology Vincent Dureau as stating at the Cable Europe Congress, "The web infrastructure and even Google's doesn't scale. It's not going to offer the quality of service that consumers expect."); John Windhausen Jr., A Blueprint for Big Broadband 7-11 (EDUCAUSE White Paper Jan 2008) (also quoting studies by Jupiter Research and Technology Futures), available at <http://www.educause.edu/ir/library/pdf/EPO0801.pdf>.

50 The term, "exaflood," to describe the prospect of a video-driven acceleration in the growth rate of internet traffic appears to have been coined initially by Bret Swanson and George Gilder. Bret Swanson, *The Coming Exaflood*, Wall St J A11 (Feb 20, 2007); Bret Swanson and George Gilder, *Unleashing the "Exaflood"*, Wall St J A15 (Feb 22, 2008); see also Bruce Mehlman and Larry Irving, *Bring on the Exaflood!: Broadband Needs a Boost*, Wash Post A31 (May 24, 2007).

Furthermore, although the increase in traffic is an important development, as I have detailed in my previous work, congestion is a complex phenomenon that depends on more than just total volume.[51] As I shall discuss in greater detail below, it also depends on the timing, location, and pattern of overall network traffic. In addition, networks' ability to compensate for increases in demand by rerouting traffic can make network performance quite unpredictable. Thus, a disruption in one portion of the network can increase congestion in areas of the network located far from the point of disruption.

The uncertainty over the rate and location of traffic growth has placed greater importance on network management. Specifically, network management represents an important alternative to expanding capacity that serves as a safety valve to relieve network congestion when expanding capacity is not an option. In this sense, capacity expansion and network management are more properly regarded as alternative approaches to deal with the problem of congestion. Which will be preferable in any particular case will vary with the circumstances and with their relative costs. It is difficult, if not impossible, to determine *a priori* which will prove the better solution at any particular moment. The relative costs of each solution are also likely to change over time, so any precommitment to one approach over the other would likely have to undergo constant oversight and revision as the underlying technology evolves.

B. The Growth of Peer-to-Peer Technologies

Another force driving the network neutrality debate is the growing importance of peer-to-peer technologies. Although the term "peer-to-peer" is often viewed as synonymous with file sharing or user-generated content, it actually embodies a more fundamental distinction. In the

---

[51] See Daniel F. Spulber and Christopher S. Yoo, *On the Regulation of Networks as Complex Systems: A Graph Theory Approach*, 99 Nw U L Rev 1687 (2005).

traditional internet architecture, content and other files are stored in large computers at centralized locations, known as “servers”. End users, known as “clients,” request files from those servers, usually by submitting a short bit of code such as a website address, also known as a uniform resource locator (URL). The server that hosts the requested files then transmits the requested files through the internet to the client.

In a peer-to-peer architecture, files are not stored in centralized locations, and the computers that are connected to the edge of the network are not divided into clients requesting files and servers hosting files. Instead, files are distributed across the network, and edge computers simultaneously request files and serve files. It is this less hierarchical structure that leads these types of edge computers to be called “peers” and this type of service to be called peer-to-peer. That peer-to-peer and user-generated content are analytically distinct is underscored by the fact that YouTube and many other repositories of user-generated content employ client-server architectures, while Vuze and other distributors of commercial media content employ peer-to-peer architectures.

Whether a network is comprised primarily of clients and servers or of peers has major architectural implications. If a network is organized around a client-server architecture, the traffic flowing from the server to the client tends to be greater than the traffic flowing in the other direction. As a result, it usually makes sense to divide the available bandwidth asymmetrically by devoting a greater proportion of the available bandwidth to downloads and a smaller proportion to uploads. Such asymmetry makes less sense if a network is organized around a peer-to-peer architecture, since each end user represents an important source of upload traffic as well as download traffic.

At the time that network providers established the basic architectures for the major broadband technologies in the late 1990s, the internet was dominated by applications such as web browsing and email that adhered to a client-server architecture. As a result, most network providers assigned bandwidth asymmetrically, devoting a greater proportion of the available bandwidth to downloading rather than uploading. For example, the dominant telephone-based technology is asymmetric digital subscriber line ("ADSL") service, which initially supported theoretical speeds of up to 8 Mbps for downloading and 768 kbps for uploading.[52] More recent versions of ADSL support higher bandwidth but still allocate it asymmetrically.[53] The initial cable modem architecture, designed around DOCSIS 1.0, supported maximum theoretical speeds of 27 Mbps downstream and 10 Mbps upstream.[54] Finally, the service offered by wireless providers deploying EV-DO technologies is similarly asymmetrical, with download rates exceeding upload rates by a ratio of eleven to one.[55]

Although some network neutrality proponents have criticized those decisions as "short-sighted" or "poor network design decisions,"[56] those decisions were quite rational at the time they were made. Since that time, network providers have begun developing new symmetric technologies, such as DOCSIS 2.0 for cable modem systems and symmetric DSL (SDSL) for wireline systems. DOCSIS 3.0 retains a degree of asymmetry, but to a lesser degree than

---

[52] See Inquiry Concerning the Deployment of Advanced Telecommunications Capability to All Americans in a Reasonable and Timely Fashion, and Possible Steps to Accelerate Such Deployment Pursuant to Section 706 of Telecommunications Act of 1996, Fourth Report, 19 FCC Rec 20540, 20558 (2004) [hereinafter Fourth Section 706 Report].

[53] See DSL Forum, About ADSL, available at <http://www.dslforum.org/learndsl/adslfaq.shtml> (last visited Feb 24, 2008).

[54] See Inquiry Concerning the Deployment of Advanced Telecommunications Capability to All Americans in a Reasonable and Timely Fashion, and Possible Steps to Accelerate Such Deployment Pursuant to Section 706 of Telecommunications Act of 1996, Third Report, 17 FCC Rec 2844, 2917-18 ¶ 21 (2002).

[55] See Advanced Wireless Services in the 2155-2175 MHz Band, Notice of Proposed Rulemaking, 22 FCC Rec 17035, 17046 n 32 (2007).

[56] Comments of Free Press et al, WC Docket No 07-52, at 21, 22 (2008), available at <http://fjallfoss.fcc.gov/prod/ecfs/retrieve.cgi?native_or_pdf=pdf&id_document=6519841216> (last visited May 9, 2008).

DOCSIS 1.0. Very-High-Data-Rate DSL (VDSL) supports both symmetric and asymmetric services.

Indeed, even now it is far from clear whether a symmetric or an asymmetric architecture will eventually prove to be the better choice. For the four years preceding 2007, peer-to-peer traffic surpassed client-server traffic in terms of percentage of total bandwidth.[57] A remarkable change occurred in 2007. Client-server traffic began to reassert itself, driven primarily by the expansion of streaming video services, such as YouTube. Thus, in 2007, client-server traffic has retaken the lead from peer-to-peer, constituting 45 percent of all internet traffic as compared with 37 percent of all traffic devoted to peer-to-peer.[58]

The growing importance of peer-to-peer technologies affects the shape as well as the volume of the upload traffic. In many ways, the effect is similar to the transformation that occurred when internet users began to use dial-up modems attached to conventional telephone lines. Before the emergence of the internet, the typical telephone call lasted only three to five minutes.[59] Because calls were relatively short and different customers tended to make calls at different times, telephone companies were able to provide adequate service by providing enough switching capacity to accommodate one fourth to one eighth of all customers at any particular time. Dial-up internet calls, in contrast, tend to be longer, lasting approximately fifteen to twenty minutes.[60] Although different parties disagreed as to the magnitude of the problem, no one

---

57 See Ellacoya Networks, Press Release (cited in note 44).

58 See id.

59 Kevin Werbach, Digital Tornado: The Internet and Telecommunications Policy 58 (Fed Commc'ns Comm'n Off of Plans & Pol Working Paper No. 29, Mar 1997), available at <http://www.fcc.gov/Bureaus/OPP/working_papers/oppwp29.pdf> (last visited May 16, 2008).

60 Id at 59.

doubted that increasing the length of internet calls required greater switching and trunk line capacity.[61]

The same problem arises as traffic shifts to peer-to-peer technologies. In a client-server architecture, where upload traffic from end users consists of URLs requesting files, short duration of the traffic relative to its frequency allows network providers to easily serve a large number of customers with only limited upload capacity. But the longer session times associated with peer-to-peer traffic reduce the ability of multiple end users to share the same bandwidth. The effect is to require networks to dedicate more capacity to serve the same number of users. In addition, in a peer-to-peer architecture end users do not simply upload URLs. Instead, they upload as well as download files. Because broadband networks allocate more bandwidth to downloading than to uploading, the emergence of peer-to-peer architectures is making it more difficult for last-mile broadband providers to ensure adequate upload speeds and quality of service.

The shift to peer-to-peer also effectively increases the number of hours in a day that any particular computer can generate upload traffic. In a client-server architecture, the amount of time that any individual could sit in front of a computer placed a natural limit on the amount of upload bandwidth that any one subscriber could consume. In a peer-to-peer architecture, however, any computer that is left running can continue to generate upload traffic even when no person is present. The result is that the lion's share of upload traffic is generated by a small

---

61 Dennis W. Moore, Jr., Note, *Regulation for the Internet and Internet Telephony Through the Imposition of Access Charges*, 76 Tex L Rev 183, 196-97 (1997).

number of superheavy peer-to-peer users. As few as 5 percent of end users may be responsible for generating 50 percent of all internet traffic.[62]

The emergence of peer-to-peer has also placed increased pressure on network providers to change their pricing models. Under the business model that currently dominates the internet, end users are generally charged on an "all you can eat" basis, in which end users can consume an unlimited amount of services for a flat monthly fee.[63] Content and applications providers are charged prices that vary with the amount of traffic they generate, typically pegged to the peak traffic that they generate over a thirty-day period.[64] Under a client-server architecture, this pricing regime did provide some basis for charging users for the amount of congestion they contributed to the network. Since every single download required action by a content provider, the amount of traffic downloaded by any particular content provider's server represented a somewhat effective measure of the amount of congestion that that particular content provider was imposing on the overall network.

This is not the case under a peer-to-peer architecture, in which a single download from a content provider could generate an untold number of additional downloads without increasing the amount that the content provider would have to pay. As a result, the amount of traffic generated directly from the content provider no longer represents an accurate reflection of the amount of congestion imposed on the entire network. Instead, much of the download traffic is

---

62 See David Vorhaus, *Confronting the Albatross of P2P*, at 1 (Yankee Group, May 31, 2007) (noting that 5% of users account for 50% of all traffic); Steven Levy, *Pay per Gig*, Wash Post D1 (Jan 30, 2008) (quoting Time Warner Cable spokesman offering similar statistics); See Comments of CTIA – The Wireless Association, WC Docket No 07-52, *12 (Feb 13, 2008) available at <http://gullfoss2.fcc.gov/prod/ecfs/retrieve.cgi?native_or_pdf=pdf&id_document=6519841180> (noting that one wireless provider reports that less than 5% of customers generate over 50% of traffic) (last visited May 16, 2008).

63 See J. Gregory Sidak and Daniel F. Spulber, *Cyberjam: The Law and Economics of Internet Congestion of the Telephone Network*, 21 Harv J L & Pub Pol 327, 340 (1998) (describing how America Online's introduction of all you can eat pricing in December 1996 shifted the industry away from metered pricing).

64 *95th Percentile Explained*, Innovative Network Concepts, available at <http://inconcepts.biz/cr/95th.html> (last visited May 9, 2008).

shifted to end users located throughout the network. The obvious solution to this problem is to make the prices paid by end users somewhat reflective of the upload traffic that they generate. However, when Time Warner attempted to impose such a regime in January 2008, it was greeted by a torrent of criticism from the network neutrality community.[65]

The emergence of new peer-to-peer technologies, such as BitTorrent, changes the calculus still further. Rather than retrieving a requested file from a single location, BitTorrent retrieves portions of the requested file from multiple computers. By reducing the size of the file that must be uploaded from any particular computer, this ingenious approach has the potential to improve the efficiency of bandwidth use dramatically by lessening the amount of capacity required from any particular location. In addition, BitTorrent readjusts the locations from which it receives files dynamically. If one particular location is running slowly, it can readjust its request to obtain the portion of the file requested from that location to another location.

This process of dynamic readjustment also gives BitTorrent a "swarming" quality that places the biggest burden on the locations with the fastest connections. As I shall explain in further detail in the next Part, this burden falls particularly heavily on technologies such as cable modem and wireless broadband providers in which end users share bandwidth with their neighbors from the moment their traffic leaves their house. Indeed, studies indicate that congestion becomes problematic when as few as fifteen of the five hundred or so cable modem subscribers sharing the same fiber node run peer-to-peer filesharing programs.[66] It is for this

---

[65] See note 9 and accompanying text.

[66] See James J. Martin and James M. Westall, *Assessing the Impact of BitTorrent on DOCSIS Networks*, in Proceedings of the 2007 IEEE Broadnets (Sept 2007), available at <http://people.clemson.edu/~jmarty/papers/bittorrentBroadnets.pdf> (last visited Apr 19, 2008). See also Leslie Ellis, *BitTorrent's Swarms Have a Deadly Bite on Broadband Nets*, Multichannel News (May 8, 2006), available at <http://www.multichannel.com/article/CA6332098.html> (last visited Apr 19, 2008).

reason that dozens of internet service providers ("ISPs") around the world restrict BitTorrent traffic in some way.[67]

Network providers thus confront a difficult decision. Not only must they determine the size and the location of the capacity to add. They must also determine the extent to which they should continue to embrace an asymmetric architecture based on their projections of the likely future success of applications such as BitTorrent and YouTube. Any imperfections in their projections are likely to have significant economic consequences.

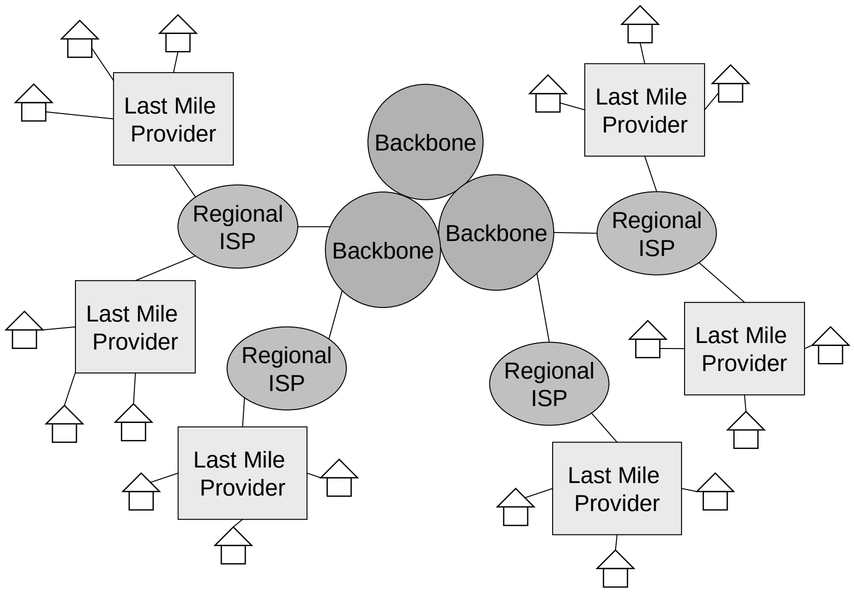
C. The Increasing Heterogeneity in Business Relationships Among Network Providers

Internet service providers have traditionally been divided into three categories. *Backbone providers* occupy the center of the network and offer high-speed transport between roughly a dozen locations spread throughout the country.[68] *Regional ISPs* carry traffic from the network access points served by backbone providers to the local distribution facilities maintained by last-mile providers in individual cities (which in the case of DSL is usually called a central office and in the case of cable modem systems is usually called a headend).[69] The final connection is provided by *last-mile providers*, which use grids of wires or local networks of wireless spectrum to carry the traffic from those central facilities to end users.[70]

---

[67] Azureus Wiki, available at <http://www.azureuswiki.com/index.php/Bad_ISPs> (last visited Apr 19, 2008).
[68] Yoo, 94 Georgetown L J at 1860 (cited in note 1).
[69] Id.
[70] Id.

Figure 1
The Original Structure of the Internet



During these early days, each regional ISP maintained a business relationship with precisely one backbone, through which it exchanged all of its traffic that it could not terminate locally. The one-to-one relationship dictated that there was typically only one path for connecting any two points. The uniqueness of the connections made the network quite vulnerable to congestion. It also made the internet quite hierarchical, with the backbones playing a role in transmitting the vast majority of traffic, which in turn provided backbones with a potential source of market power.

In addition, the business relationships were relatively simple. The largest backbones exchanged traffic through a system known as peering. Rather than metering and billing each other for the traffic they exchanged, top level backbones exchanged traffic on a settlement free basis in which no money changed hands. So long as the volume of traffic passing in each direction is roughly symmetrical, both backbones will be in roughly the same economic condition as they would have been had they metered and billed each other for the traffic they exchanged.[71]

Peering is not economical in cases where the value of the traffic being terminated is not reciprocal. As a result, smaller-volume backbones are often required to enter into "transit" arrangements in which they must pay larger backbones to terminate their traffic.[72]

Over time, these business relationships began to become more heterogeneous. Backbones began to enter into private interconnection agreements. This allowed them avoid the congestion at the network access points. The bilateral nature of the exchange also made it easier for them to manage quality of service.[73] At the same time, backbones began entering into paid peering relationships to compensate networks that were providing greater value.[74]

---

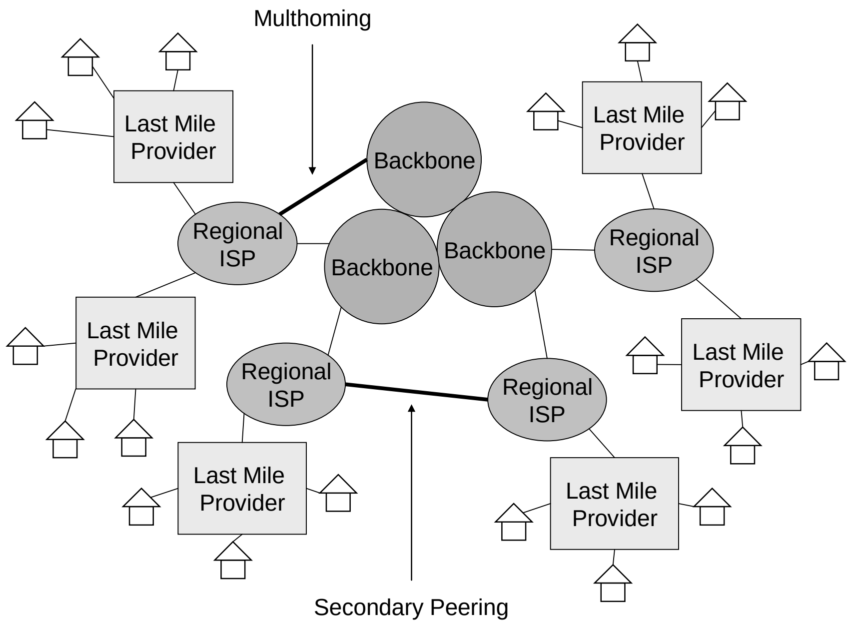
71 Id at 1877. Peering also involves a number of characteristics aside from settlement free termination. For example, peering partners engage in *hot potato routing*, in which they hand off traffic at the first mutual point of interconnection. Id at 1871 n 113. In addition, backbones can only peer traffic that they can terminate themselves. Any traffic that must be handed off to another backbone for termination must travel via transit. Michael Kende, The Digital Handshake: Connecting Internet Backbones 5 (FCC Off of Plans and Pol Working Paper No 32, Sept 2000), available at <http://www.fcc.gov/Bureaus/OPP/working_papers/oppwp32.pdf> (last visited May 9, 2008).

72 Yoo, 94 Georgetown L J at 1877 (cited in note 1).

73 P. Faratin et al, *Complexity of Internet Interconnections: Technology Incentives and Implications for Policy* 28 (paper presented at the 35th Annual Telecommunications Policy Research Conference), available at <http://web.si.umich.edu/tprc/papers/2007/797/Clark%20Lehr%20Faratin%20Complexity%20Interconnection%20TPRC%202007.pdf> (last visited Mar 6, 2008).

74 Id at 14; OECD Working Party on Telecommunications and Information Services Policies, Internet Traffic Exchange: Market Developments and Measurement of Growth 21-22 (Apr 5, 2006), available at <http://icttoolkit.infodev.org/en/Publication.3081.html> (last visited May 9, 2008).

Figure 2
The Emergence of Secondary Peering and Multihoming



Regional ISPs also began to connect to more than one backbone, a practice that became known as multihoming.[75] Regional ISPs that were too small to peer with the top-tier backbones also began to economize on transit charges by entering into secondary peering relationships, in which regional ISPs bypass the tier-1 backbones altogether and exchange traffic with each other on a settlement free basis.[76] These changes had several benefits. The avoidance of transit charges reduced the costs borne by end users. Secondary peering and multihoming also made the network more robust by creating additional paths connecting particular points.[77] In fact, as much as seventy percent of the nodes in the internet can communicate with one another without passing through the public backbone.[78] This had the additional benefit of weakening the market

---

75 See Yoo, 94 Georgetown L J at 1871 n 111 (cited in note 1).

76 See id at 1872.

77 See Shai Carmi et al, *A Model of Internet Topology Using* k-*Shell Decomposition*, 104 Proc of the Natl Acad of Sci 11150, 11151 (2007).

78 See id.

position of the top-tier backbones.[79] It did mean greater variance in the price paid by different types of traffic.

Furthermore, some content and application providers began to use content delivery networks like Akamai, which reportedly handles over fifteen percent of the world's web traffic.[80] Akamai caches web content at over fourteen thousand locations throughout the internet. When an end user sends a request for a webpage, the last-mile broadband provider checks to see whether that webpage is hosted by Akamai. If so, the last-mile provider redirects the query to the cache maintained by Akamai. This process often allows the resulting traffic to bypass the public backbone altogether.[81]

The sheer number of caches all but guarantees that the closest Akamai cache will be located closer to the end user than the server hosting the primary webpage. As a result, content served by Akamai is less likely to be plagued by problems of latency.[82] In addition, the redundancy in Akamai's server network not only insulates the content Akamai hosts from denial of service attacks; it also allows the system to redirect queries to other caches when particular caches are overly congested.[83] All of these developments represent innovative solutions to adjust to the realities of the internet. It means, however, that different providers often pay different amounts for similar services depending on the precise path taken through the network.

---

79 See Yoo, 94 Georgetown L J at 1872 (cited in note 1).

80 See id at 1882-83 ("The leading content delivery network, known as Akamai, reportedly maintains more than fourteen thousand servers and handles more than fifteen percent of the world's web content.").

81 Id.

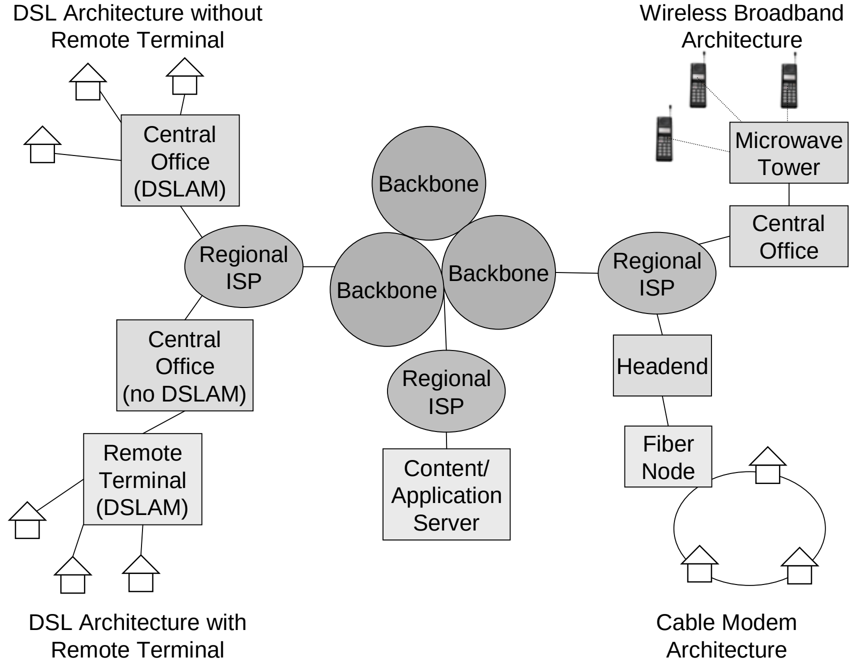
82 Id.

83 See Yoo, 94 Georgetown L J at 1872 (cited in note 1).

D. Variations in the Ways Congestion Can Arise in Different Transmission Technologies

Network providers must base their investment plans on their projections of the magnitude, location, and shape of the traffic that they will have to support. In the local transmission portions of the network, moreover, the various broadband technologies differ widely in their susceptibility to congestion.

Figure 3
Architectures of the Major Broadband Transmission Technologies



Consider first the architecture of DSL. DSL customers typically use a pair of copper wires dedicated exclusively to them to connect to the nearest central office, in which the telephone company maintains a piece of equipment known as a DSL access multiplexer (DSLAM) to separate the voice traffic from the data traffic. Because DSL customers connect to the DSLAM through a dedicated connection, their traffic is not typically aggregated with other traffic until it reaches the central office.[84] As a result, the local connection between DSL customers' premises and the central office is not subject to congestion at the neighborhood level. The primary constraint is that modern ADSL can only serve customers located within eighteen thousand feet (roughly three and a half miles) of a DSLAM.[85] To serve customers located more than three and a half miles from a central office, local telephone companies sometimes deploy DSLAMs in satellite facilities known as remote terminals, which are in turn connected to the central office through optical fiber.[86] AT&T is deploying a higher speed DSL technology known as very-high-speed DSL (VDSL) that requires the placement of remote terminals within two to four thousand feet of every customer.[87] Because DSL customers have dedicated connections to the DSLAM, their traffic is not aggregated with other traffic until it reaches the remote terminal. As a result, DSL customers do not share bandwidth with other customers in the link between their premises and the remote terminal, and thus that portion of the network is not subject to congestion.[88]

The situation is quite different in cable modem systems, which are based on a hybrid fiber coaxial (HFC) architecture. Under an HFC architecture, the copper coaxial cables connecting individual customers' premises are reconfigured into a ring configuration and connected to a satellite facility known as a neighborhood node. The node is in turn connected by optical fiber to the headend.[89] Unlike under DSL, traffic generated by individual cable modem customers shares bandwidth with the traffic generated by their neighbors from the moment it

---

84 See Daniel F. Spulber and Christopher S. Yoo, *Access to Networks: Economic and Constitutional Connections*, 88 Cornell L Rev 885, 1003-04 (2003).

85 See Yoo, 19 Yale J Reg at 255 (cited in note 25).

86 See William P. Rogerson, *The Regulation of Broadband Telecommunications, the Principle of Regulating Narrowly Defined Input Bottlenecks, and Incentives for Investment and Innovation*, 2000 U Chi Legal F 119, 141-42.

87 See Sean Buckley, *There's Gold in That Copper*, Telecommun Intl 19 (Jan 1, 2007).

88 See Yoo, 94 Georgetown L J at 1862 n 71 (cited in note 1).

89 See Spulber and Yoo, 88 Cornell L Rev at 1014-15 (cited in note 84).

leaves their house.[90] As a result, the quality of service that any particular cable modem customer receives is considerably more sensitive to the bandwidth consumption of their immediate neighbors.

The congestion problems confronted by wireless broadband providers are even more severe. Wireless broadband providers connect to the internet through transponders located on microwave towers and other high-altitude locations. Because the capacity of any one transponder is limited, customers attempting to connect to the same tower compete for bandwidth with their neighbors.[91] Thus, like cable modem service, wireless broadband service is sensitive to local congestion.

This problem is exacerbated in the case of wireless broadband by two other considerations. First, wireless broadband operates under bandwidth constraints that are much more restrictive than those faced by DSL or cable modem systems.[92] Second, in DSL and cable modem systems, broadband traffic is carried in a different channel than traffic associated with the other services provided by the company. For example, in the case of DSL, conventional voice traffic is transmitted through a different channel than data traffic.[93] Similarly, in a cable network, conventional video traffic is transmitted through a different channel than data traffic.[94] Thus, broadband traffic cannot degrade the quality of service of telephone and cable companies' core businesses no matter how much it increases. This is not true in the case of wireless. Wireless broadband shares bandwidth with the voice services offered by wireless companies. Consequently, any congestion that may arise in a wireless network degrades not only the quality

---

90 See Yoo, 94 Georgetown L J at 1862 n 71 (cited in note 1).
91 See Comments of CTIA – The Wireless Association, WC Docket No 07-52, at 7 (cited in note 62).
92 See id at 9.
93 See id at 7.
94 See id.

of internet broadband services provided; it also degrades the conventional voice services that represent the wireless providers' core business.[95]

It should thus come as no surprise that different types of providers vary in their tolerance for local congestion, with some taking more aggressive efforts to manage it and some taking less. It should also come as no surprise that different types of providers would manage congestion on a different geographic scale, depending on the nature of their technology. These technological realities caution strongly against adopting a one-size-fits-all approach to network management. Indeed, any regulatory solution that might be imposed must be carefully tailored to take these important variations into account.

II. POTENTIAL CONSUMER BENEFITS FROM DEVIATING FROM NETWORK NEUTRALITY

An examination of the literature on the economics of regulation provides both supply-side and demand-side considerations showing how deviations from network neutrality might provide net benefits to consumers. I begin by reviewing the supply-side justifications and follow that by examining the demand-side justifications. This section concludes by examining the relevance of the literature on two-sided markets, which incorporates both demand-side and supply-side considerations.

95 See Comments of CTIA – The Wireless Association, WC Docket No 07-52, at 7 (cited in note 62).

A. Supply-Side Justifications for Prioritization and Differential Pricing

1. Maximizing Consumer Welfare in the Presence of Congestion

a. The Role of Congestion-Based Pricing

When the internet first emerged, it provided only a single class of service and employed relatively simple pricing schemes with respect to both end users and content and applications providers. With respect to end users, although some internet service providers initially charged end users on a per-minute basis, "all you can eat" pricing, in which end users could consume an unlimited amount of services for a single monthly fee, soon emerged as the industry standard.[96] Network providers typically charge content and application providers fees related to their usage. In particular, they typically sample the bandwidth used every five minutes and charge the content or applications provider based on their peak usage over a thirty-day period.[97] In order to avoid penalizing content and applications providers for short-run, transient surges in traffic, they typically base the charge on the ninety-fifth percentile of traffic, which effectively excuses the thirty-six hours with the heaviest bandwidth use.[98]

The relative simplicity of these pricing schemes harms consumers in at least two ways. First, a network that only charges end users a uniform, all you can eat price is likely to set its price to reflect the amount of bandwidth consumed by the average user. Such a regime represents a windfall to end users with above-average levels of consumption of network services. At the same time, it overcharges end users whose consumption of network services falls below the average. The net effect of having a single class of service is to force low-volume users to pay

---

96 See note 63 and accompanying text.
97 See note 64 and accompanying text.
98 Id.

for more bandwidth than they need, which may force some of them to forego subscribing to the internet even though the benefits they would derive from doing so would exceed the costs. It also has the effect of forcing low-volume users to cross subsidize high-volume users.[99]

Second, as I have discussed elsewhere at length, all you can eat pricing schemes tend to induce excessive levels of congestion.[100] Congestion can arise at any one of a number of points in the network. As an initial matter, congestion can arise in the last-mile broadband network that connects the end users' premises to the central facilities maintained by local broadband providers. In the case of a cable modem system, the facility is called a headend, and in the case of DSL, it is called a central office.[101] Congestion can also arise within the regional ISP that connects the local network to the public backbone, the backbone itself, or the regional ISP or last-mile provider on the terminating end. Lastly, congestion can arise in the content server being accessed.[102]

The congestibility of the internet dictates that network performance depends in no small part on the volume being generated by other end users at any particular time. Put a different way, every end user's usage imposes congestion costs on all other end users. If the network is operating well-below capacity, the congestion costs may be negligible. If the network is operating close to capacity, the congestion costs may be significant.

---

99 Yoo, 94 Georgetown L J at 1853-54, 1855, 1877 (cited in note 1) (describing how flat-rate pricing "forces low-volume users to cross subsidize those who place more intensive demands on the Internet").

100 Id at 1864-65 (showing how "flat-rate pricing results in excessive consumption of club resources, which arises because the congestion costs represent a negative externality that individual club members responsible for causing the congestion are not forced to bear").

101 Yoo, 94 Georgetown L J at 1860-62 (cited in note 1)
Cable modem systems are particularly susceptible to local congestion, given that local traffic in cable modem systems is first aggregated in a neighborhood facility known as a fiber node. As many as five hundred households share bandwidth provided by a fiber node.

102 See id at 1862-63 (cited in note 1) (describing how congestion can arise at each of these points in the network).

Aggregate consumer welfare increases when end users increase their usage levels if and only if the benefits they would derive from doing so exceed the congestion costs they would impose on other users. If the costs exceed the benefits, consumers would be better off if end users would refrain from increasing their usage.[103]

The problem is that the pricing regimes that dominate the internet fail to give end users an incentive to behave in the way that maximizes consumer welfare. Specifically, under all you can eat pricing, the cost of increasing usage is always zero. End users thus have the incentive to continue increasing their consumption even when the benefits they derive begin to approach zero. The problem is that the congestion costs associated with that increased usage imposed on other end users are greater than zero. The fact that individual end users do not internalize the congestion costs they impose on others causes them to continue to increase their consumption even when doing so would reduce consumer welfare. Although the amount of time any one person could spend in front of a computer once placed a natural limit on the amount of bandwidth that any particular end user could consume, modern peer-to-peer technologies can adjust dynamically to allow network usage to expand to fill all available capacity.[104]

---

[103] Frischmann and van Schewick criticize me for failing to recognize that internet usage creates positive as well as negative externalities. Frischmann and van Schewick, 47 Jurimetrics J at 398-403 (cited in note 12). The focus of their criticism misunderstands the nature of my argument as well as the proper role of exemplifying theory. As I noted earlier, in showing how the negative externalities is associated with congestion can cause excessive consumption of network resources, I was simply presenting an exemplifying theory to rebut calls for per se rules categorically prohibiting certain practices by showing circumstances under which they actually benefit consumers. As I have noted earlier, offering exemplifying theories pointing in the other direction does not reestablish the case for per se illegality. Absent empirical evidence regarding the likelihood of which externalities will dominate, the existence of potential, offsetting externalities and the possibility that the market is not in general equilibrium renders the net impact of those externalities on consumer welfare ambiguous, making it impossible to draw any a priori inferences about the practice's likely consumer impact. See Christopher S. Yoo, *Copyright and Public Good Economics: A Misunderstood Relation*, 155 U Pa L Rev 635, 685-86 & n 163 (2007). Thus, if anything, the existence of exemplifying theories pointing in both directions actually supports the case-by-case approach that I have advanced over the categorical prohibitions favored by network neutrality proponents.

[104] See notes 66-67 and accompanying text.

The classic solution to this problem is to set the cost of incremental usage equal to the congestion costs that that usage would impose on other users. Perfect congestion-based pricing would cause users to internalize the costs they impose on others and in so doing would provide them with the incentives to calibrate their network usage at the level that would maximize the welfare of all consumers. As a theoretical matter, perfect congestion-based pricing would maximize the aggregate benefits enjoyed by all consumers.

Congestion problems can also arise from the way certain content providers design their websites. For example, ESPN has configured its website to download video content in the background automatically.[105] The result is that the website will consume significant bandwidth completely outside the end user's control. Again, these problems could theoretically be solved by imposing congestion-based pricing on content and applications providers as well.

b. Difficulties in Implementing Congestion-Based Pricing

The problem is that true congestion-based pricing is difficult to implement. Consider first one approach suggested by some network neutrality proponents: offering consumers different service tiers.[106] Under this approach, network providers meter each end user's usage and charge for the tier of service that reflects that user's total bandwidth consumption. Network neutrality advocates have adopted different positions with respect to whether consumer tiering is consistent with network neutrality. While some recognize consumer tiering as an acceptable way

---

105 ESPN Motion Frequently Asked Questions, available at <http://espn.go.com/motion/faq.html#gen5> (last visited May 9, 2008).

106 See Tim Wu, *Network Neutrality, Broadband Discrimination*, 2 J on Telecommun & High Tech L 141, 154 (2003) (arguing in favor of offering different tiers of service instead of discriminating against particular applications).

to manage network traffic,[107] others have greeted attempts to introduced metered billing with sharp criticism.[108]

The fundamental problem is that the amount of congestion generated by any particular end user depends on more than just the total amount of bandwidth consumed. It also depends on the timing of that usage in relation to the usage patterns of all other end users. Thus, heavy bandwidth users might impose minimal congestion if they confine their usage to times when few other users are on the network. Conversely, a light bandwidth user might nonetheless become a significant source of congestion should that user choose to use the network at a time of heavy network usage. Thus, merely counting bits may represent a poor measure of congestion costs and thus may not provide sufficient incentive for individual end users to behave in a way that maximizes consumer welfare.

Another classic solution to the problems posed by congestion of timing is time-of-day or peak-load pricing.[109] Under this approach, individual end users face higher usage charges during those times of day when the overall network usage is likely to be highest. Indeed, peak-load pricing schemes should be quite familiar to those who pay lower rates for long distance calls placed in the evening and at night and to those with wireless plans that offer free night and weekend minutes.

The need to reduce congestion costs once led network providers to experiment with peak-load pricing in local telephone service, which is another service typically priced on an all you

---

107 See id; Lessig testimony 2, 9-10 (cited in note 10).

108 See note 9 and accompanying text.

109 Frischmann and van Schewick argue that peak-load pricing may represent a better second-best solution to the problems of congestion than the proxies I propose, such as restrictions on applications and content, see Frischmann and van Schewick, 47 Jurimetrics J at 396 (cited in note 12) ("Even a crude system of peak load pricing based on the time of day might suffice to effectively limit congestion; the objective from an efficiency perspective is not necessarily to internalize all congestion externalities."); id at 405 ("There are numerous ways to implement imperfect usage-sensitive pricing based on the existing technology for metering usage: peak-load pricing based on time of day may be one of them.").

can eat basis. Some network elements, most notably the copper loop connecting individual customers to the telephone company's central office, are not shared with other end users and are thus not subject to congestion. Other network elements, such as switching, are shared with other customers and thus are subject to congestion.

The presence of these congestible elements led many analysts to speculate that charging higher rates at times of high congestion would yield substantial consumer benefits and convinced some local telephone companies to experiment with a form of imposing peak-load pricing known as "local measured service." Empirical studies indicate that local telephone companies' experiments with local measured service either harmed consumers or yielded benefits that were so small that they were likely to be consumed by transaction costs of administering the system.[110]

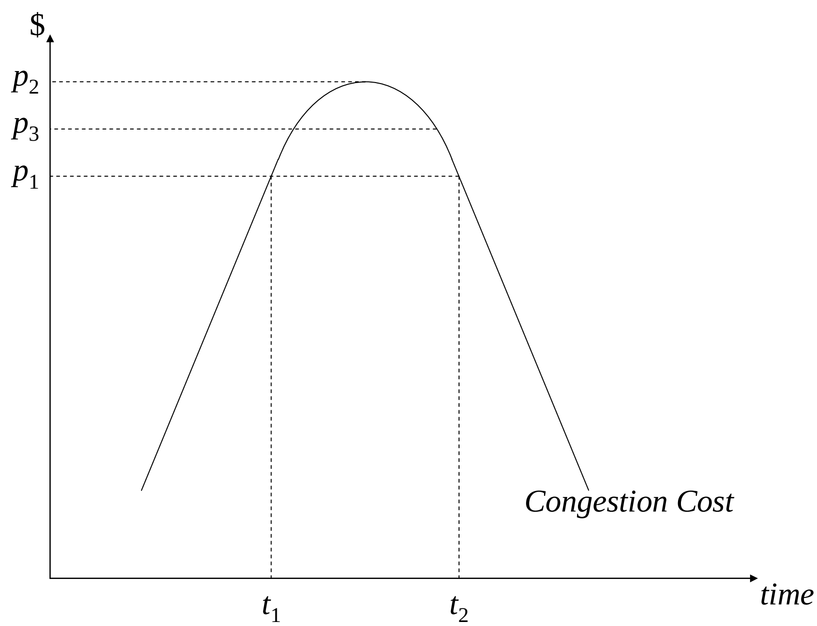
Understanding why local measured service failed to deliver the expected welfare benefits provides insights into the inherent limitations of peak-load pricing. The problem is that peak-load pricing schemes cause inefficiencies of their own.[111] Consider the peak-load pricing scheme represented in Figure 4, in which the time of day is represented on the horizontal access and the total congestion (measured in congestion cost) is represented on the vertical access. Assume that the goal is to impose a peak-load price during the busiest time of the day, represented in Figure 4 as the interval between t_1 and t_2.

---

110 See Rolla Edward Park and Bridger M. Mitchell, *Optimal Peak-Load Pricing for Local Telephone Call* 6, 32 (Rand Paper No. R-3404-1-RC Mar 1987) (concluding that local measured service is unlikely to increase economic efficiency because the modest welfare gains from discouraging excessive calls at peak times were more than offset by costs of administering the system and the inefficiency of deterring efficient calling); Lewis Perl, *Impacts of Local Measured Service in South Central Bell's Service Area in Kentucky* (May 21, 1985) (finding that imposition of local measured service in Kentucky yielded welfare gains of only 0.11%), cited by Alfred E. Kahn and William B. Shew, *Current Issues in Telecommunications Regulation: Pricing*, 4 Yale J Reg 191, 237 n 10 (1987); Bridger Mitchell, *Optimal Pricing of Local Telephone Service*, 68 Am Econ Rev 517, 531-32 (1978) (estimating the welfare changes from imposing local measured service as ranging between -1.6% and 6.0). For an overview, see Alfred E. Kahn and William B. Shew, *Current Issues in Telecommunications Regulation: Pricing*, 4 Yale J Reg 191, 237-38 & n 110 (1987) (reviewing the empirical literature assessing the welfare impact of local measured service).

111 For a similar argument, see Frischmann and van Schewick, 47 Jurimetrics J at 406 (cited in note 12).

Figure 4
Inefficiencies of Peak Load Pricing



Some degree of inefficiency will result regardless of whether the network sets price at the lowest congestion cost during this period (represented by p_1), the highest congestion cost during this period (represented by p_2), or a price set somewhere in between (represented by p_3). Consider first price, p_1. Because p_1 falls below the congestion costs created by incremental usage at every point during the peak-load period, setting price at p_1 would encourage end users to increase their consumption of network resources even when the congestion costs of doing so would exceed the benefits. On the other hand, because p_2 exceeds the congestion cost created by incremental usage at every point during the peak load period, pricing at p_2 would deter usage even though increasing usage would increase consumer welfare. Setting the price in between at p_3 gives rise to both of these problems during different portions of the peak load period. During the middle of the peak-load period, p_3 would fall below the congestion costs associated with incremental usage and thus would provide end users with the incentive to increase their consumption even when the congestion costs imposed on others would exceed the benefits that that end user would derive from doing so. At the beginning and concluding portions of the peak-load period, p_3 would exceed the congestion cost, in which case pricing at p_3 would deter additional usage even when increasing consumption would cause consumer welfare to increase.

An additional problem is that end users inevitably respond to the imposition of peak-load pricing by shifting some of their usage to the periods immediately preceding and following the peak-load period. The result is to create "shoulders" in the distribution of traffic on either side of the peak-load period. If this reallocation is sufficiently large, it can cause congestion costs outside the peak-load period to rise to welfare-reducing levels. As a result, networks that use peak-load pricing typically find it necessary also to impose near-peak rates (sometimes also called "shoulder rates") during the period immediately preceding and following the peak-load period. Near-peak rates suffer from the same consumer welfare problems discussed above associated with peak-load rates, albeit to a smaller degree.

The resulting pricing scheme also increases the complexity of the decision confronting consumers, requiring them to incur the costs of keeping track of the price at any particular time of day and adjusting their behavior accordingly. Consumers generally show considerable resistance to complex pricing schemes.[112] As a result, although these problems could be mitigated by dividing the peak-load pricing regime into a larger number of segments, analysts of local measured service concluded that consumers would not accept any more than three pricing periods over the course of a day.[113] In addition, if end users are allowed to choose between a metered pricing plan and an all you can eat pricing plan, high-volume users have the strategic

---

[112] Werbach, *Digital Tornado* 63 (cited in note 59).
[113] Park and Mitchell, *Optimal Peak-Load Pricing for Local Telephone Calls* 23-31 (cited in note 110).

incentive to opt for the latter. Together these problems dissipated the predicted welfare benefits from imposing peak-load pricing on local telephone service.

The failure of local measured service provides a real-world demonstration of the challenges posed by peak-load pricing. The migration of the wireless telephone pricing away from pricing that was metered on a per-minute basis to a form of peak-load pricing based on buckets of minutes to the more recent movement toward all you can eat pricing plans attests to the difficulties and consumer resistance confronted by any attempt to implement any form of congestion-based pricing.

There are aspects to internet traffic likely to make peak-load pricing of broadband service even less likely to benefit consumers. As an initial matter, internet traffic is much more variable than telephone traffic. For example, web browsing tends to generate sharp peaks of bandwidth usage followed by long periods of inactivity while the end user reads the webpage that has just been loaded. The result is that congestion on the internet is likely to arise much more abruptly and be much more transient than on telephone networks, which makes it much more difficult to determine whether and to what degree additional usage by one consumer will adversely affect other consumers.[114]

Congestion on the internet can also often be quite localized in technologies, such as cable modem and wireless broadband service, in which subscribers share bandwidth with their immediate neighbors. When that is the case, the network performance that any particular subscriber receives is acutely sensitive to the amount of traffic being generated by a very small number of closely situated fellow users. As a result, it is possible that congestion might be very high in one neighborhood while simultaneously being very low in the adjacent neighborhood,

---

[114] See Spulber and Yoo, 99 Nw U L Rev at 1700 (cited in note 51) (showing how greater variability in demand can make congestion more difficult to manage).

depending on the size and bandwidth intensity of the traffic being generated by end users in each neighborhood at any given time. This means that a properly functioning congestion-pricing scheme would have to do more than impose different prices during different times of the day. It would also have to adjust prices to the conditions arising in different portions of the network, depending on the local conditions in any particular neighborhood node.[115]

Lastly, any congestion-based pricing system would have to take into account packet switched networks' ability to compensate for surges in demand by routing around areas of congestion in ways that circuit switched traffic associated with conventional telephone service cannot. While the ability to reroute traffic may mean that increases in congestion need not necessarily degrade network performance, the ability to route around trouble spots can also have the effect of transferring congestion to areas of the network that are geographically distant from where network flows are increasing. This can make determining the effect that any particular increase in traffic will have on the size and location of congestion very difficult to determine.[116]

Fully deployed congestion-based pricing must thus incorporate information about the precise level of network flows and capacity in all portions of the network at any particular time in order to determine the magnitude of the congestion cost caused by any particular increase in network traffic. Such information was relatively easy to collect in local telephone systems, which have historically been dominated by incumbent local exchange carriers well positioned to collect such information. The internet, however, operates on very different principles. Indeed,

---

115 See id at 1700, 1709-11 (showing how congestion can affect different portions of the network in different ways and discussing the difficulties of creating geographically targeted approaches to managing congestion).

116 See id at 1703-07, 1711 (demonstrating how networks' ability to compensate for increases in demand by routing traffic along alternative paths can transfer congestion to other portions of the network located quite far from the locations where demand is increasing, which can make congestion particularly hard to manage).

the decentralized nature of the internet necessarily dictates that no player has access to all of the information needed to assess all of the systemic effects.[117]

Thus, although imposing bandwidth tiering or peak-load pricing would capture some of the aspects of congestion pricing, institutional considerations may well force the outcomes under both regimes to fall short of the ideal.[118] This is not to say that peak-load pricing schemes are necessarily doomed to failure. On the contrary, it is quite possible that reduced transaction costs associated with simple pricing schemes may well offset any imperfections in the ability to account for congestion costs.[119] For the purposes of this Article, we need not resolve this ambiguity. The existence of plausible circumstances under which peak-load pricing is likely to fail effectively rebuts suggestions that peak-load pricing represents a sufficient alternative to justifying treating any solution to the problems of congestion as illegal per se. Simply put, the

---

[117] See Daniel F. Spulber and Christopher S. Yoo, *Rethinking Access to the Broadband Internet*, 22 Harv J L & Tech (forthcoming 2008) (contrasting the difficulties in applying graph theoretical models to the internet with relatively successful efforts to apply graph theoretical models to electric power through institutions like PJM).

[118] Although Frischmann and van Schewick acknowledge that any system of usage-sensitive pricing would be imperfect, they nonetheless assert that an imperfect system would nonetheless assert that "the social costs of reasonably imperfect usage-sensitive pricing seem to be lower than the social costs associated with use restrictions." Frischmann and van Schewick, 47 Jurimetrics J at 406 (cited in note 12). They provide no empirical support for this assertion, basing their argument on theoretical models. The scant empirical evidence that exists, which is based on efforts to impose congestion-based pricing in local telephone and wireless telephone service, raises at least some doubts about whether usage-sensitive pricing will in fact yield benefits. Furthermore, even accepted for all it is worth, their argument suggests that the welfare implications are ambiguous and depend on the particular circumstances and the relative costs of implementing each institutional approach. When that is the case, the general thrust of competition policy articulated by the Supreme Court is not to prohibit any particular practice categorically. Instead, actors should be permitted to experiment with available second-best solutions to the problems of congestion unless and until consumer harm is shown.

In some ways, their argument parallels arguments about price discrimination. Like perfect congestion-based pricing, perfect price discrimination is always welfare enhancing. The problem is that perfect price discrimination is never possible. The ambiguousness of the welfare implications of imperfect price discrimination led Frischmann to be reticent about embracing price discrimination. See Brett M. Frischmann, *An Economic Theory of Infrastructure and Commons Management*, 89 Minn L Rev 917, 979 (2005). This is despite the fact that leading economic textbooks generally conclude that, despite the theoretical ambiguity, imperfect price discrimination is more likely to cause consumer welfare to increase. See Richard G. Lipsey et al, *Economics* 241 (HarperCollins, 8th ed 1987); F.M. Scherer and David Ross, *Industrial Market Structure and Economic Performance* 495 (Houghton Mifflin, 3d ed 1990). Absent some empirical evidence that the inefficiencies associated with imperfect congestion pricing are likely to exceed the inefficiencies associated with use restrictions, one would have expected that the imperfections in congestion-based pricing would have made him equally hesitant.

[119] See Jeffrey K. MacKie-Mason and Hal R. Varian, *Pricing Congestible Network Resources*, 13 IEEE J on Selected Areas Comm 1141, 1145 (1995)

existence of multiple exemplifying theories pointing in different directions undercuts categorically any particular practice and instead simply underscores the propriety of adopting a case-by-case approach.

2. Consumer Benefits from Network Diversity

In addition, as I have discussed extensively in my prior work,[120] permitting network providers to differentiate their services can benefit consumers by increasing the degree of competition between last mile services. The classic source of market concentration in markets for last-mile services is the supply-side economies of scale that arise when entry requires the incurrence of significant, up-front fixed costs. The presence of large, up-front capital investments gives the largest firms a decisive economic advantage. The ability to spread those investments over a larger customer base allows them to underprice their smaller competitors until they drive them out of business.[121]

What has been largely overlooked is how allowing networks to differentiate themselves can counterbalance the economies of scale created by large, up-front fixed costs. It is the fact that price is the only dimension along which firms can compete that gives the largest players

---

[120] For my more comprehensive statements of this argument, see Yoo, 19 Harv J L & Tech at 27-33 (cited in note 22); Yoo, 3 J on Telecommun & High Tech L at 60-63 (cited in note 34).

[121] For example, if a producer must incur $1,000 in up-front costs to enter the market, the up-front costs would contribute the following amounts toward unit (i.e., average) cost:

Quantity	Contribution to Unit Cost	Quantity	Contribution to Unit Cost
100	$10.00	600	$1.67
200	$5.00	700	$1.43
300	$3.33	800	$1.25
400	$2.50	900	$1.11
500	$2.00	1000	$1.00

If the impact from the amortization up-front costs dominates the impact of variable costs, average cost will decline. Note that the impact of up-front costs tends to decay exponentially as the quantity over which the up-front costs are spread increases.

their decisive advantage. A different equilibrium can result if competitors are allowed to compete along dimensions other than price. If so, a smaller player would be able to survive, notwithstanding lower sales volumes and higher unit costs (and thus higher prices), by tailoring its network towards services that a subsegment of the market values particularly highly. The greater value provided by the differentiated network allows a specialized provider to generate sufficient revenue to cover its up-front costs even though its volume is significantly smaller than that of the leading players.

The result is an equilibrium in which multiple players co-exist despite the presence of unexhausted economies of scale. Even though entrants may operate at a cost disadvantage vis-à-vis their larger rivals, they are able to survive by offering products designed to appeal to discrete subsegments of the customer base. Conversely, preventing product differentiation would cause the market to devolve into a natural monopoly.

How could such differentiation occur in the context of broadband? One way is through protocol nonstandardization, such as through the adoption of a different routing protocol. If discrete subgroups of end users place sufficiently different valuations on different types of applications, multiple networks may be able to coexist simply by targeting their networks towards the needs of different subgroups. If demand is sufficiently heterogeneous, the greater utility derived from allowing end users to consume services that they value more highly can more than compensate for any cost disadvantages resulting from the reduction in volume. For example, it is conceivable that network diversity might make it possible for three different last-mile broadband networks to coexist: one optimized for traditional internet applications such as e-mail and website access, another incorporating security features to facilitate e-commerce and to guard against malware, and a third that prioritizes packets in the manner needed to facilitate time-sensitive applications such as streaming media and virtual worlds. I will subsequently discuss in Part III.D, exclusivity arrangements with particular content or application providers can provide another basis for differentiating network services.

These examples illustrate how deviations from network neutrality may benefit consumers by facilitating greater competition in the last mile. This suggests that public policy may well be better served if Congress and the FCC were to embrace a network diversity principle that would allow networks to experiment with differentiating their services in precisely this manner. Conversely, mandating network neutrality can have the perverse effect of reinforcing this source of market failure by limiting networks to competing on price and network size, factors that favor the largest providers. If true, this raises the possibility that mandating network neutrality could turn into the source of, rather than the solution to, market failure.

3. Alternative Institutional Solutions

The fact that metered pricing and peak-load pricing schemes inevitably require the incurrence of transaction costs has led network providers to experiment with different institutional solutions. One particularly interesting solution to the problems of congestion is content delivery networks like Akamai.[122] As noted earlier, content served by Akamai often bypasses the public backbone altogether, which in turn protects the query from any backbone congestion that may exist.[123] The proximity and redundancy of the caches permits Akamai to serve content faster and to redirect queries to other caches when particular caches are overly congested.[124] Although the dynamic way that Akamai reallocates queries can improve network

---

122 See notes 80-83 and accompanying text.
123 Id.
124 Id.

performance, it can also make congestion less predictable and can make it more difficult to identify sources of congestion.

The problem from the standpoint of network neutrality is that Akamai is a commercial service that is only available to content and applications providers willing to pay a premium above and beyond the basic internet access fees that all content and applications providers pay.[125] It thus violates the basic network neutrality principles that all like traffic travel at the same speed and that network providers be prohibited from charging content and applications providers more for higher-speed service.

On some occasions, network providers have taken to blocking access to websites when proven to be harmful. The best known of these examples is the practice of denying computers sending suspiciously large volumes access to port 25, which is the port that plays a key role in spam. Some networks estimate that this practice reduces the total amount of spam by as much as twenty percent.[126] Again, blocking port 25 violates the principle of treating all like content alike and may well have the effect of blocking legitimate emails. And yet, the practice of blocking port 25 is relatively uncontroversial.

In addition, ISPs that detect end users using applications that consume large amounts of bandwidth (such as leaving their browser open to the ESPN website or engaging in large amounts of peer-to-peer file sharing), will suggest to the end users that they change their practices or purchase a higher-bandwidth service that more accurately reflects the amount of

---

[125] To the extent that Akamai's price structure contains imperfections in the internalization of congestion costs, it may give rise to welfare losses similar to those caused by the imperfections in congestion pricing discussed above.

[126] See Jim Hu, *Comcast Takes Hard Line Against Spam*, CNET News.com (June 10, 2004), available at <http://news.com.com/2100-1038_3-5230615.html> (last visited Mar 6, 2008) ("Already, Comcast has noticed a 20 percent reduction in spam since the blocks began and a 75 percent decline in the past two months.").

congestion they are imposing on other end users.[127] If the end user is unwilling to change, the ISP may choose to cease doing business with the customer.

I recount other examples of alternative institutional solutions short of imposing full-fledged congestion-based pricing elsewhere.[128] All of these practices are to some degree

---

127 Dan Mitchell, *Say Good Night, Bandwidth Hog*, NY Times C5 (Apr 14, 2007).

128 See Yoo, 94 Georgetown L J at 1874-85 (cited note 1) (institutional solutions include prohibiting the resale of bandwidth or acting as an internet service provider; imposing restrictions on home networking, attaching devices, and operating file servers; and discriminating against particular applications and against particular content.). Frischmann and van Schewick take issue with some of these examples. For example, they argue that bans on online games are overinclusive because many online games do not require much bandwidth. At the same time, they acknowledge that some online games are bandwidth intensive. Frischmann and van Schewick, 47 Jurimetrics J at 408-09 (cited in note 12). Whether a ban on an application is a good proxy for congestion is thus an empirical question that cannot be answered a priori and thus seems better suited to case-by-case analysis.

In fact, prohibitions on particular applications and content are quite common. For example, the State of Arkansas has banned the use of peer-to-peer applications because they "were utilizing a significant amount of bandwidth at the State's public schools"; as a result the State found the ban necessary "[t]o ensure that teachers and students had a high level of network quality speed, and availability." Allot Communications, Press Release, Allot Helps Arkansas Provide Government Agencies and Public Schools with Guaranteed Bandwidth and Improved Network Quality (Jan 3, 2007), available at <http://www.allot.com/index.php?option=com_content&task=view&id=449&Itemid=18> (last visited May 27, 2008). For similar reasons, the Department of Defense has banned YouTube "in an effort to boost its network efficiency," noting, "This is a bandwidth and network management issue. We've got to have the networks open to do our mission. They have to be reliable, timely and secure." Leo Shane III and T.D. Flack, *DOD Blocking YouTube*, *Others*, Stars and Stripes (May 13, 2007), available at <http://www.stripes.com/article.asp?section=104&article=53421&archive=true> (last visited May 27, 2008). These are end-user networks, rather than internet access networks. Nonetheless, the basic intuition is the same. Although in an ideal world it might be preferable always to take a content and application neutral approach to bandwidth management, sometimes use restrictions targeted at particular content or applications can represent a quick and cost-effective (albeit imperfect) way to manage congestion.

Similarly, Frischmann and van Schewick argue that a ban on WiFi routers or home networking "ha[s] no predictive power with respect to the bandwidth intensity of the corresponding uses." Frischmann and van Schewick, 47 Jurimetrics J at 408 (cited in note 12). While not perfectly predictive of the amount of congestion generated, it strikes me that use of a technology that enables subscribers to attach multiple computers to the network is at least somewhat (albeit imperfectly) probative of the amount of congestion that the average subscriber will generate. Resolving which of us is correct is ultimately an empirical question, but the record to date does not offer any concrete evidence indicating either that use of WiFi routers either is strongly associated with higher bandwidth usage (in which case the prohibition would be reasonable) or is completely unrelated to higher bandwidth usage (in which case preventing network owners from prohibiting WiFi routers would be reasonable). In the absence of a clear policy inference, the approach to competition policy laid out by the Supreme Court would support eschewing any categorical prohibitions or mandates in favor of the type of case-by-case approach that I have advocated.

Lastly, Frischmann and van Schewick acknowledge that operating a fileserver does represent a valid Coasean proxy for congestion. Id at 409. Their recognition that at least one use restriction represents a good proxy for congestion concedes the existence of at least one exemplifying theory suggesting that permitting network providers to impose some use restrictions may in fact benefit consumers to the point that a categorical ban on all use restrictions may be socially harmful. More to the point, the entire controversy between Comcast and BitTorrent centers on the fact that peer-to-peer architectures require end users to operate servers. Recognizing that a ban on operating servers represents a good proxy for heavy bandwidth usage provides some support for the idea that permitting network providers to ban servers may represent one of the institutional alternatives for managing congestion that should be given serious consideration. Of course, it is impossible to tell a priori whether it will be inconsistent with the principles advocated by network neutrality proponents. In pointing out these practices, I make no attempt to show that any particular practice is always beneficial or always harmful or to make any assessment of which is likely to prove best. Indeed, the rapid pace of change in terms of cost and functionality would make any such assessment too ephemeral a basis for policymaking. My point is that policymakers will find it difficult, if not impossible, to determine the relative merits of any of these alternative institutional solutions at any particular time, let alone keep up with the rapid pace of technological change. So long as some plausible argument exists that a practice might be socially beneficial,[129] the better course is to establish rules that give network operators the flexibility to experiment with that practice until its precise impact on consumers can be determined.

B. Demand-Side Justifications for Differential Pricing

The academic literature on the economics of regulation offers demand-side as well as supply-side justifications for charging differential prices. Although the modern literature on price discrimination is vast,[130] the key insight can be traced to the 1927 article by Frank Ramsey that proposed an innovative solution to the classic pricing problem confronted by telecommunications networks.[131]

---

the best of the available second-best alternatives. That said, the plausibility does provide a strong argument against rules categorically prohibiting the practice.

129 Note that practices do exist that are so likely to be socially harmful and so unlikely to convey any plausible benefits that there is general agreement they should categorically prohibited. The classic example is horizontal price fixing by a cartel. See, for example, Catalano, Inc v Target Sales, Inc, 446 US 643, 646-47 (1980) (calling horizontal price fixing the "archetypical example" of a practice that is so "plainly anticompetitive" that it is conclusively presumed illegal without any exploration of any offsetting benefits").

130 For surveys of the literature on price discrimination, see Hal R. Varian, *Price Discrimination*, in 1 Richard Schmalensee and Robert D. Willig eds, *Handbook of Industrial Organization* 597 (Elsevier 1989); and Lars A. Stole, *Price Discrimination and Competition*, in Mark Armstrong and Robert K. Porter eds, 3 *Handbook of Industrial Organization* 2221 (Elsevier 2007).

131 See F.P. Ramsey, *A Contribution to the Theory of Taxation*, 37 Econ J 47 (1927) (offering the seminal statement of the pricing scheme that would ultimately bear his name).

The nature of the problem can be most easily understood by examining the relationship between the average cost curve and the marginal cost curve, which are the two cost curves that receive the most attention from scholars of industrial organization. The production of most goods and services requires the incurrence of two types of costs: fixed costs and variable costs. Fixed costs are capital costs that are incurred once and do not increase as production increases. Variable costs are those that increase with the production of additional units. For example, in the case of shoe manufacturing, the cost of setting up the shoe factory would represent a one-time cost that does not vary with the number of shoes produced and thus constitutes part of the fixed cost. The costs of leather, labor, and electricity increase as the number of units increases and thus constitute part of the variable cost. Average cost is determined by adding the fixed costs and variable costs to determine total costs and dividing the total costs by the total quantity produced. Marginal cost focuses on the cost of the last unit produced.[132] Thus, if a production process is subject to unexhausted economies of scale, the cost of the last unit produced (that is, marginal cost) may fall below average cost. Conversely, if a production process is subject to diseconomies of scale, the cost of the last unit produced should exceed average cost.

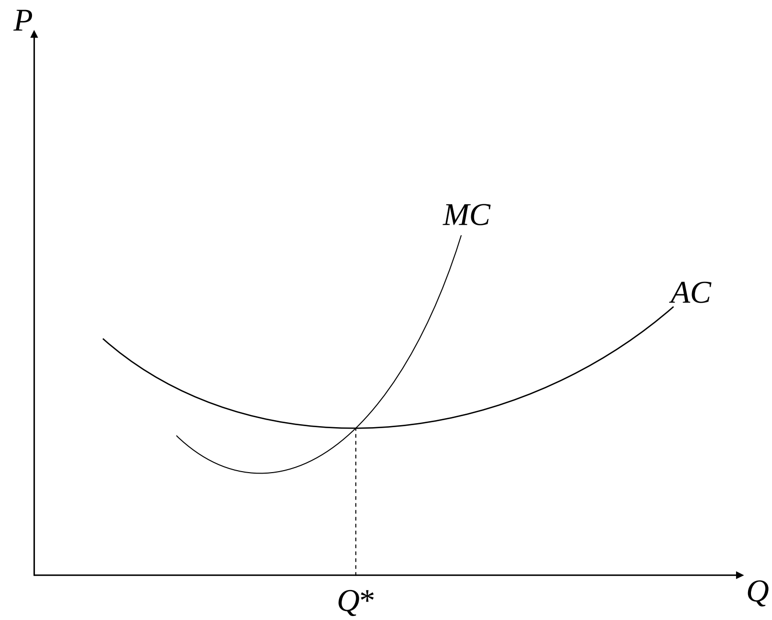
As depicted in Figure 5, the interaction of both fixed and variable cost determines the relative position of the average and marginal cost curves and gives both curves their characteristic "U" shape. Because fixed cost does not vary with production, it has no impact on marginal cost. Fixed cost does exert downward pressure on average cost as the upfront cost is spread across increasingly larger volumes. The impact of variable cost is somewhat more complex. In general, most production processes benefit from a degree of economies of scale,

---

[132] Or, more properly, the production of one additional unit beyond current levels of production.

which allows firms to produce units of a good more cheaply as production increases.[133] Together these effects cause both average and marginal cost to decline at low levels of production, with the marginal cost curve lying below the average cost curve.

Figure 5
The Relationship Between Average and Marginal Cost



The situation begins to change as production increases. The downward impact of fixed cost on average cost decays exponentially as the fixed cost is amortized over increasingly large volumes.[134] One presumes that the firm initially turned to the lowest cost inputs that it could find. Once those supplies are exhausted, any further increases in production inevitably require the firm to turn to sources of inputs that are more expensive. As a result, variable cost begins to

---

133 Some of the scale economies result from specialization, as demonstrated most eloquently by Adam Smith's example of a pin factory. See Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations* 4-5 (Edwin Canaan ed, Modern Library 1965) (1776) ("[A] workman not educated to th[e] business [of pin making] … could scarce, perhaps, with his utmost industry, make one pin a day…. But in the way in which this business is now carried on, … the important business of making a pin is … divided into about eighteen distinct operations…." In this way, a pin factory can make "about twelve pounds of pins in a day."). Other scale economies are technological. Higher volumes make it cost effective for manufacturers to use more capital equipment that requires higher up-front costs, but lowers unit costs.

134 See note 121.

rise. This causes the downward pressure on marginal cost caused by the initial economies of scale in variable cost to dissipate and eventually begin to be replaced by upward pressure. As the upward pressure associated with variable cost increases and the downward pressure associated with fixed cost decreases, eventually the former dominates the latter and average cost begins to increase. The marginal and average cost curves will cross at the lowest point of the average cost curve, indicated in Figure 5 by Q^*.

As noted earlier, consumers benefit the most if production is increased whenever the benefit to consumers from producing another unit (reflected by the marginal consumer's willingness to pay) exceeds the cost of producing another unit (reflected by the marginal cost curve). Maximizing consumer welfare thus requires increasing production whenever price exceeds marginal cost. It is thus when consumer benefits no longer exceed the costs to society (that is, when price equals marginal cost) that no further gains are possible. This leads to the familiar economic principle that consumer welfare is maximized when price is set equal to marginal cost. At the same time, producing firms must break even for an industry to be viable over the long run, which means that price must also be set at or above the average cost curve.

Finding prices that simultaneously equal marginal cost and equal or exceed average cost is quite feasible if the overall demand exceeds Q^*. For any quantity greater than Q^*, any price that maximizes consumer welfare by being set equal to marginal cost necessarily exceeds average cost and thus is sustainable. The classic pricing problem occurs if the total market demand falls below Q^*. In that case, any price set along the marginal cost curve necessarily falls below average cost and thus is unsustainable. Thus, if producing firms are to break even, they must be allowed to charge prices that represent both the variable cost plus a share of the fixed cost. The allocation of fixed cost inevitably causes price to exceed marginal cost, which in turn necessarily reduces consumer welfare by excluding some consumers who would derive benefits from being allowed to consume an additional unit that would exceed the cost of permitting them to do so. Thus any price that is sustainable fails to maximize consumer welfare, and any price that maximizes consumer welfare is inherently unsustainable.

Ramsey identified an ingenious solution to this conundrum based on the fact that consumers vary in their sensitivity to price changes. Some consumers are very price sensitive and will drastically reduce their purchases in response to any increase in price. These consumers are said to have relatively elastic demands. Other consumers are less price sensitive and will respond to price increases by reducing their purchases only minimally. These consumers are said to have relatively inelastic demands. Ramsey realized that loss in consumer welfare that arises when producers are forced to price above marginal cost would be minimized if the firm allocated a larger proportion of the fixed cost to consumers who are the least price sensitive (and thus are more likely to maintain high purchase levels even if price increases above marginal cost) and allocated a smaller proportion of the fixed costs to those who are the most price sensitive (and thus are more likely to curtail their purchases sharply in response to increasing price above marginal cost). The additional revenue made possible by this form of price discrimination enables the firm to be sustainable even in the presence of significant upfront fixed cost. In addition, if fixed cost is allocated in strict inverse proportion to every consumer's elasticity demand, Ramsey pricing can enable the firm to produce the quantity that maximizes consumer welfare.

As a result, economic commentators from a wide variety of perspectives have embraced demand-side price discrimination as a way to maximize aggregate consumer benefits in industries that require substantial fixed-cost investments.[135] Note that this solution depends on the ability to charge customers different amounts for the exact same product. The price differential is based not on supply-side considerations, such as variations in the cost of providing of the product, but rather solely on demand-side considerations, specifically variations in the elasticity of the customers' demand for the product.

The problem is that, as is the case with any scheme of perfect price discrimination, the informational requirements for a fully implemented Ramsey pricing scheme are unrealistically demanding. It requires the firm to know each consumer's elasticity of demand and to devise a pricing scheme that makes sure that each consumer pays the exact price implied by the elasticity of their demand. Effective Ramsey pricing would also require a system for preventing high-elasticity consumers who pay low prices from reselling their purchases to low-elasticity consumers who are asked to pay higher prices. Thus, any real world attempt to implement Ramsey pricing would necessarily be imperfect. This renders its welfare impact ambiguous,

---

135 See, for example, Jean-Jacques Laffont and Jean Tirole, *Competition in Telecommunications* xv (MIT 2000) ("Marginal-cost pricing for all services is not viable in telecom industries (at least in certain important segments involving large joint and common costs), so the relevant benchmark requires some markups. Allowing at least some price discrimination can therefore reduce the pricing distortion."); Scherer and Ross, *Industrial Market Structure* at 496-502 (cited in note 118) ("Price discrimination can also provide ways out of an efficiency dilemma encountered in regulated natural monopoly industries."); William J. Baumol and Daniel G. Swanson, *The New Economy and Ubiquitous Competitive Price Discrimination: Identifying Defensible Criteria of Market Power*, 70 Antitrust L J 661, 671-73 (2003) (showing how price discrimination is essential in high fixed cost industries and occurs even when those industries are competitive); Harold Demsetz, *The Private Production of Public Goods*, 13 J L & Econ 293, 301-03 (1970) (showing how price discrimination permits the production of public goods in which all of the costs are fixed and allocates resources efficiently); Benjamin Klein and John Shepard Wiley, Jr., *Competitive Price Discrimination as an Antitrust Justification for Intellectual Property Refusals to Deal*, 70 Antitrust L J 599, 611-15 (2003) (showing how price discrimination in industries characterized by high fixed cost and lower marginal cost is likely to enhance consumer welfare and economic efficiency); Michael E. Levine, *Price Discrimination Without Market Power*, 19 Yale J Reg 1, 9-17 (2002) (showing how Ramsey pricing and other forms of price discrimination can promote economic efficiency in industries in which fixed costs are shared by multiple consumers).

although many leading economic theorists suggest that imperfect price discrimination is more likely than not to increase consumer welfare.[136]

Ramsey pricing thus offers a plausible demand-side justification for charging similarly situated consumers and content/application providers different amounts for the same service. Doing so could well benefit consumers by allowing more of them to purchase the product than would be possible under a pricing regime that requires charging all end users who consume the same product the same amount. Ramsey-style price discrimination can be exercised on the server side as well by charging content and applications providers different amounts based on their elasticity of demand. The enhanced ability to recover fixed cost made possible by Ramsey pricing can also enable high fixed-cost providers to exist when they would not otherwise be able to do so. However, it is precisely this type of differential pricing that network neutrality would prohibit.

C. The Relevance of Two-Sided Markets

As noted earlier, some network neutrality proponents have taken the position that network providers should be allowed to offer tiered pricing to consumers (that is, consumer tiering), but not to content and applications providers (that is, access tiering).[137] Determining whether or not this would represent good policy requires taking into account the fact that the internet is a two-sided market.[138] For a market to be two-sided requires more than just the

---

136 See note 118.

137 See text accompanying note 10.

138 For some of the leading contributions to the field, see Mark Armstrong, *Competition in Two-Sided Markets*, 37 RAND J Econ 668 (2006); Bernard Caillaud and Bruno Julien, *Chicken & Egg: Competition Among Intermediation Service Providers*, 34 RAND J Econ 309 (2003); David S. Evans, *The Antitrust Economics of Multi-Sided Platform Markets*, 20 Yale J Reg 325 (2003); and Jean-Charles Rochet and Jean Tirole, *Platform Competition in Two-Sided Markets*, 1 J Eur Econ Assn 990 (2003). For a recent survey, see Roberto Roson, *Two-Sided Markets: A Tentative Survey*, 4 Rev Network Econ 142 (2005).

existence of a buyer and a seller. Two-sided markets arise when network economic effects create demand interdependencies that cause the value that any one party derives from participating in the platform to depend not only on price, but also on the number of other platform participants.[139] Unlike the conventional approach to network economics, in which the size of the network economic effect depends on the number of participants within the same group, in two-sided markets the network economic effect is determined by the number of participants in a different group located on the other side of the market.[140]

Pricing in two-sided markets depends on a wide variety of factors including the elasticities of demand, the magnitude of the indirect network effects, and the marginal costs borne by each side of the market, among others.[141] A survey of practices revealed that prices on

---

139 In the absence of network economic effects, price is the only determinant of value, in which case markets face few obstacles to functioning properly.

140 A classic example of the kind of network usually analyzed through the lens of network economics is the telephone system, in which the value of the network is determined by the number of similarly situated users that use the platform in the same way. See, for example, Neil Gandal, *Compatibility, Standardization, and Network Effects: Some Policy Implications*, 18 Oxford Rev Econ Pol 80, 80 (2002) (listing telephone systems as among "[t]he most common examples" of a network subject to direct network effects); Herbert Hovenkamp, *Post-Chicago Antitrust: A Review and Critique*, 2001 Colum Bus L Rev 257, 300 (calling the telephone system "the classic example of the positive network externality"); Michael L. Katz and Carl Shapiro, *Technology Adoption in the Presence of Network Externalities*, 94 J Pol Econ 822, 823 (1986) (noting that network externalities "have long been recognized of physical networks such as the telephone … industr[y]"); Mark A. Lemley and David McGowan, *Legal Implications of Network Economic Effects*, 86 Cal L Rev 479, 488-89 (1998) (citing telephones as one of the "classic examples of actual network goods"); S.J. Liebowitz and Stephen E. Margolis, *Network Externality: An Uncommon Tragedy*, J Econ Persp, Spring 1994, at 133, 139-40 (calling the telephone network "[t]he paradigmatic case of a direct network effect"); Peter Menell, *Tailoring Legal Protection for Computer Software*, 39 Stan L Rev 1329, 1340 (1987) (calling the telephone "a classic example of a product for which there are network externalities").

In a two-sided market like the internet, value to content and application providers is not determined just by price or by how many other content and applications providers participate in the platform. Instead, the demand interdependency created by advertising-based business models makes the value depend on a combination of price and the number of end users.

141 David S. Evans and Richard Schmalensee, *The Industrial Organization of Markets with Two-Sided Platforms* 11 (Nat'l Bur Econ Res Working Paper 11603, Sept 2005), available at <http://www.nber.org/papers/w11603> (last visited Mar 6, 2008) ("The optimal prices depend in a complex way on the price elasticities of demand on both sides, the nature and intensity of the indirect network effects between each side, and the marginal costs that result from changing output of each side."); see also Jean-Charles Rochet and Jean Tirole, *Two-Sided Markets: An Overview* 34-35 (Mar 12, 2004), available at <http://faculty.haas.berkeley.edu/hermalin/rochet_tirole.pdf> (last visited Apr 10, 2008) ("[F]actors that affect prices charged to end-users" include "[e]lasticities," the "[r]elative market power of service providers," the "[s]urplus on the other side," "[p]latform competition and multi-homing" and "[b]undling.").

different sides of two-sided markets tend to be asymmetric, with end users often paying little or nothing.[142]

Two features bear highlighting. First, the fact that participation on one side of a two-sided market creates network economic effects on the other side creates a positive externality that is so large that, if internalized through prices, leads end user prices to be set below marginal cost.[143] Second, any pricing regime implicitly incorporates a proportion of the fixed cost to recover from each side of the two-sided market. The logic of Ramsey pricing would suggest that the allocation depends on the relative elasticities of demand. This results in prices that are not purely cost-based. This reasoning receives an interesting twist in the case of a two-sided market, as the prices must also reflect the contributions to the other side's surplus created by network economic effects.[144]

The theoretical literature on two-sided markets suggests that preventing network providers from imposing discriminatory prices against content and applications providers may harm consumers in two ways. First, the reduced ability to cover fixed costs will cause fewer new networks to be created. This is particularly important as the internet makes the transition from

---

142 See Evans and Schmalensee, *The Industrial Organization of Markets with Two-Sided Platforms* at 8, 12-13 (cited in note 141) ("The empirical evidence suggests that prices that are at or below marginal cost are common for [two-sided platforms]."); Rochet and Tirole, 1 J Eur Econ Assn at 1013-17 (cited in note 138) (examining seven case studies in which fees tended to be highly asymmetric and often charged consumers little or nothing).

143 See Wilko Bolt and Alexander Tieman, *A Note on Social Welfare and Cost Recovery in Two-Sided Markets* 6, 9 (DNB Working Paper No 24, Dec 2004), available at <http://www.dnb.nl/dnb/home/file/Working%20Paper%20No.%2024-2004_tcm47-146681.pdf> (last visited Apr 10, 2008) ("[S]ocially optimal pricing in two-sided markets leads to an inherent cost recovery problem, inducing losses for the monopoly platform. The result is driven by the positive externality on users on one side of the market, which originates from network participation on the other side of the market. The contribution of this externality to social welfare is larger than the individual market side's price, which leads pricing below marginal cost to be socially optimal."); Evans and Schmalensee, *The Industrial Organization of Markets with Two-Sided Platforms* at 11-12 (cited in note 141) (For two-sided platforms, "[t]he profit-maximizing, non-predatory prices may be below the marginal cost of supply for that side or even negative."); Roson, *Two-Sided Markets*, 4 Rev Network Econ at 147-48 (2005)(cited in note 138) (showing that prices may turn out to be zero or even be negative if the network economic effects on the other side of the market are sufficiently strong).

144 Rochet and Tirole, 1 J Eur Econ Assn at 991 (cited in note 138) (noting that socially optimal prices differ from classic Ramsey in that they "are not driven solely by superelasticity formulae but also reflect each side's contribution to the other side's surplus").

converting legacy networks, such as cable television and local television networks, to creating new networks, such as Verizon's new fiber-based FiOS service. Second, limiting network providers' ability to recover fixed cost from the server side of the two-sided market may increase the proportion of fixed costs they must recover from the end user side.[145]

The potential for consumer harm is particularly problematic in light of the fact that content and applications providers are increasingly turning to advertising-based business models, in which the bulk of the revenue flows into the network on the server side.[146] The historic pattern of cash flows in advertising-driven industries suggests that the increasing emphasis on advertising increases the value of expanding the end user base, which in turn should put downward pressure on the prices charged to end users. This in turn implicates what Jean-Charles Rochet and Jean Tirole have called the "topsy-turvy principle," in which any factor that tends to increase prices on one side of a two-sided market tends to lower prices on the other side, because the increased margin provided on the higher prices charged on the first side increases the benefits of increasing participation on the second side.[147] Conversely, limiting network providers' ability

---

[145] For related arguments, see Larry F. Darby and Joseph P. Fuhr, Jr., *Consumer Welfare, Capital Formation and Net Neutrality: Paying for Next Generation Broadband Networks*, 16 Media L & Pol 122, 133 (2007) (noting that the socially optimal prices vary on each side of the market and that "for multisided markets, the optimal market solutions cannot generally be achieved by charging only consumers"); J. Gregory Sidak, *A Consumer-Welfare Approach to Network Neutrality Regulation of the Internet*, 2 J Competition L & Econ 349, 361-62 (2006) ("In short, each party in a two-sided market can contribute to the recovery of the sunk costs required to build a broadband network. There is certainly no basis in economic theory to presume that it would be socially optimal for end-users to pay for all of the cost….").

[146] See Faratin et al, *Complexity of Internet Interconnections* at 13 (cited in note 73) ("[T]he emergence of the commercial Internet with high-volume, high value (e.g. with advertising) content has triggered a pragmatic conclusion that value flow is the same as packet flow. Money flows in at the content end (e.g. via the advertising or merchant revenues)….").

[147] Rochet and Tirole, *Two-Sided Markets: An Overview* at 34 (cited in note 141) (The "'topsy-turvy principle': A factor that is conducive to a high price on one side, to the extent that it raises the platform's margin on that side, tends also to call for a low price on the other side as attracting members on that other side becomes more profitable."). Indeed, if the advertising revenue collected by content and applications and providers is sufficiently large, end user prices may even become negative. See Mark Armstrong and Julian Wright, *Two-Sided Markets, Competitive Bottlenecks and Exclusive Contracts*, 32 Econ Theory 353, 354 (2007) ("[B]uyers may be charged nothing to subscribe, while sellers will be charged a price that decreases in the extent to which the platform would like to set a negative price for buyers.").

to charge different prices to content and applications providers may have the side effect of forcing consumers to bear a greater proportion of the fixed cost.

This insight underscores the fallacy in an argument often advanced by network neutrality proponents that allowing network providers to employ access tiering will cause consumers to pay twice.[148] On the contrary, the literature on two-sided markets suggests that suppressing access tiering can increase the prices consumers pay by forcing them to bear costs that would otherwise be borne by content and application providers.

Finally, the theoretical literature indicates that optimal prices in two-sided markets are determined by demand elasticities and network economic effects as well as marginal cost.[149] Thus, any attempt to regulate the terms of access in a two-sided market will require that prices reflect more than just the cost data that regulatory authorities usually take into account. It will also require regulators to consider each customer's elasticity of demand as well as the magnitude of the network economic effects on the other side of the two-sided market. This task is far more complex than the task usually faced by public utility regulators. The informational requirements even surpass those required under Ramsey pricing, which requires demand elasticity information

---

148 See Consumers Union, Press Release, Importance of the Internet Public Support for Net Neutrality, Consumers Union (Jan 18, 2006), available at <http://www.consumersunion.org/pub/press_releases/003060.html> (last visited Mar 6, 2008) (quoting Ben Scott, policy director of Free Press, stating "Requiring Internet companies to pay for high-speed access to the Internet when they're already charging consumers for the same service means consumers will ultimately pay twice."); Arshad Mohammed, *SBC Head Ignites Access Debate*, Wash Post D1 (Nov 4, 2005) (quoting Vonage Chairman Jeffrey Citron as stating, "Any notion that SBC or anyone else … can get paid twice on the same service is a bit ludicrous," adding that it would be like UPS demanding the sender and recipient of a package both pay for delivery."); Marguerite Reardon, *Without "Net Neutrality," Will Consumers Pay Twice?*, CNET News.com (Feb 7, 2006), available at <http://www.news.com/Without-Net-neutrality%2C-will-consumers-pay-twice/2100-1034_3-6035906.html?tag=item> (last visited Mar 6, 2008) (framing the issue as whether consumers will have to pay twice); Dionne Searcey and Amy Schatz, *Phone Companies Set Off a Battle Over Internet Fees: Content Providers May Face Charges for Fast Access; Billing the Consumer Twice?*, Wall St J A1 (Jan 6, 2006) (quoting Vonage Chairman Jeffrey Citron as stating, "They want to charge us for the bandwidth the customer has already paid for," and that customers who already pay for high-speed internet access will end up paying even more if online services pass the new access charges to consumers).

149 See Evans and Schmalensee, *The Industrial Organization of Markets with Two-Sided Platforms* at 11 (cited in note 43); Rochet and Tirole, *Two-Sided Markets* at 34-35 (cited in note 43).

in addition to cost. Thus any attempt to regulate access in two-sided markets would require the development of sophisticated new regulatory tools.[150]

I do not mean to suggest that allowing network providers to vary the prices they charge content and applications providers will necessarily benefit consumers. It is quite possible as a theoretical matter that this type of pricing flexibility might cause consumers to pay higher prices if end users' demand is sufficiently inelastic and the benefits they generate to content and applications providers in advertising revenue is sufficiently small. That said, the increasing emphasis on advertising would appear to make it more likely that consumers would benefit. In any event, my argument does not depend on a definitive resolution of this empirical ambiguity. Simply showing a realistic possibility that permitting greater pricing flexibility might yield consumer benefits represents an exemplifying theory sufficient to rebut calls for a ban on access tiers and differential pricing even if exemplifying theories were to exist pointing in the other direction. The possibility that particular practices might be either harmful or beneficial does not justify prohibiting them categorically. Instead, it favors a case-by-case approach that allows network providers to experiment with different pricing regimes unless and until a concrete harm to competition can be shown.

* * *

Preventing network providers from prioritizing traffic, restricting the use of certain applications, or varying the prices they charge to their customers thus has the potential to reduce consumer welfare, not only by limiting network providers' ability to induce end users to

---

150 See David S. Evans and Richard Schmalensee, *The Economics of Interchange Fees and Their Regulation: An Overview* 34-35 (AEI-Brookings Joint Center for Regulatory Studies 2005), available at <http://aei-brookings.org/admin/authorpdfs/redirect-safely.php?fname=../pdffiles/php6A.pdf> (last visited Mar 6, 2008) (noting the difficulties confronted by traditional price regulation, that price regulation in two-sided markets would "require far more empirical information than classic public utility regulation," and that applying the cost-based tools associated with traditional rate regulation would not be socially optimal, and that any attempt to apply a better approach "could at best yield highly imprecise estimates.").

rationalize their consumption, but also by preventing them from engaging in pricing mechanisms that require content and applications providers to bear a greater proportion of the fixed costs. But these are not the only potential harms that might flow from mandating network neutrality. As I will discuss in the next Part, network neutrality can also harm consumers by reducing innovation.

III. POTENTIAL INNOVATION BENEFITS FROM DEVIATING FROM NETWORK NEUTRALITY

Network neutrality proponents generally oppose giving priority to traffic originating from particular users or associated with particular applications. They argue that innovation in content and applications depends on the ability to reach as wide a universe of potential customers as possible. They insist that innovation will suffer if all content and application providers cannot access the internet on the same terms.[151]

A close analysis of the economics of innovation raises serious doubts about the position taken by network neutrality proponents. Deviations from network neutrality can in fact enhance innovation. Conversely, preventing such deviations can forestall many new applications from emerging.

A. The Role of Prioritization

Many network neutrality proponents suggest that ensuring that all internet content and applications can interconnect on equal terms is essential for the internet to remain the internet.[152]

---

[151] See, for example, Lessig testimony 4 (cited in note 10) ("By minimizing the control by the network itself, the 'end-to-end' design maximizes the range of competitors who can innovate for the network. Rather than concentrating the right to innovate in a few network owners, the right to innovate is open to anyone, anywhere. That architecture, in turn, has created an astonishing range of important and economically valuable innovation.")

[152] Id at 3 ("The Internet has inspired a wide range of innovation. Because of its particular architectural design, that innovation has come primarily from the 'edge' or 'end' of the network through application competition.

This argument implicitly suggests that the entire internet should be governed by a single protocol.

As I have pointed out elsewhere, this suggestion overlooks the fact that every protocol inherently favors some applications over others.[153] Consider TCP/IP, which is the protocol that currently dominates the internet. TCP/IP has two distinctive features. First, it routes traffic on a "best efforts basis." This means the protocol makes no guarantees that any particular packet will arrive at its destination. For example, suppose that packets arrive at a router more quickly than the router can clear them, either because of limitations in its own processing capacity or because its outbound links are busy. When that occurs, under TCP/IP the router has no choice but to discard packets. That leaves the receiving computer to signal the sending computer that it is missing a packet so that the sending computer can resend it. Second, TCP/IP routes traffic on a first-come, first-served basis.

The delays associated with dropped packets did not represent a significant problem for the applications that dominated the early internet, such as email and web browsing, which focused on text and in which delays of a fraction of a second were essentially unnoticeable. As the number and diversity of users has grown and applications technology has improved, users have begun to use newer applications that are increasingly bandwidth-intensive and less tolerant of delay. Leading examples are internet telephony (also known as voice over internet protocol or

… One consequence of this design is that early network providers could not easily control the application innovation that happened upon their networks. … That architecture, in turn, has created an astonishing range of important and economically valuable innovation. Here, as in many other contexts, competition has produced growth. And that competition was assured by the network's design.")

153 See Yoo, 19 Harv J L & Tech at 20-22, 25 (cited in note 22) ("Simply put, any choice of standardized protocol has the inevitable effect of favoring certain applications and disfavoring others, just as TCP/IP discriminates against applications that are time sensitive and end-to-end favors innovation at the edge over innovation in the core.").

VoIP),[154] streaming audio and video, real-time graphics-intensive games and virtual worlds (such as World of Warcraft and Second Life), and heart monitors and other forms of telemedicine. When combined with the dramatic increase in total network traffic, the advent of these new bandwidth-intensive and delay-intolerant applications has greatly increased the pressure on the network to take additional steps to manage congestion and to guarantee quality of service. Without some types of adjustment, these forms of innovative new content and applications, as well as others that are not yet deployed, may not be able to survive.

1. The Limitations of Increasing Bandwidth as a Solution

Network providers can respond to the increasing diversity of applications, some of which are more tolerant of delay, some of which are not, in one of several ways. If they do not want to distinguish between different applications, they can continue to expand capacity until they provide enough capacity to carry the peak-load volume of traffic at speeds that satisfy the delay-intolerant applications. In fact, many network neutrality advocates contend that congestion could be alleviated (and network management rendered unnecessary) if network providers would simply increase bandwidth.[155]

As an initial matter, relying on the expansion of bandwidth as the only solution to congestion presumes that network providers are able to anticipate how much additional

---

154 Interestingly, bandwidth may have increased to the point where VoIP is able to function well without any quality of service guarantees. Edward W. Felten, Nuts and Bolts of Network Neutrality 9 (July 6, 2006), available at <http://itpolicy.princeton.edu/pub/neutrality.pdf> (last visited May 9, 2008). In addition, VoIP generally employs an alternative protocol known as user datagram protocol (UDP) that increases transmission speeds by omitting attempts to resend any packets that do not arrive or arrive with errors.

155 See Lessig, *Future of Ideas* at 47 (cited in note 14) ("But proponents of these changes often overlook another relatively obvious solution—-increasing capacity.").

bandwidth will be needed and precisely where it will be needed.[156] The reality is that some degree of misestimation is inevitable. The geographic distribution of end users may deviate from forecasted patterns. New applications may produce unanticipated demand. In addition, unanticipated developments in complementary technologies can cause congestion to increase so quickly that increasing bandwidth is not a feasible solution. Given that bandwidth cannot be expanded instantaneously, any underestimation of demand will lead to congestion for which increasing bandwidth may not be an available solution. The question from the standpoint of network neutrality is whether policymakers should categorically exclude some forms of network management as options even in the absence of demonstrated harm.

A classic example of this problem, which I have discussed in my earlier work, arose on the NSFNET in 1987. Prior to that time, end users connected to the NSFNET via dumb terminals.[157] As a result, the bandwidth that any one end user could consume was limited by the speed with which that end user could enter keystrokes. All of this began to change with the introduction of the personal computer. Once personal computers could be connected to the network, end users could use the network to transfer files. The resulting congestion caused terminal sessions to run unacceptably slowly. NSFNET's solution was to reprogram the network to give terminal sessions priority over file transfer sessions, based largely on the fact that end users found delays in file transfer sessions more tolerable than delays in terminal sessions. The emergence of the PC represents precisely the type of unanticipated exogenous shock in a

---

[156] The ability of networks to reroute traffic discussed above can alleviate congestion somewhat. See note 116 and accompanying text. But such rerouting may simply transfer congestion in unpredictable ways, particularly when no single entity is in a position to monitor and control the entire network. In any event, networks' ability to mitigate congestion in this way is ultimately limited by total network capacity. Traffic growth will eventually overwhelm the networks' ability to make such adjustments.

[157] A dumb terminal is a device consisting solely of a keyboard and a display screen without any independent processing capacity that depends on a mainframe or other remotely located computer for its processing power. *In re Am Acad of Sci Tech Ctr*, 367 F3d 1359, 1362 (Fed Cir 2004).

complementary technology to which increases in bandwidth constitute an inadequate response. It remains quite possible that the emergence of internet video technologies like YouTube, the development of new peer-to-peer technologies like BitTorrent, the development of complementary technology such as faster computer chips,[158] or some other as yet unanticipated technological change will eventually require a similar response.

In addition, network management need not be merely a transitional response to congestion while the network builds more pipes. In essence, this solution requires the maintenance of excess capacity to protect against the degradation of quality of service during short-term traffic surges. As I have argued earlier, building new bandwidth and network management represent alternative responses to the problems caused by congestion. In a world in which the relative cost of each solution is constantly changing, there is a good argument against regarding either alternative as being off the table.[159]

Indeed, the FCC explicitly recognized as much when establishing a public/private partnership to govern the public safety spectrum allocated through the 700 MHz auction. The FCC's solution was to allow public safety and commercial traffic to share bandwidth, but to give the former priority over the latter when the network becomes congested. As the FCC noted, such prioritized sharing should "both help to defray the costs of build-out and ensure that the spectrum is used efficiently."[160] This decision acknowledged that prioritization of higher value traffic represents an effective way to lower the cost of providing service, while at the same time representing a creative solution to extant bandwidth limitations. Furthermore, a recent study

---

158 Michael Fitzgerald, *Trying to Put New Zip into Moore's Law*, NY Times 4 (Feb 24, 2008).

159 See Yoo, 19 Harv J L & Tech at 22-23, 70-71 (cited in note 22) ("[O]ne would expect that the relative costs of different types of solutions to change over time. Sometimes increases in bandwidth would be cheaper than reliance on network management techniques, and vice versa. It would thus be short-sighted to tie network managers' hands by limiting their flexibility in their choice of network management solutions.").

160 Service Rules for the 698-746, 747-762 and 777-792 MHz Bands, Second Report and Order, 22 FCC Rec 15289, 15431 ¶ 396 (2007).

estimated that a network without prioritization might have to maintain up to 60 percent more capacity than a network offering prioritized service, assuming that links operate on average at 60 percent of capacity and that delay-sensitive traffic represents 20 percent of all traffic.[161]

This is not to say that prioritization will be either easy or perfect. Network providers may find it difficult to distinguish between high value and low value traffic, particularly given the strategic incentives to misrepresent the appropriate level of priority. I only seek to make the more limited point that there is no reason to assume a priori that maintaining excess bandwidth will always be the most cost-effective solution.

2. Prioritization as a Way to Promote Innovation

An alternative is to allow differential service and to raise prices for only those end users and content and applications providers who need the enhanced services. That way the content and applications providers who need the higher service can get it, while those who do not need it will not have to pay more. Prioritization on the internet can be analogized to the provision of different classes of mail service.[162] Overnight mail gets there overnight and costs roughly $10. First class mail may take as long as two to three days and costs forty-two cents.

What would happen if regulation forced all classes of mail to travel at the same speed? One option would be for the provider to make everything overnight mail, in which case everyone would have to pay more for mail service, even those who did not need their letters and packages to arrive the next day. Another option would be to make everything first-class mail, in which case those who need to get items to another city overnight could not do so even if they were

---

161 See Joseph D. Houle et al, *The Evolving Internet—Traffic, Engineering, and Roles* 5 (paper presented at the 35th Annual Telecommunications Policy Research Conference), available at <http://www.cse.unr.edu/~yuksem/my-papers/2007-tprc.pdf> (last visited Apr 10, 2008).

162 See Tim Wu and Christopher S. Yoo, *Keeping the Internet Neutral?: Tim Wu and Christopher Yoo Debate*, 59 Fed Comm L J 575, 578-79 (2007) (offering the same analogy).

willing to pay, which in turn would render business models that depend on faster service impossible. The provider could also split the difference and provide service that lies somewhere in between, in which case people who need the guarantee of truly fast service could not get what they need, and people who were fine with first-class mail would still end up paying more than forty-two cents. The problem becomes even more complicated once lower classes of mail service are taken into account. The more straightforward solution is to allow multiple classes of service. That way, those and only those who need the improved service bear the costs of the service.

This analogy underscores the mistake in arguing that creating tiers of service means that only the rich will get the fast lane. That would make as much sense as saying that only the rich get to use express mail. On the contrary, the rich will continue to send some letters and packages via first class mail, while those of lesser means will make some use of overnight mail to the extent that they can afford it. Instead, it means that only those who need their mail to get there overnight will use overnight mail. Creating different classes of service ensures that people will continue to have the choice. Although the government might intervene to ensure that end users of modest means have access to more advanced services, as I have argued at length elsewhere, such goals are better served by a targeted subsidy program than an untargeted system that attempts to preserve access by enforcing uniform prices across the board.[163]

The internet is already employing a wide variety of arrangements that guarantee quality of service, most of which are uncontroversial. For example, virtual private networks (VPNs)

---

[163] See Christopher S. Yoo, *The Rise and Demise of the Technology-Specific Approach to the First Amendment*, 91 Georgetown L J 245, 354 (2003) (similarly arguing in the context of television service that a direct subsidy targeted at low-income users would be far more cost effective than an untargeted system that attempts to preserve access by keeping prices down).

provide dedicated bandwidth. Indeed, more than 50 percent of VPNs guarantee some minimal level of quality of service.

Increasingly, ISPs have begun to deviate from the traditional dichotomy of types of contracts (peering and transit) and have begun to employ alternative arrangements, such as paid peering.[164] Although the primary explanation is asymmetry, another factor is the desire for quality of service.

In addition, many wireless platforms have begun to use prioritization techniques to compensate for the inherent bandwidth limits and propagation characteristics of spectrum-based communication. For example, some networks are attempting to leverage the fact that the physics of wave propagation dictates that the available bandwidth can vary as a person walks across the room.[165] Some networks have begun to experiment with protocols that give priority to latency-sensitive applications, such as VoIP, and hold delay-tolerant applications, such as email, until the end user reaches a location where the available bandwidth is relatively large, at which point the network will download all of the email at once.[166] This is an innovative solution to a real problem. Because those protocols discriminate on the basis of application, it is precisely the type of solution that network neutrality would prohibit.

Formal models indicate that allowing networks to offer premium services can stimulate innovation at the edges of the network, particularly among smaller content providers.[167] The

---

164 See note 74 and accompanying text.

165 See E.H. Choi et al., *Throughput of the 1x EV-DO System with Various Scheduling Algorithms*, 2004 IEEE International Symposium on Spread Spectrum Techniques and Applications 359 (Sept 4, 2004); Jinho Hwang et al., *Policy-Based QoS-Aware Packet Scheduling for CDMA 1x EV-DO* (Nov 28, 2006), available at <http://cs.seas.gwu.edu/research/reports-detail.php?trnumber=TR-GWU-CS-06-005>.

166 Id.

167 See Jamison and Hauge, *Getting What You Pay For* at 15, 19 (cited in note 39) (finding that "[t]he variety of content at the digs of the network increases when the network provider optimally chooses to offer premium transmission services," that "the value that consumers receive from the sites that purchase the premium transmission service is greater than the value they would receive if the premium transmission service were not offered," and that

primary opponents of these changes are those industry players who already have the most invested in the way the network is configured today, as well as those innovators on the verge of entering the market. These content and applications providers are and will no doubt continue to be important sources of innovation. At the same time, there are a number of other innovators whose innovations depend on the emergence of a network with greater capabilities. In preserving the benefits provided by the former, policymakers should take care not to ignore the innovative potential of the latter.

B. The Role of Price Signals

The flexibility to charge different content and applications providers different prices also promotes innovation in another way. The best way to understand this point is by focusing on the role that prices play in signaling to others the value people place on goods.

Allowing the prices to vary would allow network providers to reward those content and applications providers offering higher quality content. Indeed, such variations in pricing are quite common in the cable television industry, a model that is becoming increasingly important as Verizon's FiOS network and AT&T's VDSL-based U-Verse network draw an increasing proportion of their revenue from multichannel video.[168] Such video services are made possible by the fact that both Verizon and AT&T give video programming priority over other traffic that

---

"offering premium service stimulates innovation on the edges of the network because lower-value content sites are better able to compete with higher-value sites with the availability of the premium service" .").

[168] See Todd Spangler, *Verizon's FiOS TV Now at 1.2 Million Subs*, Multichannel News 20 (May 5 2008) (reporting the rapid growth of FiOS TV); Saul Hansell, *Wireless Business Helps AT&T's Profit Climb 22%*, NY Times C8 (Apr 23, 2008) (reporting that AT&T added 148,000 U-Verse television customers during the first quarter of 2008 for a total of 379,000 subscribers and that AT&T hopes to have 1 million subscribers by the end of the year).

is less sensitive to latency.[169] The content providers offering the highest quality and most unique content get better financial terms than those that do not. In cable, the cash flows the other way, with the network providers paying the content providers, but the principle of flexibility is the same.

Conversely, the lack of pricing flexibility would dampen the incentives for content and applications providers to innovate in content. No matter how attractive the content provider made its content, the price charged by the network would not vary. Although the opportunities for greater revenues for advertising and sales of other goods through the website do provide some incentive, the incentive would be weaker than if content and applications providers were paid directly for content. The absence of pricing flexibility will have a particularly strong negative impact on programming that is intensely preferred by a small segment of the overall audience.

In this respect, the FCC can draw on the insights from its experience with advertising-supported television.[170] In conventional markets, consumers use prices to signal the intensity of their preferences. To use a somewhat fanciful example, suppose that there is a group of one thousand extremely loyal fans of an out-of-town football team living in Philadelphia. In fact, this group is so loyal that they would each be willing to pay $10,000 each to have their games

---

[169] See AT&T and BellSouth Corp., Memorandum Opinion and Order, 22 FCC Rec 5662, 5814 (2007) (excluding AT&T's IPTV services from its network neutrality commitment).

[170] See Christopher S. Yoo, *Rethinking the Commitment to Free, Local Television*, 52 Emory L J 1579,1677-82 (2003) ("[R]eliance on pay television made possible the production of programming that would not have existed had advertising support represented the only option."); Christopher S. Yoo, *Architectural Censorship and the FCC*, 78 S Cal L Rev 669, 676-85 (2005) ("[E]fforts by Congress and the FCC to promote free [(advertising-supported)] radio and television" have "had a hidden, deleterious effect on the quantity, quality, and diversity of programming provided.").

televised locally.[171] If these fans are able to pay directly for programming, they can use the intensity of their preferences to get what they want. As a result, the programming should appear so long as revenue of $10 million is sufficient to cover the cost of televising the games to Philadelphia.

Contrast this with a world in which consumers are not allowed to vary the amount that they pay for different programs, as was the case when television was advertising supported.[172] Under advertising support, the revenue generated by programming was determined by audiences' responsiveness to the advertising contained in those programs. This responsiveness did not vary much from program to program. When prices do not vary, consumers only have one degree of freedom with which to express the intensity of their preferences: viewing versus nonviewing. This fact made revenue almost entirely a function of audience size, which rendered programming that appealed only to a small audience infeasible no matter how much those viewers would have been willing to pay for it. The relevance of this argument is demonstrated by the tremendous success of HBO, which is able to generate more than half the revenue of CBS even though its audience is nearly fifteen times smaller.[173] Put another way, allowing consumers to use prices to signal the intensity of their preferences directly allows HBO to generate eight times more

---

171 In a previous paper, I built the example around a small group of opera lovers and relegated sports fans to a footnote. See Yoo, 78 S Cal L Rev at 682-83 & n 44 (cited in note 170). I thought it only appropriate to reverse the priority here.

172 When the television industry first emerged, the technology did not exist to support pay television. When the technology needed to support direct payments for programming was invented, the FCC regulations largely blocked the practice. It was not until the D.C. Circuit overturned most of those regulations in 1977 that premium cable services and other forms of pay television began to emerge. See id at 676-78 ("The hostility toward subscription media services is also manifest in U.S. television policy. Then the development of scrambling technology made subscription television feasible, the FCC acted fairly quickly to stifle the industry's growth."); Yoo, 52 Emory L J at 1669-71 & n 298 (cited in note 170) ("[F]ollowing the 1977 judicial invalidation of the parallel restrictions on pay cable …, the FCC eventually repealed the program restrictions on [subscription television].").

173 Yoo, 52 Emory L J at 1679-80 (cited in note 170).

revenue per viewer than CBS.[174] HBO uses part of this additional revenue to offer higher quality programming, as demonstrated by its recent dominance of the Emmy awards.[175] At the same time, the higher revenue per customer lowers the minimum audience size that HBO needs to break even, which in turn makes it feasible for HBO to offer programming that appeals to smaller audiences.

The television industry thus provides an example of how enabling consumers to use prices to signal the intensity of their preferences can promote innovation in the quality and diversity of content. Such signals are particularly important when the quality of the content varies widely. It thus comes as no surprise that the magnitude of cash flows to various cable networks varies widely, with those programmers offering the highest quality and most distinctive programming faring the best and with those offering low quality, me-too programming faring the worst. It is these price signals that indicate to content and application providers which additional investments in developing additional content and applications will provide the greatest consumer value. Cutting off these price signals will make it difficult, if not impossible, for content and applications providers to learn in the areas which they need to make greater investments. The problem is that under the pricing regime that dominates the internet, the fact that backbones peer with one another on a settlement-free basis prevents such price signals from being transmitted directly through the network. Preventing network providers from charging content and applications providers different prices for different levels of service would guarantee that such signals would never be transmitted through the network itself.

Indeed, although the pricing regimes that currently dominate the internet may seem inevitable that content and applications providers will pay network providers, there is nothing

---

174 Id.

175 Yoo, 78 S Cal L Rev at 682 (cited in note 170).

inevitable about it. The cash can flow in either direction. To use an example again from television, in broadcasting, networks that provided content to local television stations generally paid those stations to carry their content. In recent years, however, weaker stations have seen the direction of the cash flow turn around, as the networks have begun to ask stations with weaker signals and smaller audiences to pay instead of being paid.

That said, most observers expect the direction to remain the same. Content providers are increasingly adopting business models based on advertising revenue. At the same time, the network providers providing the last-mile connection to content and applications providers face far lower costs than the network providers providing the last-mile connection to end user. The former only need to provide a high-capacity from a commercial location to an interexchange point maintained by a backbone or a regional ISP. The latter must typically establish extensive networks of wires and other electronics blanketing the area they intend to serve. Since revenue is flowing into the network on the content side and the costs are on the eyeball side, many expect cash to flow from content networks to eyeball networks.[176] As a theoretical matter, the direction is ambiguous, and it remains theoretically possible that the cash will flow in the other direction. Indeed, that is exactly what is happening with a new product called ESPN 360. In that case, end users cannot get access to ESPN 360 unless the network provider pays ESPN a fee. We may see more of these solutions in the years to come.

The ambiguity of the direction of the cash flow raises some question of whether the leading content and applications providers who support network neutrality might be making a mistake and that allowing greater pricing flexibility might ultimately work to their benefit. On the other hand, the magnitude of the advertising revenue flowing to the content and applications

---

[176] See Faratin et al, *Complexity of Internet Interconnections* at 13 (cited in note 73) ("[M]oney flows in at the content end … and the content provider places more value on the eyeball than the eyeball does on the content.").

providers may make it more likely that locking in prices at current levels will ultimately redound to the benefit of content and applications providers.

The argument I am advancing does not depend on determining the precise direction and magnitude of the various cash flows. Indeed, the dynamic nature of the underlying technology and consumer tastes implies that both are constantly changing. The broader point is that giving network providers greater pricing flexibility might well benefit consumers by ensuring that prices can serve as the signal that allocates resources in a more efficient way. While insufficient to establish that such pricing flexibility would always benefit consumers, the plausibility of such consumer benefits represents an exemplifying theory sufficient to rebut arguments in favor of categorically prohibiting such pricing flexibility and provides strong support for permitting a variety of pricing practices in the absence of any affirmative showing that those practices actually harm consumers. The possibility of exemplifying theories showing potential consumer harm would justify a case-by-case approach rather than the type of categorical prohibition that network neutrality proponents seek.

C. Short-Term Deadlock as an Inevitable Part of Economic Bargaining

Pricing mechanisms require a certain degree of give and take if they are to function properly. New developments arise constantly to which the market needs time to adjust. For example, when they first arose, network providers prohibited the use of VPNs and home networking devices.[177] This restriction ultimately proved short-lived. Consumer pressure soon induced the network providers to change course. Although some observers have pointed to this

---

[177] For an analysis, Yoo, 94 Georgetown L J at 1856 (cited note 1) ("Other end user restrictions included prohibitions on … engaging in home networking … and employing commercial applications such as virtual private networks (VPNs).").").

development as demonstrating the need to impose network neutrality regulation,[178] I think it demonstrates the opposite. It shows how competitive pressures and consumer preferences revealed through individuals' purchasing decisions can force openness in ways that render regulation unnecessary.

At times, the network providers and content and applications providers may disagree regarding the value of their products. Rather than jump in, regulators should allow the give and take of the bargaining process to work its way through. The point is illustrated by two recent examples involving the cable and satellite television industries. In both industries, broadcast stations have a choice. They can either invoke their right to free carriage on a cable or satellite system, or, if the broadcaster thinks they can obtain some form of economic compensation from the cable or satellite system, they can forego those rights and instead negotiate their own carriage arrangements through arms-length negotiations.[179] Most are resolved amicably enough. But sometimes, differences of opinion about the relative value the other party is providing and the strength of one's bargaining position can lead to temporary deadlock.

The television stations owned by Disney/ABC had long opted not to exercise their rights to free carriage on local cable systems and instead opted to negotiate their own retransmission consent agreements. Just before the retransmission consent agreement between Disney and Time Warner Cable was to expire in December 1999, the parties negotiated a $1 billion follow-on

---

[178] See Tim Wu, *Why Have a Telecommunications Law? Anti-Discrimination Norms in Communications*, 5 J on Telecommun & High Tech L 15, 34-35 (2006) (listing network providers "disciplin[ing of] users of Virtual Private Networks" among the early instances of broadband discrimination that prompted calls for network neutrality); Windhausen, *Good Fences Make Bad Broadband* i, 17-18 (cited in note 13) (holding up blocking of VPNs as an example of the type blocking or discrimination by network operators that should be barred by network neutrality regulation).

[179] See Yoo, 52 Emory L J at 1648, 1651-52 (cited in note 170) (providing an overview of these regulatory frameworks).

agreement that would have lasted ten years.[180] But as that deal was about to be consummated, America Online announced its agreement to acquire Time Warner. ABC immediately asked for a fee increase of an additional $300 million, and Time Warner refused.[181] After five months of short-term extensions of the previous agreement and additional negotiations failed to yield an agreement, Time Warner dropped all of the ABC-owned stations on May 1, 2000.[182] ABC filed a complaint with the FCC the same day, and FCC Chairman William Kennard warned that "[t]he television sets of average consumers should never be held hostage in these disputes" and criticized "[t]he game of brinkmanship" being played by the parties.[183] In the shadow of impending FCC action, Time Warner capitulated and put the stations back on after only one day. The FCC ruled Time Warner's actions illegal,[184] with Chairman Kennard again warning, "No company should use consumers as pawns in a private contract dispute."[185] Time Warner reached an agreement with ABC later that month, in which ABC received the $300 million increase it sought.[186]

---

180 See Harry Berkowitz, *Back on the Tube: Time Warner Sets New Deadline with Walt Disney Co.*, Newsday A7 (May 3 2000).

181 See Bill Carter, *Blackout of ABC on Cable Affects Millions of Homes*, NY Times C1 (May 2, 2000).

182 See Marc Gunther, *Dumb and Dumber*, Fortune 140 (May 29, 2000) (describing the dispute between Time Warner and ABC). Though Time Warner walked away from ABC, "[t]his isn't to say that Time Warner executives weren't provoked. Disney CEO Michael Eisner was, in effect, greenmailing his biggest competitor, threatening to turn his lobbyists loose unless Time Warner agreed to pay substantially more for Disney's cable channels…." Id.

183 Statement of FCC Chairman William E. Kennard Regarding Disney/ABC and Time Warner Dispute (May 2, 2000), available at <http://www.fcc.gov/Speeches/Kennard/Statements/2000/stwek035.html> (last visited Apr 10, 2008).

184 Time Warner Cable, Emergency Petition of ABC, Inc. for Declaratory Ruling and Enforcement Order, 15 FCC Rec 7882 (2000) (ruling that "the removal of the ABC Stations' signals from its systems by Time Warner Cable was in violation of Section 614(b)(9) of the Communications Act and Section 76.58 of the Commission's rules.").

185 Statement of FCC Chairman William E. Kennard on Ruling in Time Warner-Disney Dispute (May 3, 2000), available at <http://www.fcc.gov/ZSpeehces/Kennard/Statements/2000/stwek036.html> (last visited Mar 6, 2008).

186 *Time Warner, Disney Sign Long-Term Deal*, Television Digest (May 29, 2000) (available at 2000 WLNR 4874807) ("Disney will receive about 30% more in compensation, or roughly extra $300 million, than it would have under original $1 billion deal….").

A similar dispute arose in March 2004, when DBS provider EchoStar was unable to reach a carriage agreement with Viacom/CBS. EchoStar cut Viacom programming on March 9, which left 9 million subscribers without MTV programming and 2 million subscribers without CBS. This time the FCC followed a very different course of action. Rather than criticizing the parties, FCC Chairman Michael Powell simply acknowledged, "That's what sometimes happens in the market. Consumers usually lose and so do both parties. It usually doesn't happen very long."[187] The parties settled the dispute two days later, largely on Viacom's terms. Although both sides claimed victory,[188] the ultimate terms were essentially what Viacom initially sought.[189]

These two episodes illustrate the breathing room that is needed if real economic bargaining is to occur. Interfering with the give and take inherent in negotiations would eliminate the equilibrating and incentive effects of price signals. Policymakers must thus be careful not to regard the inability to reach agreement as a definitive sign of market failure or the necessity of government intervention. On the contrary, a certain amount of deadlock is the sign of a properly functioning economic market.

We should be even more tolerant with respect to the internet. With the increasing complexity of interconnections (for example, secondary peering and multihoming[190]), it is less likely that the failure of any particular bargaining relationship will prevent any end user from being able to access any content. Instead, the level of redundancy should allow them to still reach it, although it may be at higher cost.

---

187 Jonathan D. Salant, *EchoStar, Viacom Expect Quick Resolution*, Kan City Star C3 (Mar 11, 2004).

188 See Kris Hudson, *EchoStar, Viacom Claim Win: Angry Viewers Cited as Motivating Factor*, Denver Post C1 (Mar 12, 2004) ("Both companies claimed victory on the rate issue, with EchoStar saying the smaller-than-demanded increase will allow it to keep its customers' bills low; Viacom claims it demanded a 6-cent increase all along.").

189 See Phyllis Furman, *Viacom Seen Getting Best of EchoStar*, NY Daily News 76 (Mar 12, 2004) ("Wall Street immediately branded Viacom the victor, claiming its president, Mel Karmazin, got just what he wanted from EchoStar mogul Charlie Ergen.").

190 See notes 75-79 and accompanying text.

D. The Role of Exclusivity

The economic literature also underscores the potential innovation-enhancing effects of exclusivity arrangements.[191] Fledgling new transmission platforms often rely on exclusivity arrangements to make their offerings more attractive. Perhaps the best recent demonstration of this insight is the key role that exclusive access to the programming package known as NFL Sunday Ticket has played in helping DirecTV attract subscribers away from cable.[192] DBS's emergence as an effective competitor to cable has significantly improved consumer welfare.[193] Similarly, a recent empirical study of the video game industry indicates that exclusive access to certain hit games played a key role in helping Microsoft's Xbox enter as a competitor in an industry dominated by Sony's Playstation 2 and Nintendo's Gamecube.[194]

The primary objection raised by network neutrality proponents is that exclusivity arrangements limit consumer choice. Exclusivity arrangements may also serve to deter entry or foreclose rivals, which can also reduce consumer welfare.[195] The impact on consumers thus depends on whether the positive or negative effects dominate.

---

191 See Howard P. Marvel, *Exclusive Dealing*, 25 J L & Econ 1 (1982); Benjamin Klein, *Vertical Integration as Organizational Ownership: The Fisher Body-General Motors Relationship Revisited*, 4 J L Econ & Org 199 (1988); David Besanko and Martin K. Perry, *Equilibrium Incentives for Exclusive Dealing in a Differentiated Products Oligopoly*, 24 RAND J Econ 646 (1993); Carl Shapiro, *Exclusivity in Network Industries*, 7 Geo Mason L Rev 673 (1999); Ilya Segal and Michael D. Whinston, *Exclusive Contracts and Protection of Investments*, 31 RAND J Econ 603 (2000).

192 Yoo, 19 Harv J L & Tech at 32 (cited in note 22).

193 See Austin Goolsbee and Amil Petrin, *The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV*, 72 Econometrica 351, 351 (2004) ("Estimates of the supply response of cable suggest that without [Direct Broadcasting Satellite] entry cable prices would be about 15 percent higher and cable quality would fall. We find a welfare gain of between $127 and $190 per year (aggregate $2.5 billion) for satellite buyers, and about $50 (aggregate $3 billion) for cable subscribers.").

194 See Robin S. Lee, *Vertical Integration and Exclusivity in Platform and Two-Sided Markets*, 34 (Oct 15, 2007), available at <http://ssrn.com/abstract=1022682> (last visited Mar 6, 2008) (empirical study showing how exclusivity over popular games was a key driver in promoting sales of the new Xbox gaming platform).

195 See G. Frank Mathewson and Ralph A. Winter, *The Competitive Effects of Vertical Agreements: Comment*, 77 Am Econ Rev 1057 (1987) ("Exclusive dealing imposed by the dominant manufacturer eliminates its rival from the market—exclusive dealing both reduces actual competition and restricts the consumers' choice set."); Eric B. Rasmusen, J. Mark Ramseyer, and John S. Wiley, Jr., *Naked Exclusion*, 81 Am Econ Rev 1137 (1991)

We need not resolve this debate for the purposes of this Article. The theoretical literature showing the possibility that exclusivity can enhance consumer welfare as well as the empirical literature indicating that those benefits have been realized in other industries suggests that the claim is sufficiently plausible not to prohibit exclusivity arrangements categorically, as network neutrality proponents propose. Absent concrete empirical support, exemplifying theory pointing in the other direction would not establish that anticompetitive effects are more likely and would only underscore the ambiguousness of the welfare calculus. In so doing, it would reinforce the justifications for a case-by-case approach rather than offer support for a blanket network neutrality mandate.

E. The Ambiguous Role of Network Economic Effects

Network neutrality proponents argue that innovation depends on content and applications providers having access to the entire market. Absent perfect interconnectivity, they will not be able to innovate. In effect, these arguments are either explicitly or implicitly based on network economic effects.[196]

(“[M]onopolists may be able to exploit customer disorganization so as to exclude potential rivals.”); B. Douglas Bernheim and Michael D. Whinston, *Exclusive Dealing*, 106 J Pol Econ 64 (1998) (“We provide formal theoretical foundations for the view that exclusive dealing may be adopted for anticompetitive reasons (to enhance market power in noncoincident markets) ….”).

[196] See Lessig, *Future of Ideas* at 171 (cited in note 14) (“eBay benefits greatly from a network that is open and where access is free. It is this general feature of the Net that makes the Net so valuable to users and a source of innovation. And to the extent that individual sites begin to impose their own rules of exclusion, the value of the network as a network declines.”); Jerry A. Hausman et al, *Residential Demand for Broadband Telecommunications and Consumer Access to Unaffiliated Internet Content Providers*, 18 Yale J Reg 129, 161-62, 163-65 (2001) (arguing that “[t]he academic literature on … network externalities provides theoretical and empirical support for the conjecture that AT&T could impose proprietary standards that would raise the switching costs for its subscribers and stifle competition in vertically related software markets” and that once established, “the early leader in any broadband Internet access may enjoy a ‘lock-in’ of customers and content providers”); Wu, 2 J on Telecommun & High Tech L at 151 (cited in note 106) (“If broadband operators were to ban IP chat programs, other network applications, like file-exchange programs, may be hurt. Applications that depend on a critical mass of users may also be hurt if users will not pay for broadband when chat programs are not available.”).

If consumer preferences are sufficiently homogenous that all end users effectively want the same thing from the network, the result is standardization on a single network, and consumers and content and applications providers all benefit from being part of the largest network possible. A different situation obtains when what consumers want from the network varies. As I have pointed out in my previous work, when consumer preferences are heterogeneous, standardization on any particular protocol involves a tradeoff.[197] Consider the decision faced by two groups of end users that each prefers a different network standard. A group could adopt its preferred standard, in which case it would enjoy the benefits of employing the standard best suited to its preferences, but would forego the benefits of being part of a larger network. Or it could adopt the standard preferred by the other group, in which case it would enjoy the benefits of being part of a larger network, but would forego the benefits of employing the standard it prefers. This is why the leading network theorists regard the loss of product variety as one of the primary costs of standardization.[198] These opposing considerations provide a basis for determining the optimal level of variety. Where the market will reach equilibrium depends on whether the benefits from being part of a larger network dominate the benefits from employing the standard best suited to a particular group's preferences or vice versa.[199]

---

[197] Yoo, 19 Harv J L & Tech at 34-36 (cited in note 22) (showing how standardization involves a tradeoff between the benefits of consuming a product incorporating a consumer's preferred standard and the benefits of belonging to a larger network).

[198] Michael L. Katz and Carl Shapiro, *Systems Competition and Network Effects*, 8 J Econ Persp 93, 110 (1994) (noting that "the primary cost of standardization is loss of variety: consumers have fewer differentiated products to pick from"); Joseph Farrell and Garth Saloner, *Standardization, Compatibility, and Innovation*, 16 RAND J Econ 70, 71 (1985) (counting "reduction in variety" as one of the "important social costs" of standardization).

[199] See Joseph Farrell and Garth Saloner, *Standardization and Variety*, 20 Econ Letters 71 (1986) (offering a simple formal model capturing this tradeoff). Frischmann and van Schewick assert that some of the productive activities in which internet users engage when using the internet generates social benefits that market transactions do not take into account. As I have noted earlier, it is impossible to determine a priori whether any such positive externalities are creating systematic underproduction or instead they are offset by countervailing negative externalities. See note 103.

The ultimate ambiguity of arguments based on network economic effects is demonstrated by how the network neutrality debate inverts the usual economic arguments. The usual argument is that once markets become "tipped," network economic effects cause markets to become locked into the technologies long after they have become obsolete. This is because new technologies are unable to generate sufficient critical mass to displace the existing technology. The concern is that network participants will become too reluctant to innovate away from the standard that currently dominates. On this logic, the concern should be that the internet is a tipped market that has become locked in to an obsolete technology.[200] If so, policymakers should encourage, rather than discourage, network players to experiment with deviations from the current standard.

Network neutrality proponents take precisely the opposite position by advocating that all network operators maintain perfect interoperability, presumably on the current standard. There is no theoretical justification for such a position. The economic literature suggests that it is just as possible for markets subject to network economic effects to exhibit "excess momentum" as well as "excess inertia."[201] As a theoretical matter, it is possible for network providers to be too eager to deviate from the current standard, in which case requiring them to adhere to the current standard might benefit consumers. It is also possible for network providers to be too reluctant to deviate from the current standard, in which case experimentation with different protocols should be encouraged. Indeed, to the extent that markets are subject to excess inertia, formal models indicate that competition between a proprietary and a nonproprietary standard or between two proprietary standards may compensate (or even overcompensate) for inefficiencies in technology

---

[200] Indeed, some leading technologists have suggested that TCP/IP may be obsolete. See Carol Wilson, *Point of No Return*, Telephony (Apr 3, 2006), available at <http://blog.tmcnet.com/blog/rich-tehrani/voip/point-of-no-return.html> (quoting former FCC Chief Technologist and Carnegie Mellon Professor David Farber that the internet is "getting old" and needs new functionality, especially to handle new services such as video).

[201] See Joseph Farrell and Garth Saloner, *Installed Base and Compatibility: Innovation, Product Preannouncements, and Predation*, 76 Am Econ Rev 940, 942 (1986); Katz and Shapiro, 94 J Pol Econ 822 (cited in note 140).

adoption caused by network economic effects..[202] There is no theoretical reason *a priori* to pick one over the other, and the extant empirical evidence tends to contradict claims that network economic effects cause inefficient technology adoption decisions.[203]

IV. NETWORK DIVERSITY AS AN APPROPRIATE MIDDLE GROUND WHEN THE IMPACT ON CONSUMERS AND INNOVATION IS AMBIGUOUS

The preceding Parts have shown ways in which deviating from network neutrality might benefit consumers by reducing the congestion costs and by creating flexible pricing regimes that increase aggregate consumer welfare. They have also shown ways in which deviating from network neutrality might promote innovation by deploying innovative new protocols, using prices to provide incentives to create new content and applications, and by facilitating entry by new transmission platforms. The question is what policy inferences should be drawn from these facts?

As I have argued at length elsewhere, I believe that the Supreme Court's antitrust jurisprudence provides a useful framework for answering this question.[204] Turning to antitrust law, and in particular the law on vertical restraints, is particularly appropriate because the concern is analytically the same as the concern motivating network neutrality proponents, that is, that a firm operating at one level of a chain of production will exercise its market power to reduce the competitiveness of an adjacent level of production.

---

202 See Katz and Shapiro, 94 J Pol Econ at 835, 838-39 (cited in note 140) (showing how competition between a proprietary and a nonproprietary standard or between two nonproprietary standards can compensate for the tendency toward inefficient nonstandardization associated with competition between two nonproprietary standards); Michael L. Katz and Carl Shapiro, *Product Introduction with Network Externalities*, 40 J Indus Econ 55, 73 (1992) (showing how sponsorship of technologies can overcome the bias toward incumbent technologies and instead exhibit a bias toward new technologies).

203 See Spulber and Yoo, 88 Cornell L Rev at 930-31 (cited in note 84).

204 See Christopher S. Yoo, *What Can Antitrust Contribute to the Network Neutrality Debate*, 1 Intl J Comm 493, 508-17 (2007) (reviewing the implications of antitrust law for the network neutrality debate); Yoo, 19 Harv J L & Tech at 69-70 (cited in note 22) (same).

The Supreme Court restated the law governing vertical restraints in *State Oil Co v Khan*,[205] in which the Court recognized that "most antitrust claims are analyzed under a 'rule of reason,'" under which courts evaluate the competitive impact of a particular practice on a case-by-case basis in light of all of the facts.[206] If, however, a court has sufficient experience with a particular vertical restraint to conclude with confidence that it evinces "such predictable and pernicious anticompetitive effect, and such limited potential for procompetitive benefit" that nothing would be lost by prohibiting it without any detailed inquiry into the specific facts, it should be categorically prohibited and declared illegal per se.[207] If, on the other hand, "'the economic impact of [the challenged] practices is not immediately obvious,'" courts should refrain from imposing a per se rule and continue to apply the rule of reason.[208]

The Court reiterated these same principles just last year in *Leegin Creative Leather Products, Inc v PSKS, Inc*.[209] The Court began by noting, "The rule of reason is the accepted standard for testing whether a practice restrains trade in violation of" the antitrust laws and that courts should declare a practice illegal per se only if it evinces "'manifestly anticompetitive'" effects and a "'lack [of] any redeeming virtue.'"[210] The Court continued:

> As a consequence, the *per se* rule is appropriate only after courts have had considerable experience with the type of restraint at issue and only if courts can predict with confidence that it would be invalidated in all or almost all instances under the rule of reason. It should come as no surprise, then, that "we have expressed reluctance to adopt *per se* rules with regard to restraints imposed in the context of business relationships where the economic impact of certain practices

---

205 522 US 3 (1997).
206 Id at 10.
207 Id.
208 Id.
209 127 S Ct 2705 (2007).
210 Id at 2712, quoting *Continental TV, Inc v GTE Sylvania, Inc*, 433 US 36, 50 (1977); and *Nw Wholesale Stationers, Inc v Pac Stationery & Printing Co*, 472 US 284, 289 (1985)).

> is not immediately obvious." And, as we have stated, a "departure from the rule-of-reason standard must be based upon demonstrable economic effect…."[211]

The analytical framework implicit in this approach is clear: If a practice is always harmful (or so nearly always harmful that any inquiry into its competitive impact would likely prove fruitless), the practice should be categorically prohibited. Conversely, if a practice is always or nearly always beneficial, it should always be categorically declared legal, as indeed some Chicago School theorists once advocated with respect to vertical restraints.[212] If the precise competitive impact of a practice is as-yet unclear, or if it is theoretically possible for a practice to be harmful in some instances and beneficial in others, its precise competitive effect should be determined through an ex post, case-by-case analysis.

Applying this framework to the network neutrality debate provides a powerful argument against mandating network neutrality. The adoption of network neutrality regulations would be tantamount to declaring practices that deviate from network neutrality illegal per se. In the absence of a clear indication of what the competitive impact of practices deviating from network neutrality might be, those practices are better analyzed under the type of ex post, case-by-case approach that characterizes the rule of reason rather than the ex ante, categorical approach that characterizes per se illegality.

To justify their calls for a more categorical approach, network neutrality proponents invoke a series of exemplifying theories showing the existence of circumstances under which

---

211 Id (citations omitted and quoting *Khan*, 522 US at 10; and *Sylvania*, 433 US at 58-59).
212 See note 26 and accompanying text.

exclusionary practices can harm consumers.[213] For the most part, these theories are based on the game theoretic models that characterize the post-Chicago economic literature.[214]

Commentators on the post-Chicago literature, both sympathetic and skeptical, have long recognized that these game theoretic models depend on highly restrictive assumptions in order to yield results.[215] Ensuring that these preconditions are met is made all-the-more essential by a key aspect of game theory that makes its results less robust than traditional economic analyses. In neoclassical economics, small changes to the empirical parameters underlying the model lead only to small changes in the underlying equilibrium. The equilibria in game theoretic models, in contrast, are quite sensitive to changes in assumptions and often exhibit large, discontinuous changes to small changes to the underlying parameters.[216] Thus, before drawing any inferences, policymakers must be careful that all of the factual predicates of each model are satisfied,[217] even though those facts are often very difficult to verify.[218]

---

213 van Schewick, 5 J on Telecommun & High Tech L at 340-78 (cited in note 15); Frischmann and van Schewick, 47 Jurimetrics J at 412-20 (cited in note 12).

214 Joseph Farrell and Michael L. Katz, *Innovation, Rent Extraction, and Integration in Systems Markets*, 48 J Indus Econ 413 (2000); Michael D. Whinston, *Tying Foreclosure, and Exclusion*, 80 Am Econ Rev 837 (1990).

215 See Ian Ayres, *Playing Games with the Law*, 42 Stan L Rev 1291, 1310-11 (1990) (noting the tendency of game theoretic models to yield no equilibria or multiple equilibria without restrictive assumptions); Robert O. Keohane, *Rational Choice Theory and International Law: Insights and Limitations*, 31 J Legal Stud S307, S314 (2002) (noting that the well-known Folk Theorem indicates "that strict game-theoretic analysis, without restrictive assumptions, typically predicts multiple equilibria, sometimes involving cycling between different states of the world"); *Interview with Dennis A. Yao, Former FTC Commissioner*, Antitrust 12, 16 (Fall 1994) (observing that "[g]ame-theoretic models generally become unwieldy unless they adopt restrictive assumptions").

216 Ayres, 42 Stan L Rev at 1313-14 (cited in note 215).

217 Id at 1317 (observing that "particularized games are not neatly applicable to a broad spectrum of markets" and how the existence of "a number of 'possibility' theorems … force both sides toward empirical analysis"); *Interview with Dennis A. Yao,*, Antitrust at 16 (cited in note 215) (noting the importance of ensuring that the restrictive assumptions needed to produce game theoretic results are "descriptive or reality").

218 Herbert Hovenkamp, *Antitrust Policy After Chicago*, 84 Mich L Rev 213, 261 (1985) (noting that the factual issues central to post-Chicago theory may be "too complex to be dealt with in antitrust litigation" and acknowledging then-Chief Judge Breyer's conclusion that anticompetitive strategic behavior exists but is too complex for courts to sort out (citing *Barry Wright Corp. v. ITT Grinnell Corp.*, 724 F2d 227, 230-36 (1st Cir 1983)); Keith N. Hylton and Michael Salinger, *Tying Law and Policy: A Decision-Theoretic Approach*, 69 Antitrust L J 469, 470 (2001) ("[T]he game theory underpinning the literature rests on highly stylized assumptions that re difficult to apply to the factual settings courts confront."); Michael S. Jacobs, *An Essay on the Normative Foundations of Antitrust Economics*, 74 NC L Rev 219, 254 (1995) ("Moreover, many post-Chicago theories are

Even more fundamentally, game theoretic models at issue here make no attempt to formalize the overall impact on consumer welfare, either by offering a metric for determining optimal innovation or by taking into account potential efficiencies. The problem is that these models assume precisely the type of market structure that is likely to give rise to these efficiencies.[219]

A brief review of some of the models invoked by network neutrality proponents will help illustrate my point. Consider first Michael Whinston's seminal tying model.[220] Whinston effectively rebutted calls to treat tying as legal per se[221] by showing the existence of circumstances under which tying can be profitable and can exclude competitors. Whinston's base model turns on three key factual assumptions. First, it assumes that the firm engaging in the exclusionary conduct is a monopolist in one market, called the primary market, *A*, which consists of a_m potential buyers. Second, it assumes the existence of a market for a complementary product, *B*. Although everyone who purchases *A* also purchases *B*, there are also consumers who purchase only *B*, but not *A*. Third, the market for *B* must be characterized by economies of scale and that the minimum efficient scale for *B* must be larger than the number of *B* consumers who do not purchase *A*.

---

constructed so restrictively … that testing them is unproductive. Others depend on speculation about the reputational effects of apparently irrational strategic behavior … making empirical validation impossible.").

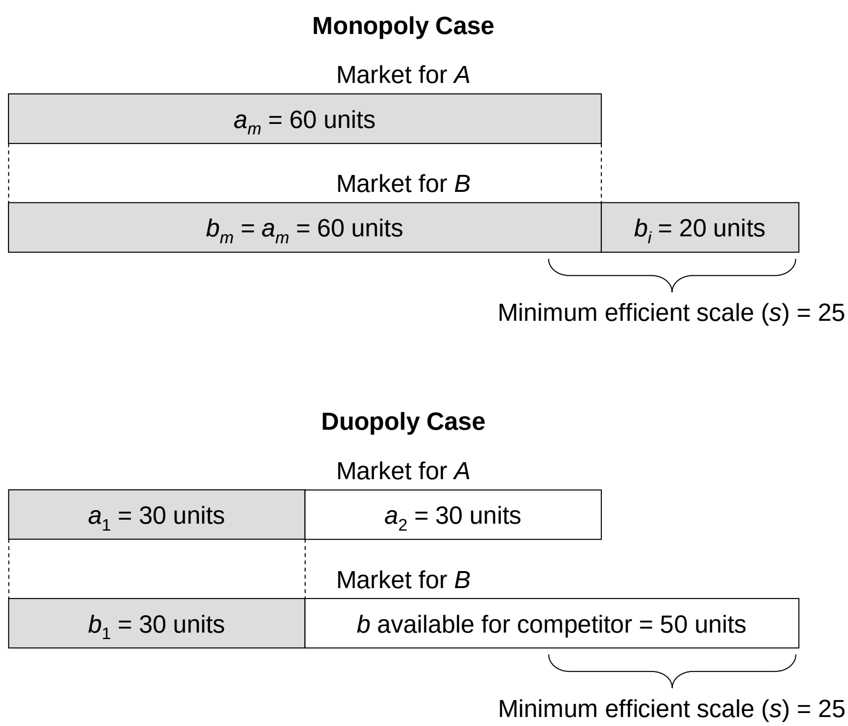
219 Hylton and Salinger, 69 Antitrust L J at 471 (cited in note 218) ("[T]he most plausible post-Chicago theory of anticompetitive tying is based on the assumption that the tying and tied goods are complementary and that they are both susceptible to market power and, indeed, monopoly. It is a long-established principle of economics, however, that integrated complementary monopoly results in lower prices than distinct complementary monopolies. A public policy that imparts a bias toward interdependent complementary monopolies instead of integrated complementary monopolies has the predictable consequence of raising prices and reducing consumer welfare."); Edward A. Snyder and Thomas E. Kauper, *Misuse of the Antitrust Laws: The Competitor Plaintiff*, 90 Mich L Rev 551, 589-91 (1991) (tracing the potential efficiency benefits and noting that "meeting the necessary conditions for anticompetitive exclusion is likely to ensure that some efficiency benefits are realized form the business practices I question.").

220 Whinston, 80 Am Econ Rev at 837 (cited in note 214), cited in van Schewick, 5 J on Telecommun & High Tech L at 342-52 (cited in note 15); and Frischmann and van Schewick, 47 Jurimetrics J at 412-14 (cited in note 12).

221 See, for example, Robert H. Bork, *The Antitrust Paradox* 380 (Free Press 1978) ("[A] logically consistent law would have to accept the legality of all tying arrangements….").

If the *A* monopolist does not tie its product to *B*, it can only exercise monopoly power over the number of consumers who purchase *A* (a_m). If, on the other hand, the *A* monopolist forces all *B* consumers who also purchase *A* (b_m) to purchase from it by tying *A* to *B*, it can increase its profits by foreclosing the market for *B*. If the minimum efficient scale (s) is larger than the number of consumers who buy only *B* (b_i), any independent producer of *B* will operate at a cost disadvantage versus the *A* monopolist and will not be able to compete. This in turn allows the *A* monopolist to drive all independent producers of *B* out of the market and to exercise market power over all *B* consumers (including both b_m and b_i) and not just those who also purchase *A* (b_m). In this manner, the *A* monopolist can exclude rival producers of *B* and thereby increase the number of consumers over which it can exert monopoly power.

Figure 6
A Numerical Example of Michael Whinston's Theory of Tying



The following numerical example may help illustrate the intuitions underlying Whinston's model. Suppose that the market for good *A* is 60 units and that that market is dominated by a monopolist. Suppose further that all 60 people who purchase good *A* also purchase good *B*, and there are 20 other consumers who only purchase good *B*. Whinston pointed out that if minimum efficient scale for producing *B* is greater than 20 units (say 25 units), the *A* monopolist can obtain a monopoly in the market for *B* simply by tying *A* and *B* together. This is because tying *A* and *B* only leaves 20 uncommitted consumers of *B*. Because the minimum efficient scale for producing *B* is greater than 20, every other potential producer of *B* will operate at a cost disadvantage in *B* production vis-à-vis the monopolist and thus will be unable to compete effectively. This allows the producer of *A* to exercise monopoly power over not only the 60 consumers who purchase both *A* and *B*, but also the 20 additional consumers who purchase only *B*.

A close analysis of the model reveals how much its results depend on its assumptions. Consider first the fact that the model depends on a very precise relationship between the number of consumers who purchase only *B* (b_i) and the minimum efficient scale for producing *B* (s). The model strictly requires that $s > b_i$. This is because the *A* monopolist can only tie up those *B* consumers who also purchase *A*. If the number of consumers that purchase only *B* exceeds the minimum efficient scale for producing *B* (that is, if $s < b_i$), a rival producer of *B* will not operate at a cost disadvantage and will not be driven from the market. Thus, Whinston's model in effect requires that the number of consumers who purchase only *B* (b_i) be relatively small and that the minimum efficient scale for producing *B* (s) be relatively large. Any deviation from this relationship prevents the effect that Whinston has identified from materializing.

In the context of network neutrality, the primary market is the market for last-mile internet access, and the secondary market is the market for content and applications. This raises the empirical questions of (1) whether there are consumers of internet content and applications who do not also purchase internet access services and (2) how large is the minimum efficient scale for providing internet content and applications. Note also that the number of consumers of internet content and applications who do not also purchase internet access services must be small relative to minimum efficient scale; otherwise the effect that Whinston identifies will not arise.

Network neutrality proponents assert that the first condition is met (that is, there are consumers of internet content and applications who do not also purchase internet access services) because the market for internet access is local. For example, the relevant market for last-mile internet services to my home is the city of Philadelphia, which for the time being we shall assume is a monopoly. The existence of consumers in other cities outside of Philadelphia represent consumers of the secondary good (internet content and applications) who do not also consume the primary good (last-mile internet access in Philadelphia).

I have long suggested that the relevant market for internet content and applications is national, not local. As I noted elsewhere:

> Major web-based providers, such as Amazon.com or eBay, are focused more on the total number of customers they are able to reach nationwide than they are on their ability to reach customers located in any specific metropolitan area. The fact that they may be unable to reach certain customers is of no greater concern, however, than the fact that manufacturers of particular brands of cars, shoes, or other conventional goods are not always able to gain distribution in all parts of the country. … The proper question is thus not whether the broadband transport provider wields market power vis-à-vis broadband users in any particular city, but rather whether that provider has market power in the national market for obtaining broadband content. In short, it is national reach, not local reach, that matters.[222]

Some network neutrality proponents assert that customers' interest in local news and local yellow pages is enough to render the relevant geographic market local.[223] Although that

---

222 Yoo, 19 Harv J L & Tech at 72-73 (cited in note 22); accord Yoo, 94 Georgetown L J at 1892-93 (cited in note 1); Yoo, 3 J on Telecommun & High Tech L at 52 (cited in note 34); Yoo, 19 Yale J Reg at 254 (cited in note 25).

223 See Frischmann and van Schewick, 47 Jurimetrics J at 412-16 (cited in note 12).

observation is true as far as it goes, acknowledging that the proper geographic scope for a small number websites is local does not undercut the fact that the relevant geographic market for the vast majority of websites and applications is national rather than local.[224] If so, it is not clear whether there is any residuum of purchasers of *B* (internet content and applications) who do not also purchase *A* (last-mile internet services).

But even if one were to accept the assertions of network neutrality proponents and assume that the relevant market is local, that would render the Whinston model inapposite. This is because leverage under the Whinston model is impossible if the number of consumers who purchase *B* without also purchasing *A* is relatively small. If not, the large number of uncommitted consumers of *B* makes it easy for independent purchasers of *B* to achieve minimum efficient scale. If the relevant market is local, the number of consumers who purchase *B* without purchasing *A* (in the example above, the number of consumers of internet content and applications outside of Philadelphia) will be very large, which renders it increasingly unlikely that this will exceed the economies of scale.

Consider next the assumption that the market for *B* (internet content and applications) is characterized by large economies of scale. The fact that most content and applications markets have not collapsed into natural monopolies suggests either that the scale economies created by fixed costs are particularly large[225] or that there is some other consideration, such as congestion

---

[224] See Yoo, 52 Emory L J at 1657-58 (cited in note 170) (making a similar point in the context of local-interest vs. national-interest television).

[225] As I noted earlier, the scale economies associated with high fixed costs decays exponentially as volume increases. See note 121. Thus, even when scale economies remain unexhausted in the strict sense, as volume increases they eventually become so small and the average cost curve becomes so flat as to no longer confer any competitive advantage. At most, it would allow the dominant player to engage in a very weak form of limit pricing that was only slightly above competitive levels, in which case the consumer harm would be miniscule.

costs[226] or product differentiation,[227] that serves to counterbalance the scale economies.[228] Indeed, the high fixed costs associated with other content-oriented industries, such as television, have not created scale economies so large as to prevent those markets from functioning properly.[229]

Finally, we come to the prerequisite that the market is a monopoly.[230] This is a very strong assumption in the model, as can be seen if one instead explores what happens if this assumption is relaxed only slightly and the market for *A* is assumed to be a duopoly. If the market is a duopoly with two firms dividing the *A* market into two segments (a_1 and a_2), it is impossible for the producer of a_1 to use tying to exert any leverage over the market for *B*. This is because any independent producer would be able to compete for not only the *B* consumers who did not also purchase *A* (b_i), but also *B* customers who purchased *A* from the firm that did not tie (b_2). In the numerical example given above, if we assume that the duopolists divide the market equally, an independent producers would have available to it not only the 20 consumers who purchase only *B*, but also the 30 consumers who purchase both *A* and *B*, but did not purchase *A*

---

226 See Yoo, 155 U Pa L Rev at 678-80 (cited in note 103) (showing how congestion costs can prevent markets from collapsing into natural monopolies despite unexhausted economies of scale).

227 See Christopher S. Yoo, *Copyright and Product Differentiation*, 79 NYU L Rev 212, 248-49 (2004) (explaining how product differentiation can prevent markets from collapsing into natural monopolies despite unexhausted economies of scale).

228 There may be particular internet services for which this may not be true. For example, auction sites that depend on bringing together large numbers of buyers and sellers may exhibit a degree of demand-side scale economies. In addition, search engines and other regimes that learn from the participation of other users may also exhibit the same quality. Whether these advantages are sufficiently large to offset any accompanying diseconomies of scale and confer a competitive advantage is an empirical question that cannot be answered a priori. The existence of such providers would not, however, support a blanket network neutrality rule covering all content and applications. It would instead simply support a narrow rule targeted only at those content and applications for which this effect holds true.

229 See Implementation of Section 11(c) of the Cable Television Consumer Protection and Competition Act of 1992, Third Report and Order, 14 FCC Rec 19098, 19115-16 ¶¶ 41-42 (1999) (estimating minimum efficient scale for cable television programming at 15 million homes or 18.56% of the market); Yoo, 52 Emory L J at 1603-07 (cited in note 170) (showing how product differentiation can offset scale economies created by high fixed costs).

230 Whinston also analyzes a scenario in which the tying firm faces competition from an inferior provider. Whinston, 80 Am Econ Rev at 852-54 (cited in note 214). Network neutrality proponents do not appear to rely on this particular scenario.

from the tying firm. Any independent producer of *B* would have an open market of 50 units, a number well in excess of the minimum efficient scale. One could posit that the duopolists divide the market unevenly, but in that case one firm would have to control roughly 92 percent of market *A* (55/60 units) before it could exert any unilateral leverage over market *B*. Alternatively, one could instead posit that the minimum efficient scale might be 50 units. Although this would make leverage feasible, the fact that minimum efficient scale exceeded half of the market would also imply that market *B* was already a natural monopoly, in which case not only would the market not be a duopoly: tying the two products together would create efficiencies by eliminating double marginalization.[231]   Of course, both duopolists could engage in tying, which would exclude independent producers. But if so, the duopolists would still engage in a degree of competition for consumers who do not also purchase *A* (b_i). Thus, whether the duopolists would be able to profit by this depends on the conjectures about duopoly pricing.

The assumption that the market is a monopoly is thus a strong one that drives much of Whinston's result. If the proper geographic market is national, this assumption is clearly not met, as no provider would control more than 22 percent of the market.[232] Moreover, the assumption does not hold even if the relevant geographic market is assumed to be local. Published reports indicate that DSL is available over 80 percent of all households that can receive phone service and that cable modem service is available in over 96 percent of households that can receive cable television service.[233] Given the near ubiquity of both telephone and cable

---

[231] See Spulber and Yoo, 107 Colum L Rev at 1838-39 & n 72 (cited in note 21) (collecting sources showing how integration between successive monopolies, successive monopoly and oligopoly, and successive oligopolies enhances consumer welfare).

[232] See Yoo, 1 Intl J Comm at 514 (cited in note 204).

[233] Fed Communications Comm'n, Industry Analysis and Technology Division, High-Speed Services of Internet Access: Status as of June 30, 2007, at 3 (Mar 2008), available at <http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-280906A1.pdf> (last visited May 9, 2008); Leicthman Research Group, Inc., Research Notes 1Q 2008, at 5, available at <http://www.leichtmanresearch.com/research/notes03_2008.pdf> (last visited May 27, 2008)

service, these numbers suggest that roughly 80 percent of the nation is served by at least two last-mile broadband internet providers.[234] The rapid growth of wireless as a broadband platform promises to increase the competitiveness of this space still further.[235] The Whinston model by itself thus provides only weak support for a network neutrality mandate. Even under the most expansive approach to market definition, the prerequisite that the market be a monopoly limits it to at most 20 percent of the country. The other preconditions limit it still further.

The foregoing discussion highlights the narrowness of the factual preconditions that must be satisfied for Whinston's model to apply. It is for this reason that Stan Liebowitz and Stephen Margolis call Whinston's theory an example of the "Goldilocks theory of tie-in sales."[236] Simply put, everything must be just right for anticompetitive effects that Whinston identified to arise. As such, it is a classic example of exemplifying theory.[237] Even more importantly, Whinston acknowledges that his model does not consider whether tying might give rise to efficiencies. Thus, Whinston himself cautioned that his model would not support the broad per se rules.[238]

This is not to understate Whinston's contribution. His work represents perhaps the most important exemplifying theory used to rebut Chicago School scholars' calls for treating tying as per se legal. At the same time, it would be a mistake to attempt to use Whinston's theory as the basis for an argument in favor of a per se rule cutting in the opposite direction. As noted above,

---

(reporting that broadband internet is available in 89% of homes passed by telephone wires and 99% of homes passed by cable television wires); see also Nat'l Cable and Telecommun Assn, Industry Statistics, available at <http://www.ncta.com/Statistic/Statistic/Statistics.aspx> (reporting that 117.7 million of 123.4 million houses in which cable television is available can receive cable modem service or roughly 95.4%).

[234] See Alfred E. Kahn, *Telecommunications: The Transition from Regulation to Antitrust*, 5 J on Telecommun & High Tech L 159, 160 n 2 (2006) (performing a similar estimate based on earlier data).

[235] Fed Communications Comm'n, High-Speed Services of Internet Access tbl 1 (cited n note 233) (indicating that wireless has skyrocketed from having no subscribers as of the beginning on of 2005 to controlling 35 million subscribers and 35% of the market for high-speed lines as of June 2007).

[236] Stan J. Liebowitz and Stephen E. Margolis, *Bundles of Joy: The Ubiquity and Efficiency of Bundles in New Technology Markets* 10 (Dec 2007), available at <http://ssrn.com/abstract=1069421> (last visited Mar 6, 2008).

[237] See Hylton and Salinger, 69 Antitrust L J at 497 (cited in note 218) (describing Whinston's model as an exemplifying theory).

[238] Whinston, 80 Am Econ Rev at 855-56 (cited in note 214).

a rule declaring tying illegal per se would depend on a showing that the practice nearly always harms consumers. Whinston's analysis is an exemplifying theory that simply identifies one set of circumstances in which tying can harm consumers without making any attempt to determine the frequency with which tying would actually harm consumers.

Consider next the systems integration model by Joseph Farrell and Michael Katz.[239] Farrell and Katz examine markets in which consumers buy two components of a system (such as a computer and a monitor) and in which one producer has a monopoly over one component and competes in the market for the other component only not in terms of price, but also in terms of research and development (R&D), which provides competitive advantages in the market for the other component by increasing the quality of the version offered.

Farrell and Katz show that the producer of the monopoly component will invest more in R&D if it integrates into printers than would a stand-alone producer of printers. The reason is that the monopolist has the incentive to induce producers of the other component to offer consumers as much surplus as possible. This in turn allows the monopolist to extract that surplus in the price it charges for the other component.[240] The effect of integration is to raise the monopolist's investment in R&D.[241] Integration may also lower competitors' incentives to invest in R&D.[242]

This model is subject to a number of limitations. As an initial matter, like the Whinston model, it requires the existence of a monopoly in the primary market. In addition, the welfare implications are subject to more than one interpretation. The welfare losses result from the

---

239 Farrell and Katz, 48 J Indus Econ at 413 (cited in note 214), cited by van Schewick, 5 J on Telecommun & High Tech L at 343 n 45 (cited in note 15).

240 Farrell and Katz, 48 J Indus Econ at 414 (cited in note 214).

241 Id at 419.

242 Id at 420-21.

increase in R&D by the monopolist and the decrease in R&D by competitors. The welfare impact of increases in R&D by the monopolist is difficult to interpret. In general, larger investments in R&D are procompetitive.[243] Indeed, determining whether the investments in R&D undertaken by the monopolist are excessive requires some measure of optimal R&D investment as well as an assessment of whether there are inefficiencies pointing in the other direction stemming from firms' well-recognized inability to internalize the full benefits created by their R&D investments.[244] Nor is it inevitable that systems integration will cause competitors' investment incentives to fall. Although the increase in investments by the monopolist may cause the competitors to lower their own investments, it may well cause them to attempt to compete with the monopolist by increasing their investments instead.[245] The dependence on the underlying factual assumptions and the ambiguity of the welfare implications of the equilibria dictate that the Farrell and Katz model be regarded as another exemplifying theory suggesting that anticompetitive consequences *may* arise without providing any insight into how likely they are to occur.[246]

Other theories advanced in support of network neutrality depend on the assumption that the market for internet access services is dominated by a single, national firm.[247] Network neutrality proponents offer a number of other exemplifying theories, including (most intriguingly) models that do not depend on the assumption that the last-mile internet access provider is a monopolist.[248] Unfortunately, these models are sketched only briefly and are too

---

243 Hylton and Salinger, 69 Antitrust L J at 496 (cited in note 218).

244 Id (citing Jean Tirole, *The Theory of Industrial Organization* 389-401 (1988)).

245 Id.

246 Id at 497.

247 Gerald Faulhaber, *Network Effects and Merger Analysis: Instant Messaging and the AOL-Time Warner Case*, 26 Telecommun Pol 311 (2002), cited in van Schewick, 5 J on Telecommun & High Tech L at 347 & n 58 (cited in note 15).

248 van Schewick, 5 J on Telecommun & High Tech L at 368-77 (cited in note 15).

incompletely specified to permit rigorous analysis of the market features that give rise to these effects.

The existence of multiple exemplifying theories strengthens the case for a broad, ex ante network neutrality mandate, but only slightly. The problem is that there is no concrete evidence that the restrictive assumptions of any particular theory have been satisfied. In addition, as I have attempted to show in my work, there are also a substantial number of equally plausible exemplifying theories pointing in the other direction. Absent some compelling empirical reason to presume that one group is more likely than the other, core principles of competition policy support eschewing broad categorical solutions in either direction and instead subjecting such practices to case-by-case analysis.

The only remaining question is which way to put the burden of proof. The real impact of the burden of proof is on what will happen in ambiguous cases, in which the firm adopting a practice cannot prove whether the practice will be harmful or beneficial. If the burden rests on the party advancing the new practice, ambiguous practices will not be permitted to go forward. If the burden rests on the party opposing the new practice, ambiguous cases will be permitted to go forward.

The need to preserve businesses' ability to experiment with new business models favors placing the burden on those opposing the practice. The need to preserve this room for experimentation is underscored still further by the difficulty in determining which business models will ultimately succeed. Perhaps the most dramatic example of this problem is America Online's acquisition of Time Warner. It is hard to understate the furor that surrounded the merger. At $165 billion, it was by far the biggest merger the world had ever seen at the time. People talked about the merger as if it were the end of history, warning that the combination of content and conduit would eliminate the openness that characterized the internet and force all broadband users into AOL's "walled garden."[249]

Time would ultimately render all of these dire warnings moot. AOL was forced to adapt to the internet, instead of the other way around. After initially attempting to keep its users within its world, AOL soon capitulated and gave its users free access to the entire internet. Thus, AOL went down the same path as CompuServe, Prodigy, and numerous others ISPs who attempted to focus on proprietary content. Rather than enabling the combined entity to force proprietary content down consumers' throats, the merger simply enabled Time Warner shareholders to lose approximately $200 billion in value.[250] The final acknowledgement of failure came in February 2008, when Time Warner announced that it was separating its America Online's internet-access business from its web-portal and advertising businesses in apparent preparation to divesting both.[251]

My point is not to engage in 20/20 hindsight. Rather, my point is that the brightest minds in the world find it difficult to predict which business models will succeed and which will fail. These difficulties suggest that policymakers should hesitate before adopting regulations that rule any particular business model off the table. The better course would be to give providers the latitude to experiment with a wide range of business strategies and to forego intervening until a particular strategy proves to be harmful to consumers.

---

249 See, for example, François Bar et al, *Access and Innovation Policy for the Third-Generation Internet*, 24 Telecomm Poly 489 (2000); Mark A. Lemley and Lawrence Lessig, *The End of End-to-End: Preserving the Architecture of the Internet in the Broadband Era*, 48 UCLA L Rev 925 (2001); Daniel L. Rubinfeld and Hal J. Singer, *Open Access to Broadband Networks: A Case Study of the AOL/Time Warner Merger*, 16 Berkeley Tech L J 631 (2001).

250 See Matthew T. Bodie, *AOL Time Warner and the False God of Shareholder Primacy*, 31 J Corp L 975, 975, 999 (2006) (estimating that the merger cost shareholders over $200 billion in value in just three years).

251 See Merissa Marr, *Time Warner's CEO Spells Out Changes*, Wall St J A3 (Feb 7, 2008) ("Mr. Bewkes, who succeeded Richard Parsons in the top job [at Time Warner] at the start of the year, also confirmed he is separating AOL's shrinking Internet-access operation from its Web-portal and advertising business. Such a move could pave the way for the divestiture of both businesses.").

My point is illustrated by the following hypothetical. Suppose that a community were served by four last-mile broadband providers. Suppose further that one of them would like to deploy a proprietary protocol that allows it to give certain applications priority over others, enter into exclusivity arrangement with particular content providers, or make some other change that renders its network imperfectly interoperable with the others' networks. There are three possible outcomes. First, the network that would like to make the change may have found a better business model, in which case society would be better off if the provider were permitted to try out its new approach. Second, consumers may completely reject the change, in which case the network making the change would lose customers to its competitors until it once again returned to the fold. Third, the change may appeal to some consumers, but not to others. As I have argued at length elsewhere, permitting this kind of diversification can make it possible for smaller players to survive even if they operate with fewer subscribers and higher costs.[252] In this way, smaller players can use niche strategies to survive in much the same way that boutiques survive in the face of competition with low-cost discounters and full-line department stores.[253]

The argument in favor of placing the burden of proof on those opposing a particular practice is based on more than just technological humility. There is a solid empirical foundation supporting the inference that vertical restraints and other forms of exclusivity tend to promote consumer welfare. For example, a recent study conducted by four members of the FTC's staff surveying twenty-two published empirical studies found "a paucity of support for the proposition

---

[252] See Yoo, 19 Harv J L & Tech at 27-33 (cited in note 22) (showing how network diversity can allow competitors to survive despite the presence of unexhausted economies of scale).

[253] See Wu and Yoo, 59 Fed Comm L J at 576 (cited in note 162) ("Employing different protocols might also provide more competition among network platforms by permitting multiple networks to survive by targeting subsegments of the overall market, in much the same way that specialty stores survive in a world dominated by low-cost, mass-market retailers.").

that vertical restraints/vertical integration are likely to harm consumers."[254] Indeed, only one study unambiguously found that vertical integration harmed consumers, and "in this instance, the losses are miniscule ($0.60 per cable subscriber per year)".[255] On the other hand, "a far greater number of studies found that the use of vertical restraints in the particular context studied improved welfare unambiguously."[256] The survey thus concluded that "[m]ost studies find evidence that vertical restraints/vertical integration are pro-competitive."[257] The weight of the evidence thus "suggests that vertical restraints are likely to be benign or welfare enhancing,"[258] which in turn provides empirical support for placing the burden on those opposing the practice.

Another survey published in the *Handbook of Antitrust Economics* similarly reviewed twenty-three published empirical studies of vertical restraints. Despite the relatively small sample size, the authors found the empirical evidence to be "quite striking," "surprisingly consistent," "consistent and convincing," and even "compelling."[259] As a general matter, "privately imposed vertical restraints benefit consumers or at least do not harm them," while government mandates or prohibitions of vertical restraints "systematically reduce consumer welfare or at least do not improve it."[260] Together "[t]he evidence … supports the conclusion that in these markets, manufacturer and consumer interests are apt to be aligned, while interference in the market [by the government] is accomplished at the expense of consumers (and

---

254 See James C. Cooper et al, *Vertical Antitrust Policy as a Problem of Inference*, 23 Intl J Indus Org 639, 648 (2005).

255 Id at 648.

256 Id.

257 Id at 658.

258 Id at 662.

259 Francine LaFontaine and Margaret Slade, *Exclusive Contracts and Vertical Restraints: Empirical Evidence and Public Policy*, in Paolo Buccirossi ed, *Handbook of Antitrust Economics* 392, 408, 409 (MIT 2008).

260 Id at 408.

of course manufacturers)."[261] The authors conclude that "the empirical evidence suggests that in fact a relaxed antitrust attitude towards [vertical] restraints may well be warranted."[262]

It is for this reason that I have long advocated an approach that I call network diversity, which would permit network providers to experiment with various practices until such time as those practices can be shown to harm competition. Enforcement of this approach will occur through a case-by-case, ex post analysis of the impact on competition, rather than through a categorical, ex ante prohibition.[263] It would thus be a mistake to call my approach deregulatory.[264] Instead, I believe it represents an important middle-ground between deregulation (that is, per se legality) and ex ante, categorical prohibition (that is, per se illegality).

Conclusion

I began writing about network neutrality when there was a fairly large consensus in favor of network neutrality and significant momentum in Congress and the FCC toward enacting some form of categorical prohibition against deviating from TCP/IP. My arguments were offered as exemplifying theory to rebut those calls. What is sometimes overlooked was that my arguments did not go to the other extreme of calling for nonregulation. Instead, I proposed a middle course that guarded against the possibility of consumer harm while leaving ample room for experimentation.

The plausibility of the benefits from deviating from the status quo convinces me that this is the appropriate course to follow at this time. The internet is becoming more complex. The

261 Id at 409.
262 Id.
263 See Yoo, 19 Harv J L & Tech at 75-76 (cited in note 22) (arguing in favor of an ex post, case-by-case approach that places the burden of proof on the party opposing the practice over ex ante, categorical prohibitions).
264 See note 40 and accompanying text.

number of users is expanding, and the variety of ways that they are using the internet is expanding even faster. As the user base grows and the industry matures, it is inevitable that networks will experiment with different pricing regimes as well as services that operate on principles other than TCP/IP. The FCC's experience in overseeing access regimes also serves as a cautionary tale about the likelihood of any attempt to regulate the terms and conditions of access.[265] As a result, an ex post, case-by-case approach remains the best way to promote investment in and competition among last-mile networks while at the same time providing meaningful protection against any demonstrated anticompetitive harm that may arise.

---

265 Yoo, 19 Harv J L & Tech at 40-43 (cited in note 22) (detailing the difficulties the FCC has faced in implementing other access regimes).