Public Knowledge

September 21, 2009

Marlene H. Dortch
Secretary
Federal Communications Commission
445 12th St. SW
Washington, DC 20554

RE: Notice of *Ex Parte* presentation in:

GN Docket No. 09-51
WC Docket No. 05-25
WC Docket No. RM-10593
WC Docket No. 07-52

Dear Ms. Dortch:

On behalf of Public Knowledge, this letter is to provide information relating to discussions between Public Knowledge and members of the Commission's staff on September 18, 2009.

Present at the meeting were: Gigi Sohn, President of Public Knowledge; Harold Feld, Legal Director of Public Knowledge; Michael Weinberg, Law Clerk, Public Knowledge; Sharon Gillett, Bureau Chief, Wireline Competition Bureau, FCC; Donald Stockdale, Deputy Bureau Chief, Wireline Competition Bureau, FCC; Randy Clarke, Legal Advisor, Wireline Competition Bureau, FCC; Alex Minard, Acting Legal Advisor, Wireline Competition Bureau, FCC; and Jennifer Prime, Acting Legal Advisor, Wireline Competition Bureau, FCC.

Public Knowledge repeated points made in the testimony of Gigi Sohn at the FCC Workshop on the role of content in the broadband ecosystem held on September 17, 2009 (see attachment). Public Knowledge urged the Commission to look towards the ATT/Bell South merger conditions to form a basis for a fifth nondiscrimination broadband principle.

Public Knowledge also discussed the importance of, if consumer broadband tiering is allowed, giving consumers exclusive control over the setting of those tiers. Such a decision would allow application and service competition at the consumer level, instead of limiting it to competition at the provider level. Whatever decision the Commission ultimately makes on tiering, it should insist that all providers disclose all relevant network practices to consumers.

Additionally, Public Knowledge urged expeditious treatment of the Special Access reform issue. Public Knowledge stressed that the Commission had enough information to begin the process. However, if the Commission feels that more information would be helpful, Public Knowledge suggested that an active collection process would be most beneficial to the


1875 Connecticut Ave, NW
Suite 650
Washington, DC 20009
t 202.518.0020
f 202.986.2539
publicknowledge.org

decision-making process. The most helpful reforms that the Commission could consider would be those designed to increase transparency in the Special Access market.

Finally, Public Knowledge reiterated suggestions and concerns raised previously in National Broadband Plan comments. Specifically, Public Knowledge discussed our proposal to reform the Universal Service Fund by requiring all eligible services to be provided through a broadband platform. Public Knowledge also advocated for the reclassification of broadband as a Title II service, and for the creation of a map of all national fiber assets, both public and private. The unbundling of services was also discussed. This discussion applied to the unbundling of consumer services, such as a requirement to offer stand-alone internet service, as well as unbundling to create wholesale access to networks.

In accordance with the FCC's *ex parte* rules, this document is being electronically filed in the above-referenced docket today.

Sincerely,

/s/
Michael Weinberg
Law Clerk
Public Knowledge

CC: Sharon Gillett
Donald Stockdale
Randy Clarke
Alex Minard
Jennifer Prime

Public Knowledge

September 16, 2009

Marlene H. Dortch
Secretary
Federal Communications Commission
445 12th St. SW
Washington, DC 20554

RE: Notice of *Ex Parte* presentation in: GN Docket No. 09-51

Dear Ms. Dortch:

On behalf of Public Knowledge, this letter is to provide information relating to discussions between Public Knowledge and a member of the Commission's staff on September 15, 2009.

Present at the meeting were: Gigi Sohn, President of Public Knowledge; Harold Feld, Legal Director of Public Knowledge; Michael Weinberg, Law Clerk, Public Knowledge; and Blair Levin, coordinator and Executive Director, FCC Omnibus Broadband Initiative.

The parties discussed the importance of Universal Service Fund (USF) reform to an effective National Broadband Plan. Public Knowledge suggested that the Commission adopt a new rule that would allow the USF to be used to expand broadband access. The rule would require that, in order to be eligible for the USF, a provider would have to provide an eligible service through a broadband platform. This rule would be phased in over a number of years, and would allow USF funds to be used to help pay for the necessary infrastructure upgrades.

Additionally, Public Knowledge urged the Commission to take steps to map existing government fiber optic assets, and to request that Congress act to remove any existing legal barriers to such a mapping.

The remainder of the discussion is contained in our earlier filing in the above-referenced docket. In accordance with the FCC's *ex parte* rules, this document is being electronically filed in the above-referenced docket today.

Sincerely,

/s/
Michael Weinberg
Law Clerk


1875 Connecticut Ave, NW
Suite 650
Washington, DC 20009
t 202.518.0020
f 202.986.2539
publicknowledge.org

Public Knowledge

**PUBLIC KNOWLEDGE PRINCIPLES FOR A SUCCESSFUL FCC BROADBAND AGENDA**

Access to broadband Internet services is no longer a luxury – it is vital to full participation in American society. Not only can these services make it easier and cheaper to communicate both next door and around the world, they can bring high quality health care to the underserved, provide an education to those with time and distance limitations, and create new job opportunities.

Although the FCC must fully engage on the hard questions and broad policy concerns such as Network Neutrality and Universal Service, the FCC must also attend to the numerous smaller decisions that form the building blocks of policy. Pending matters such as the pending classification of text messaging; special access; wireless roaming; cable set-top box interoperability; details around the deployment of "white space devices;" and numerous other "technical" proceedings may lack the attraction of larger proceedings, or the deadlines of the national broadband plan. But they are each in their own way essential building blocks of the larger policies and principles set forth below.

• **Open Competition Among Broadband Providers**. Every American should be able to choose among multiple, competing broadband networks, services, applications and content providers, including municipalities. To facilitate competition, the US should reinstate line sharing and access to broadband telephone company facilities and grant access to cable facilities that provide broadband Internet service. The US should lift barriers to deployment of broadband networks by municipalities.
*Examples of relevant proceedings:* roaming, special access reform, broadcast white spaces, spectrum caps, unbundling.

• **Open Network to All Non-Harmful Equipment.** Consumers must have the right to attach to broadband networks any equipment that does not harm the operation of the network. The Federal Communication Commission's landmark 1968 *Carterfone* decision requiring AT&T to permit the use of any non-harmful device on its network resulted in great innovation, including eventually the creation of the PC modem. The FCC should extend the *Carterfone* principle to wireless and broadband networks by granting the February 2007 Petition for Declaratory Ruling filed by Skype. In addition to extending this principle to wireless devices, the FCC must vigorously enforce its existing rules on equipment attachment, such as CableCard and other regulation of set-top boxes in compliance with Section 629. It also requires the FCC to deny efforts to interfere with end user rights, such as the MPAA *Petition* for waiver of the selectable output control rules.
*Examples of relevant proceedings:* Skype *Petition*, hand-set exclusivity, Set-Top Box Interoperability, Selectable Output Control *Petition*.

• **Open Network for All Applications and Content.** Consumers and applications providers must have the opportunity to use wireline and wireless networks without restriction or degradation in quality, except when authorized by a court for law enforcement purposes or where necessary to protect against technical interference and guarantee signal quality. Any necessary "network management" by service providers should not discriminate against any content, service or application based on its source, ownership or destination. Consumers have a right to access information and ideas from a diversity of sources, and have the right to disseminate their own

ideas to the public in any manner they desire. Finally, every broadband network should be able to interconnect with every other broadband network. The FCC should therefore 1) enact rules embodying its four broadband Internet principles as well as a fifth non-discrimination principle; 2) grant the December 2007 Public Knowledge, *et al.* petition seeking to have text messaging services declared Title II common carrier services; and 3) require interconnection between broadband providers.
*Examples of relevant proceedings:* Text Messaging *Petition*, Skype *Petition*/Wireless *Carterfone*. Network Neutrality, Selectable Output Control *Petition*, possible interference with "over the top" video providers.

• **Open Spectrum for Commercial and Non-Commercial Uses.** Unlicensed services should have the benefit of a presumption that they be authorized in any spectrum band as long as they do not cause interference with existing licensees. Any future allocation of spectrum should be presumed to be for unlicensed use unless parties seeking exclusive licensing can show that the public will benefit more from exclusive access than public access. The U.S. should identify and make available unutilized and underutilized spectrum for unlicensed and licensed uses.
*Examples of relevant proceedings:* Broadcast "White Spaces," NOI on variable power for rural spectrum users, spectrum inventory, new proceedings around cognitive radios.

• **Open to All Users at Affordable Prices.** All Americans, regardless of income, race, geographic location or disability, should have access to affordable broadband connectivity by the end of 2012. The U.S. should spur greater deployment of new broadband services through, among other things, low interest loan programs, tax credits and technology grants. It should either reorient part of the $7 billion Universal Service Fund for broadband deployment or create a new, time limited fund for broadband. Finally, the FCC should grant a pending petition to expand Lifeline and Link-Up programs to include access to broadband.
*Examples of relevant proceedings:* USF Reform, Lifeline Service.

• **Open FCC, A Transparent Market, And Data Driven Policy**. Creation and implementation of a successful broadband policy requires fundamental reforms in how the FCC does business. The processes by which the FCC collects data, makes raw data available to the public, conducts research, and deals with the public must become more open and transparent. The FCC should take steps to minimize the influence of traditional players by reforming the *ex parte* process to prevent negotiation of "backroom deals;" invest in the development of its own research to reduce the dependence on information provided by interested parties; and reach out to researchers in a broad variety of disciplines. Finally, the FCC must recognize that providers themselves are often the only reliable source for data, and that the development of an efficient market and an informed national policy depends on access to information. The FCC should reevaluate the deference shown to industry concerns about the burden of information collection and confidentiality with the public interest in information collection and disclosure.
*Examples of relevant proceedings*: Broadband Data Reporting/Section 706 Inquiry, Annual competition reports, various *Trienniel* reviews (*e.g.*, Section 257)

**Before the**
**Federal Communications Commission**
**Washington, D.C. 20554**

| | | |
|---|---|---|
| In the Matter of | **)** | |
| | **)** | |
| A National Broadband Plan for Our Future | **)** | GN Docket No. 09-51 |

# COMMENTS OF PUBLIC KNOWLEDGE, MEDIA ACCESS PROJECT, THE NEW AMERICA FOUNDATION, AND U.S. PIRG


Jeffrey Pearlman
Art Brodsky
Alex Curtis
Mehan Jayasuriya
Sherwin Siy
Robb Topolski
Public Knowledge
1875 Connecticut Ave. NW, Suite 650
Washington, DC 20009
(202) 518-0020

Andrew Schwartzman
Parul Desai
Media Access Project
1625 K Street, NW
Suite 1000
Washington, DC 20006
(202) 232-4300

Michael Calabrese
Benjamin Lennett
New America Foundation
1899 L St., N.W., Suite 400
Washington, DC 20036
(202) 986-2700

Amina Fazlullah
U.S. PIRG
218 D Street SE
Washington, DC 20003-1900
(202) 546-9707


June 8, 2009

## TABLE OF CONTENTS

**COMMENTS OF**
**PUBLIC KNOWLEDGE, MEDIA ACCESS PROJECT,**
**THE NEW AMERICA FOUNDATION, AND U.S. PIRG**

Public Knowledge (PK), the Media Access Project (MAP) the New America Foundation (NAF), and U.S. PIRG (collectively "PK, *et al.*") submit the following comments in the above captioned proceeding.

We approach the National Broadband Plan with a new, fundamental understanding of the role of broadband in our economy and in our society. Plainly put, access to broadband has become an essential utility, as much as water and electricity are essential utilities. Broadband fits into that category because through a broadband connection to the Internet, businesses large and small can reach new markets and make their enterprises more efficient. Students have at their fingertips educational resources not conceivable a few years ago. Some sources of news and information, once confined to the printed page, are to be found online only. For far too long, however, policymakers treated broadband as a service available for the privileged, much like a high-priced model vehicle.

In short, broadband is not a luxury, and the policy of the United States, particularly the National Broadband Plan, should reflect that reality. Every aspect of U.S. telecommunications policy needs not only to be re-examined and revised, but in some cases done away with in favor of policies that are appropriate to the reality of not only the broadband market of today, but to make sure that the broadband market of tomorrow will exist to serve the needs of ordinary citizens and businesses.

To that end, PK, *et al.*, believe that the foundation of the National Broadband Plan must rest on the belief that certain fundamental principles are too important to be left to the marketplace, as the government has done. As a result, broadband policy has strayed significantly

from the Communications Act principles which are based not on the good works of "market forces," but on the rights of Americans to receive telecommunications services without discrimination at reasonable rates.

To correct the failures of our recent broadband policy, we suggest several elements that should be part of a new policy:

- An open Internet should be the foundation of the National Broadband Plan. The FCC should move quickly to adopt a non-discrimination principle, which will allow the Internet to operate as an open system as it has from the start. Activities such as monitoring Internet connections for copyrighted materials must not be allowed, just as opening of mail is not allowed to be part of a widespread fishing expedition on behalf of a private industry.

- User privacy must be protected in areas of content and customer records.

- Consumer rights must be rigorously enforced, with Internet Service Providers required to provide the services they advertise, without hidden charges or unfair practices.

- The Universal Service Fund and Lifeline programs must be restructured to aid in the deployment of broadband networks. Broadband, not voice communications, is the "must have" utility of the 21st century, and a broadband plan should address continuing funding needs for upgrades of networks and demand-side outreach and training.

The regulatory regime should reflect the failures of the marketplace, which have resulted in minimal competition, and higher rates for lower speeds than in many other industrialized nations. Our lack of competition can be traced to the reclassification of high-speed broadband services into Title I. As a result, we propose a range of options for the Commission to consider. Among them:

- The FCC should reclassify broadband services into Title II; impose structural separations on the offering of wholesale and retail broadband services; impose functional separations or divestiture on the companies offering wholesale and retail broadband services.

- The FCC should conduct a thorough review of competition and prices in the special access market, with the goal of making certain that incumbents make bandwidth available at reasonable and non-discriminatory rates. The absence of

competition for this crucial building-block of the broadband economy makes its regulation necessary.

- Spectrum reform is essential, starting with an inventory of what spectrum is being used, and continuing through reallocation of unused spectrum. In addition, the Commission should reimpose spectrum caps, which are necessary to further competition. As part of the spectrum issue, the Commission must make certain consumers have no restrictions on the devices they can use, while examining the anti-competitive aspects of handsets tied to one or two carriers.

Outside of the regulatory realm, the National Broadband Plan must examine the nature of demand for services, and the nature of traffic on the Internet. A complete broadband map, including "middle mile" facilities is essential. We must also study the nature of broadband traffic, to find out not only where the service is available, but how it is used. This data will ensure that we invest in networks that can withstand the needs of the future, and that all Americans enjoy the benefits of our national broadband infrastructure. We must also better understand the crucial role that non-commercial providers can play in the provision of services, and make certain there are no barriers to their entry to provide service to their residents.

These are wide-ranging and comprehensive proposals. They are certainly controversial. But at this stage of our economic and technological development, we can suggest no less, as the challenges to our economy become greater and our capacity for innovation shrinks due to the failure of the market to provide a vibrant, competitive and open market for the broadband services essential to our economic survival.

## ARGUMENT

### I. BASIC PRINCIPLES CRITICAL TO THE SUCCESS OF OUR NATIONAL BROADBAND PLAN ARE TOO IMPORTANT TO LEAVE TO THE MARKET

The mandate to create a National Broadband Plan represents a long-overdue acknowledgement by Congress that broadband has become a critical component of our national infrastructure. As Congress invests billions of dollars in broadband technology and increasingly integrates this new technology into our economy and our everyday lives, Congress and Federal agencies must begin with the fundamental principles and values that have formed the bedrock of our communications policy for more than 75 years, and allowed the Internet to become an invaluable tool for promoting education, economic growth, free speech and civic engagement.

We begin therefore by calling on the National Broadband Plan to recognize that certain fundamental principles are too important to leave to the marketplace. For far too long, the debate around broadband has centered exclusively on the use of market forces and commercial providers. Rather than begin with first principles such as the fundamental freedom of users to speak and innovate without permission from either the government or network operators, policy has looked to competition alone to protect these freedoms or protect traditional consumer rights such as privacy. Debates therefore centered on whether or not competition existed, and if not how to coax it into being.

But a vibrant broadband network, placed at the center of our economy and our ability to communicate with one another, must follow in the footsteps of the Communications Act of 1934. This begins not with a discussion of competition, but a broad statement of the rights of all Americans to provide "all people of the United States" without discrimination access to "adequate facilities at reasonable charges." The first statutes in Title II gave meaning to these

principles by outlawing unreasonable rates and practices.[1] The National Broadband Plan should likewise begin with a commitment to principles too important to leave to the marketplace, followed by rules that give these principles meaning.

Chief among these is the principle of open networks. Since its inception, the Internet has operated as an open system, allowing end users to connect the devices and applications of their choice to the network and to access the content of their choosing[2]. It is this very 'openness' that has allowed the Internet to grow and flourish, to build and enrich communities and to connect citizens with educational resources, their government and each other. At present, however, this openness faces a grave threat from proprietary systems that seek to regulate the flow of information on the Internet in a discriminatory manner. Recognizing that the FCC's four principles have proven insufficient in ensuring that networks operate on a non-discriminatory basis[3], we urge the FCC to adopt and enforce a strict principle of non-discrimination as part of the National Broadband Plan.

In order to ensure the full utilization of broadband resources as well as robust protections for citizens and end-users, the National Broadband Plan should also address the issue of user privacy and other consumer protections. The Plan should review existing statutes and regulations governing electronic privacy and should recommend improvements to ensure that user privacy is respected across all digital communications media. Additionally, the Federal government should draw up requirements regarding the type of notice that service providers are required to provide users with in the event of a change in that provider's terms of service or acceptable use policies.

[1] *See, e.g.,* 47 U.S.C. §§201, 202.
[2] Insofar as none of those choices violated Federal, State or local laws or harmed the network.
[3] For example, the principles do not explicitly address cases where a network provider prioritizes or favors certain content, applications and services over others.

Subscribers should receive the service that network providers advertise, without hidden charges or subject to other unfair practices.

Finally, Congress and the FCC must recognize that access to broadband is not a luxury or privilege. Broadband has become an essential utility, and we must ensure that every American has not merely the opportunity to subscribe to broadband access service, but has ***meaningful*** access that includes both consideration of affordability and training and equipment to use broadband connectivity to its full potential. Accordingly, the National Broadband Plan should propose to reform and overhaul existing programs to provide long-term material support for broadband deployment initiatives.

Both Lifeline and the Universal Service Fund (USF) should be restructured so as to aid in the deployment of broadband services nationwide. These programs might also be leveraged to promote and enable the deployment of tools that will increase the adoption and meaningful impact of broadband, including hardware, software and education.

In order to fully realize universal access as well as the benefits of connectivity, we must recognize that one of the most important roles played by the Federal government in the National Broadband Plan will be to protect those principles that are too important to entrust to the market.

### A. Open Networks – Users Must Have Freedom to Speak, Download and Innovate Without Interference From Gatekeepers

As the Commission has recognized, the "[n]ew, innovative broadband products and applications" brought to us by the Internet are "fundamentally changing the way Americans communicate at work, … how they are educated and entertained, and care for themselves and each other."[4] The success of the Internet as a world-changing communications medium and its

[4] Federal Communications Commission, *A National Broadband Plan for Our Future, Notice of Inquiry* ¶ 4, FCC 09-31, GN Docket No. 09-51 (April 8, 2009) [hereinafter *National Broadband Plan NOI*].

ability to offer us new, unforeseen uses for communications is largely a product of its openness. In the context of access to the Internet, "openness" refers to the equal ability of any person or organization to reach any other person or organization, without interference from the entity providing that access. It is this lack of interference which has allowed countless innovators to create new services and applications without permission from gatekeepers in whose interest it might be to disadvantage, prevent, or control those innovations.

The Commission has recognized the need to protect users and innovators on the Internet by ensuring this openness. This recognition is embodied, in part, by its Internet Policy Statement.[5] And as the Act has recognized, adherence to the Internet Policy Statement is a *minimum* for ensuring open networks as we move forward as a broadband-connected nation.[6]

*1. The Fifth Broadband Principle*

However, as demonstrated by the past activities of Comcast[7] and the recent reported activities of other ISPs, it is clear that service providers still have the ability and the incentive to engage in discrimination against types of traffic they deem to be less important. Thus, we need a fifth "nondiscrimination" principle: "consumers are entitled to communicate any lawful data with the destination of their choice without any degradation or preference the consumer has affirmatively requested."

---

[5] *See National Broadband Plan NOI* ¶ 48; *Appropriate Framework for Broadband Access to the Internet over Wireline Facilities; Review of Regulatory Requirements for Incumbent LEC Broadband Telecommunications Services; Computer III Further Remand Proceedings: Bell Operating Company Provision of Enhanced Services; 1998 Biennial Regulatory Review – Review of Computer III and ONA Safeguards and Requirements; Inquiry Concerning High-Speed Access to the Internet Over Cable and Other Facilities Internet Over Cable Declaratory Ruling; Appropriate Regulatory Treatment for Broadband Access to the Internet Over Cable Facilities, Policy Statement*, 20 F.C.C.R. 14986 (2005) [hereinafter *Internet Policy Statement*].

[6] *See American Recovery and Reinvestment Act of 2009*, Pub. L. No. 111-5, 123 Stat. 115 (2009) [hereinafter *Recovery Act*]; *National Broadband Plan NOI* ¶ 48.

[7] *See In the Matters of Formal Complaint of Free Press and Public Knowledge Against Comcast Corporation for Secretly Degrading Peer-to-Peer Applications*
*Broadband Industry Practices Petition of Free Press et al. for Declaratory Ruling*
*that Degrading an Internet Application Violates the FCC's Internet Policy Statement and Does Not Meet an Exception for "Reasonable Network Management"*, 23 F.C.C.R. 13028 (2008) [hereinafter *Comcast Order*].

At it's heart, nondiscrimination in the Internet context means that no piece of data is preferred over another piece of data based on *anything* other than which user that data it is from or specific preferences users have affirmatively requests (including requests based on user-driven QoS standards). This means that an ISP may not alter how it treats a piece of data based on where on the Internet it is being sent, what type of protocol it uses, what type of data it contains unless the consumer has affirmatively requested it, and in that case, that data's treatment may only be altered with respect to that user's other data.

It is not the role of an ISP to decide which of user's data is more important – only the user can make that determination, and the ability of the user to choose is what makes an open Internet a near-perfect competitive platform. Acting Chairman Copps has repeatedly recognized the need for a fifth nondiscrimination principle.[8] And while such a principle has not yet been adopted by the commission for all U.S. networks, it should be made a part of any plan to expand the reach of real broadband access to all Americans.

*2. Platform-Neutral Openness*

It is also critical that openness principles be applies to all broadband Internet access platforms, regardless of the technology used. The differing technical challenges brought by wireless or other communications technologies should not be allowed to provide an excuse to tilt the otherwise level playing field that the Internet provides or enable some providers of Internet access to leverage their use of a particular technology into an anticompetitive force in markets for other services. Only network management which is necessary to the basic functioning of the network should be allowed.

[8] *See, e.g., Comcast Order, Separate Statement of Commisison Michael J Copps* 3, *available at* http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-08-183A3.pdf; *Copps: FCC Needs Fifth Net Neutrality Principle,* Ars Technica (Apr. 6, 2009), *available at* http://arstechnica.com/tech-policy/news/2009/04/a-pre-history-of-the-fccs-fifth-internet-rule.ars.

*3. Free Speech and Open Networks*

The idea of open networks embodied by "net neutrality" and the five principles is akin to that of free speech embodied by the First Amendment. Without placing strict limits on what ways gatekeepers can interfere with, block, or control communications, the party who controls the medium gains the ability to change the message. The Internet is becoming part and parcel of speech itself. We must ensure that our National Broadband Plan is truly planning for a national, democratic, free communications medium.

*4. "Managed" Networks*

In some cases, service providers offer non-Internet "managed" services to their customers. These services may include video, voice, direct-to-provider health care monitoring, or a host of other services. In most cases, these services will compete with similar Internet-based offerings from third parties. While managed networks may offer value to consumers or business customers, in order to fully realize the benefits of competitive service offerings promised by the Internet, a national broadband plan must ensure that these managed services are not used to competitively disadvantage Internet-based solutions. A National Broadband Plan must not allow such managed networks to provide an end-run around the openness principles which have always and should continue to drive Internet innovation.

**B. Neither Copyright Filtering Nor Any Other "Good Cause" Justifies Networks Acting as Gatekeepers**

As discussed above, for the Internet to be useful and worthwhile to all Americans, access to the Internet should be content-neutral, and application and device agnostic. The success of the Internet is owed to the ease with which information can be copied from one computer to another and to its flexibility as a medium through which innumerable applications can communicate

freely.[9] Not only have these features enhanced our democracy, but they have acted as catalysts for innovation by lowering market barriers for new entrants, which can threaten competitors and force them to adapt to new business models[10].

While some adapt, others prefer to slow progress by changing how the Internet functions. The Commission must resist the suggestions from incumbents who wish to change the rules of the Internet and require Internet Service Providers (ISPs) to police their networks for alleged illicit behavior. While unlawful behavior is not to be condoned, expanding the definition of "reasonable network management" to provide a way for non-governmental third parties to enforce legal regimes would drastically change the nature of the Internet.

Today, Internet connections are as important as yesterday's access to the postal, electric, or telephone service. As our recent experience with the phone company monitoring of domestic calls following the attacks of 9/11 clearly demonstrate, many Americans find the idea of using third parties to monitor our conversations to prevent wrongdoing deeply abhorrent to our values as a free society. If the threat to national security does not – in the minds of many – justify recruiting phone companies to monitor our national phone system indiscriminately, surely we cannot justify granting (or requiring) even greater intrusions on our personal privacy that comes from monitoring our broadband traffic.

Nevertheless, industry participants and their supporters have asked that we submit to non-stop monitoring of what we view, what we say, what we buy, and whatever else we do online for the purpose of maintaining their existing business models. For example, in previous proceedings

[9] Reply Comments of Public Knowledge, *et al.* WC Docket No. 07-52, at 1 (July 16, 2007).

[10] See Letter from Wireless Founders Coalition to Chairman Kevin Martin, WT Docket Nos. 06-150, 96-86, PS Docket No. 06-229, at 3 (June 7, 2007) ("What makes the Internet so friendly from an entrepreneur's perspective is its Openness. One does not have to ask … permission to launch a new product, service, or device. To borrow the Nike slogan, you can 'just do it.'"), available at http://gullfoss2.fcc.gov/prod/ecfs/retrieve.cgi?native_or_pdf=pdf&id_document=6519535073.

on the issue of network neutrality, NBC Universal asked the Commission to require that broadband providers "use readily available means to prevent the use of their broadband networks to transfer pirated content."[11] Although some in Congress have echoed the idea of requiring ISPs to filter and police traffic for what seem good causes – such as "deterring unlawful activity, including child pornography and copyright infringement"[12] – Congress has consistently rejected these attempts. The National Broadband Plan should state unequivocally that we reject the concept that we should outsource Big Brother and subject the traffic of Americans to unwavering scrutiny by ISPs however noble the cause.

Even assuming we considered it reasonable to subject internet traffic to unwarranted search in a manner we would never permit of our physical mail, telephone conversations, and physical premises, imposing such an obligation on ISPs (or allowing them to assume it voluntarily) would engender enormous cost for little gain. Employing techniques to "filter" broadband connections to look for unlawful behavior would require ISPs to examine every bit of information a subscriber puts on the web, in an email, in an instant message, or remixed into a video, in order to find an illicit act, all without probable cause. Technologies that enable filtering, like "deep packet inspection" are flawed for being over-inclusive (and thus unconstitutional for their chilling effect on free speech) and under-inclusive (and thus a waste of government and private resources).[13] If the alleged act is copyright infringement, not only may

[11] Comments of NBC Universal, Inc., WC Docket No. 07-52, at 8 (Jun. 15, 2007).
[12] Senate Amendment 417 to Senate Amendment 98 to H.R.1, proposed by U.S. Senator Diane Feinstein, directs the Assistant Secretary of the National Telecommunications and Information Administration to "allow for reasonable network management practices such as deterring unlawful activity, including child pornography and copyright infringement." http://thomas.loc.gov/cgi-bin/bdquery/z?d111:SP00417:
[13] Reply Comments of Public Knowledge, *et al*. at 3.

unauthorized copying be lawful and critical to free speech,[14] but only a court can determine whether a use is infringing. Additionally, monitoring subscribers' network traffic would be an invasion of privacy, done at the request of a minority industry, and violating the rights of everyone who is online.

C. The Law Must Continue To Protect Privacy and Other Consumer Rights

Nowhere is the protection of privacy and other traditional consumer protections more urgently needed than in the broadband world. As broadband becomes more central in our daily lives, from our most significant financial transactions to our most personal communications, the ability of network operators and service providers to collect user information, or engage in unfair or deceptive practices, grows exponentially. But existing law provides no clarity on the protections broadband subscribers can expect, or even what federal, state or local agency will provide them.

For too long, users have suffered while regulators have waited to see if "the market" will provide adequate protection. But a competitive market does not ensure consumer protection. The used car market is fiercely competitive, but we still need "lemon laws" to protect consumers. The National Broadband Plan must make appropriate recommendations on privacy and consumer protection rather than simply wait to see what "the market" will provide.

*1. Privacy*

Broadband users' privacy rights are a paramount consideration as penetration increases and more aspects of individuals' lives and business are conducted online. Ensuring privacy protection increases trust and use of networks, encourages free expression and enterprise, and most importantly, protects the safety and civil rights of users.

[14] *Eldred v. Ashcroft*, 537 U.S. 186, 197 (2003): ("Copyright … does not impermissibly restrict free speech, for it grants the author an exclusive right only to the specific form of expression … and it allows for "fair use" even of the expression itself.")

One overarching concern regarding communications privacy is that the laws intended to protect user privacy form a complex patchwork across different technologies, differ according to specific applications, and make geographic distinctions that are all decreasingly relevant to online use of networks. A user whose email is stored locally on her hard drive is entitled to Fourth Amendment protections that disappear once that same information is stored on a remote server. Nevertheless, the user is most likely to formulate her expectations of privacy based not upon the working of the underlying technology, but upon the functions they serve and the interface placed upon them. Internet-based services, operating under one set of regulations, might imitate differently-regulated services (such as VoIP imitating traditional telephony).[15] Likewise, a given application or protocol might run over a variety of networks, which could be subject to different regimes. For example, a consumer may receive the same programming via a cable system or an Internet application designed by the same provider, and yet his personal information in each instance could be subject to different subpoena and disclosure rules[16]. Depending upon the particular information being disclosed and the particular parties initiating disclosure, very different laws may apply. These complications apply not only to the privacy and confidentiality of communications, but also to the privacy interest in customer records. While Customer Proprietary Network Information (CPNI) may be protected for the customers of telecommunications carriers,[17] the protections of section 222 are unavailable for customers of other types of service providers.

The Commission's findings that privacy should be protected equally among telecommunications and Internet services are therefore key to protecting and advancing privacy

[15] *See* In the Matters of IP-Related Services and E911 Requirements for IP-Enabled Service Providers, Final Rule 70 Fed. Reg. 37273, 37274 (June 29, 2005) ("The record clearly indicates, however, that consumers expect that VoIP services that are interconnected with the PSTN will function in some ways like a 'regular telephone' service").
[16] *See, e.g., Interscope Records v. Does 1-7*, 494 F.Supp.2d 388 (E.D. Va. 2007).
[17] 47 U.S.C. §222.

rights. In the areas of both content and customer records, the Commission should work actively with Congress to ensure that consumers and citizens may expect their information to be well protected regardless of the technical underpinnings of the services they use. Information services' collection of customer information should be subject to at least as rigorous a standard as CPNI for telecommunications services, and the confidentiality of information received should not be dependent upon whether that information is received as part of cable service or some other electronic transmission of information.

The interactive nature of the Internet provides a tempting mechanism for advertisers to collect targeting information from users. Likewise, ISPs have incentives to discriminate against traffic on their networks that may compete with their own content offerings. The fruits of these incentives have been evident in two recent controversies involving Deep Packet Inspection (DPI): the NebuAd program[18] and Comcast's blocking of BitTorrent.[19] Each of these incidents implicates an unwanted intrusion into a user's information. Despite this, questions of jurisdiction and authority are raised in attempts to excuse this behavior. The Commission's rules and the applicable statutes should make clear that such intrusions should not occur on any platform.

Despite consumers' significant concern about their online behavior being tracked,[20] the opaque nature of information gathering and the ease with which contractual terms are imposed make online privacy a poor fit for market- or self-regulation. Recent studies have found that consumers widely misunderstand the nature of online privacy policies, thinking that uses of their

[18] Robb Topolski, *NebuAd and Partner ISPs: Wiretapping, Forgery, and Browser Hijacking*, June 18, 2008, http://www.publicknowledge.org/pdf/nebuad-report-20080618.pdf.
[19] FCC press release, *Commission Orders Comcast to End Discriminatory Network Management Practices*, August 1, 2008, *available at* http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-284286A1.pdf.
[20] *See, e.g.*, Center for the Digital Future, Univ. of S. Cal., *Surveying the Digital Future: Survey Highlights*, 6, April 28, 2009, *available at* http://www.digitalcenter.org/pdf/2009_Digital_Future_Project_Release_Highlights.pdf; Consumers Union, *Consumer Reports Poll: Americans Extremely Concerned About Internet Privacy*, Sept. 25, 2008, *available at* http://www.consumersunion.org/pub/core_telecom_and_utilities/006189.html.

information is restricted when it is not.[21] The lack of clear information for consumers mean that they are often unaware of practices that might otherwise cause them to migrate to other providers. Compounding the inability for market forces to encourage better privacy protection is the marked lack of competition in many areas for broadband Internet service.

*2. Transparency and Other Consumer Rights*

Getting accurate information to consumers is critical to developing a plan for national broadband deployment, higher speeds, competition, and innovation. This means that those offering network services should be obligated to provide real, useful information to consumers about what their service provides before the consumer makes a service choice. This information should include not only the maximum speed allowed, but more detailed information about likely average speeds, times of congestion, the extent a network is overprovisioned, and any minimum speed guarantees. Most importantly, while more technical details should be available to technology-savvy users, the basic information provided must be accessible to and understandable by the average customer. The average customer should be able to compare advertised speeds and terms of service with what he or she actually receives, and should have clear recourse where the provider does not deliver the promised service.

Broadband providers are increasingly being found using false advertising or for not being transparent in declaring their terms and services in an effort to lure customers.[22] Some providers

[21] Joseph Turow, Deirdre K. Mulligan & Chris Jay Hoofnagle, Univ. of Pa.'s Annenberg Sch. for Commc'n & U.C.-Berkeley Law's Samuelson Law, Tech. and Pub. Policy Clinic, *Research Report: Consumers Fundamentally Misunderstand The Online Advertising Marketplace*, 1, Oct. 2007, *available at* http://groups.ischool.berkeley.edu/samuelsonclinic/files/annenberg_samuelson_advertising.pdf.

[22] *See, e.g.*, BBR, "U.S. Broadband Price Comparisons," [May 22, 2007], http://text.broadbandreports.com/shownews/83886 [accessed: June 2, 2009].

are providing offers at low costs with a hidden catch of long-term contracts. [23] The length of these contracts means that the customer would miss out on better upcoming offers in the market and stay tied to the network provider for a long time. Just a few months ago, Clearwire was charged on behalf of five customers for deceptively advertising and then charging early termination fees when the customers sought to cancel its services.[24] The customers alleged that Clearwire had advertised its service as "fast, reliable and always on alternative to cable and DSL" while it was in fact, "far inferior to cable Internet and DSL, as consumers frequently experienced service disruptions, including dial-up speeds and lack of service entirely". [25]The Commission should require that broadband providers adhere to transparency in declaring their service contracts and abide by them in deployment of services. Such an action will warrant fair competition among broadband providers and in turn secure consumer welfare.

Disclosure should also extend to all forms of network management or monitoring which are used by a service provider. In those rare cases where a service provider may lawfully limit speeds, observe user activity, or in any way alter user communications, then these must be clearly and publicly disclosed. Further, unless such activity is necessary to the basic functioning of the network, the provider should be required to secure an affirmative, informed, "opt-in" consent from users.[26] Critically, federal law must ensure that consent is real. Service providers cannot be allowed to use consent – especially in the absence of a real competitive alternative --

[23] *See, e.g.*, New York Computer Help Blog, " Verizon is Jumping on The Bandwagon…With a Hidden Cost," [May 14, 2009], http://www.newyorkcomputerhelp.com/blog/2009/05/14/verizon-is-jumping-in-on-the-netbook-bandwagonwith-a-hidden-cost/ [accessed: June 2, 2009].
[24] Multichannel News, "Clearwire Sued Over Early-Termination Fees," Todd Spangler, [April 24, 2009], http://www.multichannel.com/article/209951-Clearwire_Sued_Over_Early_Termination_Fees.php [accessed: June 1, 2009]
[25] Id.
[26] *See, e.g., Testimony of Gigi B. Sohn Before the U.S. Senate Committee on Commerce, Science & Transportation, Hearing On Broadband Providers and Consumer Privacy* (Sept. 25, 2008), *available at* http://commerce.senate.gov/public/_files/SohnTestimony.pdf.

as a method to evade the internet principles, nondiscrimination rules, and consumer protections discussed above.

Finally, we reiterate that disclosure and consent are not sufficient safeguards to ensure competition and innovation on the Internet. Rather, ***in addition*** to affirmative rules on openness and consumer protection, federal law must enhance transparency as a means of providing accountability. The boundaries of how service providers may treat user data should not be pushed back by any ISP capable of getting a user to click "Agree" in a web form. Whether the activity in question conforms to openness principles, privacy concerns, and the law must be considered independently of any notice given to the user.

**D. Reform of USF Can Give All Americans Meaningful Access To Broadband**

Congress has understood since 1996 that "advanced telecommunications capabilities" would be essential for Americans in the 21st Century[27] and critical to preserving "favoring diversity of media voices, vigorous economic competition, technological advancement, and promotion of the public interest, convenience, and necessity."[28] At the same time, Congress also embraced the idea of providing broadband to schools, libraries and rural healthcare institutions.[29] Nevertheless, recognizing the state of existing technology in 1996, Congress did not disturb legacy programs, such as lifeline/lifelink and the high cost fund, designed to bring voice service to the urban poor and high cost rural communities.[30]

But technology has vastly improved. We can preserve existing subsidized service by providing them over an IP platform while expanding the concept of universal broadband to include meaningful access for all Americans. The time has come to recognize that broadband,

[27] Telecommunications Act of 1996, Pub. L. 104-104, Section 706 (codified at 47 U.S.C. §157 nt).
[28] 47 U.S.C. §257(b).
[29] *See* 47 U.S.C. §254.
[30] *See* 47 U.S.C. §254(j).

not voice, has become the "must have" utility for the 21st Century. A National Broadband Plan should have as its centerpiece a plan to reform USF to address the continuing funding needs pertaining to both build out/upgrades and demand side outreach/training. It should also provide funding for customer premise equipment (CPE) such as computers to make the dream of meaningful universal access a reality.[31]

The annual budget collected by USF nearly matches the entire broadband stimulus package for NTIA and RUS approved in the ARRA. It is a credit to the power of incumbency and bureaucratic inertia that so much money does so little good. Every study and proceeding declares that USF wastes billions of dollars through artificial stovepipes breaking services into irrational categories, each with its own arcane rules and set of incumbents resistant to change. Simply rationalizing this fund to eliminate the bureaucratic waste and outright fraud that has bedeviled the program can provide the equivalent of an annual broadband stimulus package without drawing on new sources of revenue.

This is not to say that funding should not also be rationalized. To the contrary, a shift to a broadband-oriented fund should entail rationalization of the funding mechanism to distribute the burden more effectively. But none of this will happen while the Commission and Congress entertain endless debate without making the hard decisions. We learn from the recent DTV transition, which began in 1996 and is now scheduled to end on June 12, 2009. What happens when Congress leaves a complicated conversion to industry stakeholders without setting firm deadlines.

Until 2005, progress on transition was glacial. Conversion to digital relied on a slow shifting of stations and consumers to digital television, triggering a shut off when 85% of the

[31] We address the need for a policy to address "demand side" generally below in Part III.

country converted to digital television. This created a market paralysis where broadcasters had no market incentive to convert to DTV or invest in developing expensive DTV programming that no one could receive with analog television sets. This, in turn, stalled development of the production of digital tuners, since manufacturers had no reason to build them and retailers had no reason to stock them. And, if they would have existed, consumers had no reason to buy them. Congress finally broke this logjam by mandating a conversion to digital and an end to full-power analog broadcasting. Even then, the initial failure of Congress to fund outreach and sufficiently fund converter boxes nearly created a disaster for the transition.

Applying this lesson to broadband, we should seek to phase out the stovepipes around USF and convert the fund to one that subsidizes broadband (wireline and wireless), services deliver by broadband, and equipment and training so that consumers can take full advantage of broadband access. Pure broadband connectivity is capable of providing all the services currently supported by USF, in addition to the expanded capabilities of an open broadband connection. The ultimate goal of USF reform should be to ensure that all existing services are supported by the mechanism of USF paying into a single, coherent fund that subsidizes a broadband connection, CPE, outreach, and education on the enhanced capabilities of broadband. This can be accomplished by requiring all recipients of funds for any purpose to provide the subsidized service ***via an open broadband connection***, with the broadband capacity also made available to the recipient of the service. In other words, plain old telephone (POTS) providers in high-cost areas will only receive funding if they provide POTS service via VOIP, and make the broadband connection available as part of the subsidized service.

## II. THE NATIONAL BROADBAND PLAN SHOULD TAKE STEPS TO ENSURE A VIBRANT AND COMPETITIVE MARKET EXISTS FOR BROADBAND ACCESS AND OTHER SERVICES DELIVERED BY BROADBAND.

Although we begin with principles too important to trust to the vagaries of the marketplace, markets will continue to play a critical role in broadband policy. As a first step, however, we must distinguish "market forces" and "competition" from an unregulated marketplace. Since the passage of the Telecommunications Act of 1996, both the Commission and the federal courts have at time confused a preference for the use of market forces over classic "command and control" regulation with a religious conviction that competition occurs only in the absence of regulation.[32] For example, former FCC Chairman Michael Powell opined at his first press conference that:

> I don't believe deregulation is like the dessert that you serve after people have fed on their vegetables, like a reward for competition. I believe deregulation is instead a critical ingredient to facilitating competition, not something to be handed out after there is a substantial number of players and competitors in the market.[33]

Powell's insistence on deregulating as a prerequisite to competition, rather than considering whether specific regulations actively promoted the development of competitive markets, created a wave of deregulation that created our current levels of market concentration in both wireline and wireless broadband, the lack of any meaningful consumer protection measures, and the imposition of practices such as early termination fees (ETFs) that make it harder for consumers to "vote with their feet" by switching provider. The number of wireline ISPs providing service rapidly dropped from over 6,000 to a comparative handful, and supposed "third pipe" competition from broadband over powerlines and wireless providers never materialized. Rather than a prerequisite to competition, mindless deregulation proved an unmitigated market disaster – a result utterly predictable (and predicted) by application of sound economic principles rather than ideology.

---

[32] *See, e.g.*, Fox Television Stations, Inc. v. FCC, 280 F.3d 1027, 1044 (D.C. Cir. 2002) (equating 1996 Act preference for market mechanisms with a mandate to deregulate).

[33] *See* Rodney L. Pringle, "Powell Wants A Less Intrusive FCC," Communications Today (February 7, 2001), *available at* http://findarticles.com/p/articles/mi_m0BMD/is_25_7/ai_71060655/ (last viewed June 8, 2009).

Commentors here propose regulatory structures that promote competition. These proposals recognize the economic realities of communications markets will require significant regulatory structure to create opportunities for market forces. Communications markets are large, complex and interrelated, with significant opportunity for market participants to create lock in through increased switching cost, information asymmetry, and leveraging market dominance in related markets. The difficulty for regulators in monitoring market behavior creates further challenges, as does the need to respond quickly to changes in technology. Finally, the need for massive investment in networks, coupled with the low marginal cost of providing service, and the government restriction on certain inputs, such as spectrum, create further opportunities for market actors to thwart competition in a deregulated environment.

Rather than conclude that these circumstances make regulation impossible, a study of what has worked historically and works currently overseas should inform the economic portion of the National Broadband Plan. As an initial matter, some form of structural separation between consumer services and bandwidth wholesale should be considered for companies that provide services to both consumers and competitors. Alternately, service providers could be required to completely divest themselves of any holdings that are seen to conflict with those services offered by the core business. It is only through some form of structural separation that the FCC and Congress can truly ensure that wholesale access to bandwidth is offered at reasonable, nondiscriminatory rates. Subsequently, reasonable pricing at the wholesale level will encourage both competition and affordability at the consumer level. Alternatively, the Commission should consider reclassifying broadband access as a Title II service. Finally, at a minimum, the Commission should place a high priority on special access reform to facilitate competition even under the existing structure.

In order to aid broadband deployment, especially in rural areas where extensive physical infrastructure may not prove cost-effective, the FCC and Congress should work to ensure that spectrum is being used as fully and efficiently as is possible. To this end, both the FCC and Congress should seek to reform those statues governing the use of spectrum and should seek to make more spectrum available on an unlicensed basis, so as to encourage innovation as well as competition among service providers.

Finally, costs that discourage users from switching providers – either direct costs such as early termination fees (ETFs) or indirect costs of switching such as the inability to move equipment or applications from one network to another – should, to the greatest extent possible, be addressed and eliminated. Unless consumers are allowed to freely switch between providers in fact as well as in name, we will never have a truly competitive market for broadband access and related services.

**A. Structural Separation**

While the National Broadband Plan will have a great many components, the heart of any competitive plan must include a regulatory structure that provides for the most benefits for consumers. Without a regulatory structure that sets the ground rules for true competition, consumers will be denied the benefits of lower prices, increased choices and more innovative services.

The course the Commission has followed over the past eight years has turned out to be spectacularly wrong in all of those aspects. There is little to no competition for broadband services in the residential and "middle mile" markets. As a result, U.S. consumers pay higher rates for services with slower speeds than do consumers in other industrialized nations. Our record of online innovation has slowed to a crawl. The U.S.'s standing in the world ranking of

broadband adoption falls continually. (One can look at various rankings and dispute any given position, but the trend in all of them is clear. America is clearly falling behind.)

The reason the U.S. is falling behind can be traced directly to the decisions the Commission made over the past 10 years to reclassify broadband service, taking it out of the environment of Title II while moving it into the more legally murky area of Title I by classifying broadband as an "information service" instead of as a "telecommunications service." Now is the time to recognize that this deliberate decision to deregulate by redefinition failed to produce the promised land of "intermodal competition" and reverse that decision.

Current Acting Chairman (then-Commissioner) Copps set out the benchmarks for such a renewed look at the situation in his concurring statement on the DSL reclassification order released Sept. 23, 2005:

> Let me sum up by reminding the Commission that we are saying today that we take the dramatic step of reclassifying DSL in order to spur broadband deployment and to help consumers. I want us to test that proposition a year from now. If by next year consumers have more broadband options, lower prices, higher speeds and better services, maybe this proposition holds true. If our broadband take-rate reverses course and the United States begins to climb up the ladder of broadband penetration rather than falling further behind so many other nations, then we'll have something to crow about. If we get no complaints about higher bills, loss of privacy and diminished access for the disability communities, we can take a bow. And critically, if we make progress on public safety and homeland security, we can be proud of our actions. So I hope next year the Commission will put its money where its mouth is and check to see if its theory yields real world results for American consumers. And if it doesn't achieve these results, I hope we'll admit it.[34]

By the Copps Index of Broadband Progress, the results since the reclassification have been dismal. Our broadband remains expensive. According to the Organisation for Economic Cooperation and Development (OECD), U.S. consumers pay an average $10.02 per mbps.

[34] *Appropriate Framework for Broadband Access to the internet Over Wireline Facilities*, 20 FCC Rcd 14853 (2005) (statement of Commissioner Copps).

Fourteen nations have cheaper rates, starting with Korea at $0.85/mbps, France at $3.30 and the U.K. At $4.08.

Not surprisingly, most of the nations which rank above us have a different, if familiar, regulatory structure. It is the structure abandoned by the Commission in the 2005 *Wireline Classification Order*. The structure rejected by the FCC and adopted by many European countries allows for line sharing and wholesale access for competitors with strong enforcement of anti-discrimination rules. The time has come for the rest of the Commission to respond to Acting Chairman Copps' call for an honest assessment of this result.

But the European Union, at the urging of European Commissioner Viviane Reding, is pursuing an even more aggressive course – to require network operators to separate their operations by function, the wholesale from the retail. As she said in a speech last year:

> I see it as entirely appropriate for regulators to allow infrastructure providers to make a reliable return on next generation access investments in return for testable guarantees of non-discrimination and an agreed plan for infrastructural investment that will lead to an open, high speed infrastructure. By the way, one of the potential attractions of functionally separating access networks is to make this incentive structure clearer and more operational.

As the U.S. moves forward to reclaim the broadband momentum that has eluded it for the past four years, it faces a range of policy choices.

1) Do nothing. Maintaining the current regulatory environment will only strengthen the role of the incumbents, stifle competition and lead to a further trend in the decline of U.S. leadership and economic development. It is the least desirable option.

2) Reclassify broadband services. Putting broadband services offered by carriers back under Title II carries a wide range of benefits. For one thing, the Commission can stop trying to shoehorn those services into some Title II responsibilities that the Commission has determined should remain regulated.

More importantly, under Title II, more competitors will have access to networks they wouldn't have otherwise, leading to more competition, lower prices and more innovation – but only if the regime is strictly enforced. As we saw previously in the long-distance, competitive context, incumbents will constantly try to manipulate any lever to force the failure of the "level playing field," including serial challenges to any pricing mechanism. A negotiated rate level, with public disclosure and most-favored nation clauses might help alleviate some of the angst.

The Supreme Curt's recent decision in *FCC v. Fox Television Stations, Inc.*, No. 07-582 (U.S. April 28, 2009) has made it clear that the Commission can revisit its decision and reclassify broadband access as a Title II service in light of market developments since 2005. In *Fox*, the Court explained that a policy reversal required no greater standard of review than an initial decision. "We find no basis in the Administrative Procedure Act or in our opinions for a requirement that all agency change be subjected to more searching review. The Act mentions no such heightened standard."

3) Impose structural separation. In Computer II and Computer III, the Commission attempted to separate the provision of basic and other services. The structure, particularly in Computer III, was a tentative attempt at deregulation and set the stage for the current fiasco. Structural separation may well be a compromise that deregulation advocates may endorse, if only because such a structure is easily subverted. A system based on structural separation, therefore, must carefully examine how to prevent discrimination by vertically integrated network operators. Furthermore, given the history of FCC lack of enforcement during the period when the FCC determined to deregulate broadband in 2001 and the final elimination of access regulation in 2005, it would be imperative for a structural separation policy to

provide a means whereby new entrants can secure speedy relief without reliance on the political will of the FCC to enforce the law.

4) Impose functional separation/divestiture. In order to create the right incentives for a new broadband explosion, functional separation should be seriously considered. Every economic entity exists to further its own enlightened self-interest. If the self-interest of a company is to sell wholesale access to broadband services, then it will try to sell as much of that product as possible, and will invest what is necessary to grow the business. If its self-interest depends on restricting access to competitors in order to further a retail industry, then it will do so, whether by lack of investment, throttling, bandwidth caps or other means. European companies are pursuing this option at the direction of their governments and the European Union.

5) Finally, although not a regulatory solution, local, state or federal government construction of infrastructure can provide a means of enhancing competition and consumer choice. Here also, however, Commentors acknowledge potential risks. Incumbents have proven successful in the past in persuading the governments to subsequently divest physical facilities even after successful build out. For example, in the early 1990s, the National Science Foundation (NSF) built out the first national backbone and facilitated the development of a robust "middle mile" market by providing free traffic exchange (peering) and national network access points (NAPs). In the mid-1990s, under pressure to avoid "competing with the private sector" and convinced of the inherent superirority of private sector management, NSF sold the national backbone – including the NAPs – to the RBOCs for the regions in which the NAPS were located. The RBOCs preferred to build out their private infrastructure, not subject to peering requirements. In a relatively brief time,

industry-wide peering essentially disappeared, altering and consolidating the middle mile market.

Recognizing the diversity of the broadband industry, we should expect that a combination of approaches will prove necessary to strike the right balance of ensuring competitive service while maintaining private investment and innovation. Our current experience tells us whatever the flaws of these regulatory alternatives, the current deregulatory environment has utterly failed to produce the residential service or middle mile infrastructure we must have to meet our national broadband needs. The time has come to engage in the assessment called for by then-Commissioner Copps when we adopted deregulation in 2005, and consider the value of other approaches.

In sum, our regulatory structure has failed. It is time for a new one that will serve a greater variety of interests than simply those of the carriers.

**B. Competition Requires Special Access Reform**

As part of the Commission's analysis of effective mechanisms to achieve the goals of the Recovery Act, there is a need to review existing regulatory structures for "special access" services. Special access is a high capacity transmission path that connects the Internet backbone to local facilities (*i.e.*, cell phone towers, local area networks, etc.). In most parts of the nation, these paths are provided only by incumbent local exchange carriers. The local facilities, so connected, include carriers that do not own physical broadband infrastructure and therefore rely on special access to support their telecommunication services. Large businesses that have higher and advanced bandwidth requirements also need direct access to the Internet and therefore also make use of special access.

Absence of competition amongst the providers of special access makes its regulation necessary. There is building consensus that for incumbent local exchange carriers, providing

special access to wholesale customers is a primary business course which they use to leverage increasing special access revenues via inflated special access prices.[35] According to a study, in 2007 Verizon extracted a rate of return of 700 percent for its special access transmission path.[36] The Commission should ensure that incumbent providers charge fair access rates so that carriers competing with them in the "last mile" can offer consumers a choice of broadband services. The theory of special access price regulation is simple: (1) small carriers and consumers are protected from unreasonable rates and (2) carriers are encouraged to operate with increased efficiency by utilizing the cost savings obtained as a result of lowered special access prices. For wireless service providers, this would mean lower costs for backhaul to their cell towers and enable them to concentrate on increasing "cell sites" in underserved areas.

For determining whether regulation on special access pricing is required, the Commission currently assesses competition by measuring the degree of "collocation" in Metropolitan Statistical Areas. Collocation refers to the number of competitive carriers who set up equipment in a facility provided by the incumbent carriers. The Commission believes that the degree of collocation of equipment is an indication that competitive carriers have made investments in the special access facility and that this will thwart any market power that the incumbents possess.[37]

In 2006, the Government and Accountability Office reviewed this methodology and found it flawed. According to the report, after the Commission establishes competition in an area (on the basis of collocation) and deregulates prices, it does not re-examine whether the competing carriers survive or are folded within the incumbents. Additionally, the report

[35] *See, e.g.*, "Report Reignites Fights over Special Access Rates", Ars Technica, [January 26, 2009], http://arstechnica.com/tech-policy/news/2009/01/report-reignites-fight-over-special-access-rates.ars [accessed: May 29, 2009].
[36] *See* Derek Turner, "Dismantling Digital Deregulation: Towards a National Broadband Study," Freepress, http://www.freepress.net/files/Dismantling_Digital_Deregulation.pdf [accessed: June 1, 2009].
[37] "Report Reignites Fights over Special Access Rates", Ars Technica.

established that after price deregulation, special access rates did increase noticeably. The report recommended the use of market share and pricing data as determinants of competition.[38] A recently released report, by the National Association of Regulatory Utility Commissioners (NARUC)[39], also suggests that the Commission's methodology of measuring competition in these special access markets is inaccurate. It even points out that the incumbent local exchange carriers have strong market power in most areas. Although, the report does not call for a reduction in special access rates, it proposes a cap on these prices.

All in all, the Commission must conduct a thorough review of its policies regarding competition and prices in the special access market and ensure that incumbent providers make bandwidth available at reasonable and non-discriminatory prices. This will guarantee that consumers are benefited in terms of choice, price and quality of broadband services and that efficient investment in broadband infrastructure is encouraged.

### C. Spectrum Access Is A Critical Input For A Vibrant Broadband Market.

Spectrum reform is essential to make more effective use of public spectrum, enable increased broadband access and stimulate competition among service providers. Advances in the engineering of wireless systems are changing the way spectrum should be regulated. Thus, there is a strong need for the Commission to metamorphose the governance of the spectrum from a "property-right" model to a "public-good" model which will help promote national access to broadband service.

---

[38] U.S. Government Accountability Office," Telecommunications: FCC Needs to Improve Its Ability to Monitor and Determine the Extent of Competition in Dedicated Access Services," [November 29, 2006], http://www.gao.gov/products/GAO-07-80 [accessed: June 1, 2009].

[39] Peter Bluhm and Robert Loube, "Competitive Issues in Special Access Markets," National Regulatory Research Institute, [January 21, 2009], http://www.naruc.org/Publications/09%200121%20NARUC%20NRRI_spcl_access_mkts_jan09-02%20_2_.pdf [accessed: June 1, 2009].

The traditional concept of interference in the spectrum is no longer a constraint on the efficient and flexible use of wireless communication. Use of spread-spectrum radio and novel "smart" devices that can coordinate with each other at lower power levels and no interference have transformed our understanding of the spectrum. Wifi (currently using three small slices of open spectrum: 900 MHz, 2.5 GHz and 5.7 GHz)[40] has already demonstrated that this form of spectrum use can open doors to myriad innovative "prosumer" (consumer who acts both as a consumer as well as a producer)[41] innovations.

Spectrum is currently thought of as scarce since it has been treated as permanent property and this scarcity makes it highly "valuable." Re-assignment of spectrum from licensed to unlicensed spectrum reduces this value and faces opposition from those who have held licenses and especially those who have paid billions of dollars for those licenses. Of course, the approach of expanding the use of unlicensed spectrum will require that detailed communication protocols be established to guard against sources of inefficiency. One such source of inefficiency could be introduced by a device that does not have any incentive to conserve the shared spectrum. It could become "greedy" and use greater bandwidth for transmission or transmit for prolonged intervals. This would make the spectrum a rivaled "commons" model and in the extreme lead to the "tragedy of the commons" with many greedy devices degrading the performance of other devices. Thus there is a need to implement technical rules (*i.e.*, modulation, back-off schemes, etc.) to address these failures.[42]

---

[40] Free103point9 Newsroom, "Congress considers inventory of spectrum use in America," Blog, [March 25, 2009], http://blog.free103point9.org/2009/03/congress-considers-inventory-of.html [accessed: May 28, 2009].
[41] Term coined by futurist Alvin Toffler.
[42] Jon Peha, "Emerging Technology and Spectrum Policy Reform," Carnegie Mellon University, [January 2007], http://www.itu.int/osg/spu/stn/spectrum/workshop_proceedings/Background_Papers_Final/Jon%20Peha%20ITU%20spectrum%20workshop.pdf [accessed: May 28, 2009].

There is also speculation that a large part of the spectrum is currently not fully utilized.[43] The radio spectrum is an asset of the nation and its effective utilization for the benefit of the citizens is long overdue. Without a clear understanding of the current use of spectrum, it is difficult to have an informed decision on how to allocate spectrum more efficiently. As a first step, the FCC should get a full account of the country's spectrum assets. To this effect a bill, the Radio Spectrum Inventory Act, has already been introduced before the Congress earlier this year.[44]

As a next step, the Commission should reallocate these unused spectrum resources as shared and unlicensed spectrum capacity; in particular from private markets that have simply "borrowed" spectrum frequencies. Auctioning off the fallow spectrum for exclusive use by a corporation or entity might deliver immediate monetary gains but will strangle long-term economic benefits of an open spectrum.

It should also be noted that although unlicensed spectrum bands will enable greater spectrum utilization, licensed spectrum bands are more appropriate than unlicensed ones for applications that require a guaranteed and higher quality of service viz. public safety information transmissions and broadcast television etc.[45] Therefore, maintaining a proper balance between licensed and unlicensed use is necessary. With that said, when the Commission provides licensed spectrum, it should recoup some compensation for the use of this public good. This compensation could be in the form of either an "annual user fee" or provisioning for services over this spectrum that would foster the public interest.

[43] Free103point9 Newsroom, "Congress considers inventory of spectrum use in America."
[44] OpenCongress, "Radio Spectrum Inventory Act," [March 19, 2009], http://www.opencongress.org/bill/111-s649/text [accessed: May 29, 2009].
[45] Jon Peha, "Emerging Technology and Spectrum Policy Reform."

The Commission's propitious decision on opening white spaces for unlicensed use is laudable, yet more pliancy is needed as far as the power levels are concerned. The Commission has established that transmission be restricted to low power levels while using white spaces adjacent to the broadcast channels.[46] Rural areas would have more white spaces compared to urban regions due to presence of fewer broadcasting channels there. Thus, they should be able to take better advantage of this mode of wireless broadband access as compared to urban areas. However, the requirement that the white space devices use low-power transmission will be a problem.[47] Signals will dampen and become weaker as distances increase and this means use of more repeaters to keep up the signal strength. Promoting transmission using a higher power level in these underserved areas would make wireless broadband more accessible to the rural community.[48]

Regarding the low-power devices to be used for transmission in the white spaces, there are a few noteworthy issues that require to be addressed by the Commission, to advance fair competition amidst service providers. The low-power white space devices will use geolocation sensing to avoid interference with other broadcasting signals. This device will access a geolocation database via the Internet to detect available free channels in its vicinity. The Commission should ensure that these databases are developed using open formats and protocols that are non-proprietary. The Commission has already asserted that the database will be administered by a third party selected via a "public notice process."[49] In this regard, the

[46] *See* ET Docket No. 04-186 and ET Docket No. 02-380.
[47] Gigaom, "We're Gonna Have to Wait a Year for White Spaces," Stacey Higginbotham, [November 5, 2008], http://gigaom.com/2008/11/05/were-gonna-have-to-wait-a-year-for-white-spaces/ [accessed: June 2, 2009].
[48] It is also important to remember that a white space device would still need a backhaul to the Internet and that the speed of the device would depend on the backhaul technology (*i.e.,* telephone modem, cable, fiber). Therefore, to facilitate broadband access, especially in rural areas, the Commission should adopt measures to complement wireless broadband with other broadband platforms like cable and fiber.
[49] *See,*\ ET Docket No. 04-186 and ET Docket No. 02-380.

Commission should also safeguard that the administration of the database takes place in an open and non-exclusive manner to foster healthy competition.

### D. Without A Spectrum Cap, The Mobile Market Will Experience Further Consolidation

Spectrum caps are necessary to facilitate competition in the market for wireless broadband. Spectrum caps were introduced in 1994[50] to stimulate competition and ensure that incumbent wireless providers did not gain substantial "first-mover advantages". In the following years, however, there was Intense lobbying by incumbent carriers for the termination of these caps.[51] In 2003, the Commission finally eliminated the spectrum cap on the belief that no single carriers had significant market power to be considered a threat to competition.[52]

Since the Commission lifted the spectrum cap, the wireless market has undergone a great change from a competitive model to one where just a few wireless providers have most of the market power.[53] These handful providers have acquired control over large chunks of the spectrum either at auctions for licensed spectrum, where these providers have been able to use their purchasing power to ward off new-entrants in the market, or in subsequent mergers between two providers with large market power or in deals between these providers, such as negotiations over roaming provisions. This anti-competitive environment has been created due to the absence of spectrum caps.

For example, the absence of any spectrum cap in the recent 700 MHz spectrum auction led to greater market consolidation, with the wireless market getting divided into two parts –

[50] The New York Times, "F.C.C Is Expected to Lift Airwave Spectrum Cap," Simon Romero, [November 8, 2001], http://www.nytimes.com/2001/11/08/business/fcc-is-expected-to-lift-airwave-spectrum-cap.html [accessed: June 3, 2009].
[51] Id.
[52] FCC News, "FCC Anounces Wireless Spectrum Cap To Sunset Effective January 1, 2003," [November 8, 2001], http://www.fcc.gov/Bureaus/Wireless/News_Releases/2001/nrwl0129.html [accessed: June 2, 2009].
[53] Wired, "In Spectrum Auction, Winners Are AT&T, Verizon and Openness," Bryan Gardiner, [March 20, 2008], http://www.wired.com/epicenter/2008/03/fcc-releases-70/ [accessed: May 29, 2009].

AT&T and Verizon and the rest of the providers.[54] Another example of growing anti-competition due to absence of a spectrum cap is the merger between Sprint and Clearwire. The Commission voted in favor of the merger without even assessing the spectrum holdings of the two companies. Presence of a spectrum cap would have subjected the spectrum holdings of the two companies to close examination and could have led to divestiture of some of their holdings, if they exceeded the cap.[55]

The Commission should re-adopt a spectrum cap to mitigate the problems associated with the abuse of market power in an anti-competitive wireless market. The Commission should set a spectrum cap at 95 MHz and should also implement a "spectrum screen" (on spectrum above 2.3 GHz) as a guideline and rule for spectrum holdings.[56] Both the cap and the screen will help to promote competition by limiting the acquisition power of incumbents and by allowing smaller providers access to additional spectrum in the higher spectrum bands.

**E. Roaming, Wireless Carterfone & Handset Exclusivity**

Spectrum holdings become an important asset, especially in the last mile, which sets the pace of competition in the wireless broadband market. To sell services in a particular region, the network service provider has to have rights to the spectrum in that region or rely on roaming deals with the holder of the spectrum for that region. The absence of substantial competition in the current market has led to an abuse of market power with regard to these roaming agreements for small providers. Unable to bargain on these roaming contracts with incumbent providers, the competitors have to either stop operating or merge with the incumbent providers. The Commission should enforce reasonable and non-discriminatory terms on established operators to

[54] Wired, "In Spectrum Auction, Winners Are AT&T, Verizon and Openness."
[55] Market Watch, " FCC Approves Sprint, Clearwire Merger," [November 5, 2008], http://www.marketwatch.com/story/fcc-approves-sprint-clearwire-merger [accessed: June 3, 2009].
[56] New America Foundation, "Reply Comments of The Public Interest Spectrum Coalition," [December 22, 2008], http://www.newamerica.net/files/PISC_RTG_Reply_Comments_122208.pdf [accessed: June 2, 2009].

provide roaming and infrastructure facilities to new entrants in the market. This will enable more effective use of the spectrum and encourage competition among service providers.

To further foster competition between service providers, the Commission must address the anti-competitive issues that arise due to handset exclusivity between handset phone companies and the wireless service providers. Ideally, and in keeping with the principle of openness discussed in Part I, the Commission would impose the same rule on wireless networks that it imposes on wireline networks – the so-called "Cartefone" rule requiring that network operators create a standard that permits anyone to attach a any device to the network and run any application over the network that does not harm the network.[57] Even absent adoption of a "wireless Cartefone" rule, the Commission should still ban exclusive contracts for hand-held devices as a simple matter of competition policy.

Many popular devices (*e.g.*, smartphones) in the market are making exclusive arrangements with the incumbent national carriers and creating a large barrier to entry into the wireless markets. Apple's iPhone exclusivity to AT&T combined with AT&T's policy of limiting its users to no more than 40% of roaming time off the AT&T network,[58] exemplifies how handset exclusivity is driving users away from small wireless providers. Consequently, handset exclusivity leads to greater market consolidation and such arrangements should be annulled. Termination of handset exclusivity contracts by the Commission will advance economical use of the spectrum and increase competition in the wireless market, ultimately benefiting the consumers.

---

[57] *See* Tim wu, "Wireless Net Neutrality: Cellular *Cartefone* and Consumer Choice in Mobile Broadband," New America Working Paper #17 (2007).

[58] Barbara Esbin and Berin Szoka, "Exclusive Handset Prohibitions: Should the FCC Kill the Goose that Laid the Golden iPhone," The Progress and Freedom Foundation, [June 2008], http://www.scribd.com/doc/3253827/Exclusive-Handset-Prohibitions-Should-the-FCC-Kill-the-Goose-that-Laid-the-Golden-iPhone-PFF-EsbinSzoka [accessed: May 29, 2009].

### F. Commission Policy Should Explicitly Identify and Seek To Reduce Switching Costs of All Kind to Enhance Competition

Switching costs aggravate the problem of anti-competition in the broadband market. Early termination fee is one form of switching costs which discourages consumers from switching amongst network providers in search of better quality of service, costs, etc. The current early termination fee in the market is inappropriately high which is why consumers are forced to stay with their service provider despite dissatisfaction with the services or cost. As a result, these costs discourage competition.

This early termination fee has been a familiar feature in wireless phone contracts for several years now. In the cellular market, several lawsuits, introduction of bills[59] and proposals by the Commission[60] have led major mobile carriers to pro-rate their early termination fees.[61] However, the fee is still quite high. Despite the long drawn debate on the early termination fee issue in the cellular market, consumers continue to pay unfair wireless phone penalties. A troubling phenomenon is that a parallel situation is developing in the broadband market. The Commission should ensure that such a tedious and protracted early termination fee dispute, like the one in the cellular market, does not occur in the broadband market.

---

[59] *See* Cell Phone Consumer Empowerment Act of 2007, Bill No. S.2033, [filed September 7, 2007]. The Bill was introduced to allow subscribers to cancel their cell phone carrier's contract before the expiry of the contract, if they were dissatisfied by carrier's services. It also called for pro-rating the early termination fees.

[60] The Commission proposed that the early termination fee should be reasonably related to the cost of the equipment provided to the consumer. However, since phone companies do not reveal such information, estimating the true cost of the handset subsidies will be a challenge. Under the same proposal, the early termination fee would be pro-rated and would not be re-evaluated on a new phone plan unless, the consumer desired a new phone. Additionally, consumers would get up to 30 days after signing the contract (or until 10 days after the first bill is issued), to evaluate the phone they purchased. Returning the phone within this time period would entail no early termination fees. *See* Market Watch, "Wireless-phone penalties unfair to consumers," Jeffry Bartash, [June 14, 2008], http://www.marketwatch.com/story/early-termination-fees-on-wireless-customers-fail-fairness-test-20086140100?pagenumber=1 [accessed: June 2, 2009].

61 Cnet News, "Sprint, T-Mobile to Pro-Rate Early Termination Fees," Kent German, [November 9, 2007], http://news.cnet.com/8301-17938_105-9814184-1.html [accessed: June 1, 2009].

In the recent years, broadband Internet providers have started increasingly imposing early termination fee to prevent their consumers from straying away. In the broadband market, many low-priced promotional services are offered along with the early termination fee to entice customers to sign up.[62] While broadband providers may justify these fees on the pretext of discounted equipment, reduced installation charges and monthly subscription fees, they deprive customers of the benefits of competition especially in the underserved markets. Additionally, these different promotional offers and plans are constantly changing.[63] This makes it very difficult to asses the actual worth of the offerings and creates ambiguity on whether a customer actually benefits from purchasing a plan with early termination fee provisions or not. For customers who move from one place to another, payment of early termination fee is a huge cost, especially if it is not pro-rated. The Commission should ensure that broadband providers focus on providing better quality of service to ensure the retention of their customers and not on vexing contract terms that shackle their customers. It must assess, slash and where needed terminate these expensive penalties. Doing this will foster greater competition in the mobile wireless market.

### G. The Lack of Available Information Prevents The Development of an Efficient Market

As discussed in Part IV below, we as a nation find ourselves without the data necessary to make truly informed broadband policy choices. In considering what information to collect and what data to require industry participants to disclose, the Commission must consider how the lack of available information warps the existing wholesale services market and inhibits the development of competition.

[62] hearusnow.org, "The Next Big Thing in Broadband: Early Termination Penalties", Frayam, [April, 9, 2007], http://www.consumersunion.org/blogs/hun/2007/04/the_next_big_thing_in_broadban.html [accessed: June 1, 2009].
[63] Id.

It is a truism to the point of cliché to observe that competitive markets require information to operate efficiently. Unless potential buyers can compare prices and services, they cannot make informed choices among vendors and so provide discipline to the market. Unless vendors know the prices paid by buyers, they cannot tell how to price their own services. In the consumer market, the need to advertise to attract customers provides consumers with at least some information. But in commercial markets, this need not be the case. Especially where, as here, market concentration produces a handful of providers capable of imposing non-disclosure agreements.

Although all parties suffer from the information asymmetries produced, the burden falls hardest on small businesses. Larger businesses with greater resources, and engaged in a greater number of transactions, can accumulate their own knowledge as to how the market operates. But smaller businesses are unlikely to have the resources to invest in information gathering such as soliciting multiple offers. Nor are the same number of vendors or purchasers likely to approach smaller market participants.

To the extent the National Broadband Plan relies on market forces, it must carefully consider what disclosures to require to ensure a functioning market. At one extreme, the Commission may require providers to tariff certain services, even if the Commission does not regulate the price of the tariff. At the other end, the Commission may continue to do nothing and hope that market forces drive participants to disclose sufficient information. In between these two extremes lies a range of possibilities such as conducting regional price surveys and prohibiting nondisclosure agreements. In all cases, the National Broadband Plan should carefully consider the impact on competition of allowing market participants, particularly in markets dominated by a few large actors, to withhold or suppress information.

## III. THE NATIONAL BROADBAND PLAN MUST ADDRESS THE "DEMAND SIDE" AS WELL AS THE "PRODUCTION SIDE" OF BROADBAND ADOPTION

As part of a National Broadband Plan, the Commission should consider also addressing in a limited fashion the demand side of the broadband equation. As a starting point, the Commission should not consider itself as a marketing arm for the service providers. At the same time, the Commission has an obligation to monitor not only the deployment of broadband facilities, but to track barriers to deployment and adoption and to help to make the benefits of broadband widely known.

The research about why potential customers don't subscribe to broadband is very thin. One study, by the Pew Internet and American Life Project, found that dial-up users don't want give up their connections, and don't believe broadband is a good value.[64] Other, more anecdotal reports from Free Press, e-NC and others, suggest there is a great demand for broadband in rural areas that isn't being satisfied.[65] Indeed, a recent article in the Wall Street Journal chronicled how the homeless regarded broadband access as a lifeline to keep from drifting permanently into poverty.[66]

The Commission's role should be to track and, if necessary, contribute to research on broadband take-up by continually examining pricing data, take rates for the service and build-out progress, and reporting that research to the public. The Commission could also support locally

[64] John Horrigan, "Stimulating Broadband: If Obama Builds It, Will They Log On?" Pew Internet and American Life Project (2009). Available at http://www.pewinternet.org/Reports/2009/Stimulating-Broadband-If-Obama-builds-it-will-they-log-on.aspx (last viewed June 8, 2009).

[65] *See, e.g.*, http://www.internetforeveryone.org/americaoffline/nc (video interview with the Foushee Family); e-NC.org, "Bigger Vision, Bolder Action, Brighter Future: Capturing the Promise of Broadband for North Carolina and America (The Baller Herbst Report)", June 2004, *available at* http://e-nc.org/Baller-Herbst_Report.asp (last visited June 8, 2009).

[66] Phred Dvorak, "On the Street and On Facebook: the Homeless Stay Wired," Wall St. J. Page 1 (May 30, 2009), available at http://online.wsj.com/article/SB124363359881267523.html (last viewed June 18, 2009).

based research, from universities or other institutions, which tracks why, and why not, broadband might be lagging in any particular area.

Implementing other parts of the broadband plan, of course, could help adoption rates to increase. By allowing for more competition, consumers would have more choice in services and features, and at lower prices than offered in today's quasi-duopoly environment.

In raising awareness of the value of broadband, the Commission should endeavor to support local groups already engaged in such activities. Experience has shown that the most effective means of raising awareness about the potential for broadband is through locally based technology evangelists, whether based in state or local governments, non-profits or the private sector.

Having an outside group come into a locality, conduct a few focus group sessions and then leave, is not a workable model for sustained broadband adoption. The Commission could lend its support to existing, local organizations by, among other items, serving as a clearinghouse for good practices, using social networking tools to encourage information-sharing in how to communicate the benefits of broadband, outlining specific benefits that have accrued to each community, by sponsoring national or regional conferences as a means of information exchange.

Similarly, the Commission could help with ancillary activities such as computer training and access to computers. Computer training has been an ongoing activity for decades, although usually targeted toward use of software programs. The Commission could support activities in communities to showcase the use of computers and online services as provided by local organizations.

## IV. THE NATIONAL BROADBAND POLICY MUST CAPTURE ALL NECESSARY DATA TO PROPERLY INFORM THE POLICY

Since Congress first ordered the Commission to ensure deployment of "advanced telecommunications capability" to all Americans, the FCC has focused on the simple question of whether residential subscribers have access to "broadband" – defined initially as 200 kbps and recently raised to 768 kbps. Recently, Congress and the Commission have attempted to augment this concept of a "broadband map" with comparisons with other countries and with improved data collection requirements. The focus, however, remains the same: can residential end users get access to broadband, at what speeds, and – possibly – at what cost.

A data driven broadband plan requires more than information on availability of services to end users. As an initial matter, we cannot fully address any perceived shortages and deficiencies in the provision of broadband unless we know what our broadband network looks like. This includes not merely the "last mile," but the "middle mile" and the nature of the traffic that rides the existing network. We need a greater understand of potential resources available, such as dark fiber and government fiber which we could harness to alleviate shortages in transport.

Furthermore, while evidence to date bears out the policy that universal broadband increases access to economic and educational resources, encourages civic engagement, and allows individuals to develop needed social capital, we have no deep science of "qualitative metrics" that can tell us whether broadband is, in fact, delivering these benefits – and if not why not. The numerous conflicting studies to date are worse than the proverbial apples to oranges comparison. For example, on the pivotal policy question of "does local government investment in broadband work," we cannot even agree on what constitutes success or failure – or even what constitutes government investment. In this absence of any sort of qualitative metrics about what

broadband does and how it impacts people's lives, we cannot even know if our national broadband policy will have succeeded in its goals of improving the lives of all Americans.

### A. Complete The National Broadband Map From Beginning to Middle To End

A National Broadband Plan worthy of the name requires far greater data then this simple map of the final product. Too many questions that should inform our policy of ubiquitous, affordable broadband go unanswered. As an initial matter, we know absolutely nothing about the "middle mile" market in this country. Carriers have repeatedly complained about a need for backhaul for both wireless and broadband services. But we have no way to evaluate whether this is a regional issue related to cost of deployment, an antitrust issue related to vertical integration, a market inefficiency from the lack of market transparency, or a misallocation of resources such as spectrum access. Indeed, we do not know with certainty whether the complaints of carriers are real, whether they represent systemic problems, or are simply attempts to invoke Commission authority to further their own commercial interests.

Before we can begin to address any of these issues, we must remedy our ignorance. We must go beyond previous mapping exercises and begin to cast our net much wider. This must include information the Commission has previously refused to collect on the grounds that carriers preferred to keep this information proprietary. It must also include information from non-commercial actors, such as federal and local government agencies. As discussed in Part V below, these agencies can potentially provide significant resources if freed from the artificial restrictions placed on them.

In short, our national broadband map, which must inform our continued investment in our broadband infrastructure, must be complete. We can no longer tolerate a broadband map ending in a drawing of a cloud with the notation "here be servers." In no other area of our critical

infrastructure are we content to allow such utter ignorance—never mind make policy in the absence of fundamental data.

### B. We Need A Better Understanding of The Nature of Internet Traffic

In addition to a thorough map of the "pipes," we need a better understanding of the traffic that rides the pipes. In the early days of the Internet, Worldcom succeeded in perpetrating a fraud of devastating proportions on the backbone market by claiming to experience phenomenal traffic growth. No one could document this growth or disprove it, so investors and competing providers accepted Worldcom's statements. By the time the truth became too difficult to hide, Worldcom had driven itself into bankruptcy and had created such over investment in "dark fiber" that carriers experienced depressed returns and avoided future investment for years.

This is but a single, blatant example of how our ignorance of the nature of Internet traffic can hurt us. This ignorance about the nature of traffic impacts every aspect of broadband policy. Experts constantly debate the reality of the supposed "exaflood" that commercial actors use to justify any behavior from blocking peer-to-peer traffic to demanding third-party payments for delivery of high-bandwidth content. But no one can even say with certainty what percentage of Internet traffic is peer-to-peer, streaming video, routing information, or even security patches for Microsoft Windows.

It is imperative for the development and sustainability of a robust national broadband system that we gain a better understanding of the traffic that passes daily through our critical infrastructure. Collecting this information, however, presents numerous difficulties. Even if broadband access providers and other network operators wanted to provide this information, the collection of the information could give rise to significant concerns about user privacy and cybersecurity.

The Commission should therefore carefully consider what data network operators and access providers already collect, and how to collect this information with sufficient granularity to be useful while obscuring personal information or creating new network weaknesses. While the need for information is real, we must balance this need against the real concerns of both network operators and network subscribers.

### C. The Data Collected Should Include Qualitative Metrics, Create Standardized Methodologies, and Set Triggers For Policy Intervention.

In the last fifty years, we have developed numerous qualitative metrics to measure the economy, education, and other areas critical to our national well being. We can track consumer confidence, look at the relationship between high school education and earning power, and analyze whether a market is competitive. But we have yet to agree upon a standard set of indicators by which we can judge the success of our broadband policy in improving people's lives.

As part of the National Broadband Plan, we should at least begin to develop a set of "qualitative metrics," that will indicate not merely where broadband is available, but what it does for people. We must consider this a critical element not merely in assessing the "demand side" of the broadband equation, as discussed in Part III above, but as necessary data to develop a complete broadband policy. The results should inform federal funding, and related federal policy, as discussed further below.

## V. THE ROLE OF NON-COMMERCIAL ACTORS: FEDERAL, STATE, LOCAL AND NON-PROFITS.

As noted in the beginning, federal policy until now has viewed broadband deployment as a matter of bribing or coercing commercial actors. Only recently, with the passage of the ARRA, has Congress explicitly recognized the important role of non-commercial players in providing

sustainable broadband to communities.[67] Indeed, rather than examine how non-commercial entities could play a productive role in our National Broadband Plan, federal and state policy has too often sought to limit the role of non-commercial actors on the spurious ground that such entities should "not compete with the private sector." The National Broadband Plan should discard this outmoded and dangerous notion and should instead consider how best to incorporate the resources non-commercial actors have to offer in developing our critical broadband infrastructure.

The Federal government has a crucial role to play in the National Broadband Plan beyond writing blank checks to commercial entities. It is the job of the Federal government to coordinate across Federal, State and Local agencies and to fully leverage the assets and expertise of local communities and non-profit organizations as we work toward the goal of universal broadband access. In order to do this, the Federal government will have to design and adopt a "mindful" policy across all agencies that incorporates broadband into government-wide objectives, in order to fully leverage and promote the benefits of broadband access.

Physical connectivity, of course, is but one piece in a larger puzzle. In order to enjoy the benefits of universal access, we will need a national strategy for education. To this end, the Federal government should call upon community institutions such as libraries, educational institutions and community centers to assist with deployment and training. These institutions are uniquely positioned to lend a hand in these efforts, as they understand the needs of their local communities and can serve as intermediaries between citizens and the executors of the National Broadband Plan.

[67] ARRA §§6001(b)(3), 6001(e).

Finally, either the FCC or Congress should consider enacting taking steps to preempt any legal impediments to broadband deployment, such as State or Federal acceptable use agreements and policy restrictions on "competing with the private sector". As long as such impediments remain, the National Broadband Plan will be unable to live up to its promise of universal access for all Americans.

### A. The National Broadband Plan Should Explicitly Call for the Removal of Legal Barriers to Deployment By Non-Commercial Entities.

Many municipalities have attempted to meet consumer and citizen demand for high-speed connectivity in areas where incumbent providers have failed to provide it.[68] In many of these cases, consumers are denied these services because of restrictive "acceptable use policies" or because of protectionist laws lobbied for by incumbent providers. If a broadband plan is to benefit from the diverse methods available for providing connectivity, such counterproductive and anticompetitive barriers to connectivity must be eliminated.

The Commission itself has recognized in the telecommunications sector that municipally-owned services would serve the goals of enhancing access and competition.[69] The Supreme Court has likewise recognized that state laws restricting municipal and local provision of services is purely the result of anticompetitive lobbying.[70] When existing service providers fail to deliver the speeds that residents demand, those residents should have the choice, as expressed through their votes in both ballots and dollars, to have their municipalities offer them a much-needed utility. The Commission should therefore act to preempt state and local regulations that

[68] *See* Fiona Morgan, *Mighty, mighty broadband*, INDYWEEK, June 18, 2008, *available at* http://www.indyweek.com/gyrobase/Content?oid=oid%3A259848 (industry estimates 44 publicly owned fiber networks in the US).
[69] *See Nixon v. Missouri Municipal League*, 541 U.S. 125, 131-32 (2004) (quoting statements from the then-Chairman, Commissioner Tristani, and Commissioner Ness)
[70] *Nixon v. Missouri Municipal League*, 541 U.S. at 138.

would prevent citizens and consumers from choosing to create access where existing providers have failed to create a competitive market—or any market at all.

The FCC possesses the authority—and perhaps the obligation—to do so.[71] Current precedent does not create any burden to such preemptive action. To the extent that state laws prohibiting publicly-offered broadband Internet services may be based on laws limiting telecommunications services like those in *Nixon v. Missouri Municipal League*,[72] it should be noted that the Supreme Court's limitation of preemption explicitly did not extend to municipally-owned, independently chartered corporations.[73] Section 253(a) therefore provides the Commission with a ready authority to preempt laws that so restrict independently chartered corporations owned by municipalities. The Commission is also obligated by section 706 of the 1996 Act to actively encourage the deployment of broadband Internet services, and is authorized to use "measures that promote competition in the local telecommunications market, or other regulating methods that remove barriers to infrastructure development."[74]

This obligation, coupled with the Commission's broader mission to encourage competition and connectivity, militates towards the proper exercise of its preemption powers to let communities meet their own demand for broadband.

---

[71] Matthew Dunne, *Note: Let My People Go (Online): The Power of the FCC to Preempt State Laws that Prohibit Municipal Broadband*, 107 COLUM. L. REV. 1126 (2007).
[72] Notably, since the ruling in *Nat'l Cable & Telecomm. Ass'n v. Brand X Internet Services*, cable broadband providers (and presumably a number of over broadband providers) are not telecommunications services. 545 U. S. 967 (2005). Thus, laws prohibiting the offering of telecommunications services would not preclude the offering of broadband Internet.
[73] *Nixon v. Missouri Municipal League,* 541 U.S. at 131 n.2.
[74] Telecommunications Act of 1996, Pub. L. No. 104-104 §706(a), 110 Stat. 56 (1996).

### B. All Federal Agencies Should Seek to Leverage Broadband To Meet Their Objectives and Be Mindful Whether Programs Accidentally Create Barriers To Broadband Adoption

If the Federal government is to promote the benefits of broadband connectivity to its citizens, it must itself become an exemplar of these benefits. As part of the National Broadband Plan, the Federal government should be an evangelist for broadband, by embracing new media technologies at all levels of government. In order to achieve this goal, the architects of the plan will have to craft a "mindful" Federal policy, one that incorporates broadband into government-wide policy initiatives and which leverages and promotes the advantages of broadband access.

During the 2008 Presidential campaign, President Barack Obama successfully utilized a number of new media tools--including Twitter and YouTube--to organize and mobilize online communities. The Federal government now has before it an opportunity to use many of these same tools to provide citizens with information, increase transparency and encourage engagement and participation in government processes. We recommend that Congress establish an executive agency to craft and execute a Federal policy for adopting and evangelizing these tools and others across all Federal agencies.

### C. Use of Community "Anchor Institutions" Such As Libraries and Community Colleges

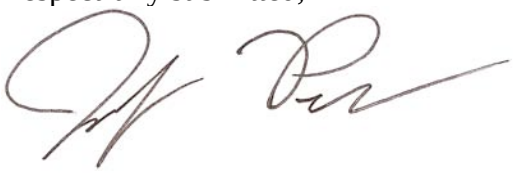
With regard to both the physical deployment of broadband and the education of new broadband users, community organizations such as libraries, community colleges and community centers have an important role to play. First, they can serve as "anchor organizations," organizations that provide free broadband service and training, thereby serving as a beachhead into their community. Additionally, if these organizations receive high-speed Internet access as part of the National Broadband Plan, they could serve as resellers of bandwidth, thereby encouraging competition and incentivizing new entrants into the market at a local level.

Second, these institutions can serve as important allies in the effort to provide information and training in local communities. As broadband becomes available in communities where it was previously unavailable or unaffordable, there will be a great need for institutions that can provide citizens with information regarding how best to acquire, set up and utilize a broadband Internet connection. Furthermore, these citizens could be trained in new media literacy, on how best to use their broadband connection for educational, economic and creative pursuits and on how to use the web to become more engaged in their communities and government. If we succeed in our goal of making broadband ubiquitous in the United States, such training will be crucial in ensuring that all citizens are able to take full advantage of the tremendous opportunities that broadband connectivity can bring to each and every American.

## CONCLUSION

The creation of a comprehensive National Broadband Plan provides a unique opportunity for us as a nation to choose how our digital destiny will unfold. Congress, and the American people, expect bold action. The principles set forth in these comments provide a road map for the agency to formulate a broadband plan that can secure for us a rich and productive future.

Respectfully submitted,



Jeffrey Pearlman
*Counsel for PK et al.*


Jeffrey Pearlman
Art Brodsky
Alex Curtis
Mehan Jayasuriya
Sherwin Siy
Robb Topolski
Public Knowledge
1875 Connecticut Ave. NW, Suite 650
Washington, DC 20009
(202) 518-0020

Andrew Schwartzman
Parul Desai
Media Access Project
1625 K Street, NW
Suite 1000
Washington, DC 20006
(202) 232-4300

Michael Calabrese
Benjamin Lennett
New America Foundation
1899 L St., N.W., Suite 400
Washington, DC 20036
(202) 986-2700

Amina Fazlullah
U.S. PIRG
218 D Street SE
Washington, DC 20003-1900
(202) 546-9707


June 8, 2009

# Forcing the Net Through a Sieve: Why Copyright Filtering is Not a Viable Solution for U.S. ISPs

Mehan Jayasuriya, Jef Pearlman, Robb Topolski, Michael Weinberg, Sherwin Siy


www.PublicKnowledge.org

## Acknowledgements

The authors of this paper would like to thank the following advocates for their assistance and suggestions during the editing process: Public Knowledge president Gigi Sohn, Public Knowledge staff attorney Rashmi Rangnath, Public Knowledge legal consultant Adam Thomas, Public Knowledge intern Katy Tasker and Electronic Frontier Foundation staff technologist Peter Eckersley.

## Table of Contents

## Executive Summary

Copyright filtering, the latest proposed "magic bullet" solution from the major music and movie studios and industry trade groups, poses a number of dangers to Internet users, legitimate businesses and U.S. federal government initiatives to increase the speed, affordability and utilization of broadband Internet services. The following whitepaper presents a number of reasons why the use of copyright filters should not be allowed, encouraged or mandated on U.S. Internet Service Provider (ISP) networks. Among them:

1. **Copyright filters are both underinclusive and overinclusive.** A copyright filter will fail to identify all unlawful or unwanted content while harming lawful uses of content.
2. **Copyright filter processing will add latency.** Copyright filters will slow ISP networks, discouraging use, innovation and investment and harming users, businesses and technology policy initiatives.
3. **The implementation of copyright filters will result in a technological arms race.** Users will act to circumvent the filters and the architects of the filters will find themselves caught in a costly, unwinnable arms race.
4. **Copyright filters do not make economic sense.** The monetary costs associated with copyright filtering far outweigh any perceived benefits.
5. **Copyright filters will discourage investment in the Internet economy.** Copyright filters will disrupt the Internet ecosystem, severely undermining our most promising engine for economic growth.
6. **Copyright filters will harm free speech.** Due to technological limitations, copyright filters will harm lawful, protected forms of speech such as parody and satire.
7. **Copyright filters could undermine the safe harbor provisions that shield ISPs from liability.** Under the Digital Millennium Copyright Act (DMCA), ISPs are shielded from liability for their users' actions. Copyright filters could undermine these safe harbors, which have allowed the Internet to become the most important communications medium of the modern era.
8. **Copyright filtering could violate the Electronic Communications and Privacy Act.** Copyright filtering could constitute unlawful interception under the Electronic Communications and Privacy Act (ECPA).

# 1. Introduction

Ever since the advent of mainstream file-sharing networks in the late 1990s, the major music and movie studios and industry trade groups (henceforth referred to as "the content industry") have been searching for a "magic bullet" solution to the problem of online copyright infringement—one that would simultaneously eradicate the problem of online file sharing while breathing new life into pre-digital business models. Obviously, this magic bullet has eluded the industry thus far, though not due to a lack of effort. In the years since the emergence of Napster, the content industry has taken legal action against service providers and end users, shut down centralized file-sharing networks and flooded the web with bogus copies of commonly traded files. Yet, as even a casual observer can attest, illicit file sharing persists online.

The industry's latest silver bullet comes not in the form of a legal campaign or disruption strategy but rather in the form of a technology that falsely promises to automatically and effectively eradicate copyright infringement online. Copyright filtering, as it is called, is a method whereby network appliances use a technology known as Deep Packet Inspection (DPI) to inspect the data that travels over an Internet Service Provider's (ISP's) network, identifying content as it passes through the filter and then dealing with that content accordingly. Unsurprisingly, the content industry has become an outspoken advocate for the use of this technology, pushing governments the world over to pass legislation requiring its use on ISP access networks.

Here in the United States, the content industry was almost successful in forcing language into the American Recovery and Reinvestment Act of 2009[1] that would have allowed ISPs to engage in copyright filtering under the auspices of "reasonable network management," an act that would have had far-reaching consequences for citizens, businesses and the entire Internet ecosystem.[2] Meanwhile, based on information publicly available,[3] a secretive, multilateral, international trade agreement known as the Anti-Counterfeiting Trade Agreement (ACTA)[4] could also pave the way for copyright filtering in the United States, the European Union, Australia, Japan, Canada and a number of other nations.

Before we rush to implement a technology—much less require its use—we must objectively examine that technology and question not only its efficacy but also its costs, consequences and drawbacks. It is in this spirit that Public Knowledge, a non-profit public interest advocacy group, has embarked upon an analysis of copyright filtering technologies. The results of that analysis are presented in this whitepaper.

In the following paper, we will take a close look at the technology behind copyright filtering, analyze the legal ramifications of filtering and discuss the policy implications of condoning or mandating the use of copyright filters. We will consider the impact that filtering is likely to have on user privacy, examine the costs associated with filtering and contemplate the potential impact copyright filtering will have on network security for both users and service providers.

---

[1] *See* the Recovery.gov website (http://www.recovery.gov/?q=content/act).
[2] "Senator Feinstein Trying to Sneak ISP Copyright Filtering Into Broadband Stimulus Bill," *TechDirt*, Febuary 10, 2009 (http://www.techdirt.com/articles/20090210/1050313726.shtml).
[3] "Secret ACTA Treaty May Include ISP Filtering," *Ars Technica*, June 4, 2008 (http://arstechnica.com/tech-policy/news/2008/06/secret-acta-treaty-may-include-filtering-provisions.ars).
[4] For more information on ACTA, see the Electronic Frontier Foundation website (http://www.eff.org/issues/acta).

Regardless of how they are implemented, copyright filters will alter the fundamental behavior of the Internet and in so doing, will likely disrupt the Internet ecosystem in ways that we cannot predict. The Internet was designed to be an open system from end-to-end, which is to say, a system that moves content between hosts and clients as quickly as possible on a first-come-first-served basis—regardless of the nature of that content. Copyright filters will inject delay into this system, make automated judgments regarding the legality of content and will then degrade or discard that traffic accordingly. The Internet was not designed to support this type of activity and for this reason, the implementation of copyright filters will assuredly result in a variety of technical problems for all parties involved in the Internet ecosystem. Much like traffic lights on our interstate highway system, copyright filters on our open, high-speed networks will be a poor fit.

Technical issues aside, there are other pressing questions that must be answered before we can determine whether copyright filters should be installed on the networks of Internet service providers. Chief among them is the question of efficacy: will copyright filters solve the problem that they purport to solve? Based on our technical analysis, the answer appears to be no. By virtue of their design, Internet filters are doomed to be both underinclusive and overinclusive—they will fail to identify all illegal uses of content while simultaneously blocking legal content. The filters will be underinclusive because their technology is not advanced enough—and will likely never be advanced enough—to identify every instance of prohibited content on the network. Filters will also be overinclusive; as a filter will never be able to distinguish between fair, legal uses of content and illegal uses of content with 100 percent accuracy. Given that even legal scholars and courts are often unable to reach a consensus on questions of fair use, this is not surprising—the question of whether or not a piece of content

constitutes fair use is often a difficult one for even human beings to answer. Furthermore, as history attests, users will work to actively circumvent the filter, thereby luring the architects of the filter into a fruitless technological arms race.

Ultimately, we will demonstrate that copyright filtering does not constitute "reasonable network management," as some proponents of the technology would have policymakers believe; rather, copyright filtering is *content* management. Instead of making determinations based on how data moves over the network, copyright filters attempt to ascertain *what* that data constitutes, in order to block, degrade or delay certain types of content. Copyright filtering is not necessary for a network to operate properly or efficiently and therefore, should not be considered a form of network management at all.

The content industry would like to convince policymakers and the general public that copyright filtering is the most effective means by which to combat online copyright infringement and protect America's creative economy. This could not be further from the truth. In practice, copyright filtering is likely to harm innovators, end users, online service providers and Internet service providers alike. What's more, it will compromise the privacy of all American Internet users for the perceived benefit of one industry. As such, copyright filtering will discourage investment in the Internet economy—our most promising engine for economic growth—and will harm American competitiveness in the global market.

Finally, copyright filtering holds the potential to undermine the goals of the National Broadband Plan (NBP), a Federal initiative to increase the speed, adoption and affordability of broadband Internet services nationwide. As we will see, copyright filtering, if implemented on U.S. networks, will chip away at each of the NBP's stated goals. The network appliances that act

as filters will slow traffic on the networks on which they are installed, an unavoidable consequence of their traffic analysis. The costs associated with filtering—most notably, the purchasing and maintenance of filtering hardware and software—are likely to be passed on to consumers, decreasing the affordability of broadband services. And in compromising the privacy of Internet users, copyright filters will discourage Internet use, even as the Federal government works to promote the educational, economic and civic benefits of broadband access.

If we are serious about combating copyright infringement online, we should invest our time and resources in developing methods that will effectively discourage illegal activity, without harming fair users, innovators and the Internet economy. If we rush to require the use of a technology that we have not fully considered, we will find ourselves trapped by a burdensome and even dangerous mandate, one that will have far-reaching consequences that will affect the flow of information, knowledge, political discourse and capital. To date, the content industry has presented the debate surrounding copyright filtering as a choice between a specific technology and rampant online “piracy”. In the paper that follows, we will demonstrate that this is a false choice and that copyright filtering should not be considered as a viable solution for U.S. ISPs.

## 2. Technological Analysis: The Anatomy of a Copyright Filter

The term "filter" refers to a technology that can be employed by an ISP or end-user to automatically detect a specific type of traffic or content and then take action based on the nature of the data detected. Historically, filters have been used at the network level on private networks to block content that is presumed to be illegal (child pornography, unlawfully traded movies and music, etc.) or unwanted (spam, personal communications, social networking content, etc.). In such cases, the filtering hardware or software is installed at the point where the private network connects to the public Internet. If a filter were installed on an ISP network, however, it would likely be installed at the transit provider level—either at points along the Internet backbone or at the point where the private ISP connects to the backbone.

The architecture of the Internet is such that in order to provide Internet service, a network operator needs only to concern itself with the Internet Protocol (IP) headers of data packets that traverse its network—that is to say, the outer layer of the most basic unit of Internet traffic. This is because all of the information required by a provider to do its job—receiving data packets and then forwarding them to their next stop en route to their destinations—is contained in the packet's IP header. ISPs are occasionally required to inspect the traffic of individual users, pursuant to regulatory requirements and the needs of law enforcement but otherwise generally do not analyze any part of the packet other than the header.

However, if an ISP chose to implement a copyright filter, that ISP would be electing to scrutinize all traffic from all subscribers, in real time. The provider would then act upon the information gleaned by the filter, in accordance with its policy aims. Needless to say, this system

represents a significant departure from the traditional role of the ISP, which is to forward traffic as quickly and efficiently as possible.

When designing an Internet filter of any sort, one must first determine the policy that will define the filter's function. In this section, we will consider two policy definitions of theoretical copyright filters: a "strict" policy definition and a "permissive" policy definition. We will then address the different mechanisms whereby a filter might identify different types of content. Next, we will explore the filter's policy enforcement mechanism, whereby it takes action once a piece of content has been isolated and identified. Finally, we will look at examples of technologies on the market that are currently marketed as real-time filtering solutions.

## I. DETERMINING FILTER POLICY

### A. Strict Policy Definition

As all works created in the United States are essentially "born" copyrighted[5], a strict copyright filter would have to operate under the assumption that all Internet traffic contains copyrighted work. As such, a filter based on a strict policy definition would analyze each packet, in an attempt to identify data that pertains to a copyrighted work. If a packet were determined to be a piece of such a work, the filter would then attempt to verify the identity of the work. Once the work's identity is determined, the filter would verify that both the server and client are properly licensed to distribute and receive the content in question and if not, would then presumably take action to delay, degrade or discard that packet.

[5] A work is "created" when it is fixed in a copy or phonorecord for the first time (17 U.S.C. § 101); "Copyright protection subsists, in accordance with this title, in original works of authorship fixed in any tangible medium of expression, now known or later developed, from which they can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device." (17 U.S.C. § 102(a)).

This, of course, will be an extremely complicated process. Questions of content identification aside, the filter must consider a number of factors when making judgment calls regarding licenses. This is because certain content licenses (mechanical licenses, for example) might only authorize a certain number of reproductions of a work or might only allow distribution to certain users in certain geographies or within certain time frames. The complex nature of licensing in the United States adds an added layer of complexity to the copyright filter.

Here is an abstracted version of how a strict filter might operate:

```
SomeServer is sending SomeContent to SomeUser

{

if SomeContent is authorized on ContentAllowedList;

And SomeServer is authorized on the AuthorizedToSend list for
SomeContent to SomeUser;

Then, allow the transfer;

};

Otherwise deny the transfer;
```

### B. Permissive Policy Definition

Unlike a strict copyright filter, a permissive copyright filter would operate under the assumption that any traffic passing through it does not represent an unauthorized transfer of copyrighted work. In this sense, a permissive filter could be said to take an "innocent until proven guilty" approach, rather than the "guilty until proven innocent" approach taken by the strict filter. Like the strict filter, the permissive filter would still have to identify works, the identities of the parties involved in the transfer of the works and any applicable licenses that authorize that transaction. Unlike the strict filter, however, the permissive filter would cross-check the content against a database of "blacklisted" content—that is to say, content that is not

allowed to be transferred on the web, in accordance with the wishes of the copyright owner. The permissive filter might also record the IP address of the end user and/or cross-check that IP address against a database of known offenders.

Here is an abstracted example of a permissive filter:

```
SomeServer is sending SomeContent to SomeUser

{

if SomeContent is not listed on RegisteredContentList;

Or SomeServer is authorized on the AuthorizedToSend list for
SomeContent to SomeUser;

Then, allow the transfer;

};

Otherwise deny the transfer;
```

Most copyright filters will fall into either the strict or permissive category, with regard to policy definition. It should be noted that neither type of filter is compatible with existing copyright law, as neither filter allows for the unauthorized uses permitted under copyright law. While the permissive filter appears to be a more practical implementation (as it would likely halt/slow the flow of traffic far less than the strict filter), it still fails to consider those uses that are not subject to the exclusive rights of the content owner and hence permitted under copyright law. These limitations are discussed in depth in the legal analysis section of this paper.

## II. IDENTIFYING TYPES OF CONTENT

As part of the policy definition process, the designers of a copyright filter would have to craft a policy that would allow the filter to identify different types of content. In order to determine the type of content being transmitted, the filter, on behalf of the network operator, would analyze the bits that travel over the network and then make a determination as to what

type of content is in transit. Is it audio? Is it video? Is it both audio and video? Is it a document? If so, is it a book? Is it a magazine article? Is it sheet music? Is it a photograph?

Once the filter has made a first-level content determination (type of content), it will then attempt to identify the exact identity of the work being transferred. Needless to say, this part of the process will significantly increase the complexity of the determination process. A filter cannot simply block music as a broad category; it must only block specific pieces of music that match a certain profile. In generic terms, a filter might seek to block *SomeRestrictedPerformance* of *SomeRestrictedTitle* by *SomeSignedBand* which was released *SomeYear* by *SomeLabel*. It might also be required to block *SomePublicDomainMelody* as sung by *SomeSinger* as arranged by *SomeRestrictedArranger*, for example.

In order to have enough data to make these determinations, the filter must either permit the user to download a significant portion of the file in question (perhaps the entire file), or it must itself download a significant portion of the file, before passing that data on to the user, if that data is determined to not run afoul of the filter's policy. Given that the former method undermines the effectiveness of the filter, we must assume that the latter method would be used and that Internet traffic would be significantly slowed, as each file would effectively have to be downloaded twice—once by the filter and a second time by the user. The only obvious solution to this problem is to require a license that authorizes the transfer of a work in advance of the file transfer. Such a system would, of course, require that all legitimate transfers on the Internet be accompanied by a license—a solution that seems unrealistic at best.

Regardless of how transfers are regulated, the most basic task that the filter performs will be identifying a piece of content. Technologically speaking, there are a number of different ways

that this might be accomplished, many of which are discussed below. These methods have been divided into three categories: traffic inspection, content analysis and instrumentation of the end-user's device.

### A. Traffic Inspection

The term traffic inspection refers to an identification mechanism that does not attempt to identify specific pieces of content online. Rather, a traffic inspection scheme would simply analyze the nature of traffic that travels over the network and would then make assumptions about the content carried by that traffic. This was the method used by Comcast, whereby it clandestinely degraded all traffic that used the BitTorrent protocol, under the flawed assumption that all of the traffic that uses that protocol is illicit in nature. Comcast has since been reprimanded by the Federal Communications Commission (FCC) for its actions, as it was found that Comcast's content management technique violated the four principles outlined in the FCC's Internet Policy Statement.[6] Comcast has since ceased degrading traffic related to BitTorrent.

#### *i. Identifying Content Through Associated Transactions or Through the Identification of Connection Patterns*

One traffic analysis method that might be used for content identification is the analysis of either packet exchange or geographic patterns. The phrase "packet exchange" refers to the conversational patter of an Internet connection—the back-and-forth exchange of data between the host and the client. Connection pattern analysis methods look at who sits on either end of a packet exchange transmission and how many connections those hosts and clients are opening and uses this data to speculate as to what the nature of the transmission might be.

[6] See "Formal Complaint of Free Press and Public Knowledge Against Comcast Corporation For Secretly Degrading Peer-to-Peer Applications" (http://www.publicknowledge.org/pdf/fp_pk_comcast_complaint.pdf).

While there are existing technologies that can analyze Internet traffic using these methods, some of which are marketed for copyright protection, the fact of the matter is that these technologies do not identify instances of copyrighted content. Rather, they block certain types of traffic—file transfers, for example—regardless of the nature of the content in transit. In other words, they interrupt all traffic, lawful and unlawful.

### *ii. Identification Through Analysis of Traffic Patterns*

Similar to the technique of identifying works through packet exchange patterns, traffic pattern analysis identifies traffic patterns popular with—but not exclusive to—users who illegally share copyrighted works online. Like with packet exchange analysis, this method does not attempt to directly distinguish between legal and illegal content and instead, blocks all traffic that matches a certain pattern. However, unlike connection pattern analysis, which looks only at who is sending and receiving data, traffic exchange analysis also examines the size and timing of the packets exchanged. Commercial network products such as ArborNetworks eSeries[7] and RedLambda Integrity[8] use this method.[9]

## III. CONTENT ANALYSIS

Unlike traffic inspection techniques, content analysis methods look beyond the headers of packets, in an attempt to analyze the content of those packets. While content analysis techniques avoid some of the pitfalls of traffic inspection, they exhibit a different set of shortcomings and, as such, are similarly flawed.

[7] See Arbor Networks' website (http://www.arbornetworks.com/index.php?option=com_content&task=view&id=1466&Itemid=693).
[8] See Red Lambda website (http://redlambda.com/integrity_overview.php).
[9] Examples of filtering technologies will be discussed in-depth later in this section.

In order to identify a work using a content analysis technique, the filter would compare the data in hand to a registry containing the identifying characteristics of all known works or, failing that, all works of concern. Considering that it will be impossible to create a database containing information on all works, we must assume that any such database will contain a subset of those works, all of which will likely be works protected under copyright.

### A. Identifying Content Through the Use of Metadata

Most digital works contain what is called metadata: data about data that is used for purposes of categorization. One of the most recognizable forms of metadata are the ID3 tags that come attached to many digital music files. In this case, the metadata provides information including the performer, songwriter, title, album, year of recording and file quality of a file to the software and hardware that decodes that file. This is how devices like iPods and software like iTunes are able to display the title and artist when playing a song. Most digital media files including movies, photographs and digital books, contain similar metadata.

Given this wealth of data, you might assume that identification would be a relatively simple task for a copyright filter. Unfortunately for the filter, this is not the case. Metadata can be easily edited by users and as such, often contains incomplete or imperfect information. If you have ever seen two songs by the same artist in your music library that were classified correctly but differently (for example, “John Lennon” vs. “Lennon, John”), you have already observed this phenomenon. If we allow for the possibility of mislabeling (“John Lennon” vs. “The Beatles”), this problem is compounded.

Even if the metadata in question were complete and reliable, however, a filter would still encounter a number of problems when attempting to identify a work based on that data. Much of

this has to do with the complex nature of the type of data that is found in metadata and the possibility for overlap. Take for example, a song entitled "Happy Birthday". Given only this title, a filter would be unable to determine whether the work in question is the copyrighted song "Happy Birthday," a different song of the same name or something else entirely that has been mislabeled. If the filter attempts to identify the song based on its listed performer, it will run into a similar set of problems. Does the artist listed perform the song in question or is it an amateur cover of a song by that performer? Or, is the file in question simply a song performed in the style of that performer?

In attempting to positively identify a piece of content based on its metadata, a copyright filter will likely have no choice but to view the metadata holistically. Given that none of the data will be certifiably reliable, however, the filter will never be able to identify a piece of content based solely on its metadata with any degree of certainty. You might say that the filter would have trouble judging a book by its cover.

### B. Identification of Content Through Digital Watermarks

The term "digital watermark" refers to a technique whereby content producers embed an invisible digital signature into a file intended for physical or digital distribution. Once the watermark has been embedded, the watermarked content can be easily identified, tracked, managed and secured by technology that can read the watermark. Such technology is currently being used by a number of major movie studios, record labels, television broadcasters and enterprises with digital image assets.[10] While a copyright filter could theoretically filter only for watermarked content, this would result in an underinclusive filter. While an unauthorized

[10] See "Descriptions of Demonstrators at the ICAC 2007 Tech Exhibition," Congressional Internet Caucus Advisory Committee (http://www.netcaucus.org/events/2008/kickoff/demonstrators.shtml).

transfer of a file containing the watermark would be halted, a non-watermarked version of the same file would be allowed to pass. Considering that many of the unauthorized transfers of copyrighted files that take place online involve the transfer of content created by users through means such as camcording and the ripping of insufficiently secured CDs and DVDs, it is likely that a great deal of copyrighted content would pass through the filter undetected. Furthermore, if watermark filters were widely deployed, users would likely develop techniques for stripping the watermarks out of files.[11] This would only exacerbate the content-protection arms race (discussed at length in the "limitations and consequences" section of this paper) and inject additional costs into Internet services, increasing costs for the consumer and providing no discernable benefit.

### C. Identification of Content Through Acoustic or Visual Characteristics ("Fingerprinting")

Metadata aside, every copyrighted work contains a great deal of identifiable information. A song is not known only by its title but also by a succession of notes and words. A book isn't simply known by its title but by its characters, setting and plot. The term "fingerprinting" refers to a technique whereby audio, video or literary patterns in the work are used to generate a unique fingerprint, which can then be used to identify other instances of that same work. Network products like Vobile VideoDNA[12] and Audible Magic[13] use this method to identify content.

While fingerprinting might be an effective technique for content identification, it is not a practical solution for Internet filtering. In order to positively identify a work, a fingerprint filter

[11] For an in-depth account of one research team's successful effort to crack digital watermarks, see Craver, Scott A. et al. "Reading Between the Lines: Lessons Learned from the SDMI Challenge," August 13, 2001 (http://www.usenix.org/events/sec01/craver.pdf).
[12] See Vobile website (http://www.vobileinc.com/technology.html).
[13] See Audible Magic website (http://www.audiblemagic.com/index.asp).

would have to download a considerable portion of the work in question. This would mean that either all Internet traffic would have to be considerably delayed or that, in many cases, most or all of a transfer would have to be completed before the filter would be able to make a determination regarding the nature of the content.

Even if it were practical to use digital fingerprinting in conjunction with a copyright filter, such a filter would fail to recognize legal uses of copyrighted content--for example, a film review containing brief movie clips, a transformative remix of a song or a parody or satire of a copyrighted work--and would therefore encroach on the rights of Internet users. This matter is discussed in depth in this paper's legal analysis section.

### D. Instrumentation of the End-User's Device

While this paper interprets the term copyright filter to mean an in-network technology (that is to say, hardware or software that is implemented by an ISP and which sits on the ISP's network), some have suggested a different method, whereby filtering software would be installed on the machines of end-users, ostensibly as the result of a legal mandate. Certain content industry representatives, including Recording Industry Association of America (RIAA) president Cary Sherman,[14] have advocated this method and some have interpreted this advocacy as a tacit admission that in-network copyright filtering will be ineffective.[15]

While a full analysis of filtering techniques that require instrumentation of the end-user's device is outside the scope of this paper, it should be apparent that such instrumentation would raise a number of questions regarding legality, practicality and feasibility. While instrumentation

[14] "RIAA boss: Move copyright filtering from ISPs to users' PCs," *Ars Technica*, February 7, 2008 (http://arstechnica.com/old/content/2008/02/riaa-boss-spyware-could-solve-the-encryption-problem.ars).
[15] *Ibid.*

of the end-user's device would offer some advantages to those attempting to filter, such a solution would still suffer from most, if not all of the shortcomings discussed in this paper. Additionally, the installation of software on end-users' machines raises an entirely new set of security questions.[16] When the Chinese government issued a mandate that the "Green Dam" filtering software be installed on all machines sold in the Chinese market, major security vulnerabilities were discovered within days of the software's release.[17] Some of these flaws could have been exploited remotely, allowing malicious programmers to compromise the security of any machine running Green Dam over a network connection. As a result, the Chinese government rescinded the mandate, making the use of Green Dam optional.[18]

## IV. POLICY ENFORCEMENT

Once a piece of content has been isolated and identified, the filter must act pursuant to the policy definition that it was created to enforce. This stage of the filter's operation is referred to as policy enforcement.

By default, the sending and receiving of all content is allowed on the Internet. Currently, the exceptions to this rule are filters installed by users, businesses and governments on their private networks, some of which allow no traffic to pass as a default and then allow data through, bit-by-bit, after it has been determined to meet a certain set of criteria.

[16] See also the Sony BMG CD copy protection scandal, wherein Sony encoded certain audio CDs with a copy protection mechanism that would surreptitiously install a "rootkit" on users' PCs, thereby compromising the security of those PCs while sending user information to Sony. As a result, lawsuits were filed against Sony BMG in the states of California, New York and Texas. For more information, see "Real Story of the Rogue Rootkit," *Wired*, November 11, 2005 (http://www.wired.com/politics/security/commentary/securitymatters/2005/11/69601).
[17] "Big Vulnerabilities in China's Mandatory Filtering Software," *Ars Technica*, June 1, 2009 (http://arstechnica.com/tech-policy/news/2009/06/big-vulnerabilities-in-chinas-mandatory-filtering-software.ars).
[18] "China Caves, Says Green Dam is Optional," *ChannelWeb*, June 16, 2009 (http://www.crn.com/software/217900033).

Given that the Internet's default is to allow all traffic to pass, we might reasonably assume that a copyright filter would operate in the same manner, allowing any content that it cannot identify to pass. While it is possible that a copyright filter could be set up to operate like some security filters, allowing no content to pass as a default, such a policy definition would likely result in massive traffic problems on that network, as the filter would simply reject any bit that it did not understand. Given the speed with which new technologies and content types are created and propagated on the Internet and the interconnected nature of all Internet traffic, the consequences of deploying such a filter on a public ISP could be disastrous.[19]

Regardless, once the filter has captured and identified a piece of content, it must take action. The range of actions that can be taken are defined by the following broad categories:

**1. Allow:** The data is allowed to pass unmolested, as it would on the Internet.

**2. Flag:** Before being forwarded on to its destination, the traffic is 'flagged'. In most cases, this will mean that the traffic is reclassified to have a higher/lower precedence over other traffic, in order to slow the traffic, off-load the traffic to follow a different path, or change the traffic's source or destination. Flagging might also be used to track traffic (*i.e.* to identify the sender or receiver of the data).

**3. Deny:** The data is discarded and is not forwarded toward its final destination.

Within these three categories, there is much room for ingenuity. If the filter alters the path of traffic, the intended destination may never receive that traffic if the path of redirection is

[19] See Public Knowledge and Free Press' complaint against Comcast wherein it is stated that not only does Comcast's blocking of BitTorrent traffic affect Comcast users, it also causes a disproportionate amount of BitTorrent traffic to be offloaded on to other ISPs, thereby disrupting the usual balance of traffic using that protocol. "Formal Complaint of Free Press and Public Knowledge Against Comcast Corporation For Secretly Degrading Peer-to-Peer Applications" (http://www.publicknowledge.org/pdf/fp_pk_comcast_complaint.pdf).

congested or has no route to the destination host. Alternatively, that traffic might simply be dropped, as if the end host had gone offline. For example, a filter designed to falsely return empty search results may accomplish this task by routing the search queries down dead-end paths. Or, the designers of a filter could create a virtual tunnel restricted to a very low speed and could then reroute all unwanted traffic to that tunnel. In this latter scenario, while end-to-end communication would still be possible, its purpose (rapid file transfer, for example) would be defeated.

In all of these categories, the ISP would retain the ability to log what is being observed and acted upon by the filter. In some cases, the action taken by the filter might vary depending on network conditions—a policy that some U.S. ISPs claim to have adopted with regard to their network management techniques.[20]

## V. EXAMPLES OF COPYRIGHT FILTERS

While copyright filtering is rarely used at the ISP level, filtering technologies have long been used by the administrators of academic, corporate and government networks as a means to regulate the content, applications and information that can be accessed by end-users. In this section, we profile a few of the most commonly used filtering technologies—solutions that might be used by an ISP looking to implement a copyright filter.

### A. Audible Magic Copysense

[20] Both Comcast and Cox initially claimed that their network management scheme, which throttled BitTorrent traffic, was only used during times of "peak congestion". This was quickly disproven by network researchers who discovered that Comcast and COX were actually throttling that traffic 24-hours a day, regardless of network conditions. See "Study: Comcast, Cox Slowing P2P Traffic Around the Clock," *PCWorld*, May 15, 2008 (http://www.pcworld.com/businesscenter/article/145952/study_comcast_cox_slowing_p2p_traffic_around_the_clock.html).

Audible Magic develops, markets and sells content inspection filters that use fingerprinting technology. Audible Magic solutions have most commonly been used by destination websites that host user-uploaded content and are currently in use by a number of prominent Internet companies, including YouTube and Fox Interactive subsidiary MySpace.[21] As was previously described in the content identification section, fingerprinting filters analyze changes in the pitch, rhythm and relative sound level of multimedia files and then compare the results of that analysis to a database of known or "protected" works.[22]

Copysense, an on-the-network appliance sold by Audible Magic, can inspect data in transit on a live network. At present, Audible Magic markets Copysense to universities that are looking to police and/or discourage the use of peer-to-peer file trading software on their networks.[23] When a Copysense appliance detects a violation, it can take a number of actions, including logging that violation, interrupting the TCP protocol by forging reset (RST) packets,[24] or sending a series of increasingly severe notices to the user suspected of engaging in infringement (this regime, which has also been proposed for government use, is known as "graduated response" or "three strikes").[25]

Like all copyright filters, Copysense has a number of limitations. Most glaringly, it only monitors known P2P networks, which means that violators who download material using the

[21] According to the Audible Magic website (see http://audiblemagic.com/products-services/contentsvcs/customers.asp).
[22] See Audible Magic website (http://www.audiblemagic.com/index.asp).
[23] See the Audible Magic website (http://audiblemagic.com/products-services/copysense/copysense-university/).
[24] This technique, which was also employed by Comcast to sever BitTorrent connections, tricks both the host and the client machine into believing the other has reset the connection. See "Packet Forgery By ISPs: A Report on the Comcast Affair," Electronic Frontier Foundation, November 2007 (http://www.eff.org/wp/packet-forgery-isps-report-comcast-affair).
[25] For more background on "three strikes," see "Should Online Scofflaws Be Denied Web Access?," *The New York Times*, April 12, 2009 (http://www.nytimes.com/2009/04/13/technology/internet/13iht-piracy13.html). For more technical information on Copysense's specific implementation of three strikes, see Audible Magic's website (http://audiblemagic.com/products-services/copysense/copysense-university/graduatedresponse.asp).

FTP or HTTP protocols will remain undetected. Furthermore, Copysense is strictly reactive and requires a great deal of data in order to make a positive match—and this data must be provided by the copyright holder. This is to say that Copysense can only detect works that have been registered with Audible Magic.[26]

Copysense must also possess the P2P network's metadata in order to correctly identify a file as it travels over the P2P network. Not only does this metadata often vary from network to network, it also changes if any aspect of the file is altered (i.e. if a user changes the metadata contained within the ID3 tag of an MP3 file), which further undermines the filter's efficacy.

### B. Red Lambda Integrity

Despite its shortcomings, Audible Magic's Copysense seems less overinclusive than other copyright filtering technologies on the market. Other solutions utilize traffic analysis as a detection mechanism, which is to say that they simply block content at the protocol or applications level, without even attempting to make a determination as to the legality of the content that is in transit. Red Lambda's Integrity is one such popular technology. Red Lambda boasts that Integrity "Uses advanced hybrid DPI/Behavioral IDS to monitor network traffic for undesired protocols and applications,"[27] including P2P apps, IRC, FTP, IM, Skype, proxy use and application tunneling over HTTP, HTTPS, DNS and ICMP.[28] Red Lambda blocks a wide range of file transfers that take place using a wide range of protocols, presumably resulting in the blocking of untold numbers of authorized and/or non-protected works, including open-source software, independent music releases and public domain material. Even as it casts an

[26] See Audible Magic's website (http://www.audiblemagic.com/products-services/registration/).
[27] See Integrity features as listed on Red Lambda website (http://redlambda.com/integrity.php?p=features).
[28] See Red Lambda press release, "Joint House and Senate Committee Endorse Red Lambda's Integrity Solution to Fight Music and Movie Piracy on University Networks" (http://www.redlambda.com/files/press080408_red_lambda_house_senate_committee_recommend_rl.pdf).

unnecessarily wide net, Integrity is unable to detect browser-based file-sharing services, which are becoming increasingly popular with file sharers[29]. In many ways, Integrity exemplifies the shortcomings of copyright filters: it blocks a great deal of legal content, while still allowing a considerable amount of illegal content to pass.

### C. Vobile VideoDNA

Vobile's VideoDNA is a fingerprinting technology that is used to identify video and audio content. At present, Vobile does not market a copyright filtering solution; rather, its technology is used mostly for identification purposes, though it does offer a Software as a Service (SaaS) application that automates the process of identifying content and sending DMCA (Digital Millennium Copyright Act) takedown notices to the sites that are suspected of hosting infringing content.[30] According to Vobile's website, VideoDNA is currently being used by "all major Hollywood studios,"[31] to track the movement of content online and to identify opportunities for sending DMCA takedown notices. As we have seen, not only do these DMCA notices often target legal, fair uses of content, some studio representatives have explicitly stated that they don't consider fair use when sending such notices.[32] Recently, however, a court in the Northern District of California held that copyright holders are required to consider fair use when sending DMCA takedown notices.[33] On its website, Vobile hints at the complex legal issues that surround its products, though it makes no claim that its technology can be used without resulting in false positives over-indicating infringement. In fact, the language on the company's website

[29] "German Court Rules Against Rapidshare," *Billboard*, June 23 2009 (http://www.billboard.biz/bbbiz/content_display/industry/e3i6fad5a2a1d8e51328f91857dabe3e123).
[30] See Vobile's website (http://www.vobileinc.com/solutions.html).
[31] *Ibid.*
[32] "Universal Says DMCA Takedown Notices Can Ignore 'Fair Use'," *Wired*, July 18, 2008 (http://www.wired.com/threatlevel/2008/07/universal-says/).
[33] "Judge Rules That Content Owners Must Consider Fair Use Before Sending Takedowns," Electronic Frontier Foundation, August, 20 2008 (http://www.eff.org/deeplinks/2008/08/judge-rules-content-owners-must-consider-fair-use-).

suggests that the overinclusive nature of VideoDNA might be a desirable feature for content owners looking to eliminate all uses of their content online—whether legal or not ("In this digital age, content is often transcoded, mashed-up and transformed in a variety of ways before finding its way online. This creates a need for the content owner to actively monitor many online sharing sites to identify their content and decide whether to allow it to remain on the site or send a DMCA notice to the site operator asking for the copy to be taken down.").[34] It's also worth noting that U.S. ISP AT&T and entertainment companies Disney and NBC Universal are major investors in Vobile.[35]

[34] See Vobile's website (http://www.vobileinc.com/solutions.html).
[35] "AT&T, NBC and Disney invest in Vobile's VideoDNA," *Broadcast Engineering*, November 12, 2007 (http://broadcastengineering.com/news/att-nbc-disney-vobile-videodna-1112/).

# 3. Limitations and Consequences of Copyright Filtering

In the previous section, we discussed the different methods that might be used to automatically identify and filter out copyrighted content. In so doing, we also identified some of the limitations and shortcomings of copyright filters. As we have seen, depending on the technology used to identify copyrighted works, copyright filters will be underinclusive, overinclusive or both. The filter will fail to identify all copyrighted works that pass through it, will filter out legal, legitimate content or, as is the case with most filtering technologies currently on the market, the filter will fail on both counts.

In the following section, we will summarize some of the technological limitations of copyright filters and will discuss an unintended consequence that can be expected if copyright filtering is done at the ISP level: an encryption arms race.

## I. TECHNOLOGICAL LIMITATIONS OF COPYRIGHT FILTERS

In this section, we will briefly summarize the technological limitations of copyright filters. The non-technical limitations of filters will be discussed in depth in the legal analysis section.

### A. Architecture is a Poor Indicator

Both peer-to-peer and client-server architectures carry legitimate traffic. As such, it is impossible to block traffic based on architecture alone, without also blocking legal content.

### B. Protocol is a Poor Indicator

Like with specific types of architecture, a great deal of legitimate traffic is carried over protocols that are popular with peer-to-peer users. Many open-source software developers, for example, use the BitTorrent protocol to deliver their software to users efficiently.[36] Without the use of BitTorrent, which allows anyone to distribute a large file online without paying for bandwidth costs, many small and independent software developers would not have the means to distribute their software online. In addition to such small entities, even large corporations like Blizzard Entertainment[37] and public broadcasting entities like the Canadian Broadcasting Corporation and the Norwegian Broadcasting Corporation have used BitTorrent to make files available to users.[38] Finally, the similarities between P2P protocols and client-server protocols can result in the blocking of traffic that is completely unrelated to illicit file sharing. For example, when Comcast attempted to block the BitTorrent protocol, the company also inadvertently blocked traffic related to Lotus Notes, a commercial software suite used by businesses for email, scheduling and file-sharing purposes.[39]

### C. Media Type is a Poor Indicator

As is the case with architecture and protocol, media type (a category assigned to a certain type of media, *i.e.* film, music, image, etc.) is a poor indicator, as blocking any media type outright will result in the blocking of fair use, legally-shared and public domain content.

## II. FILTER PROCESSING ADDS LATENCY

[36] See "BitTorrent," *Wikipedia* (http://en.wikipedia.org/wiki/BitTorrent_(protocol)#Software).
[37] See "Blizzard Downloader," *World of Warcraft Universe Guide* (http://www.wowwiki.com/Blizzard_Downloader).
[38] See "BitTorrent," *Wikipedia* (http://en.wikipedia.org/wiki/BitTorrent_(protocol)#Broadcasters).
[39] 'Comcast Is Blocking More Than BitTorrent, Including Lotus Notes," *InformationWeek*, October 22, 2007 (http://www.informationweek.com/blog/main/archives/2007/10/comcast_is_bloc.html).

No matter how sophisticated filtering technologies eventually become, filtering will always slow the speed of traffic that travels over the network, so long as it is used to prevent unwanted or disfavored content from reaching users.[40] While the Internet was originally designed as a system that forwarded packets on toward their destination as quickly as possible, filters alter this behavior by analyzing each packet before determining how it should be treated or flagged. This process introduces delay or latency into the normal packet delivery process and as such, a filtered network will always be slower than an unfiltered network.

### A. Filtering Technologies Can Expose Networks to Security Risks

As we saw in the case of China's "Green Dam" filtering software, any piece of software that sits between users and the network or which is installed on a large number of PCs pursuant to a mandate will provide a highly visible target for malicious hackers.[41] Presumably, a great deal of resources will be required to defend the filter against those who wish to subvert, control or misuse it—a possibility that is discussed in the next sub-section.

### B. A Copyright Filter Could be Intentionally Misused for Censorship Purposes

At present, content filtering is used by a number of governments to censor Internet content, so as to control the flow of information. According to Reporters Without Borders' annual report, "Internet Enemies," China, Iran, North Korea, Syria and Cuba are among the nations who most tightly control the flow of information using some form of active content

[40] While it is possible to create a filter that would not slow traffic on a network, that filter would not be able to block or degrade traffic in real-time. Such a filter would "clone" the traffic that passed through it and then send that data elsewhere for analysis, without interrupting the flow of the original traffic. This data could then be analyzed and action could be taken after the fact (infringement notices could be sent to users, etc.).

[41] "Big Vulnerabilities in China's Mandatory Filtering Software," *Ars Technica*, June 1, 2009 (http://arstechnica.com/tech-policy/news/2009/06/big-vulnerabilities-in-chinas-mandatory-filtering-software.ars).

filtering.[42] This list also lists two additional "democracies under surveillance": Australia and South Korea.[43] In 2008, a minister in the ruling Australian Labour Party attempted to institute a nationwide filtering mandate with the aim of blocking content relating to child pornography. In late 2008 and early 2009, a series of leaked "blacklists" revealed that the filter would also block several legal sites with no relation to child pornography, including sites hosted by Wikipedia, the online encyclopedia.[44] As a result, the filtering mandate has yet to be enforced, though members of the Labour party have continued to push for implementation, most recently discussing plans to use the filter to prevent any Australian, regardless of age, from accessing video game content that has been deemed inappropriate for someone who is 15-years of age or younger.[45] Meanwhile, in South Korea, the government uses filtering technology to block access to a number of sites containing political content, most commonly sites that are deemed to have a "pro-North Korean" agenda.[46]

The timeliest example, however, comes from Iran, where filtering technologies are now being used to block access to popular websites like Twitter, YouTube and Facebook, in the wake of widespread political unrest. According to a report that appeared in *The Wall Street Journal*,[47] Iran's communications system, which serves an estimated 23 million Internet users, was designed to allow government surveillance of any and all online communications that originate or terminate in Iran. As the Iranian government maintains a monopoly on telecommunications

[42] "Internet Enemies" (2009 edition), Reporters Without Borders, April 9, 2009 (http://www.rsf.org/IMG/pdf/Internet_enemies_2009_2_.pdf).
[43] *Ibid.*
[44] "Rudd's internet blacklist includes dentist, kennel, tuckshop," *The Courier Mail*, March 20, 2009 (http://www.news.com.au/couriermail/story/0,23739,25214413-953,00.html).
[45] "Web filters to censor video games," *The Sydney Morning Herald*, June 25, 2009 (http://www.smh.com.au/digital-life/games/web-filters-to-censor-video-games-20090625-cxrx.html).
[46] "South Korean Internet Censorship," *1StopKorea.Com*, April 2007 (http://1stopkorea.com/index.htm?korean-internet-censorship.htm~mainframe).
[47] "Iran's Web Spying Aided by Western Technology," The Wall Street Journal, June 22, 2009 (http://online.wsj.com/article/SB124562668777335653.html).

services, the goal of omnipresent Internet surveillance was easily met, as it simply required the deployment of filters, powered by Deep Packet Inspection (DPI) technology, at "a single choke point" in the government network, through which all inbound and outbound traffic passes.[48] Apparently, the Iranian government first installed this hardware for the purported purpose of blocking pornography, citing "lawful intercept"—an internationally-recognized concept that "relates to intercepting data for the purposes of combating terrorism, child pornography, drug trafficking and other criminal activities carried out online."[49] This example illustrates that the act of filtering is a slippery slope. While filtering technology might be deployed to serve a legitimate purpose—be it to stem the flow of child pornography or copyrighted content—when placed in the wrong hands, that same technology can become a highly effective instrument of private or governmental censorship. It is for this reason that the Open Internet Coalition urged Congress to convene hearings to address the use of these technologies domestically, warning that policymakers "must fully understand the implications of wide deployment of deep packet inspection technology [in order to make] decisions to prevent its misuse in the United States".[50]

## III. THE IMPLEMENTATION OF COPYRIGHT FILTERS WILL RESULT IN A TECHNOLOGICAL ARMS RACE

Even if we could somehow design a filter that was 100 percent accurate, users would actively work to circumvent that filter, in order to access the content of their choice without exposing themselves to liability. The end result would be a technological arms race whereby users attempt to evade the filter while those maintaining the filter attempt to thwart those evasion attempts. As is the case in most arms races, this is a cat-and-mouse game where everyone loses:

[48] *Ibid.*
[49] *Ibid.*
[50] See Open Internet Coalition's letter to Congress, June 29, 2009 (http://www.openinternetcoalition.org/files/OIC_DPI_Iran_062909.pdf).

network protocols used by all users will be slower and less efficient and filters will become even more costly and less effective than they are today. Given that service providers are likely to pass on any additional costs to the consumer, not only will these networks be slower but subscription prices will also be higher.

A similar, if brief, arms race was observed during the FCC's proceeding regarding Comcast's blocking of the BitTorrent protocol. As the legal process wound its way through the FCC, the BitTorrent development and user community did not stand idly by. Rather, the community took matters into its own hands, devising a number of different workarounds that allowed users to bypass Comcast's traffic analysis filter.[51] Unfortunately, while these methods rendered Comcast's content management less effective, they did so at the cost of the BitTorrent protocol's efficiency.[52]

Another arms race is currently unfolding as a result of the Iranian government's politically motivated web filtering regime. Soon after the Iranian government started blocking websites in an attempt to censor news relating to the June 2009 election protests, web users outside of Iran made unfiltered connections available to users inside Iran, through the use of web proxies and Tor bridges.[53] A few weeks after the protests began, a group of software developers announced the upcoming release of Haystack, a software package for the Windows, Macintosh and Unix operating systems that was specifically designed to circumvent the Iranian

[51] "Beating Comcast's Sandvine On Linux With Iptables," *Slashdot*, June 30, 2008 (http://tech.slashdot.org/tech/08/06/30/0249249.shtml).
[52] "As Expected, BitTorrent Providers Planning To Route Around Comcast Barrier," *TechDirt*, February 18, 2008 (http://www.techdirt.com/articles/20080215/171450267.shtml).
[53] These techniques allow users to tunnel their traffic through gateways, so as to disguise that traffic's point of origin. See "Help Protestors in Iran: Run a Tor Bridge or a Tor Relay," Electronic Frontier Foundation, June 29, 2009 (http://www.eff.org/deeplinks/2009/06/help-protesters-iran-run-tor-relays-bridges).

government's web filters.[54] In contrast to previous methods, Haystack will significantly reduce the amount of technical expertise required to subvert Iran's filters, allowing any Iranian with basic computer literacy skills to access the web freely.

Meanwhile, in France, hackers have announced plans to release custom router firmware known as "the HADOPI router," named after French President Nicholas Sarkozy's famed "HADOPI" law, which would have required French ISPs to kick users off of their networks after receiving three allegations from copyright holders that that user had engaged in acts of online infringement (the HADOPI law was rejected by the French Constitutional Council).[55] The HADOPI router software allows a user to route his or her traffic at random through other nearby connections, effectively making that user's traffic untraceable by those who would enforce the HADOPI law.[56]

If an ISP attempts to implement a network-wide copyright filter, we will likely see a similar arms race, albeit on a far larger scale. There are a number of different ways that users and developers might subvert a copyright filter, thereby instigating an arms race. These methods are discussed in detail below.

### A. Encryption

Encryption is the science by which communications are scrambled so that an observer who views that communications data—for instance, an ISP—cannot determine what its contents are. To a party without the encryption "key"—the Rosetta Stone that allows encrypted

[54] "Haystack Anti-Censorship Tool Specifically For Users in Iran, to Launch Soon," *Boing Boing*, July 6, 2009 (http://www.boingboing.net/2009/07/06/haystackanticensor.html).
[55] "Top Legal Body Strikes Down Anti-Piracy Law," *France 24*, June 10, 2009 (http://www.france24.com/en/20090610-top-legal-body-strikes-down-anti-piracy-law-hadopi-constitutional-council-internet-france).
[56] "French Hackers Unveil the HADOPI Router: Cracks Nearby WiFi and Makes Your Traffic Traceable to Your Neighbors," *Boing Boing*, July 10, 2009 (http://www.boingboing.net/2009/07/10/french-hackers-unvei.html).

communications to be decrypted—the data appears as little more than a random string of characters. Encrypting a communications stream makes it very difficult for an observer to not only view the data but to also determine that data's type (music, video, etc.) or protocol (HTTP, BitTorrent, etc.). This holds the potential to circumvent both content and traffic analysis filters, as neither type of filter will be able to make a determination when presented with an encrypted data stream. Unfortunately, encryption also has drawbacks—most notably, it increases the amount of computation required on both sides of a communication, subsequently increasing the time and cost associated with that transfer. It is for this very reason that even very large corporations like Google (which offers equipment that is used for security and privacy purposes) have been hesitant to enable encryption by default in their products.[57]

Encryption has long been used in conjunction with many Internet-based applications, and is increasingly being used with many more. HTTP, the protocol associated with most browser-based web traffic, has been available in a secure form since at least 2000.[58] The BitTorrent file transfer protocol first saw an implementation of header encryption in 2005[59] and in 2006, a standardized, full-protocol "Message Stream Encryption" method was devised.[60] Transport Layer Security, meanwhile, provides a security and encryption specification that can be used in conjunction with a wide variety of protocols and is available as an easily integrated software package on many platforms.[61]

---

[57] See "Gmail to go with HTTPS by Default," *WindowsITPro*, June 17, 2009 (http://windowsitpro.com/Articles/Index.cfm?ArticleID=102316).
[58] See Internet Engineering Task Force, "Request for Comment 2818, HTTP Over TLS," May 2000 (http://tools.ietf.org/html/rfc2818).
[59] See IPP2P, "News," April 2006 (http://www.ipp2p.org/news_en.html).
[60] See "Message Stream Encryption Specification," January 27, 2006 (http://www.azureuswiki.com/index.php?title=Message_Stream_Encryption&oldid=4197).
[61] See Internet Engineering Task Force, "Request for Comment 5246, The Transport Layer Security (TLS) Protocol Version 1.2," August 2008 (http://tools.ietf.org/html/rfc5426).

Encryption, for all its benefits, does not always render a communication fully opaque. For instance, even if a filter is unable to make a direct determination regarding the nature of the content or the protocol used, it might still be able to analyze the size, sequencing, and timing of data packets, in an attempt to make a protocol determination.[62] In this way, a filter may still be able to determine with some degree of reliability that a user is using a particular protocol, though it would not be able to determine if that protocol was being used to send a licensed movie, an infringing song, or an open-source software package.[63] As such, even if a filter were able to glean information from an encrypted data stream, it would still face limitations that would likely result in either the blocking of legal content, the allowing of illegal content or both. Furthermore, such analysis can also be easily circumvented.

### B. Protocol Obfuscation

When filters target specific protocols (for example, BitTorrent), that protocol can be "obfuscated" in order to make it more difficult to detect both the protocol and what type of content is being carried by that protocol. This process generally does not require the sharing of keys (as is the case with encryption); rather, obfuscation merely rearranges data in order to create a more complicated layout that requires additional analysis if that data is to be decoded. If a user used protocol obfuscation on a copyright filtered network, the operator of that network would have to look more closely at that user's communications. This would cause the ISP to incur greater costs, would lower the efficiency of the network and would further compromise the privacy of the user, who in most cases would be engaging in only lawful activity in the first place.

[62] The particulars of such traffic and connection pattern analysis methods are discussed in this paper's technical section.

[63] "Company Cracks BitTorrent Protocol Encryption and Introduces Tracker Whitelists," *TorrentFreak*, April 27, 2007 (http://torrentfreak.com/company-cracks-bittorrent-protocol-encryption-and-introduces-tracker-whitelists/).

Protocol obfuscation can take many forms. The form of obfuscation used would likely vary depending on the type of filter and the particulars of the protocol in question. For instance, if a filter attempted to identify a protocol based on the size of the packets,[64] then an effective obfuscation technique might be to pad the data or packet headers with random amounts of extraneous blank space, in order to bypass the filter. While methods like data padding are fairly generic, others have been devised for use with specific protocols and the filters that detect those protocols. Some BitTorrent clients, for example, offer a "lazy bitfield" option, which intermittently sends phony messages suggesting that a piece of data is either incomplete or unavailable—a method devised specifically to evade certain filtering methods favored by cable ISPs.[65] Other, even more sophisticated methods masquerade targeted protocols as other data types, including email and web traffic.[66]

As filters advance in sophistication, so will the methods used for purposes of obfuscation. Filters that identify protocols by analyzing the timing between packets can be evaded by randomly altering the time at which the host sends data—a method that generally does not require the receiver to be aware of the change. Filters that work by correlating the use of multiple users who connect to the same host can be evaded by rerouting data through additional hosts.[67] In all cases, each additional level of complexity employed by a filter will result in more sophisticated methods of data obfuscation. This arms race will inevitably result in more data on the network, more resources being used by hosts attempting to communicate, greater resource

[64] This method, referred to as connection or traffic pattern analysis, is discussed in this paper's technical section.
[65] For an explanation of the use of this option to "[a]lways send a faked incomplete bitfield, and instead send remaining completed piece info via Have messages", see Azureus Wiki, "Command Line Options" (http://www.azureuswiki.com/index.php/Commandline_options). Certain ISP (i.e. Cablevision's Optimum Online) networks seem to block peer seeding via "complete" bitfield message filtering."
[66] See "A Robust Data Obfuscation Approach for Privacy Preserving Collaborative Filtering" by Rupa Parameswaran (http://etd.gatech.edu/theses/available/etd-05082006-193521/unrestricted/parameswaran_rupa_200608_PhD.pdf).
[67] See Tor website (http://www.torproject.org).

requirements and higher costs to those trying to filter and more invasive analysis of end-user communications.

## IV. THE RAMIFICATIONS OF THE ARMS RACE

While the arms race between filter architects and end users will result in inefficiencies on both ends, history suggests that end-users will have a slight advantage. This is due to the fact that there is an entire field of study devoted to ensuring that users can communicate freely without interlopers being able to read or alter their communications. As Paul Kocher, President and Chief Scientist at Cryptography Research, Inc. once put it at a NIST Physical Security Testing Workshop, "Moore's Law favors the cryptographer."[68]

Effective, efficient, and even government-approved[69] encryption standards are commonplace, have low (though not insignificant) cost barriers to use and are impractical or impossible for ISPs to circumvent. Other data obfuscation techniques, meanwhile, require relatively simple changes on the user's end, yet require a far more detailed inspection of data, more computing resources, and a greater volume of sample data on the ISP's end. In many cases, users will need only to upgrade or tweak the settings of the software on their home PCs in order to evade the newest, most expensive filtering technology.[70] Considering that the filtering arms race will result in increased costs and decreased performance both for service providers and users, we can conclude that the cost of filtering will far outweigh the benefits—especially when

[68] See "The Increasing Complexity and Need for Validation," NIST, September 26, 2005 (http://csrc.nist.gov/groups/STM/cmvp/documents/fips140-3/physec/presentations/physecpre02.pdf).
[69] See National Institute for Standards and Technology, "Announcing the Advanced Encryption Standard (AES)," November 26, 2001 (http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf).
[70] For examples of high-cost digital protection schemes that were quickly circumvented by users, see "Quick Guide: Cracked DRM Systems," *BBC News*, August 27, 2007 (http://news.bbc.co.uk/2/hi/technology/6944830.stm).

one considers the primary aims of the National Broadband Plan, which are to increase the speed and adoption of broadband services while decreasing costs.[71]

If service providers and content owners choose to go down this path, they might soon find themselves unable to identify the vast majority of data on their network as the result of users migrating to encrypted systems. In such a world, where service providers would be unable to make an accurate determination regarding the type of data traversing their networks, ISPs might decide to block or delay any and all data that cannot be identified. While this presents a predicament in and of itself, the problem will be amplified if users react by masquerading their targeted data as commonly encrypted web traffic (for example, email or traffic related to online commerce). If this happens, service providers will find themselves in an unenviable position: they will be forced to choose between blocking or delaying legitimate traffic and allowing through the very data they were attempting to filter in the first place.

In such a world, where providers might elect to block or degrade all unidentified or unreadable traffic, users will also be forced to make a difficult choice, one between privacy and access. If they choose to read their email over an encrypted connection, as nearly every service provider recommends,[72] they may find that their traffic is degraded or blocked. If users choose not to utilize encryption, they will be subjecting their data and communications to inspection by their ISP, content companies and anyone else who is able to view that traffic—including users in

[71] See "FCC Launches Development of National Broadband Plan**,**" April, 8, 2009 (http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-289900A1.pdf).
[72] See Comcast Customer Central, "Configuring Outlook Express to send and receive email while traveling" (http://customer.comcast.com/Pages/FAQViewer.aspx?Guid=0a2e36ac-2737-412b-b0ab-d2bf6a0fe54d).

the same office, on the same wireless network, or elsewhere along the data's path.[73] This choice, between security and privacy, is one that Internet users should not be forced to make.

[73] See "Using Wireless Technology Securely," US-CERT, available at (http://www.us-cert.gov/reading_room/Wireless-Security.pdf).

# 4. Economic Analysis

While a full economic analysis of the theoretical impact of copyright filtering is outside of the scope of this paper, there are certain issues regarding the economics of filtering which cannot be ignored. These issues are briefly addressed in the sections below.

## I. WHO WILL PAY FOR COPYRIGHT FILTERING?

The most obvious economic question that filtering raises is a pragmatic one. Who will pay for the deployment, maintenance and upkeep of the hardware, software and personnel that will be required to implement copyright filtering at the ISP level? Regardless of the technology used, ISPs will be required to deploy dedicated hardware devices on their network, which we can only assume will be purchased or rented from a hardware vendor. We can also assume that a certain number of manpower hours will be required to maintain this hardware and any associated software and keep it operating at maximum efficiency. And when this equipment needs to be upgraded, maintained or replaced, there will be an additional, recurring cost.

As this hardware and software will sit on the ISP's network, one might naturally assume that the ISP will cover the cost associated with filtering its network. However, seeing how there is no clear benefit and several harms to ISPs, there is no reason why a service provider should feel compelled to cover the costs associated with copyright filtering. Perhaps, as is the case with the Communications Assistance for Law Enforcement Act (CALEA), the party requesting the examination of communications data would foot the bill.[74] In the case of law enforcement

[74] See the Communications Assistance for Law Enforcement Act (CALEA) (http://en.wikisource.org/wiki/Communications_Assistance_for_Law_Enforcement_Act_of_1994).

wiretapping, that party is the United States government. In this case, that party would be the content industry.

In addition to the costs associated with filtering, ISPs have a number of non-monetary incentives to not engage in copyright filtering on their networks. Some of these incentives are legal in nature and will be discussed at length in the legal section of this paper. Others are purely technical. As we have seen, copyright filters will significantly slow any network on which they are installed, resulting in a decrease in the quality of the customer experience. This will place the ISP who engages in copyright filtering at a competitive disadvantage *vis-à-vis* its competitors.

At least one U.S. ISP has recognized that the drawbacks of copyright filtering far outweigh any perceived benefits. In January 2008, Verizon Communications publicly stated that it has no intention to install copyright filters on its network. "From a business perspective, we really don't want to assume the role of being police on the Internet," Verizon executive vice president Tom Tauke told the audience at a Washington Internet policy conference.[75] "We are leery of using these technologies on our networks."

Other U.S. ISPs, however, have expressed a strong interest in copyright filtering. At the 2008 Consumer Electronics Show (CES), AT&T senior vice president, external and legal affairs, James Cicconi revealed that AT&T was considering the possibility of installing copyright filters on its network and had engaged in discussions with technology companies, content companies like NBC Universal and content industry trade groups like the RIAA and the Motion Picture Association of America (MPAA).[76] This prompted Columbia Law School professor Tim Wu to

[75] "Verizon: We Don't Want to Play Copyright Cop on Our Network," *CNET*, January 30, 2008 (http://news.cnet.com/8301-10784_3-9861402-7.html).
[76] "AT&T and Other ISPs May be Getting Ready to Filter," *The New York Times*, January 8, 2008 (http://bits.blogs.nytimes.com/2008/01/08/att-and-other-isps-may-be-getting-ready-to-filter/).

write an opinion piece for *Slate* magazine, entitled "Has AT&T Lost its Mind?".[77] "…[T]he bizarre twist is that the proposal is such a bad idea that it would be not just a disservice to the public but probably a disaster for AT&T itself," Wu wrote in the article. "Has AT&T, after 122 years in business, simply lost its mind?"

Despite the lack of any clear motive on the part of the ISPs, some have floated theories as to why service providers might be amenable to go along with the content industry's copyright filtering agenda. In the case of AT&T, it is likely that the decision was motivated by the company's other, non-Internet offerings, namely its U-Verse TV television service. Since AT&T also acts as Multichannel Video Programming Distributor (MVPD),[78] copyright filtering would provide a means by which the company could demonstrate its commitment to copyright enforcement, thereby currying favor with the studios from whom it must license video content.

In the case of Comcast, a company that did not filter for copyright but which did block the BitTorrent protocol using a traffic inspection method of the type described earlier in this paper,[79] it was suggested that the company was simply hoping to boot high-bandwidth users off of its network, in the hope of increasing speeds on its oversubscribed network without having to invest in additional network capacity.[80] As has been noted elsewhere in this paper, this act earned Comcast a strong rebuke from the FCC.[81]

---

[77] "Has AT&T Lost its Mind?," *Slate*, January 16, 2008 (http://www.slate.com/id/2182152).
[78] The term "MVPD" includes cable television providers, direct broadcast satellite providers and wireline video providers, including those who sell IPTV services.
[79] See section 2.II.A. of this paper.
[80] See "Formal Complaint of Free Press and Public Knowledge Against Comcast Corporation for Secretly Degrading Peer-to-Peer Applications," November 1, 2007 (http://www.publicknowledge.org/pdf/fp_pk_comcast_complaint.pdf).
[81] See the FCC's Memorandum Opinion and Order, August 1, 2008 (http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-08-183A1.pdf).

Some have even gone further to suggest that these ISPs have an additional incentive to block video online for competitive reasons.[82] As was mentioned earlier in this paper, as a result of technological limitations, filters will be overinclusive. This means that a great deal of lawful video content--mashups, parodies, satire and non-copyrighted works--is likely to be blocked by any copyright filter that is implemented by an ISP. From a user's perspective, the end result will be that a significant amount of video content is no longer available online. This will, in effect, make watching online video a less attractive prospect for the user, who might turn to alternative sources for video entertainment as a result. This could offer an advantage to any MVPD service that seeks to compete with online video.

At present, the prospect of free streaming and downloadable online video clearly offers an attractive alternative to users who are used to paying an average monthly fee of $71 per month for cable television[83]—especially when one considers that the average monthly cost for a broadband Internet connection is only $39.[84] Indeed, one recent study found that 77 percent of Internet users watch online video and that that 20 percent of those viewers watch less programming on a television set as a direct result of their online viewing habits.[85]

Internet video distribution also provides an attractive proposition for content creators, by allowing those creators to beam content more directly to the consumer, cutting out the MVPD middleman while retaining the ability to derive revenue from that content through advertising. Indeed, some forms of programming are already commanding higher ad rates online than on

[82] "AT&T TOS Update Shows Network Management's True Colors," Public Knowledge Blog, September 11, 2008 (http://www.publicknowledge.org/node/1736).
[83] "Study: Average Cable TV Bill Is $71 Per Month," *Multichannel News*, April 16, 2009 (http://www.multichannel.com/article/196364-Study_Average_Cable_TV_Bill_Is_71_Per_Month.php).
[84] See "Home Broadband Adoption 2009," Pew Internet and American Life Project, June 17, 2009 (http://www.pewinternet.org/Reports/2009/10-Home-Broadband-Adoption-2009.aspx).
[85] "Magid Video Futures 2009: Opportunities in Online Video," Frank N. Magid Associates, Inc., June 2009 (http://www.magid.com/metacafe.pdf).

television. The popular prime-time television program "The Simpsons," for example, has a higher Cost Per Thousand (CPM) price on the streaming video site Hulu than it does on network television.[86]

In this light, it becomes clear why ISPs and their shareholders might want to invest in copyright filtering, especially considering how many service providers now offer video services either alongside Internet access or as part of a package containing Internet, telephony and video services (e.g. so-called "Triple Play" offerings). In fact, late last year, the RIAA announced that it was working with "major ISPs" to kick subscribers suspected of engaging in infringement off of those ISP's networks.[87]

## II. COPYRIGHT FILTERING HOLDS THE POTENTIAL TO DISRUPT THE INTERNET ECONOMY, OUR MOST PROMISING ENGINE FOR ECONOMIC GROWTH

As we will discuss in the following section, the Internet economy is a delicate ecosystem, consisting of a number of different layers of hardware, software and service providers, all working in tandem. This is a fact of which Congress was acutely aware during the process of crafting the Digital Millennium Copyright Act (DMCA).[88] Any change made to how the Internet operates will have a ripple effect that will be felt widely, at all levels of the ecosystem. Tinkering with the basic paradigm of Internet access is extremely unwise, especially at a time when the Internet's ability to promote economic growth and ability to provide citizens with economic opportunity is needed more than ever.

---

[86] "D'oh! The Simpsons Worth More on Hulu Than on FOX," *NewTeeVee*, June 29, 2009 (http://newteevee.com/2009/06/25/doh-simpsons-worth-more-on-hulu-than-on-fox/).
[87] "Music Industry to Abandon Mass Suits," *The Wall Street Journal*, December 19, 2008 (http://online.wsj.com/article/SB122966038836021137.html).
[88] See section 5.II. of this paper, "Copyright Filtering Will Undermine the Safe Harbor Provisions Granted to ISPs Under the Digital Millennium Copyright Act (DMCA)".

As was discussed in the technological analysis section of this paper, one likely consequence of copyright filters is that they will inject delay into commonplace transactions on the Internet. Without an actual implementation of a copyright filter on a U.S. ISP's network to analyze, it is difficult to predict just how disruptive these delays will be. Given that a copyright filter will likely have to a) stop multiple packets traversing the network b) analyze those packets c) cross-reference the analyzed data against a database of "known" or "protected" copyrighted works d) make a determination as to how to flag those packets (e.g. allow/delay/discard) and e) act on whatever flag has been assigned to the packets in question, it is clear that for each transaction online, the filter will be required to process a great deal of information. As such, we can only assume that there will indeed be a delay and that the delay will be noticeable, at least for certain types of traffic.[89]

A delay of a few seconds or even a few milliseconds might not seem like much, but in realm of Internet use, even such minor delays can have staggering effects. We know for a fact that Internet users with faster connections, "spend more time online, do more things, and do them more often" than users with slower connections.[90] What's more, studies have clearly demonstrated that delays of even a few milliseconds can result in appreciable changes in user behavior. One study found that 100 ms of additional load time translates into a 1 percent drop in sales for Amazon.com,[91] while another study found that 500 ms of delay resulted in a 20 percent drop in both traffic and revenue for the Google search engine.[92] These statistics attest that users are acutely aware of speed differences when using the Internet and will modify their online

[89] Unless the filter does not block, delay or degrade traffic in real-time (see 3.I.D.).
[90] "The Broadband Difference: How online behavior changes with high-speed Internet connections," Pew Internet & American Life Project, June 23, 2002 (http://www.pewinternet.org/Reports/2002/The-Broadband-Difference-How-online-behavior-changes-with-highspeed-Internet-connections.aspx).
[91] "Faster is Better," *Dries Buytaert*, January 6, 2009 (http://buytaert.net/faster-is-better).
[92] "Marissa Mayer at Web 2.0," *Geeking With Greg*, November 9, 2006 (http://glinden.blogspot.com/2006/11/marissa-mayer-at-web-20.html).

behavior accordingly.[93] So it stands to reason that not only will users potentially spend less time on the Internet, do fewer things online and do them less frequently but also that Internet commerce and advertising will be adversely impacted as a result.

Further complicating the matter is the plain fact that regardless of the technology used, any Internet filtering scheme is likely to create greater delays for certain specific types of traffic. As an example, let's consider an Internet filter that is designed with certain built-in exemptions. Such a filter might scrutinize traffic flowing to and from YouTube.com—a site where material suspected of infringing on copyrights is commonly found—but not traffic traveling to and from Hulu.com—a studio-owned streaming video site where infringing content is rarely found.

This, of course, raises a number of network neutrality[94] concerns, many of which were discussed in the context of the Comcast/BitTorrent proceeding.[95] For example, if BitTorrent is more heavily scrutinized than iTunes, thereby creating a greater delay for BitTorrent users, will major label artists be given a competitive advantage over independent artists who distribute their work using BitTorrent? If websites hosting open-source software are treated as suspect while Microsoft's website is given a speed pass, will independent software developers and the open-source community suffer? If copyright filters were instituted on the networks of major American ISPs, would the next YouTube-like service even launch, knowing full well that its traffic would be delayed and that its service would be adversely affected, resulting in a less than optimal user

[93] For additional information on the relationship between speed and user behavior see "Speed Matters for Google Web Search," Google, Inc., June 22, 2009 (http://code.google.com/speed/files/delayexp.pdf) and "The Effect of Network Delay and Media on User Perceptions of Web Resources," *Behaviour & Information Technology*, 2000, Vol. 19, No. 6, 427-439.

[94] For more information on the issue of network neutrality, see http://www.publicknowledge.org/issues/network-neutrality.

[95] See "Formal Complaint of Free Press and Public Knowledge Against Comcast Corporation for Secretly Degrading Peer-to-Peer Applications," November 1, 2007 (http://www.publicknowledge.org/pdf/fp_pk_comcast_complaint.pdf).

experience (i.e. slow, choppy video)? Depending on their design, copyright filters hold the potential to pick winners and losers on the Internet—a fact that could have serious and far-reaching economic consequences.

Even if we ignore the effects of delay, copyright filters will still harm innovators and businesses. In interfering with in-transit Internet traffic, copyright filters will inevitably cause some legal software and protocols to malfunction. A product like Red Lambda's Integrity, for example, would block all traffic that uses a protocol that is associated with P2P use,[96] thereby interfering with legal P2P services like the Octoshape P2P video player used by CNN.com.[97] While developers might be able to modify their software and services to circumvent the filter, they will have to bear the costs of doing so themselves. Of course, in order for a developer to make such modifications effectively, the filter's architects will have to operate transparently, providing legitimate developers with documentation regarding the filtering algorithms and targeting criteria. This, however, presents the architect of the filter with a Catch 22: if this documentation is made available, it will inevitably make its way into the hands of those who aid infringers, resulting in a less effective filter; if the documentation is not made available, legitimate businesses will be harmed. Therefore, it follows that copyright filters will be either ineffective or will stifle innovation. What's more, all of these problems will be compounded if the architects of the filter are involved in an arms race like the one described in the previous section, making iterative tweaks to the filter and subsequently causing recurring, persistent problems for developers.

[96] See Integrity features as listed on Red Lambda website (http://redlambda.com/integrity.php?p=features).
[97] "CNN Video Streaming Tech Raises Questions," *Ars Technica*, February 10, 2009 (http://arstechnica.com/web/news/2009/02/cnn-p2p-video-streaming-tech-raises-questions.ars).

While the long-term effects that copyright filtering would have on the American economy remain unknown, this much is clear: copyright filtering will be a disruptive force on the Internet, one that will be felt by users, online service providers, businesses and investors. The Internet has become a powerful engine for economic growth due to its open architecture, which encourages investment, by providing a platform whereby anyone can build an application or service that can compete on equal footing with incumbent offerings. It is crucial that we allow the Internet to remain a level playing field. Otherwise, we will discourage investment, competition, new entrants and innovation, dealing a blow to our long-term economic health and international competitiveness.

# 5. Legal Analysis

The laws that currently govern speech and copyright online are the result of a delicate balance that was reached between a number of competing interests and considerations. These laws have allowed ISPs and Online Service Providers (OSPs) to foster a dynamic ecosystem of speech, entertainment, and debate without fear of legal liability for the actions of users online.

If the federal government imposes a mandate requiring ISPs to filter traffic in search of copyright violations or otherwise encourages or condones the act of filtering, this delicate balance will be disrupted. The protections that innovators and citizens have relied on in building the Internet ecosystem would be substantially weakened if not outright eliminated. More troubling, such legislation would undermine current legal and constitutional protections for speech and free expression. Below, we discuss the existing laws with which a copyright filtering mandate would conflict.

## I. MANDATORY COPYRIGHT FILTERS WOULD IMPOSE AN UNCONSTITUTIONAL BURDEN ON FREE EXPRESSION, CONTRARY TO THE PRINCIPLES OF COPYRIGHT LAW

Regardless of how it is accomplished, content filtering will interfere with citizens' ability to communicate online. No matter what type of technology is used, no filter will be capable of determining if a communication is authorized, fair use or infringing. As a result, copyright filters will always be overinclusive when blocking online speech. Thus, copyright filters will inevitably interfere with and suppress completely legal forms of speech and expression online. While such interference is worrisome when practiced by a private company, it may well be unconstitutional if imposed by government mandate.

Under the law, a copyright owner is never granted complete control over a copyrighted

work.[98] Limitations on and exceptions from copyright keep copyright law from conflicting with the First Amendment rights of citizens. Fair use and other limitations such as the requirement of originality, the idea/expression dichotomy, and the doctrine of thin copyright[99] allow for free expression including protected forms of speech like parody and criticism. As the Supreme Court has explained, "Copyright … does not impermissibly restrict free speech, for it grants the author an exclusive right only to the specific form of expression … and it allows for 'fair use' even of the expression itself."[100]

A fair use of a copyrighted work is therefore protected free speech. Proponents of copyright filtering suggest that the filtering of copyrighted material would be a straightforward and entirely legal process. However, the nuances of copyright law make distinguishing between a lawful and infringing use of a piece of copyrighted content challenging even for courts. As such, no filtering technology, no matter how advanced, will ever be able to make fair use determinations with 100 percent accuracy. And as courts have held, no law should "allow any copyright owner, through a combination of contractual terms and technological measures, to repeal the fair use doctrine."[101] Because a use labeled by a filter as "unauthorized" is not necessarily an illegal use, no automated system should be allowed to give the desires of a copyright holder priority over the First Amendment. Furthermore, a government mandate requiring copyright filtering might also run afoul of case law governing prior restraint, which states that speech cannot be preemptively censored, except in extenuating circumstances (i.e.

[98] Exclusive rights "do[] not give a copyright holder control over all uses." *Fortnightly Corp. v. United Artists*, 392 U.S. 390, 393 (1963).
[99] "Thin copyright" refers to the lower level of copyright protection for compilations of fact. The doctrine was first explained by the Supreme Court in *Feist Publications, Inc. v. Rural Telephone Service Company, Inc.* 499 U.S. 340 (1991).
[100] *Eldred v. Ashcroft*, 537 U.S. 186, 197 (2003).
[101] *Chamberlain Group, Inc. v. Skylink Technologies, Inc.*, 381 F.3d 1178, 1202 (Fed. Cir. 2004).

issues of national security).[102]

As discussed in the technological analysis section of this paper, copyright filters are not only overinclusive—they will also be underinclusive. As we have seen, today's filters have proven ineffective at stopping even moderately sophisticated infringers. And there is no indication that future filters will be able to overcome this technological shortcoming. As a result, a government mandate that ISPs filter for copyright infringement would substantially interfere with free speech and non-copyrighted content while providing little or no benefit to copyright holders to justify this constriction of First Amendment rights.

By virtue of their technological limitations, copyright filters will inevitably block some protected forms of speech while allowing some infringement. The simultaneously overinclusive and underinclusive nature of filters will ultimately result in unconstitutional restrictions on free speech, just as other government attempts to block access to objectionable web sites have in the past.[103] Therefore, we can conclude that a government copyright filtering scheme that fails to specifically accommodate fair use is unlikely to be constitutional.

## II. COPYRIGHT FILTERING COULD UNDERMINE THE SAFE HARBOR PROVISIONS GRANTED TO ISPS UNDER THE DIGITAL MILLENNIUM COPYRIGHT ACT (DMCA)

Currently, the Digital Millennium Copyright Act offers ISPs certain safe harbors from copyright infringement liability for activity which occurs over their networks.[104] It is not hard to understand the value of this protection to an ISP, and to service providers in general. If ISPs could be held liable for every infringement perpetrated by their customers, they would be exposed to a flood of lawsuits from the sound recording, motion picture and software industries.

[102] *Near v. Minnesota*, 283 U.S. 697 (1931).
[103] *CDT v. Pappert*, 337 F. Supp. 2d 606 (E.D. Pa. 2004).
[104] 17 U.S.C. § 512(a).

And unlike with individual infringers, ISPs are easy to find and likely to have the resources to pay a judgment and are therefore, attractive targets for lawsuits (suing individuals has proven to be a time- and cost-intensive activity that is unlikely to result in a significant financial recovery[105]).

These ISP safe harbors under the DMCA are not the result of a happy accident. During the process of drafting of the DMCA, Congress was convinced of two points—first, that effectively monitoring every bit that travels through an ISP's network was not technologically feasible[106] and second, that if ISPs were forced to monitor their networks, that requirement would effectively cripple the nascent Internet.[107]

Congress ultimately determined that limiting the liability of ISPs would ensure "that the efficiency of the Internet [would] continue to improve and that the variety and quality of services on the Internet [would] continue to expand."[108] Essentially, Congress decided to allow ISPs to develop viable business models, fearful that the specter of liability could, in the words of one commentator, "virtually halt American participation in the emerging information society."[109] Today's Internet is a testament to the wisdom of that decision. It is hard to imagine any common activity on the Internet today—be it making a comment on a news site, posting a photograph on a social networking site, or perusing indexed links on a search engine site—that has not benefited from the DMCA safe harbors.

In order to qualify for the safe harbor provisions, an ISP has to meet a certain set of

---

[105] See, *e.g.*, *Capitol Records Inc. v. Thomas*, 579 F. Supp. 2d 1210. (D. Minn. 2008).
[106] See *Copyright Protection on the Internet*: *Hearings on H.R. 2441 – The NII Copyright Protection Act of 1995 Before The Subcommittee on Courts and Intellectual Property of The House Committee on the Judiciary*, 104th Cong. (1996) (statements of Stephen M. Heaton, General Counsel and Secretary, CompuServe Incorporated and Scott Purcell, Commercial Internet eXchange Association).
[107] *Ibid.*
[108] S. Rep. No. 105-190, at 8 (1998)
[109] *Copyright Protection on the Internet*: *Hearings on H.R. 2441 – The NII Copyright Protection Act of 1995 Before The Subcommittee on Courts and Intellectual Property of The House Committee on the Judiciary*, 104th Cong. (1996) (statements of Stephen M. Heaton, General Counsel and Secretary, CompuServe Incorporated).

requirements, which are outlined in the DMCA. The first of these requirements is that all material that travels over the network must be "initiated by or at the direction of a person other than the service provider."[110] In layman's terms, this means that the ISP must act as a simple conduit through which data travels—any transaction that takes place on the network must be initiated by someone at the edge of the network—either a client or a server—but may not be initiated by someone who sits in the middle of the network. If an ISP implemented a copyright filter, that ISP could, arguably, become an active participant in the chain of transmission, making decisions about what data to transmit and what data to discard. Instead of merely passing a bit of data along, the ISP would inspect, categorize, and possibly interrupt, delay or discard that bit of data. In so doing, the ISP could potentially be disqualified from the DMCA's safe harbor protections and therefore, would be exposed to liability for any infringement that takes place over its network.

Filtering similarly jeopardizes the second requirement that an ISP must meet in order to qualify for DMCA safe harbor protection. This requirement states that the transmission of data must occur "through an automatic technical process without selection of the material by the service provider."[111] This use of the word "selection" is not further clarified in the statute, creating an open question as to what degree of filtering would qualify as "selection". Depending on the level of sophistication of the prioritization process, certain packet management techniques could be interpreted as constituting a "selection" of material. If an ISP could be described as actively selecting what material is allowed to travel over its network, its safe harbor protection would be jeopardized. Furthermore, this selection process could quickly rise to the level of an "editorial function" (*i.e.* choosing to prioritize data from a preferred source over a non-preferred

[110] 17 U.S.C. § 512(a)(1).
[111] 17 U.S.C. § 512(a)(2).

source), which is likely to disqualify an ISP from DMCA safe harbor protection.[112]

In establishing the safe harbor provisions in the DMCA, Congress recognized that in order for the Internet to thrive, ISPs would need to be shielded from liability for their users' actions. Given that any filtering technology that actively inspects the content that flows over an ISP's network may eliminate that ISP's ability to claim protection under the DMCA's safe harbors, it should be clear that any government mandate that requires or condones copyright filtering by ISPs would undermine the safe harbors on which today's Internet was built.[113]

## III. ISPS THAT ENGAGE IN PACKET INSPECTION RISK VIOLATING THE ELECTRONIC COMMUNICATIONS PRIVACY ACT (ECPA)

The Electronic Communications Privacy Act (ECPA) extended traditional wiretap privacy protections to include communications between computers.[114] In passing this Act, Congress recognized that private communications that take place on a computer network deserve the same sorts of protections traditionally granted to communications that take place over voice networks. Congress also extended these protections to other types of private communications, including email.[115]

Specifically, ECPA prohibits the interception of electronic communications by those not party to the communication. Intercepting the packets of Internet users, whether for copyright filtering or other purposes, would certainly seem to implicate the statute. Even when law enforcement agencies in the course of their duties wish to intercept communications, the government must justify to a neutral third party why a specific intrusion of privacy is

[112] S. Rep. No. 105-190, at 42 (1998).
[113] "Has AT&T Lost its Mind?," *Slate*, January 16, 2008 (http://www.slate.com/id/2182152).
[114] See P. L. 99-508, 100 Stat.1848.
[115] See 18 U.S.C. §2701-11.

necessary.[116] While the standards of proof that the government must meet in order to intercept private communications vary, under no circumstance is the government authorized to examine all communications in the hope of stumbling upon evidence of illegal conduct.

A commercial enterprise possesses even less justification for warrantless and unwarranted surveillance of private communications.[117] Nor does it seem likely that any of the exceptions provided within ECPA allow an ISP to inspect all of its customers in order to pass judgment on the desirability or legality of their communications. While a provider may engage in activity "necessary incident to the provision of service, or to the protection of the rights or property of the provider of that service," this exception would not apply to an ISP seeking to enforce an outside party's rights. The exception exists to preserve the ability of a network to operate, not to give free reign to a carrier to act as arbiter of its users' content. In order to make use of this exception, a provider must show a "necessary incident," or at the very least, a "substantial nexus," between the monitoring and the protection of the provider's rights and property.[118] Nor would fine print or a clickthrough EULA seem to suffice for the purposes of consent to wiretapping. The necessary amount of notice for consent to wiretapping may be far higher than for consent in ordinary contract terms.[119]

Since such exceptions would likely not apply to ISP filtering, we are left with the fact that a filtering ISP would be unlawfully intercepting electronic communications between citizens, a

---

[116] See 18 U.S.C. § 2710 (15) and § 2711(2).
[117] Controversy surrounding ISP use of DPI for advertising purposes raised the question of ECPA violations by ISPs and their ad network partners. *See, e.g.*, Center for Democracy and Technology, *An Overview of the Federal Wiretap Act, Electronic Communications Privacy Act, and State Two-Party Consent Laws of Relevance to the NebuAd System and Other Uses of Internet Traffic Content from ISPs for Behavioral Advertising*, July 8, 2008 (http://www.cdt.org/privacy/20080708ISPtraffic.pdf) (commenting on the likelihood of ECPA violations by NebuAd and other DPI-based advertising systems).
[118] *Id.* at 5; Paul Ohm, *The Rise and Fall of Invasive ISP Surveillance*, U. Ill. L.R. (forthcoming 2009) (manuscript at 69-71), *available at* http://ssrn.com/abstract=1261344 (detailing the high bar providers must meet to avail themselves of the "protection of rights and property" exception).
[119] *Id.* at 72-73.

violation of the intricately constructed wiretap laws.

Allowing ISPs to filter all communications for potential copyright infringement would turn this finely calibrated system on its head. Congress would be permitting private parties to engage in privacy-violating activities that it elsewhere explicitly prevents private actors and government agencies alike from conducting. If ISPs were authorized to engage in copyright filtering, the end result would be a massive invasion of the privacy of all Internet users, carried out by those few, privileged gatekeepers through whom U.S. web traffic passes.

# 6. Conclusion

As we have seen, copyright filtering is, at its core, a technology that is ill-suited for use on U.S. broadband networks. Due to their inherently underinclusive and overinclusive nature, copyright filters will never be an effective solution to the problem of online copyright infringement and furthermore, will harm free speech, online privacy and the speed and affordability of broadband services. If implemented, a copyright filtering mandate would force ISPs to inspect all data transmitted by all Internet users, invading the privacy of hundreds of millions of citizens at the behest of one industry. This would likely have far-reaching consequences for the Internet ecosystem as a whole, disincentivizing investment, innovation and creativity and undermining the goals of the National Broadband Plan.

If we are serious about combating copyright infringement online, we should use the tools already at our disposal for identifying, trying and prosecuting infringers. The Digital Millennium Copyright Act (DMCA), for example, already contains robust provisions for copyright owners looking to have infringing content removed from the web and copyright law allows copyright holders to sue for statutory damages that are more than adequate—if not excessive—in many cases.

While existing law already provides copyright holders with a number of options for combating infringement in all its forms, the best solution for content companies will ultimately be one that has little to do with copyright enforcement. If entertainment companies and other content providers adapt to meet the needs of the digital economy, by providing consumers with access to the content that they want in the ways that they want, it is likely that copyright infringement will become a secondary concern for many of these companies. By exploring new,

innovative business models, the content industry can encourage consumers to purchase entertainment goods even when that same content is available for illegal download online. This fact is evidenced by increases in online music sales,[120] decreases in unlawful music file sharing[121] and strong sales of movie tickets[122] and high-definition physical video products like Blu-Ray discs,[123]—all despite the widespread availability of albums and full-length films on P2P file-sharing networks. Additionally, some have suggested that even if unlawful file sharing continues, artists and content companies could still be compensated if they are willing to adopt a more innovative licensing model, such as voluntary collective licensing.[124]

Ultimately, the content industry will have to work closely with technologists, innovators and policymakers to find solutions that make sense in a digital economy. Until then, however, we should not rush blindly to implement ineffective, dangerous solutions like copyright filtering. To do so would be to endanger free speech, to imperil user privacy and to undermine our efforts to deliver the promise of broadband connectivity to all Americans.

[120] According to *the New York Times*, online sales of music rose by 27 percent between 2007 and 2008. See "Music Sales Fell in 2008, but Climbed on the Web," *The New York Times*, December 31, 2008 (http://www.nytimes.com/2009/01/01/arts/music/01indu.html).
[121] "Report: UK Filesharing Drops, Even Among Teens," *Ars Technica*, July 13, 2009 (http://arstechnica.com/media/news/2009/07/report-more-uk-users-going-the-legal-route-for-music.ars).
[122] "What Piracy? Movie Biz Sees Record Box Office in 2008," *Ars Technica*, January 5, 2008 (http://arstechnica.com/media/news/2009/01/what-piracy-movie-biz-sees-record-box-office-in-2008.ars).
[123] According to the Digital Entertainment Group, sales of Blu-Ray products tripled between 2007 and 2008. See "DVD Sales Down 5.5% in '08," *Variety*, January 7, 2009 (http://www.variety.com/article/VR1117998174.html).
[124] "A Better Way Forward: Voluntary Collective Licensing of Music File Sharing," The Electronic Frontier Foundation, April 2008 (http://www.eff.org/wp/better-way-forward-voluntary-collective-licensing-music-file-sharing).

June 10, 2009

Acting Chairman Michael J. Copps
Federal Communications Commission
445 12th Street SW
Washington, DC 20554

Re: *Special Access for Price Rates for Price Cap Local Exchange Carriers*
WC Docket No. 05-25

Dear Acting Chairman Copps:

Fair prices and conditions for special access are essential for the future of broadband deployment and a vibrant wireless market. The above captioned proceeding therefore has greater significance than as an "industry food fight" between sophisticated competitors. Resolution of the special access proceeding will have significant impact on the future of the economic recovery, job creation, health care, education technology, and e-commerce. For this reason, the Commission must act quickly to impose rules that will ensure reasonable rates for special access, rather than delay this four year old proceeding with yet another data request. If the Commission does decide it must refresh the record for the second time since beginning this proceeding in 2005, it should issue a short, targeted request enforced by the power to compel full responses from incumbent and competitive providers.

The Commission's decision in 2000 to rely primarily on the market alone to develop competition in special access has resulted in higher prices and market dominance by incumbent providers such as Qwest, AT&T and Verizon. This, in turn, has seriously impeded the development of a competitive market in either broadband or wireless. The time has come for the FCC to act expeditiously to reregulate the special access market, and to reject any further requests for delay.

Specifically, the Commission should reject the effort by the US Telecom Association (USTA) to require yet another data request. The Commission began this proceeding four years ago based on disturbing reports that competition had failed to emerge in the special access market. *See, Special Access Rates for Price Cap Local Exchange Carriers and AT*

*&T Corp. Petition for Rulemaking to Reform Regulation of Incumbent Local Exchange Carrier Rates for Interstate Special Access Services*, 20 FCCRcd 1994 (2005). The Commission refreshed the record on this proceeding in 2007**,** *Parties Asked to Refresh Record in the Special Access Notice of Proposed Rulemaking*, 22 FCCRcd 13352 (2007). Nothing has happened to alleviate the competitive situation. To the contrary, all evidence in the record indicates that the market has grown less competitive since 2005.

The FCC's 2005 *Notice of Proposed Rulemaking* recognized that data provided by businesses and competitors to AT&T indicated a need to reexamine the Commission's previous deregulatory approach. Empirical data provided since then consistently show that, due to the lack of an effective system to measure and regulate competition in the special access market, incumbents continue to overcharge their competitors for wholesale access to high-speed digital lines. Steep special access rates and unreasonable terms and conditions prevent competitor carriers from building infrastructure and offering services necessary for reasonable competition, driving prices to consumers and delaying the deployment of broadband services in rural areas and the emergence of competing broadband providers even in more densely populated areas.

USTA's argument that the Commission cannot act unless it refreshes the record ***yet again***, with data from entities beyond the reach of Commission authority and precluded from disclosing terms subject to non-disclosure agreements (NDAs) should be dismissed as a transparent delaying tactic designed to prolong for as long as possible the extraction of monopoly rents by its members. To the extent the Commission requires more information, the simplest and most direct way to gather it is to require that USTA's members produce any further information required. *See, e.g.*, 47 U.S.C. §§211 (authority to order filing of contracts), 215 (authority to examine business dealings), 220 (authority to review financial records). If this information indicates that potential competitors have not been forthcoming about their services, the Commission has the power to compel responses from them as well.

As part of your efforts to improve the U.S. broadband market, we urge you to address this critical issue as quickly as possible. We believe that sufficient evidence of uncompetitive practices exists to drive the Commission to act now. If you find that a data request is necessary to gather additional information, we ask that it be short and clearly targeted, and that the Commission use its authority to compel incumbents to provide full, accurate, and timely responses to the inquiry.

Sincerely,

Harold Feld
Legal Director
Public Knowledge
1875 Connecticut Avenue, NW
Suite 650
Washington, DC 20009
(202) 518-0020

Andrew Jay Schwartzman
President
Media Access Project
1625 K Street, NW
Suite 1000
Washington, DC 20006
(202) 232-4300

Michael Calabrese
Director
Wireless Future Program
New America Foundation
1899 L St., N.W., Suite 400
Washington, DC 20036
(202) 986-2700

Amina Fazlullah
U.S. PIRG
218 D Street SE
Washington, DC 20003-1900
(202) 546-9707

Public Knowledge

**Statement of Gigi B. Sohn, President, Public Knowledge**
**Before the Federal Communications Commission**
**National Broadband Plan Workshop: The Role of Content in the Broadband Ecosystem**
**September 17, 2009**

I'd like to thank the Commission for inviting me to speak at today's workshop on the role of content in the broadband ecosystem. There's little doubt that online content produced by the entertainment industry serves as an important driver for both broadband adoption and utilization, and I expect that content to be the focus of most of my fellow panelists. In my comments, I hope to first provide a different perspective--one that highlights the important work being done by innovators, independent and amateur creators — and to then sound a note of caution. Any measures taken to protect content online must take into account all creators in the broadband ecosystem, must not obstruct the free flow of information online and must protect the privacy of end users.

Online content extends far beyond the boundaries of Hollywood. User generated content, in its many forms, has enriched the lives of Americans and rivals studio content in terms of popularity. In the online video market, for example, nearly 42 percent of all videos viewed online are hosted by YouTube.[1] Of the tens of billions of videos that YouTube streams each month, nearly one fourth of the top 100 most viewed videos are clips created by end users, not studios.[2]

As you might imagine, not only is user-generated content popular--it also serves a diverse set of needs. Amateur podcasters inform the public about events that are ignored or

[1] "U.S. Online Video Market Soars in July as Summer Vacation Drives Pickup in Entertainment and Leisure Activities Online," comScore, August 27, 2009 (http://www.comscore.com/Press_Events/Press_Releases/2009/8/U.S._Online_Video_Market_Soars_in_July_as_Summer_Vacation_Drives_Pickup_in_Entertainment_and_Leisure_Activities_Online).
[2] "YouTube's Most Popular Clips: Still Mostly Free," *All Things Digital*, August 12, 2009 (http://mediamemo.allthingsd.com/20090812/youtubes-most-popular-clips-still-mostly-ad-free/).

underreported on by the national news media.[3] Universities showcase their course materials and lectures via "OpenCourseWare" to members of the general public, many of whom would normally not possess the means to pursue higher education in a traditional setting.[4] Even Presidential candidates have harnessed the power of user-generated content, encouraging users to submit debate questions via YouTube in 2008.[5] And yes, many users, "vidders" and mashup artists use copyrighted clips to express themselves through commentary, criticism, and parody under the legal doctrine of fair use.

Meanwhile, innovative technology companies are empowering users to access and create the content of their choice in exciting new ways. Apple's iTunes provides easy access to the best of big studio content as well as user-created podcasts.[6] Boxee's "app box" allows users to automatically promote the media they enjoy to friends, putting "traditional" web video sources like Hulu, CNN, CBS, and Comedy Central on equal footing with user generated sources like YouTube, BitTorrent, and online video networks like TWIT.tv and Revision3.[7] Hardware manufacturers are embedding web interfaces into new HDTVs to allow easy-to-use online video streaming from the comfort of the consumer's couch.[8] Clearly, a great deal of creativity and innovation is taking place on the Internet, and up until now, it's been a level playing field. This innovation must be allowed to continue without undue restraints so that users can reap the

[3] "News Unfiltered: YouTube Embraces Citizen Journalism," *Ars Technica*, May 20, 2008 (http://arstechnica.com/old/content/2008/05/news-unfiltered-youtube-embraces-citizen-journalism.ars).
[4] "MIT's OpenCourseWare Project Continues Apace," *The Chronicle of Higher Education*, March 23, 2007 (http://chronicle.com/article/MIT-s-OpenCourseWare-Project/15958).
[5] "In Obama-McCain Race, YouTube Became a Serious Battleground for Presidential Politics," *U.S. News and World Report*, November 7, 2008 (http://www.usnews.com/articles/news/campaign-2008/2008/11/07/in-obama-mccain-race-youtube-became-a-serious-battleground-for-presidential-politics.html).
[6] *See* Apple Inc. website (http://www.apple.com/itunes/whats-on/).
[7] "Boxee Launches Update, 'App Box,' and Hulu RSS Support," *TUAW*, March 6, 2009 (http://www.tuaw.com/2009/03/06/boxee-launches-update-app-box-and-hulu-rss-support/).
[8] "Internet-Ready TVs Usher Web Into Living Room," *Wall Street Journal*, January 5, 2009 (http://online.wsj.com/article/SB123111603391052641.html).

benefits of increased choice, affordability and greater convenience and exercise their right to free speech and public discourse.

**Before Taking Action, the FCC Must Collect Empirical Data About Online Infringement and its Actual Effect on the Entertainment Industries**

Since assuming the office of Chairman of the Commission, Julius Genachowski has stated numerous times that he will ensure that decisions made by the FCC are data-driven.[9] With this in mind, the Commission should seek out independent empirical studies that quantify the extent of, and actual harm caused by, online copyright infringement. No policymaker can deliberate a solution unless she fully understands the problem. Though the content industry has cited studies in the past that purportedly address the financial losses resulting from online infringement, many of these industry-funded studies have proven unreliable. One widely cited study produced by research firm L.E.K. for the Motion Picture Association of America (MPAA), claimed that 44 percent of all Internet traffic on University networks was infringing in nature.[10] Three years after releasing the study, the MPAA was forced to admit that the 44 percent figure was inaccurate--the correct figure, according to the MPAA, was 15 percent.[11] EDUCAUSE, however, asserts that the correct figure is, in fact, 3 percent.[12] And while Congressman Arlen Specter asked the MPAA to provide its methodology for the study to Congress in 2006, the MPAA has yet to make this information available.[13]

---

[9] "The GigaOm Interview: FCC Chairman Julius Genachowski on Mobile, Broadband, iPhone and Innovation," *GigaOm*, August 3, 2009 (http://gigaom.com/2009/08/03/the-gigaom-interview-fcc-chair-julius-genachowski/).
[10] "MPAA Statement on Motion Picture Industry Losses Due to Piracy Among College Students," Motion Picture Association of America, January 22, 2008 (http://www.mpaa.org/press_releases/lek%20college%20student%20data_f.pdf).
[11] "Movie Industry Admits Error in Study on Campus Piracy," *The Chronicle of Higher Education*, January 22, 2008 (http://chronicle.com/blogPost/Movie-Industry-Admits-Error-in/3632).
[12] "MPAA Admits Piracy Study Flawed," *Copyright Advisory Network*, February 13, 2008 (http://librarycopyright.net/wordpress/?p=75).
[13] *Ibid.*

If the FCC is to address the matter of online copyright infringement, it must first seek out reliable, empirical studies produced by impartial third parties. In fact, many such studies are already available. In a Harvard Business School working paper on the topic of file-sharing and copyright, Felix Oberholzer-Gee and Koleman Strumpf embarked upon a literature review of studies published on the topic of online copyright infringement.[14] The study is attached to this statement. Among the studies they list, many differ from those cited by the entertainment industry and assert that the impact of infringement on sales of entertainment products is far less negative than is often acknowledged. Some of these studies even suggest that online file-sharing does not effect sales, or that online file-sharers are more likely to buy entertainment products than those who do not engage in file-sharing, which suggests a positive, rather than negative, relationship between file-sharing and sales.[15] The Commission should consider all available data when researching the impact that file-sharing has on the entertainment industry, including those studies cited by Oberholzer-Gee and Strumpf.

**Copyright Filters Are Not Suitable for Use on ISP Networks**

As the comments submitted in this proceeding attest, many in the entertainment industry are looking to automated copyright filtering technologies to solve the problems caused by the unlawful exchange of copyrighted content online.[16] While such technologies are being heralded as a fail-safe solution, they are, in fact, critically flawed. Simply put, a copyright filter is a blunt

[14] *See* "File Sharing and Copyright," Harvard Business School, May 15, 2009, p. 35-37 (attached) and "The Effect of File Sharing on Record Sales: An Emprical Analysis," Felix Oberholzer-Gee and Koleman Strumpf, December 12, 2006 (attached).
[15] *Ibid.*
[16] *See Joint Comments of American Federation of Television and Radio Artists, AFL-CIO, MPAA et al.*, In the Matter of A National Broadband Plan for Our Future, GN Docket No. 09-51 (all referenced comments below can be found in this docket unless otherwise noted); *Comments of Songwriters Guild of America*; *Comments of Arts+Labs* at 6 (urging the commission to "stand by its existing four principles, including the right to reasonably manage the networks" to among other things prevent copyright infringement); *Comments of the Entertainment Software Association* at 4 (urging the Commission to clarify that the Internet Policy Statement permits network operators to use "reasonable network management to combat online piracy"); and *Comments of Walt Disney Company*.

instrument and is not suitable for use on an ISP network. Public Knowledge has already submitted for the record a whitepaper that these flaws in great detail. I'd like to briefly address a few of the reasons why copyright filters are not a viable solution.[17]

A close analysis of copyright filtering technology reveals that ISP-level filtering will be ineffective at best and at worst, harmful to the network, end users and the goals of the National Broadband Plan. By virtue of their design, filters will be both overinclusive and underinclusive--that is to say that they will fail to identify all unlawful uses of content while blocking lawful uses.[18] As a result, automated filtering will block, delay or degrade lawful content, including but not limited to content that makes fair use of copyrighted content for the purposes of parody, satire and critical commentary. In so doing, copyright filters will discourage citizens from exercising their free speech rights online. While big studio content is important, we should recognize that user-generated content is an equally, important part of the broadband ecosystem and as such, it deserves an equal degree of protection. For this reason, copyright filters, which would automatically block all uses of copyrighted content online, including those aforementioned fair uses, are not appropriate tools to curb infringement.

*A. Copyright Filters Will Harm the Network and User Privacy*

Copyright filtering will also alter the behavior of data networks on a fundamental level, slowing down traffic, impeding the operation of high-latency applications and compromising the privacy of all Internet users.[19] In so doing, copyright filters will discourage investment in the Internet ecosystem, prevent innovators from developing exciting new applications, dissuade

[17] *See* "Forcing the Net Through a Sieve: Why Copyright Filtering is not a Viable Solution for U.S. ISPs," Public Knowledge, July 22, 2009 (http://www.publicknowledge.org/pdf/pk-filtering-whitepaper-200907.pdf).
[18] *Ibid.*, 7-24.
[19] *Ibid*., 25-37.

users from fully utilizing their broadband connections and raise the cost of access for consumers--all the while undermining some of the most important goals of the National Broadband Plan.[20]

*B. Copyright Filtering is Likely to Result in Unintended Consequences*

Copyright filtering at the ISP level could also result in unintended consequences that will likely result in greater harm than good to its proponents. Users will likely devise methods for circumventing the filter, in order to access the content of their choice without interference.[21] The methods that will likely be used--encryption and protocol obfuscation--will decrease the efficiency and speed of the network.[22] What's more, the architects of the filter will be forced to pour a tremendous amount of resources into research and development in order to thwart those users who seek to circumvent the filter, resulting in increased costs for both the user and provider.[23]

In addition, the core technology behind copyright filtering also holds the potential to be misused for purposes of censorship. The timeliest example of this sort of misuse comes from Iran, where content filtering technologies were used to block access to popular websites like Twitter, YouTube and Facebook, in the wake of widespread political unrest.[24] Apparently, the Iranian government first installed this hardware for the purported purpose of blocking pornography, citing "lawful intercept"—an internationally-recognized concept that "relates to intercepting data for the purposes of combating terrorism, child pornography, drug trafficking and other criminal activities carried out online."[25] This example illustrates that the act of filtering

[20] *See* "FCC Launches Development of National Broadband Plan," April, 8, 2009 (http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-289900A1.pdf).
[21] *Ibid*., 29-37.
[22] *Ibid*.
[23] *Ibid.*
[24] "Iran's Web Spying Aided by Western Technology," *The Wall Street Journal*, June 22, 2009 (http://online.wsj.com/article/SB124562668777335653.html).
[25] *Ibid.*

is a slippery slope. While filtering technology might be deployed to serve a legitimate purpose—be it to stem the flow of child pornography or illegally obtained copyrighted content—when placed in the wrong hands, that same technology can become a highly effective instrument of private or governmental censorship. Once such technology is installed at the ISP network level, it is all too easy to use that hardware for undemocratic ends.

*C. Copyright Filtering is Content Management, Not Network Management*

In the context of the National Broadband Plan, the entertainment industry is urging that copyright filters be considered a form of reasonable network management, and as such would not violate the FCC's broadband principles or other principles of openness. This characterization is misleading and obfuscates the manner in which filters operate. Copyright filters decide how data packets should be treated based on the content that they carry. As such, copyright filtering is a form of content management rather than network management and has no place in the National Broadband Plan.[26] Notwithstanding the question of whether the FCC has the authority to mandate copyright filters (discussed below), the Commission should not unduly burden ISPs with technology mandates and to the extent that the Commission addresses network management in the context of the National Broadband Plan, only those methods that are neutral and nondiscriminatory should be permitted.

**A Three Strikes Regime Would Allow Private Companies to Decide Who Should Have Access to the Internet**

Another blunt instrument that is being recommended by some in the entertainment industry is the so-called "graduated response" or "three strikes" regime.[27] Under this system, ISPs would be compelled to kick users off of their networks after receiving three notices from

[26] *Ibid.*, 5.
[27] "Digital Piracy Spreads, and Defies a Fix," *the New York Times*, April 6, 2009 (http://www.nytimes.com/2009/04/07/business/media/07piracy.html).

copyright holders indicating that a user was suspected of sharing content in an unlawful manner.[28] Never mind the industry's poor track record with regard to the accuracy of such notices:[29] three strikes lets the entertainment industry decide who gets kicked off the Internet based on a mere accusation of unlawful activity--without any judicial finding that infringement actually took place. Three strikes does not contemplate any judicial review and users would have no legal recourse or ability to challenge the accusations made. If we implemented such a regime devoid of due process in this country, we would be allowing private companies to cut citizens off from the most effective conduit for civic engagement, economic opportunity and education that is currently available, based purely on allegation. Indeed, the French Constitutional Council recently struck down a three strikes regime in that country, citing the Declaration of the Rights of Man and of the Citizen, a founding document of the French Revolution.[30] "…[W]hereas under section nine of the Declaration of 1789, every man is presumed innocent until he has been proven guilty, it follows that in principle the legislature does not establish a presumption of guilt in criminal matters," the Council wrote in its ruling.[31]

Needless to say, such a mandate, if instituted in this country, would be contrary to the goals of the National Broadband Plan.

[28] "Music Industry to Abandon Mass Suits," *Wall Street Journal*, December 19, 2008 (http://online.wsj.com/article/SB122966038836021137.html).
[29] "Flunking File Swappers: Inside the RIAA's Anti-P2P Machine," *Ars Technica*, May 14, 2008 (http://arstechnica.com/old/content/2008/05/flunking-file-swappers-inside-the-riaas-anti-p2p-machine.ars).
[30] "Top Legal Body Strikes Down Anti-Piracy Law," *France 24*, June 10, 2009 (http://www.france24.com/en/20090610-top-legal-body-strikes-down-anti-piracy-law-hadopi-constitutional-council-internet-france).
[31] *See Decision n° 2009-580 of June 10th 2009*, The Constitutional Council, June 10, 2009 (http://www.conseil-constitutionnel.fr/conseil-constitutionnel/root/bank/download/2009-580DC-2009_580dc.pdf).

**The FCC Does Not Possess the Authority to Mandate Either Copyright Filtering or a Three Strikes Regime**

The FCC has tried to step into the business of copyright regulation before. In its Broadcast Flag ruling,[32] the Commission required that hardware that connects to public networks contain technology designed to prevent some kinds of copying. The FCC's ruling was reversed by the D.C. Circuit, which held that the agency exceeded its authority when it adopted rules requiring digital media devices to prevent some unauthorized (but not necessarily illegal) copying.[33] Before that, the Commission was reversed when it tried to mandate that all broadcasters add video description information to their programming.[34] In both cases, the D.C. Circuit held that the FCC had overstepped the authority given to it by Congress.[35]

If the FCC were to mandate the use of copyright filters on ISP networks or a "three strikes regime," it would once again be stepping outside of its jurisdiction. Neither of these mandates would constitute the regulation of "communication by wire or radio".[36] Instead, both of these scenarios would transform the FCC into a copyright agency and would, in practice, limit rights of fair use. The FCC's authority to regulate "communication" does not give it the general

[32] Digital Broadcast Content Protection, *Report & Order & Further Notice of Proposed Rulemaking*, 18 FCC Rcd. 23,550 (2003).
[33] *American Library Ass'n. v. F.C.C.*, 406 F.3d 689 (D.C. Cir. 2005) (Commission exceeded its authority by requiring that devices respect the "broadcast flag").
[34] *Motion Picture Ass'n of Am. v. F.C.C.*, 309 F.3d 796 (D.C. Cir. 2002) (Commission exceeded its authority by implementing video description rules).
[35] *Motion Picture Ass'n.*, 309 F.3d at 801 ("An agency may not promulgate even reasonable regulations that claim a force of law without delegated authority from Congress."). Also, in *American Library Ass'n.*, 406 F.3d at 698, the court explains,

> *The FCC, like other federal agencies, 'literally has no power to act … unless and until Congress confers power upon it.' La. Pub. Serv. Comm'n v. FCC, 476 U.S. 355, 374 … (1986). The Commission 'has no constitutional or common law existence or authority, but only those authorities conferred upon it by Congress.' Michigan v. EPA, 268 F.3d 1075, 1081 (D.C. Cir. 2001). Hence, the FCC's power to promulgate legislative regulations is limited to the scope of the authority Congress has delegated to it. Id. (citing Bowen v. Georgetown Univ. Hosp., 488 U.S. 204, 208 … (1988)).*

[36] *American Library Ass'n.*, 406 F.3d at 703 ("The Federal Communications Commission may not lawfully exercise jurisdiction over activities that do not constitute communication by wire or radio.").

authority, absent an express delegation of power from Congress, to regulate content.[37] Additionally, under any mandated scheme involving copyright filtering or three strikes, disputes would undoubtedly arise hinging on whether a particular unauthorized use is in fact an illegal one. The FCC has no authority to hear or decide this kind of dispute, and cannot require that network operators become "copyright cops."[38] It has no power at all to regulate copyright absent an express delegation of power. As Sen. Patrick J. Leahy, Chairman of the Senate Judiciary Committee has written, Title 47 "grants…the FCC…no express authority…to address the complex issues of intellectual property matters[.]"[39]The Commission should not heed calls for it to again exceed its authority by instituting policies that would, in practice, constitute copyright law.

**Solutions and Remedies**

If we wish to prevent content from being shared unlawfully online, we should first look to those existing strategies and remedies that have proven successful. The most effective solution will always be the simplest one: for the entertainment industries to make content widely available online at a fair price and to allow users to access that content at the time of their choosing and on their device of choice. The current state of the market serves as a testament to this fact. Movie studios have found great success selling and renting films through services like Netflix,[40] music labels have seen consistent growth in the sales of online music via services like iTunes[41] and literary publishers have generated great excitement through their support for

[37] *Motion Picture Ass'n.*, 309 F.3d at 801.
[38] *Am. Library Ass'n*, 406 F.3d at 702 ("[T]he Commission may not invoke its ancillary jurisdiction under Title I to regulate matters outside of the compass of communication by wire or radio.").
[39] Letter from Sen. Patrick J. Leahy, Chairman, Senate Judiciary Committee and Rep. F. James Sensenbrenner, Jr., Chairman, House Committee on the Judiciary, *et al*. to Michael K. Powell, Chairman, FCC (Sept. 9, 2002).
[40] "Netflix Boss Plots Life After the DVD," *Wall Street Journal*, June 23, 2009 (http://online.wsj.com/article/SB124570665631638633.html).
[41] "iTunes Sells 25% of All Music in the US, 69% of Digital," *Ars Technica*, August 18, 2009 (http://arstechnica.com/apple/news/2009/08/itunes-sells-25-of-all-music-in-the-us-69-of-digital.ars).

devices like the Amazon Kindle.[42] Those studios who have threatened to hold back content unless certain, unproven protection mechanisms are put in place, as some did in the Broadcast Flag proceeding, only encourage users who seek that content to obtain it through unlawful means.[43] I applaud those content providers who have made their products available online at a fair price and hope that they will continue to expand their offerings as others follow suit. Within existing law, content providers also have access to a number of legal remedies for combating the unlawful use of their content online. The Digital Millennium Copyright Act's notice-and-takedown system provides a framework for removing infringing content that is hosted online. Some ISPs have agreed to pass infringement notices on to end users; a practice that the entertainment industry itself has found discourages users from engaging in unlawful conduct in the vast majority of cases.[44] And as always, the entertainment industry has the ability to combat the unlawful sharing of content by cutting that content off at the source, by targeting large-scale infringers, hard-goods counterfeiters and other unlawful providers who make infringing content widely available.

## Conclusion

Ultimately, our goal should be to encourage the creation, sale and use of content online, which in turn, will further the goals of the National Broadband Plan. Digital entertainment is a valuable driver for both broadband adoption and utilization. By fully harnessing its potential, we can increase access to broadband for all Americans, educate users on how to make full use of

[42] "The Lessons From the Kindle's Success," *the New York Times*, August 12, 2008 (http://bits.blogs.nytimes.com/2008/08/12/the-lessons-from-the-kindles-success/).
[43] *Comments of Viacom*, FCC Docket 02-230, at 1.
[44] "RIAA President: No Talk of Blacklisting File Sharers," *Cnet*, December 19, 2008 (http://news.cnet.com/8301-1023_3-10127313-93.html).

their connections, and improve the lives of citizens by delivering economic, civic and educational opportunities directly to their personal computers.

# ATTACHMENT A

# FILE SHARING AND COPYRIGHT

# FELIX OBERHOLZER-GEE AND KOLEMAN STRUMPF

HARVARD | BUSINESS | SCHOOL

# File-Sharing and Copyright

**Felix Oberholzer-Gee**
**Koleman Strumpf**

## Working Paper

**09-132**

# File-Sharing and Copyright[1]

Felix Oberholzer-Gee
Harvard University
foberholzer@hbs.edu

Koleman Strumpf
University of Kansas
cigar@ku.edu

May 15, 2009

## 1. Introduction

The advent of file-sharing technology has allowed consumers to copy music, books, video games and other protected works on an unprecedented scale at minimal cost. In this essay, we ask whether the new technology has undermined the incentives of authors and entertainment companies to create, market and distribute new works. While the empirical evidence of the effect of file sharing on sales is mixed, many studies conclude that music piracy can perhaps explain as much as one fifth of the recent decline in industry sales. A displacement of sales alone, however, is not sufficient to conclude that authors have weaker incentives to create new works. File sharing also influences the markets for concerts, electronics and communications infrastructure. For example, the technology increased concert prices, enticing artists to tour more often and, ultimately, raising their overall income.

Data on the supply of new works are consistent with our argument that file sharing did not discourage authors and publishers.[2] The publication of new books rose by 66% over the 2002-2007 period. Since 2000, the annual release of new music albums has more than doubled, and worldwide feature film production is up by more than 30%

[1] We would like to thank Josh Lerner, Scott Stern, Amitay Alter and participants in the NBER's 2009 Innovation Policy and the Economy Conference in Washington, D.C., for helpful comments.

[2] Copyright refers to a complex bundle of rights that includes the rights of authors (composers, lyricists) and publishers (for a detailed description of these contracts, see Towse 1999; Passman 2000). Throughout this essay, we use the term somewhat loosely, referring to all legal protections – including, for instance, the "neighboring rights" of performers – that encourage the creation, production, marketing, and distribution of works. Also, we neglect the tensions that exist in copyright between artist and publisher interests (see Towse, 1999; Gayer and Shy, 2006.)

since 2003. At the same time, empirical research in file sharing documents that consumer welfare increased substantially due to the new technology.

Over the past 200 years, most countries evolved their copyright regimes in one direction only: lawmakers repeatedly strengthened the legal protections of authors and publishers, raising prices for the general public and discouraging consumption.[3] Seen against this backdrop, file sharing is a unique experiment that considerably weakened copyright protections. While file sharing disrupted some traditional business models in the creative industries, foremost in music, in our reading of the evidence there is little to suggest that the new technology has discouraged artistic production. Weaker copyright protection, it seems, has benefited society.

In this essay, we discuss the currently available research that sheds light on the effects of file sharing, particularly in music where its effects have been most pronounced. We start by describing the new technology and how consumers are using it. Section 4 reviews the evidence that file sharing reduces the profitability of creating and selling new works. We discuss the importance of complements to original works in Section 5 and describe the artistic and corporate response to file sharing in section 6. The concluding section offers policy implications.

## 2. File-Sharing and Copyright

In setting copyright terms, lawmakers trade off the increased incentives to create protected works and the higher prices that consumers face when books, movies, and recordings must not be copied freely (Landes and Posner, 1989). As this description suggests, the lawmakers' task is a challenging one. Setting copyright terms in a manner that benefits society requires an answer to two questions. First, we need to know how much weaker the incentives to create new works would be in a regime with more

[3] In the United States, as elsewhere, the degree of protection has steadily expanded, from the modest Copyright Act of 1790, which offered 14 years of protection with a renewal period of 14 years, to the legislation passed in 1831 (28 years), 1909 (renewal extended to 28 years), 1976 (50 years after the author's death), 1992 (automatic renewal), and 1998 (70 years).

constrained copyright. Second, and equally important, is the question how producers would respond to weaker incentives. Would they offer fewer works? Or perhaps works of lesser quality? In this essay, we discuss what we know about these questions, using the advent of file-sharing as our example for a technology that considerably weakened copyright protection for music, movies, books and video games.

Weaker copyright is unambiguously desirable if it does not lessen the incentives of artists and entertainment companies to produce new works. To appreciate the impact of file sharing, we first need to know whether the technology did in fact reduce the profitability of creating, marketing, and distributing new works. Of course, we know that millions of consumers share billions of files without compensating artists or entertainment companies. But the fact that file sharing is popular tells us little about the impact of the technology on industry profits. At a price close to zero, many consumers will download music and movies that they would not have bought at current prices. This issue is likely to be important. In a sample of 5,600 consumers who were willing to share their iPod listening statistics, the average player held a collection of over 3,500 songs (Lamere, 2006). A full 64% of these songs had never been played, making it unlikely that these consumers would have paid much for a good portion of the music they owned. While it is difficult to say how representative this sample is, there is no doubt that trade groups such as the Business Software Alliance vastly exaggerate the impact of file sharing on industry profitability when they treat every pirated copy as a lost sale (Economist, 2005). The demand for titles is not completely price inelastic.

Weaker property rights can undermine industry profitability if consumers who would have purchased a recording obtain a free copy instead. The critical question is then whether consumers perceive protected and freely shared works as close substitutes. As the name suggests, substitutes are products that meet similar consumer demands. For two substitute goods, a price decline for one leads to a decline in the demand for the other.[4] For example, if we allowed mash-up artists to freely copy parts of an original song, consumers who regard the derivative work as a close substitute would be less likely

[4] A classic example is butter and margarine.

to buy the original.[5] However, if consumers learned to better appreciate the original through the mash-up, demand for the original work might actually increase. In this case, the two versions of the song are complements, two goods for which a decrease in the price of one leads to an increase in the demand for the other. A well-known example for two complements is music and iPods. As file-sharing eroded the effective price of music for a large group of consumers, demand for mp3-players soared, allowing Apple to benefit from consumers' increased willingness-to-pay for its line of products.[6]

In practice, it is often surprisingly difficult to predict whether new products and technologies are complements or substitutes. As a result, we can often not be sure how changes in copyright will influence demand and industry profitability. The entertainment industry's history provides many examples of the difficulties involved in distinguishing substitutes, unrelated products, and complements. Music companies fought the introduction of radio in the 1920s, fearing the new medium would provide close substitutes to buying records. Since that time, the numerous attempts to bribe radio stations in the hopes of influencing playlists suggest the industry has come to see radio as an important complement to recordings (Coase, 1979). Similarly, the entertainment industry battled home taping[7] and the introduction of the VCR, arguing the new technology "is to the American film producer and the American public as the Boston strangler is to the woman home alone" (Valenti, 1982). Once the Supreme Court decided to protect technologies like the VCR, it did not take the industry long to discover that selling videotapes (and now DVDs) presents a major business opportunity.

Similar uncertainty surrounds file-sharing technology today. Some argue that protected works and copies on file-sharing networks are substitutes because consumers who would have bought the copyrighted version now choose to download a free copy instead. Others see protected works and copies on file-sharing networks as largely

[5] A mash-up is a song created out of pieces of two or more songs, usually by overlaying the vocal track of one song over the music track of another.
[6] Leung (2008) estimates that piracy contributes 20% to iPod sales.
[7] Stanley M. Gortikov, president of the Recording Industry Association of America (RIAA), explained in hearings before a House committee on 14 April 1982: "I'm scared, and so is my industry. Changing technology today is threatening to destroy the value of our copyrights and the vitality of the music industry. Our nemesis is home taping."

unrelated because they believe that file sharers are mostly consumers who are not willing to pay $10 for Taylor Swift's latest release. Finally, protected works and copies on file sharing networks are complements if consumers rely on the new technology to discover CDs or DVDs they want to purchase. These views need not be mutually exclusive. In a recent survey among file sharers, we found some support for all three conjectures (Oberholzer-Gee and Strumpf, 2005). 65% of respondents acknowledged they did not buy an album because they had downloaded it. An even larger group (80%) claimed they bought at least one album because they sampled it first on a file-sharing network. Fortunately, there is now a body of research that studies in a more systematic manner whether copyright protected works and copies on file-sharing networks are complements or substitutes. We will discuss this literature in section 4 of this essay.

Even if a weakened copyright regime turned out to reduce industry profitability, it is not obvious whether a decline in profits would undermine the incentives to create, market and distribute artistic works. Two considerations seem particularly important. First, as copyright weakens, the effective price of music, movies, and books falls and consumer willingness-to-pay for complements increases. If artists derive income from these complements as well, the overall incentives to produce new works might not decline. For instance, as music becomes effectively available for free, the price of concerts, a complement to music, is likely to rise, and artists who earn income from concerts might not be hurt by a decline in music sales (Krueger, 2005; Mortimer and Sorensen, 2005). Similarly, authors might be better able to supplement their income from books through speaking tours if many more readers are familiar with their writings.[8]

A second reason that a decline in industry profitability might not hurt artistic production has to do with artist motivations. The remuneration of artistic talent differs from other types of labor in at least two important respects. On the one hand, artists often enjoy what they do, suggesting they might continue being creative even when the monetary incentives to do so become weaker. In addition, artists receive a significant portion of their remuneration not in monetary form – many of them enjoy fame,

[8] Author Cory Doctorow, for instance, says:"I really feel like my problem isn't piracy. It's obscurity." (Rich, 2009).

admiration, social status, and free beer in bars – suggesting a reduction in monetary incentives might possibly have a reduced impact on the quantity and quality of artistic production.

There is no doubt that file sharing substantially weakened the protection of copyrighted works. Yet, as our discussion shows, the outcome of this experiment is far from certain. Three conditions need to hold for less-certain rights to undermine the incentives for artistic production: original works and copies on file-sharing networks must be reasonably close substitutes; artists and the entertainment industry must not be able to shift from previous sources of income to the (similarly profitable) sale of complements; and falling incomes must be an important-enough motivator for artists to reduce production. Only if all three conditions hold will file sharing hurt social welfare.

It might seem curious to some of our readers that we do not consider the welfare of artists and entertainment companies in our calculus. Our approach, however, reflects the original intent of copyright protection, which was conceived not as a welfare program for authors but to encourage the creation of new works. We know that stronger copyright protection can increase the market value of companies.[9] But these gains are a mechanism to raise social welfare, not the intended consequence.[10]

## 3. A Brief History of File-Sharing

To better understand the impact of file-sharing technology on copyright protection, it is useful to review the basics of file-sharing. In this section, we will also describe recent changes in technology and review the most significant legal challenges that companies providing file-sharing software faced to date.

File sharing relies on computers forming networks to allow the transfer of data. Each computer (or node) may agree to share some files, and file-sharing software allows

[9] Baker and Cunningham (2006), for example, estimate that a statue broadening copyright adds up to $39 million to the market capitalization of a typical firm.
[10] To frame our discussion in terms of efficiency (Pareto improvements), we argue that the relevant benchmark is the welfare of groups in a situation without copyright.

users to search for and download files from other computers in the network. Individual nodes are called clients if they request information, servers if they fulfill requests, and peers if they do both.

Shawn Fanning, an 18-year-old student at Boston's Northeastern University, started the file-sharing revolution when he released Napster in June of 1999 (table 1 provides a timeline). The software first allowed the freshman to trade music with his dorm mates. Prior to Napster, fans used search engines such as Lycos and music websites to download music. However, searching for files was cumbersome because the available music indices were often out of date. Many sites offered more broken links than hits. Napster was novel in that it maintained a central, dynamic index of all available files. This index was updated every time a user logged on or off. Thanks to its user-friendly interface and seemingly unlimited supply of music, the service gained 30 million users in its first year.

Napster's legal difficulties started not long after its initial release. In December 1999, the Recording Industry Association of America (RIAA) sued Napster for contributory and vicarious copyright infringement (*A&M Records, Inc. v. Napster, Inc.*, 239 F.3d 1004 (9th Cir. 2001).[11] Two years and one appeal later, the Ninth Circuit Court of Appeals ruled against Napster, arguing the service's central directory of files gave its makers knowledge of and the ability to control user infringement. Unable to filter files from the network, Napster shut down. However, putting Napster out of business proved easier than ending file sharing. Most Napster users simply switched to second-generation peer-to-peer services, and they were joined by millions of file-sharing novices. Three major networks eventually developed: eDonkey; FastTrack, a network used by KaZaA and Grokster; and Gnutella, an open-source network for clients such as Bearshare, Gnucleus, LimeWire, and Morpheus.

The Circuit Court decision also proved influential for the further technological development of file-sharing services. If peer-to-peer companies had no direct knowledge

[11] A party is liable for contributory infringement if it knows of the infringing activity and materially contributes to it. Vicarious infringement occurs when the indirect infringer benefits financially from the infringement.

of and control over infringing activities, many in the industry believed, file-sharing services might be protected by the Supreme Court's Betamax decision (*Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417 (1984). The decision holds that companies are not liable for customers' acts of copyright infringement if their technology is capable of substantial non-infringing uses. In the Sony case, the Court estimated that about 9% of VCR recordings were of TV shows that consumers had taped to watch at a later time and that the producers of these shows did not object to time shifting. This was sufficient to shield Sony from liability.

Convinced that peer-peer technology had substantial legal uses – for example the exchange of files that were in the public domain or the sharing of documents within a company – second-generation file-sharing services eliminated centralized indices (Oberholzer-Gee, 2006). In these systems, users first connect to a single peer using a specific internet protocol. The peer then tells the software about other peers in the network, in effect decentralizing the search and download processes and making it impossible for peer-to-peer companies to know whether users trade copyrighted materials. At first, this strategy appeared to work. When the RIAA sued the makers of Grokster, a branded version of KaZaA, and Morpheus for contributory and vicarious copyright infringement, District Court Judge Stephen V. Wilson ruled that the two companies could not be held liable (*MGM Studios, Inc. v. Grokster, Ltd.,* 259 F. Supp. 2d 1029 (D. Cal. 2003): "All Napster search traffic went through, and relied upon, Napster… [But] when users search for and initiate transfers of files using the Grokster client, they do so without any information being transmitted to or through any computers owned or controlled by Grokster… If either defendant closed their doors and deactivated all computers within their control, users of their products could continue sharing files with little or no interruption."

The entertainment companies appealed the case, but the circuit court upheld the earlier decision, affirming that decentralized peer-to-peer systems met the standard set in Sony. On June 27, 2005, however, the Supreme Court overturned the Ninth Circuit, sending the case back to the district court for further consideration (*MGM Studios, Inc. v. Grokster, Ltd.*, 545 U.S. 913 (2005): "Because substantial evidence supports MGM on all

elements, summary judgment for the respondents was in error. On remand, reconsideration of MGM's summary judgment motion will be in order." The justices ruled that a company that distributed a device "with the object of promoting its use to infringe copyright" could be liable for the resulting illegal acts. The Court argued that Grokster and Morpheus had wanted to be the next Napster, showing their goal was to induce copyright infringement.

The Supreme Court's decision led most peer-to-peer companies to settle with the entertainment industry. An exception was LimeWire, a service that continues to operate to this day. LimeWire argues that its software provides substantial legal uses. For example, the company operates a digital music store that offers 500,000 songs, many of them from independent bands. And LimeWire insists that it does not induce consumers to infringe copyright. The RIAA filed a lawsuit against LimeWire in April 2006. At the time of this writing, no decision has been reached, leaving open the question whether services such as LimeWire are protected by the standard set in Sony. At the same time, several second-generation file-sharing programs such as Ares Galaxy and eMule, the former eDonkey, continue to be available as open-source software.

While pursuing the developers of peer-to-peer software in the courts, the RIAA also started suing P2P users who shared a large number of files—typically more than 1,000 tracks—starting in 2003. The association hoped its actions would help reverse the common view that file sharing was a legitimate activity. In a Pew Internet & American Life Project survey in 2000, 78% of internet users who downloaded music did not think they were stealing. A majority of the general internet population held the same view (Lenhart and Fox, 2000). By the end of 2008, the industry had brought suits against more than 35,000 file sharers. Most cases were settled, typically for a few thousand dollars.

In a surprising shift in legal tactics, however, the RIAA announced in December 2008 that it had decided to drop its campaign against individual file sharers. Instead, the industry hoped to collaborate with internet service providers (ISPs) to stop the transfer of copyrighted materials. The trade group has worked out preliminary agreements with

major ISPs under which it will send an email to the provider when it finds that customers share copyright-protected files (McBride and Smith, 2008).

While the RIAA had some success putting peer-to-peer companies out of business, file-sharing technology continued to evolve. The most important technical advance was the emergence of BitTorrent. BitTorrent file requests differ from classic full-file HTTP requests in that the client makes many small data requests, similar to internet telephony which breaks voices into small packets of data. In addition, BitTorrent downloads follow a "rarest-first" order which ensures high availability of files across the network. To start the downloading process, users first obtain a torrent, a small file that contains metadata about the file to be downloaded and information about the tracker, the computer that coordinates the file distribution. Torrents are hosted by a fairly small number of websites. The Pirate Bay is probably the best-known among them. The torrent allows the client to connect to the tracker, from which it receives a list of peers that currently transfer pieces of the file. As more peers connect to a tracker, they form a swarm and begin to trade pieces with one another.

The advent of BitTorrent is significant for a number of reasons. First, the improved technology significantly reduces download times. While the user experience varies significantly, it has now become possible to download a feature film in less than two hours. Second, the technology forces users to share the parts of files that they already own while they download the remaining bits. This procedure reduces the opportunity to free-ride that plagued older P2P systems. The protocol also rewards users who contribute more generously, for instance by allowing faster downloads for those with greater upload capacity. Sharing digital files was always non-rivalrous because the original owner of a file retained his copy. But more efficient file distribution systems such as BitTorrent have now also succeeded in reducing the negative externalities that users impose on one another when they transfer files.

### a. Size of File-sharing Activity

Measuring the extent of file sharing is challenging (Karagiannis, 2003; Pasick, 2004). Initial studies relied on surveys to determine the number of users, but this

approach is flawed because respondents are likely to understate their participation in a potentially illegal activity. More worrisome, the level of understatement likely varies over time based on the legal climate and peer effects among teens. Surveys are also unreliable because it is difficult to survey a representative population of file sharers and due to recall issues.

A better approach involves identifying the packets traversing computer networks. These studies use special hardware to classify messages that are sent along networks by source, such as web (http) traffic, email, or file sharing. This approach is taxing because of the scale of the activity (ISPs typically handle many gigabits per second), the changes in the predominant protocol file-sharing protocol, and the recent move to encryption, which makes packets unreadable to unauthorized observers. Measurement studies employ three basic approaches to deal with these technical issues: flow monitors, deep-packet inspection, and direct interface with file sharing users.

Flow monitoring analyzes unidirectional sequences of packets from one IP address to another at the router level (Shalunov and Teitelbaum, 2001). This approach inspects packets in a rather shallow way, relying primarily on header information such as IP protocol and an examination of ports. Flow monitoring can analyze a large amount of traffic, at the risk of misclassifying some of it. A detailed flow analysis of Internet2, the U.S. high-speed network which primarily connects universities, is available at the weekly level back to 2003 (Internet2 Netflow Statistics, 2009). Figure 1 shows that file sharing traffic on Internet2 has roughly grown by a factor of ten – from about 1 terabyte to about 10 terabytes – from 2003 through 2009.[12] While this growth has been fairly steady, during 2003-2005 there were large traffic dips during late spring and early summer as well as smaller drops during Christmas. These drops in file-sharing activity reflect school vacations, periods during which college students, who are among the highest file sharing users, leave their high-speed campus internet connections.

The second type of evidence comes from deep packet inspection. Rather than relying just on the packet header, this approach considers characteristics of the payload

[12]Karagainni, et al (2004) employ a similar methodology in studying Tier 1 ISP traffic. They conclude that file sharing did not decline over the period 2003-2004.

itself (Allot Communications, 2007). Packet inspection is the most accurate method of identifying file sharing, but the technique requires extremely sophisticated equipment since huge amounts of data must be analyzed. The deep-packet inspection company Sandvine has been monitoring file-sharing trends for several years. The company's reports show that file sharing accounted for between forty and sixty percent of all bandwidth usage over 2002-2008 (Sandvine, 2002-2007 and 2008ab). CacheLogic, another deep-packet inspection company, finds similar trends in global file activity (Ferguson, 2006). Figure 2 shows the growing role of file sharing over 1999-2006. By 2006 sixty percent of all consumer internet traffic was due to file sharing, a majority of which was composed of video files.

The final approach to measuring file sharing comes from studying peer-to-peer networks directly. Observers use a modified version of file-sharing software to connect to a large number of users on the network. Direct observation can provide fine-grained information such as the identity of files. A difficulty with this approach is that direct observers need to monitor an ever-changing representative sample of networks. The leading practitioner is BigChampagne, a company which monitors individual search requests as well as the content of folders that users share. Figure 3 shows BigChampagne's count of the monthly number of U.S. file-sharing users from mid-2002 through mid-2006.[13] By the end of this period there were about seven million simultaneous users in the U.S. Unfortunately, more recent figures are not publicly available. As with the earlier data on file sharing traffic, there is evidence of secular growth as well as reductions, or least a lack of growth, during summer months. The data also suggest one reason why the RIAA has abandoned its approach of suing individual file sharers. In figure 3, it is difficult to ascertain an effect of the beginning of the 2003 lawsuit campaign (Manuse, 2003). While the overall campaign may have been disappointing from the RIAA's perspective, research has documented a short-run decline in the number of files shared and in downloading activity in response to the first round of

[13] User counts from the independent file-sharing site slyck.com largely mirror these numbers.

lawsuits (Bhattacharjee et al., 2006). In contrast, the *Grokster* Supreme Court decision in 2005 does not appear to have had much impact on the user-base.[14]

The data from these disparate sources paint a similar picture for trends in U.S. file sharing. There has been secular growth in both the amount of file sharing and the number of users. This upward trend has largely been unaffected by shifts in technology and the legal environment. At the same time, figure 1 shows that the intra-year cycle in file sharing observed in the early years has started to disappear. As broadband has proliferated outside of universities and to the home, young file-sharing users no longer rely on their university connections during the school year to download files.

### b. Consumer Behavior

Three facts about consumer behavior on file-sharing networks strike us as particularly interesting: the narrow focus on a limited set of files; the truly global nature of file sharing; and the continued importance of industry marketing efforts. We discuss each of these in turn.

Users share a wide variety of files on P2P networks. Table 2 shows the distribution of a selected list of genres on a popular P2P network and compares it to store sales of these albums and downloads of songs (for a detailed description of the sample, see Oberholzer-Gee and Strumpf, 2007). Genres such as R&B, Rap and New Artists are overrepresented, while there is comparatively little country music. Looking at what users actually download, it is striking to see how dominant the Current Alternative category is. Almost one half of all downloads are transfers of songs in this genre. The data in Table 2 reflect the supply of music files in 2002, the stone age of file sharing. We don't know of any study that has systematically compared changes in content over time.

While the supply of files is vast, peer-to-peer users download only a small share of the files that are available. In our sample of 10,271 different music tracks, 60% are never downloaded over a period of 17 weeks, and 81% are downloaded less than 5 times,

[14]Similarly, Ferguson (2006) shows that eDonkey traffic levels were largely unaffected in 2006 when legal authorities forced the closure of a large network of servers.

a number that is just slightly above the mean.[15] Even in movies, where the number of available titles is far smaller, there is a notable focus on the most popular titles. Table 3 shows the availability of and the demand for movies on Mininova, a popular BitTorrent index site. Not surprisingly, the top DVD rentals are all in high demand (column 2). But demand trails off markedly for older titles, many of which are not even available. A point in case is Malin Akerman, a Swedish actress voted number one on IMDB's starmeter in early 2009. Akerman was one of the stars of the then popular movie Watchmen. As the last column in Table 3 shows, there was in fact significant demand for that release. But movie buffs with an interest in Akerman's previous films faced rather slim pickings. At the height of the popularity of Akerman, four of her last ten movies were unavailable and there was no demand for two additional films.[16] As in music, downloading activity for movies is heavily concentrated on current releases and the supply of titles is substantially broader than the demand.

A second interesting fact about consumer behavior on peer-to-peer networks is the truly global nature of file-sharing. Table 4 shows the top countries for users and downloads (from Oberholzer-Gee and Strumpf, 2007). Interactions among file sharers transcend geography and language. U.S. users download only 45.1% of their files from other U.S. users, with the remainder coming from a diverse range of countries including Germany (16.5%), Canada (6.9%) and Italy (6.1%). One implication of these interactions is that national regulations of file sharing will only have limited bite. For instance, if the RIAA and domestic ISPs discouraged U.S. users from making files available, as they currently hope to do, users in the U.S. could simply download files from other countries.

A final observation concerns the marketing efforts of the entertainment industry. In view of the vast supply of music and videos on the internet and the many electronic networks connecting individuals, it might seem reasonable to expect that the industry's

[15] Our sample is drawn from SoundScan charts, which include all commercially relevant albums. Though some of the albums in the sample had low sales, many in fact were very high sellers.
[16] The concentration of movie downloads in part reflects the current BitTorrent technology. Index sites, which list the files available for download, typically de-list a title when no one is sharing a complete copy for some length of time. As a result, less popular movies become often unavailable, as are older movies since the number of shared copies tends to decline over time.

ability to draw attention to particular products has been greatly diminished. But the data in figure 4 tell a different story. The graph shows downloads and sales of the popular Eight Mile soundtrack, a commercial success directed by Curtis Hanson, starring rapper Eminem. Note that the recording leaked about 6 weeks prior to the official album release, with Eight Mile songs becoming available on peer-to-peer networks. But, interestingly, the level of downloads remained small until the industry marketing campaign began. Unless the industry drums up support for a new release, it is apparently difficult to give it away for free. This pattern of downloads and sales is fairly typical in our data. Contrary to the view that the entertainment industry has lost its ability to create value in a networked world, these data suggest the recording industry remains unrivaled in its ability to steer consumer attention.

## 4. Does File-Sharing Reduce the Sale of Copyrighted Materials?

The sharing of information goods such as music, movies, and books has been the subject of a substantial literature, both theoretical and empirical. Theory has most often focused on two competing intuitions about the effects of file sharing. A first is obvious: copying hurts producers because consumers who would have purchased a product now obtain it for free. But there is a second effect that runs counter to this idea. Because consumers anticipate sharing products, their willingness to pay (and hence producer profits) might actually increase. For example, a family might be willing to buy an expensive videogame because the parents know that several children will enjoy playing it. The theoretical literature has successfully identified a number of factors that influence the balance of these two effects, including the relative cost of producing information goods and sharing, the variation in the size of groups that share protected works, as well as the diversity in consumer valuations and the correlation of valuations within a sharing group (Novos and Waldman, 1984; Johnson, 1985; Liebowitz, 1985; Besen and Kirby, 1989; Bakos, Brynjolfsson and Lichtman, 1999; Varian, 2000). Depending on the importance of the relevant parameters, theoretical modeling predicts that file-sharing can

either hurt or help producers (for a review of theory papers, see Peitz and Waelbroeck, 2003).

Because the theoretical results are inconclusive, the effect of file sharing on industry profitability is largely an empirical question. We summarize the findings of some of the major studies in table 5. As the list shows, the results are decidedly mixed. There are two studies that document a positive effect of file-sharing on sales: Andersen and Franz (2008) for a representative sample of Canadian consumers and, more narrowly, Gopal et al. (2006) for the effect of sampling on CD sales.[17] The majority of studies finds that file sharing reduces sales, with estimated displacement rates ranging 3.5% for movies (Rob and Waldfogel, 2007) to rates as high as 30% for music (Zentner, 2006).[18] A typical estimate is a displacement rate of about 20%. One implication of these results is that developments other than file sharing must have had a profound impact on sales. For music, the popularity of new types of (internet-based) entertainment and the end of the transition from LPs to CDs are leading explanations for the overall decline in sales (Hong, 2004; Oberholzer-Gee and Strumpf, 2007). While many studies find some displacement, an important group of papers reports that file-sharing does not hurt sales at all (Tanaka, 2004; Bhattacharjee et al., 2007; Oberholzer-Gee and Strumpf, 2007; Smith and Telang, 2008). And even among the studies that show some displacement, there tend to be important subsamples that were not affected. For example, Rob and Waldfogel (2006) find an average displacement effect of 20% but report that file sharing had no impact on hit albums.

In order to better understand why file-sharing studies come to varying conclusions, it is instructive to consider a number of challenges in the empirical literature.

*Choice of Sample* – Researchers frequently rely on convenience samples, typically students, to estimate the effect of file sharing on sales. This is problematic because surveys show high school and college students to be among the most active file

[17] Gopal et al.'s (2006) results are consistent with the theoretical findings in Peitz and Waelbroeck (2006).
[18] An outlier is Liebowitz (2008) who reports a displacement rate of more than 100% for a selection of U.S. music markets.

sharers (Pew Internet Project, 2003). As a result, the displacement rates documented in these studies are likely to lie above the true population rates. Convenience aside, we suspect that many scholars rely on unrepresentative samples of students because it used to be almost impossible, and remains often expensive, to gain access to representative sales data. For instance, U.S. sales data for music, traditionally shared among record companies, has only become available to researchers in the most recent years. And even today, short-term subscriptions to industry databases can cost thousands of dollars, excluding scholars with more limited research budgets.[19] To arrive at a more complete understanding of file sharing, increased collaboration between industry and academia – and the employment of representative samples – appears essential to us.

*Measures of piracy* – A key difficulty in interpreting the findings of many studies is that they rely on self-reported data or poor proxies for actual file sharing. As table 5 indicates, surveys with self-reported measures of piracy play a significant role in the literature. Unfortunately, we do not know much about the accuracy of survey data in the context of file sharing. As Zentner (2006) points out, some individuals might play down their file sharing because they understand it is illegal. On the other hand, if file sharing is hip, as is the case on many college campuses, students might exaggerate the activity. In Andersen and Frenz (2008), more than 10% of respondents who report having downloaded music do not provide the number of downloaded files, suggesting recall or perhaps response bias might also be an issue. In view of the popularity of survey-based measures of piracy, we consider it important for future research to establish their accuracy. If these data turn out to be reliable, they could play a major role in future research because survey data are simple and inexpensive to obtain.

Where survey data on piracy is unavailable, researchers tend to rely on crude proxies for file sharing such as internet penetration. In a number of studies, internet-related measures (penetration, user sophistication) also serve as an instrument for downloading. In our view, both usages are inappropriate. Internet penetration proxies

[19] Nielsen SoundScan, the dominant provider of record sales, offers an academic subscription for $10,000 a year. Nielsen VideoScan is even more expensive. Box office numbers for theatrical releases are freely available from Box Office Mojo, but learning about geographic variation in sales is more difficult. Fortunately, Nielsen Bookscan data are available at a reasonable cost.

for new forms of entertainment – think YouTube and World of Warcraft – that compete directly with music and traditional film consumption, yielding a negative bias in displacement studies. Given these fairly obvious shortcomings, why are there so few papers that use actual data on file sharing to measure its effect on sales? One reason, we believe, is that collecting data on file-sharing networks is labor intensive and often cumbersome. Sometimes it is necessary to gain the trust of individuals operating file-sharing servers. And automated measurement studies require considerable programming skills and knowledge of file-sharing software. These hurdles notwithstanding, it is disappointing to see how few social scientists have made the effort to collect data on actual behavior. Many scholars prefer to use widely available, but in our view inappropriate, proxies for file sharing. The resulting research is poorer for it. The situation in the social sciences is in marked contrast to the research in computer science where many studies carefully measure individual file-sharing activity (e.g. Leibowitz et al. 2002; Gummadi et al. 2003; Pouwelse et al. 2005; Liang et al. 2005a, 2005b; Dhungel, et al. 2008).

We emphasize these issues because the results in table 5 seem to suggest that measurement choices have a systematic impact on results. While the majority of papers reports some sales displacement, the four studies using actual measures of file sharing (Tanaka, 2004; Bhattacharjee et al., 2007; Oberholzer-Gee and Strumpf, 2007; Smith and Telang, 2008) find that file sharing is unrelated to changes in sales.

*Unobserved heterogeneity* – A common difficulty in studying the link between downloads and sales is that file sharing is endogenous. That is, there are factors, some of them unobserved by the econometrician, that influence both downloads and sales. For example, music lovers are likely to download more songs and they also buy a larger number of albums, making it look like there was a positive relation between file sharing and sales. To see this, consider figure 5, taken from Oberholzer-Gee and Strumpf (2005). In this graph, downloads (horizontal axis) appear to increase sales (vertical axis). But an alternative explanation is that the popularity of a release increases both file-sharing activity and sales: popular recordings are in high demand on the internet and in the store.

Difference-in-difference (DD) estimates and instrumental variable techniques are popular means by which scholars hope to break the link between unobserved factors and the estimated impact of piracy on sales. DD models yield unbiased estimates if the unobserved heterogeneity is time invariant. Unfortunately, time-varying unobserved factors appear to play a major role in file sharing. Comparing DD estimates with results that take into account how cohort characteristics change over time, Hong (2008) finds that DD estimates attribute the entire 2002 decline in record sales to Napster. Once changes in unobserved heterogeneity are taken into account, the sales displacement rate drops from 100% to 20%. Similarly, Oberholzer-Gee and Strumpf (2007) show that the combination of album and week fixed effects is insufficient to control for unobserved heterogeneity.

Instrumental variable techniques provide a potentially more promising way to identify the effect of file sharing on sales. As noted above, we are skeptical of attempts to use measures of broadband adoption or user internet sophistication as instruments. More promising identification strategies exploit technical aspects of file-sharing systems – the availability of BitTorrent indexing sites, for instance, fluctuates considerably over time for largely technical reasons – and shocks to the global supply of content. For example, Oberholzer-Gee and Strumpf (2007) exploit the fact that many files downloaded in the US come from Germany. During German school holidays, file sharing in the US becomes easier: download times are shorter, a greater fraction of searches lead to a successful download, and fewer download requests remain incomplete. Because German holidays are unrelated to U.S. music sales, the holiday shock makes a promising instrument. More generally, because file sharing is a truly global phenomenon there are many shocks that spread from country to country. Some of these will be unrelated to the domestic demand for entertainment, making them promising prospects in the quest for proper identification.

## 5. How Important Are Complementary Sources of Income?

Even if file sharing displaces sales, the weaker copyright regime need not undermine the incentives to produce new works if artists and entertainment companies can shift their earnings from selling music, games and movies to selling complements to these products. An interesting example is concerts. As Table 6 shows, concerts and merchandising have become an important source of income for major artists (Connolly and Krueger, 2006). Concerts and new recordings are complements. A recording becomes more enjoyable if one can reminisce about the time at the concert, and knowing the songs in advance might make the concert more enjoyable. In the presence of complementary goods, file sharing will have two opposing effects (for a formal model, see Mortimer and Sorenson, 2005). As the effective price of music falls close to zero, a larger number of consumers will be familiar with an album, driving up the demand for concerts. At the same time, artists have weaker incentives to tour because concerts are a less effective way to increase revenues from a new recording if a large fraction of the audience shares files. Which of these effects is more important? Figure 6 shows that concert prices rose much more quickly than the CPI, and the difference appears to have widened since the advent of file sharing (Krueger, 2005). More detailed evidence on the link between file sharing and concerts comes from Mortimer and Sorenson (2005). Studying 2,135 artists over a ten-year period, they also conclude that the demand for concerts increased due to file sharing. One way to see this is to ask how many CDs an artist needs to sell to produce $20 of concert revenue. This number fell from 8.47 in the pre-Napster era to 6.36 in the 1999 to 2002 period. Not surprisingly, artists responded to these incentives by touring more frequently. Overall, the shift in relative prices and activities led to a sharp increase in income for the typical artist included in the authors' dataset.

As these results show, income from the sale of complements can more than compensate artists for any harm that file sharing might do to their primary activity. We are not aware of empirical work that has looked at these effects in industries other than music. But the potential of complements to provide ancillary income is certainly not unique to the music industry. In film, for instance, the International Licensing Industry Merchandisers' Association (LIMA) estimates that Hollywood derives $16 billion

annually from sales of entertainment merchandise, a figure that exceeds the value of ticket sales (Film Encyclopedia, 2008).

The role of complements makes it necessary to adopt a broad view of markets when considering the impact of file sharing on the creative industries. Unfortunately, the popular press – and a good number of policy experts – often evaluate file sharing looking at a single product market. Analyzing trends in CD sales, for example, they conclude that piracy has wrecked havoc on the music business. This view confuses value creation and value capture. Record companies may find it more difficult to profitably sell CDs, but the broader industry is in a far better position. In fact, it is easy to make an argument that the business has grown considerably. Figure 7 shows spending on CDs, concerts and iPods. The decline in music sales – they fell by 15% from 1997 to 2007 – is the focus of much discussion. However, adding in concerts alone shows the industry has grown by 5% over this period. If we also consider the sale of iPods as a revenue stream, the industry is now 66% larger than in 1997. Obviously, these numbers are no more than a rough back-of-the-envelope calculation. A more serious investigation would take into account differences in profitability across music and concert sales as well as the decreased spending in other electronics categories (CD players, speakers, etc.) The point of the graph, however, remains: technological change will often lead to changes in relative prices and shifts in business opportunities. Focusing exclusively on traditional streams of revenue to arrive at a sense of how new technology changes welfare will typically be misleading.

## 6. **Does File-Sharing Undermine Artistic Production?**

In any evaluation of file sharing, a key question is whether financial incentives are needed to encourage artistic output.[20] While this is in large part an open question, several indirect pieces of evidence suggest that financial incentives play a smaller role in the

[20] In this respect, the arts are similar to the production of open source software where many programmers appear to work for little monetary gain (Lerner and Tirole, 2005).

creative industries than elsewhere in the economy.[21] For concreteness we will focus our discussion on popular music, but many ideas discussed here carry over to film, visual arts, writing, and high culture music (see Caves, 2000).

The economic prospects for the group of popular musicians as a whole are quite poor. An album selling a half million copies or more (a Gold Album) is considered successful. Typically, a few hundred albums reach this level each year. Yet over 50,000 albums are released annually, suggesting the chance of success is less than one in a hundred. Perhaps more strikingly, only 950 new albums sold more than 25,000 copies in 2007.

Moreover, it is difficult for musicians to earn substantial income from recorded music sales, regardless of the success of their album. This is in part due to the nature of recorded music contracts (Passman, 2000). Recording musicians are paid for album sales based on the product of a royalty rate and album sales. The royalty rate is quite low (usually about a dollar or two per album) and musicians are not paid this money until they recoup all expenses, primarily the advance which is typically applied to the cost of recording the album. If an earlier album did not sell well enough to pay for the advance, music companies often deduct the difference from future album payments under a system called cross-collateralization. Putting all this together, even a Gold Album may not provide a musician with an economic windfall.[22]

Given these poor prospects, why are there so many musicians? One explanation is that musicians enjoy their profession. Under this view, musicians take pleasure from creating and performing music, as well as aspects of the lifestyle such as flexible hours and the lack of an immediate boss. If this theory is correct, the economic impact of file sharing is not likely to have a major impact on music creation.

An alternative explanation is that popular music is a tournament, where a few artists collect most of the economic rewards. This view is rooted in the theory of superstars (Rosen, 1981). Superstars develop in industries with low marginal cost of

[21] The broader critique of Boldrin and Levine (2008) implies that for innovation to take place more generally, copyright and patents are not needed.
[22] For specific dollar totals from insiders in the music industry, see Albini (1994) and Love (2000).

production, little relation between output and quality, and quality-conscious consumers. This seems to be a reasonable model of popular music: it is relatively cheap to produce CDs and even cheaper to make digital albums. Each album produced provides the same quality level, and most consumers would rather listen to one very good album than a few albums of lesser quality. Under the superstar theory musicians essentially consider their job to be a lottery. With some small chance they will become a star. In 2007, the top one percent of new releases accounted for 82% of new-release sales. In a superstar environment, file sharing has a muted effect on music output. Even if the new technology had a marked negative effect on the returns to stardom, it is not likely to have big effect on the chances of becoming a star.[23]

Survey evidence (as well as the long lines of contestants hoping to be part of talent shows like American Idol) support these theoretical arguments. In a Pew study of 2,755 musicians and songwriters (Madden, 2004), over three-fourths of respondents reported having a paying non-music job.[24] These second jobs are the primary source of income for most musicians. Only 16% reported that at least sixty percent of their income derived from their music job, while 66% said they earned less than twenty percent of their income from music. The small income share is not simply due to spending few hours on music. Even among those who spent at least thirty hours a week on music-related activities, only 22% derived at least four-fifths of their income from music.

Overall production figures for the creative industries appear to be consistent with this view that file sharing has not discouraged artists and publishers. While album sales have generally fallen since 2000, the number of albums being created has exploded. In 2000, 35,516 albums were released. Seven years later, 79,695 albums (including 25,159 digital albums) were published (Nielsen SoundScan, 2008). Even if file sharing were the

[23]Consider a model in which individuals must choose between being a musician and some outside reservation job. If p is the probability of being a star, S the income (and non-pecuniary benefits) of being a star, NS the income of a non-star, and R the income from the reservation jobs, than the person decides to be a musician when,

$$pU(S) + (1-p)U(NS) \geq U(R)$$

where U(.) is a utility function and S>>R>NS. Even if file sharing has a large negative effect on S, this will only have a limited impact on the left-hand side presuming S remains large and U''<0.

[24]The musicians surveyed come from a wide range of music genres including Pop, Folk, Country, Electronic, Blues, Rock, Jazz, Christian, Punk, Dance, Bluegrass, Latin, Reggae, and Hip Hop. This wide coverage suggests the responses should incorporate a range of viewpoints.

reason that sales have fallen, the new technology does not appear to have exacted a toll on the quantity of music produced.[25] Obviously, it would be nice to adjust output for differences in quality, but we are not aware of any research that has tackled this question.

Similar trends can be seen in other creative industries. For example, the worldwide number of feature films produced each year has increased from 3,807 in 2003 to 4,989 in 2007 (Screen Digest, 2004 and 2008). Countries where film piracy is rampant have typically increased production. This is true in South Korea (80 to 124), India (877 to 1164), and China (140 to 402). During this period, U.S. feature film production has increased from 459 feature films in 2003 to 590 in 2007 (MPAA, 2007).

## 7. Policy Implications and Conclusions

File-sharing technology considerably weakened copyright protection, first of music and software and increasingly of movies, games, and books. The policy discussion surrounding file sharing has largely focused on the legality of the new technology and the question whether or not declining sales in music are due to file sharing. While these are important questions, in our view, the debate has been overly narrow. Copyright exists to encourage innovation and the creation of new works; in other words to promote social welfare. The question to ask is thus whether the new technology has undermined the incentives to create, market, and distribute entertainment. Sales displacement is a necessary but not a sufficient condition for harm to occur. We also need to know whether income from complementary products offset the decline in income from copyrighted works. And even if income fell, welfare may not suffer if artists do not respond to weaker monetary incentives.

As our survey indicates, the empirical evidence on sales displacement is mixed. While some studies find evidence of a substitution effect, other findings, in particular the

[25] Similarly, recording contracts seem to remain appealing. In 2009, 1,900 acts performed at South-by-Southwest, a large music festival that attracts musicians looking to sign their first recording contract. The artists must typically pay their own travel and lodging expenses, in addition to any foregone wages from their secondary job. Clearly a large number of musicians thought attending the festival was a worthwhile investment (Pareles, 2009).

papers using actual file-sharing data, suggest that piracy and music sales are largely unrelated. In contrast, there is clear evidence that income from complements has risen in recent years. For example, concert sales have increased more than music sales have fallen. Similarly, a fraction of consumer electronics purchases and internet-related expenditures are due to file sharing. Unfortunately, we know little about the distribution of these impacts. How markets for complimentary goods have responded to file sharing remains an area of inquiry that is largely unexplored in academic research.

The same holds true for the question how artists would respond to weaker monetary incentives. Looking at aggregate output – the number of recordings, books, and movies produced every year – we see no evidence that file sharing has discouraged the production of artistic works. However, as with income from complementary goods, aggregate statistics need to be interpreted with some care. For example, digital formats not only encouraged file sharing; digital technology also lowered the cost of producing movies and music and they allowed artists to reach their audience in novel ways. The observed increase in output is in part due to these changes. The response of artists to technology-induced changes in income is a second area that we would like to single out as important for future research.

As this essay has made clear, we do not yet have a full understanding of the mechanisms by which file sharing may have altered the incentives to produce entertainment. However, in the industry with the largest purported impact – music – consumer access to recordings has vastly improved since the advent of file haring. Since 2000, the number of recordings produced has more than doubled. In our view, this makes it difficult to argue that weaker copyright protection has had a negative impact on artists' incentives to be creative.

**Literature**


Andersen, Birgitte and Marion Frenz (2008). The Impact of Music Downloads and P2P File-Sharing on the Purchase of Music: A Study for Industry Canada. University of London Working Paper.

Albini, Steve (1994). "The Problem with Music." *Maximum Rock 'n' Roll* Issue #133. http://www.arancidamoeba.com/mrr/problemwithmusic.html.

Allot Communications (2007). "Digging Deeper Into Deep Packet Inspection (DPI)." https://www.dpacket.org/articles/digging-deeper-deep-packet-inspection-dpi.

Baker, Matthew J. and Brendan M. Cunningham (2006). Court Decisions and Equity Markets: Estimating the Value of Copyright Protection. *Journal of Law and Economics* 49: 567-596.

Bakos, Yannis, Erik Brynjolfsson, and Douglas Lichtman (1999). Shared Information Goods. Journal of Law and Economics 42: 117-155.

Besen, Stanley M. and Sheila N. Kirby (1989). Private Copying, Appropriability, and Optimal Copyright Royalties. *Journal of Law and Economics* 32: 255-280.

Bhattacharjee, Sudip, Ram D. Gopal, Kaveepan Lertwachara, James R. Marsden and Rahul Telang (2007). The Effect of Digital Sharing Technologies on Music Markets: A Survival Analysis of Albums on Ranking Charts. *Management Science* 53(9): 1359-1374.

Boldrin, Michele and David K. Levine (2008). *Against Intellectual Monopoly*. Cambridge: Cambridge University Press.

Caves, Richard (2000). *Creative Industries: Contracts Between Art and Commerce*. Cambridge, MA: Harvard University Press.

Coase, Ronald H. (1979). Payola in Radio and Television Broadcasting. *Journal of Law & Economics* 22 (2): 269-328.

Connolly, Marie and Alan Krueger (2006). Rockonomics: The Economics of Popular Music. Handbook of Arts and Culture, Amsterdam: North Holland: 667-720.

Dhungel, P., D. Wu, B. Schonhorst, K.W. Ross (2008). A Measurement Study of Attacks on BitTorrent Leechers, IPTPS.

Economist (2005). BSA or just BS? 19 May:

Edison Media Research (2003). The National Record Buyers Study III. Sponsored by Radio & Records. http://www.edisonresearch.com.

Ferguson, David (2006). "Trends and Statistics in Peer-to-Peer." CacheLogic. Presented at Workshop on Technical and legal aspects of peer-to-peer television, the Netherlands.

Film Encyclopedia (2008). Historical Background and Definitions. At http://www.filmreference.com/encyclopedia

Forrester (2002). "Downloads Save the Music Business." http://www.forrester.com.

Forrester 2004. "US Antipiracy Bill Won't Stop File Sharing."

Gayer, A. and O. Shy (2006). Publishers, Artists, and Copyright Enforcement. Information Economics & Policy 18(4): 374-384.

Gopal, Ram D. and Sudip Bhattacharjee (2006). Do Artists Benefit from Online Music Sharing? *Journal of Business* 79(3): 1503-1533.

Gummadi, Krishna, Richard Dunn, Stefan Saroiu, Steven Gribble, Henry Levy and John Zahorjan (2003). Measurement, Modeling, and Analysis of a Peer-to-Peer File-Sharing Workload. Proceedings of the 19th ACM Symposium on Operating Systems Principles (SOSP-19).

Hui, Kai-Lung and Ivan Png (2002). On the Supply of Creative Works: Evidence from the Movies. *American Economic Review* 92(2): 217-220.

Hui, Kai-Lung and Ivan Png (2003). Piracy and the Legitimate Demand for Recorded Music. *Contributions to Economic Analysis & Policy* 2(1): Article 11.

Hong, Seung-Hyun (2004). The Effect of Napster on Recorded Music Sales: Evidence from the Consumper Expenditure Survey. Working Paper, Stanford University.

Hong, Seung-Hyun (2008). Measuring the Effect of Napster on Recorded Music Sales. Working Paper, University of Illinois.

International Federation of the Phonographic Industry 2002. Recording Industry in Numbers 2001. International Federation of Phonographic Industry.

Internet2 Netflow Statistics (2009). Internet2 NetFlow: Weekly Reports. http://netflow.Internet2.edu/weekly/.

Johnson, William R. (1985). The Economics of Copying. *Journal of Political Economy* 93:

Jupiter Media Metrix (2002). "File Sharing: To Preserve Market Value Look Beyond Easy Scapegoats." http://www.jupiterresearch.com.

Karagiannis, Thomas, Andre Broido, Nevil Brownlee, k claffy, Michalis Faloutsos (2003). File-sharing in the Internet: A characterization of P2P traffic in the backbone. Technical Report, University of California, Riverside.

Karagiannis, Thomas, Andre Broido, Nevil Brownlee, k claffy, Michalis Faloutsos (2004). "Is P2P dying or just hiding?" Presented at Globecom 2004.

Khan, B. Zorina (2001). Do Property Rights Matter? Evidence from U.S. International Copyright Law, 1790-1910. Bowdoin College Working Paper.

Klein, Benjamin, Andres V. Lerner and Kevin M. Murphy (2002). The Economics of Copyright Fair Use in a Networked World. *American Economic Review* 92: 205-208.

Krueger, Alan (2005). The Economics of Real Superstars: The Market for Concerts in the Material World, *Journal of Labor Economics* 23(1): 1-30.

Lamere, Paul (2006). What's On Your iPod? At http://blogs.sun.com/plamere/entry/what_s_on_your_ipod.

Landes, William M. and Richard A. Posner (1989). An Economic Analysis of Copyright Law. *Journal of Legal Studies* 18: 325-363.

Leibowitz, Nathaniel, Aviv Bergman, Roy Ben-Shaul and Aviv Shavit (2002). Are File Swapping Networks Cacheable? Characterizing P2P Traffic. Expand Networks working paper. Presented at the 7th International Workshop on Web Content Caching and Distribution (WCW).

Lerner, Josh and Jean Tirole (2005). The Economics of Technology Sharing: Open Source and Beyond. *Journal of Economic Perspectives* 19(Spring): 99-120.

Liebowitz, Stanley J. (1985). Copying and Indirect Appropriability: Photocopying of Journals. *Journal of Political Economy* 93: 945-957.

Liebowitz, Stanley J. (2008). Testing File Sharing's Impact on Music Album Sales in Cities. Research Note. *Management Science* 54(4): 852-859.

Lenhart, Amanda and Susannah Fox (2000). Downloading Free Music: Internet Music Lovers Don't Think It's Stealing, The Pew Internet & American Life Project's Online Music Report, 28 September 2000, at: http://www.pewinternet.org/PPF/r/23/report_display.asp.

Leung, Tin Cheuk (2008). Should the Music Industry Sue Its Own Customers? Impacts of Music Piracy and Policy Suggestions. University of Minnesota Working Paper.

Liang, J., R. Kumar, and K.W. Ross (2005a). The Kazaa Overlay: A Measurement Study, Computer Networks (Special Issue on Overlays).

Liang, J., R. Kumar, Yongjian Xi, K.W. Ross (2005b). Pollution in P2P File Sharing Systems, Infocom 05, Miami.

Love, Courtney (2000). "Courtney Love does the math." *Salon*. http://dir.salon.com/tech/feature/2000/06/14/love/index.html

Madden, Mary (2004). "Artists, Musicians and the Internet." Pew Internet & American Life Project. http://www.pewinternet.org/Reports/2004/Artists-Musicians-and-the-Internet.aspx.

Manuse, Andrew J. (2003). Music Industry Threats Not Slowing File-Sharing Volume. Boston Herald, June 27.

McBride, Sarah and Ethan Smith (2008), "Music Industry to Abandon Mass Suits," The Wall Street Journal, 19 December 2008, at B1.

Mortimer, Julie Holland and Alan Sorensen (2005). Supply Responses to Digital Distribution: Recorded Music and Live Performances. Harvard University Working Paper.

MPAA (2007). Theatrical Market Statistics. At http://mpaa.org/researchstatistics.asp.

Novos, Ian E. and Michael Waldman (1984). The Effects of Increased Copyright Protection: An Analytic Approach. *Journal of Political Economy* 92: 236-246.

Oberholzer-Gee, Felix (2006). eDonkey – Deciding the Future of File Sharing. Harvard Business School Case 707-482.

Oberholzer-Gee, Felix and Koleman Strumpf 2005. "The Effect of File Sharing on Record Sales: An Empirical Analysis." Working Paper, Harvard Business School and the University of North Carolina at Chapel Hill.

Oberholzer-Gee, Felix and Koleman Strumpf (2007). The Effect of File Sharing on Record Sales: An Empirical Analysis. *Journal of Political Economy* 115(1): 1-42.

Pareles, Jon (2009). "Stoking Careers in Frenzy of South by Southwest," 22 March 2009, New York Times.

Pasick, Adam (2004). "LIVEWIRE — File-sharing network thrives beneath the radar," Yahoo! News at http://www.interesting-people.org/archives/interesting-people/200411/msg00078.html

Passman, Donald (2000). *All You Need to Know About the Music Business*. New York: Simon & Schuster.

Peitz, Martin and Patrick Waelbroeck (2003). Piracy of Digital Products: A Critical Review of the Economics Literature. CESIfo Working Paper 1071.

Peitz, Martin and Patrick Waelbroeck (2004). The Effect of Internet Piracy on Music Sales: Cross-section Evidence. *Review of Economic Research on Copyright Issues* 1(2): 71-79.

Peitz, Martin and Patrick Waelbroeck (2006). Why the Music Industry May Gain From Free Downloading: The Role of Sampling. *International Journal of Industrial Organization* 24(5): 907-913.

Pew Internet Project (2003). Music Downloading, File-Sharing and Copyright. http://www.pewtrusts.com/pubs/.

Png, Ivan and Qiu-hong Wang (2006). Copyright Duration and the Supply of Creative Work: Evidence from the Movies. National University of Singapore Working Paper.

Pouwelse, J.A., P. Garbacki, D.H.J. Epema, H.J. Sips (2005). The Bittorrent P2P File-sharing System: Measurements and Analysis", 4th International Workshop on Peer-to-Peer Systems (IPTPS'05).

Rich, Motoko (2009). Print Books Are Target of Pirates on the Web. *New York Times*, 11 May.

Rob, Rafael and Joel Waldfogel (2006). Piracy on the High C's: Music Downloading, Sales Displacement, and Social Welfare in a Sample of College Students. *Journal of Law and Economics* 49(1): 29-62.

Rosen, Sherwin (1981). "The Economics of Superstars." *American Economic Review*. 71: 845-858.

Sandvine (2002-2007). Occasional white papers and press reports on file sharing activity in North America and World-wide. http://www.sandvine.com.

Sandvine (2008a). "2008 Analysis of Traffic Demographics in North-American Broadband Networks." http://www.sandvine.com.

Sandvine (2008b). "2008 Global Broadband Phenomena." http://www.sandvine.com.

Shalunov, Stanislav and Benjamin Teitelbaum (2001). "TCP Use and Performance on Internet2." ACM SIGCOMM Internet Measurement Eorkshop 2001.

Smith, Michael D. and Rahul Telang (2006). Piracy or Promotion? The Impact of the Broadband Internet Penetration on DVD Sales. Carnegie Mellon University Working Paper.

Smith, Michael D. and Rahul Telang (2008). Competing With Free: The Impact of Movie Broadcasts on DVD Sales and Internet Piracy. Carnegie Mellon University Working Paper.

Tanaka, Tatsou (2004). Does File-sharing Reduce CD sales?: A Case of Japan, Conference Paper Prepared for Conference on IT Innovation. Hitotsubashi University, Tokyo.

Towse, Ruth (1999). Copyright and Economic Incentives: An Application to Performer's Rights in the Music Industry. *Kyklos* 52: 69-90.

Varian, Hal (2000). Buying, Sharing and Renting Information Goods. University of California at Berkeley Working Paper.

Zentner, Alejandro (2006). Measuring the Effect of Music Downloads on Music Purchases. *Journal of Law & Economics* 49(1): 63-90.

TABLE 1
KEY EVENTS IN FILE SHARING

| Date | Event |
|---|---|
| Spring 1998 | First mass-produced MP3 player |
| October 1998 | RIAA files restraining order against leading MP3 player manufacturer |
| June 1999 | Napster begins operations |
| December 1999 | RIAA sues Napster for copyright damages |
| July 2000 | US District Court rules against Napster and in favor of RIAA. Case moves to US Court of Appeals which affirms in February 2001 that Napster is liable for damages |
| Spring-Summer 2001 | Several alternative file sharing protocols are released including FastTrack/KaZaA, WinMX, Limewire, and BitTorrent |
| July 2001 | Napster effectively shut-down |
| November 2001 | RIAA and MPAA sue file sharing software distributors Morpheus and Grokster in *MGM v. Grokster* |
| Spring 2003 | FastTrack/KaZaA peaks at about 4m simultaneous users. |
| September 2003 | RIAA begins suing file sharing users. About 35,000 lawsuits have been filed by the end of 2008. |
| November 2003 | The Pirate Bay, a BitTorrent index and tracker site, is founded |
| Fall 2004 | A leading BitTorrent tracker + indexer has over 1m visits per day |
| June 2005 | Supreme Court upholds the content-holders position in *MGM v. Grokster*. By the end of the 2005 distribution companies eDonkey and WinMX shut-down after receiving cease and desist letters from the RIAA |
| May 2006 | In part due to pressure from the MPAA, Swedish police shut down The Pirate Bay and confiscate its servers. Site was operational again in three days, and servers are now spread over several countries |
| November 2008 | 25m users on leading BitTorrent tracker The Pirate Bay |

TABLE 2
FILES ON FILE-SHARING NETWORKS

| | % songs on network | % store sales | % downloads |
|---|---|---|---|
| Full sample | 100.0% | 100.0% | 100.0% |
| Catalogue | 8.0% | 9.8% | 12.6% |
| Current Alternative | 19.1% | 24.8% | 48.6% |
| Hard Music Top Overall | 3.0% | 5.9% | 5.3% |
| Jazz Current | 2.9% | 4.6% | 0.4% |
| Latin | 3.5% | 5.8% | 0.7% |
| New artists | 8.0% | 3.3% | 1.8% |
| R&B | 25.2% | 9.7% | 14.9% |
| Rap | 13.7% | 8.2% | 4.6% |
| Top Current Country | 10.2% | 18.4% | 7.3% |
| Top Soundtrack | 6.4% | 9.4% | 3.9% |

Source: Oberholzer-Gee and Strumpf (2007)

TABLE 3
AVAILABILITY OF MOVIES ON MININOVA

| RANK | TOP DVD RENTALS MARCH 2009 | # DOWNLOADS | MALIN AKERMAN MOVIES | # DOWNLOADS |
|---|---|---|---|---|
| 1 | Role Models (2008) | 10,482 | Watchmen | 53,476 |
| 2 | Transporter 3 (2008) | 11,225 | Bye Bye Sally | NA |
| 3 | Australia (2008) | 17,244 | 27 Dresses | 367 |
| 4 | Milk (2008/I) | 2,833 | Heavy Petting | 0 |
| 5 | Beverly Hills Chihuahua (2008) | 3,050 | The Heartbreak Kid | 53 |
| 6 | Rachel Getting Married (2008) | 1,705 | The Brothers Solomon | 0 |
| 7 | Body of Lies (2008) | 10,394 | The Invasion | NA |
| 8 | In the Electric Mist (2009) | 1,885 | Harold & Kumar | 382 |
| 9 | Changeling (2008) | 11,149 | The Utopian Society | NA |
| 10 | Nights in Rodanthe (2008) | 1,290 | The Circle | NA |

Sources: Internet Movie Database (http://www.imdb.com/) and Mininova (http://www.mininova.org/), accessed on 14 March 2009

TABLE 4
THE GEORGRAPHY OF FILE SHARING

| Country | Share of users | Share of downloads | Users in U.S. download from (%) | Users in U.S. upload to (%) | Share World Population | Share World Internet Users |
|---|---|---|---|---|---|---|
| United States | 30.9 | 35.7 | 45.1 | 49.0 | 4.6 | 27.4 |
| Germany | 13.5 | 14.1 | 16.5 | 8.9 | 1.3 | 5.3 |
| Italy | 11.1 | 9.9 | 6.1 | 5.7 | 0.9 | 3.2 |
| Japan | 8.4 | 2.8 | 2.5 | 1.8 | 2.0 | 9.3 |
| France | 6.9 | 6.9 | 3.8 | 4.7 | 1.0 | 2.8 |
| Canada | 5.4 | 6.1 | 6.9 | 7.9 | 0.5 | 2.8 |
| United Kingdom | 4.1 | 4.0 | 4.2 | 4.2 | 1.0 | 5.7 |
| Spain | 2.5 | 2.6 | 1.8 | 2.0 | 0.6 | 1.3 |
| Netherlands | 2.1 | 2.1 | 1.9 | 1.6 | 0.3 | 1.6 |
| Australia | 1.6 | 1.9 | 0.8 | 2.2 | 0.3 | 1.8 |
| Sweden | 1.5 | 1.7 | 1.8 | 1.5 | 0.1 | 1.0 |
| Switzerland | 1.4 | 1.5 | 0.9 | 1.0 | 0.1 | 0.6 |
| Brazil | 1.3 | 1.4 | 1.2 | 1.3 | 2.9 | 2.3 |
| Belgium | 0.9 | 1.2 | 0.5 | 1.0 | 0.2 | 0.6 |
| Austria | 0.8 | 0.6 | 0.6 | 0.4 | 0.1 | 0.6 |
| Poland | 0.5 | 0.7 | 0.7 | 0.5 | 0.6 | 1.1 |

Source: Oberholzer-Gee and Strumpf (2007)

TABLE 5
STUDIES OF THE ECONOMIC IMPACT OF FILE SHARING

| Study | Study Question, Data and Sample | Methodology | Key Findings |
|---|---|---|---|
| **Music** | | | |
| Hui and Png (2003) | Do country-level piracy rates explain the decline in music sales? Macro data, 28 countries, 1994-1998 | Sales regressions with country fixed effects; uses piracy rates for music cassettes and business computer software as instruments | For every pirated CD, sales fall by 0.42 units. Estimated effect is not robust to including year fixed effects and estimating separate displacement effects for high- and low-income countries. |
| Peitz and Waelbroeck (2004) | Do country averages in the likelihood of having downloaded music at least once predict music sales? Macro data, 16 countries, 1998-2002 | Cross-sectional analysis relating changes in sales to the level of file-sharing in 2002; no measure for the intensity of file sharing | Piracy reduced sales by 20%; effect is significant at 10% level |
| Tanaka (2004) | Do albums that are popular on file-sharing networks sell fewer copies? Observed piracy; 261 best-selling titles; 2004 | Study relates actual downloads on Winny, a popular Japanese file-sharing software, to CD sales; uses music genres as instruments | File-sharing does not reduce sales. |
| Gopal et al. (2006) | Are students who sample music they don't know more likely to purchase the CD? Survey; 200 students | Students indicate interest in buying and sampling music in a hypothetical-choice setting with set prices. | Students with faster internet connections are more likely to sample music; sampling increases the propensity to buy. |
| Rob and Waldfogel (2006) | Do students who downloaded music purchase fewer albums? Survey; 412 students; 2003/2004 | Students report purchases and downloads of 8,200 specific recordings; study uses access to broadband to instrument for downloads | For hit albums the authors find no relationship between downloading and sales. For a wider set of music, downloading five albums displaces the sale of one CD. Instrumenting for downloads results in estimates that are too imprecise to draw any firm conclusions. Using student valuations of albums, the authors conclude that file-sharing increases social welfare. |

| | | | |
|---|---|---|---|
| Zentner (2006) | Do individuals who downloaded at least once buy fewer CDs?<br>Survey; 15,000 European consumers, 2001 | Cross-sectional analysis; uses measures of Internet sophistication and access to broadband as instruments; no measure for the intensity of file sharing | Having shared files reduces the probability of purchasing music by 30%. |
| Bhattacharjee et al. (2007) | Do albums that are more frequently shared drop off the Billboard charts in a shorter period of time?<br>Observed piracy; best-selling titles; 2002-2003 | Relates the supply of files on file-sharing network (WinMx) to chart rankings; study uses RIAA announcement of lawsuits as instrument | Overall, file sharing has no statistically significant effect on survival on charts. The authors find a small negative effect for weaker releases. |
| Oberholzer-Gee and Strumpf (2007) | Do albums that are popular on file-sharing networks sell fewer copies?<br>Observed piracy; representative sample of recordings; 2002 | Relates downloads of files to CD sales; uses the supply shock due to German school holidays to instrument for downloads | File-sharing does not have a statistically significant impact on record sales. |
| Andersen and Frenz (2008) | Do individuals who obtain music for free buy fewer CDs?<br>Survey; representative sample of Canadians, 2006 | Authors have information on many forms of sharing, including P2P, ripping, promotional downloads, and copying of mp3 files; cross-sectional regressions without instruments | File sharing increases music purchases. 12 additional downloads lead to the sale of an additional 0.44 CDs. |
| Hong (2004, 2008) | Do households with internet access report lower music purchases post Napster?<br>Survey; 2000 | Two-variate propensity score matching; probability of using Napster is unobserved; needs to be imputed from UCLA survey using demographic information | The introduction of Napster explains 20% of the decline in music expenditures. 80% of the decline is due to changes in the prices of other entertainment goods and the ending of the transition from LPs to CDs (Hong 2004). Using a conventional difference-in-difference approach, the effect of Napster would be significantly overestimated, explaining the entire decline. |
| Leung (2008) | Do students who indicate they would download music intend to buy fewer songs?<br>Conjoint survey; 884 (270) students | Students report past consumption of music and make hypothetical choices between legal music, iPods, and pirated music; the study uses an assumed probability of getting caught and the size of the fine as instruments | When students pirate 10% more music, they intend to buy 0.7% fewer iTunes songs and 0.4% fewer CDs. |

| | | | |
|---|---|---|---|
| Liebowitz (2008) | Do U.S. cities with greater internet penetration have lower record sales?<br>Macro data; 89 markets, 1998-2003 | Compares changes in city-wide internet penetration with changes in record sales, controlling for demographics | Using all markets, internet penetration is unrelated to changes in music sales; for a subset of markets (60) the internet reduces per-capita-sale by 1.55, indicating file sharing explains more than 100% of the decline in record sales. |
| **Movies and TV** | | | |
| Smith and Telang (2006) | Does broadband help or hurt DVD sales?<br>Macro data; 2000-2003 | Market fixed effects specification with autoregressive errors | Broadband penetration increases DVD sales. Almost 10% of the increase in DVD sales during the study period is attributable to advances in broadband penetration. |
| Rob and Waldfogel (2007) | Are students who watch a pirated copy of a movie subsequently less likely to purchase the DVD?<br>Survey; 500 students; 2002-2005 | Students report their viewing of 50 top movies; no instrumental variables; person fixed effects control for time-invariant unobserved heterogeneity | Illegal burning of DVDs and downloading make up 5.2% of movie viewing; unpaid consumption reduces paid consumption by 3.5%. |
| Waldfogel (2007) | Do students who watch a TV series on the web less likely to watch episodes on TV?<br>Survey; 287 students; 2005-2007 | Students report the consumption of TV series on TV, YouTube and network websites; no instruments; demand for TV is estimated in first differences | Web consumption (authorized and unauthorized) reduces the number of shows that students watch frequently on TV but it increases the number of shows they watch sometimes. Additional web viewing exceeds the reduction in traditional viewing; even network-controlled viewing (excluding YouTube) increases by 1.5 hours per week. |
| Smith and Telang (2008) | Do TV broadcasts of movies and piracy reduce the sale of DVDs?<br>Observed piracy; 267 movies; 2005-2006 | The study uses TV broadcasts as shocks to identify the effect of piracy on DVD sales | Free broadcasts of movies on TV increase DVD sales on Amazon by 118% during the first week after the broadcast. Piracy does not affect this increase in demand. |

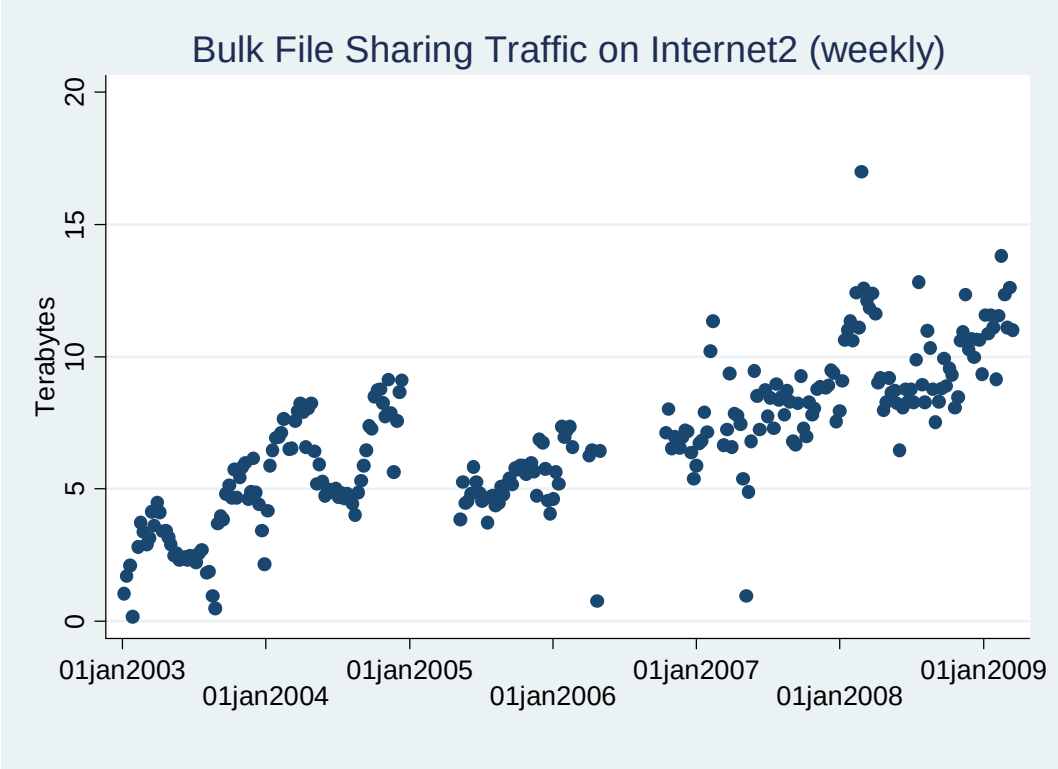
TABLE 6
ARTIST INCOMES (IN MILLIONS USD)

| Rank | Artist | Concerts | Recordings | Publishing | Total |
|---|---|---|---|---|---|
| 1 | Paul McCartney | 64.9 | 2.2 | 2.2 | 72.1 |
| 2 | The Rolling Stones | 39.6 | 0.9 | 2.2 | 44.0 |
| 3 | Dave Matthews Band | 27.9 | 0.0 | 2.5 | 31.3 |
| 4 | Celine Dion | 22.4 | 3.1 | 0.9 | 31.1 |
| 5 | Eminem | 5.5 | 10.4 | 3.8 | 28.9 |
| 6 | Cher | 26.2 | 0.5 | 0.0 | 26.7 |
| 7 | Bruce Springsteen | 17.9 | 2.2 | 4.5 | 24.8 |
| 8 | Jay-Z | 0.7 | 12.7 | 0.7 | 22.7 |
| 9 | Ozzy Osbourne | 3.8 | 0.2 | 0.5 | 22.5 |
| 10 | Elton John | 20.2 | 0.9 | 1.3 | 22.4 |
| 11 | The Eagles | 15.1 | 0.7 | 1.4 | 17.6 |
| 12 | Jimmy Buffet | 13.7 | 0.2 | 0.5 | 17.6 |
| 13 | Billy Joel | 16.0 | 0.0 | 1.0 | 17.0 |
| 14 | Neil Diamond | 16.5 | 0.0 | 0.3 | 16.8 |
| 15 | Aerosmith | 11.6 | 1.0 | 0.8 | 16.5 |
| 16 | CSNY | 15.7 | 0.0 | 0.3 | 16.0 |
| 17 | Creed | 10.9 | 1.1 | 1.6 | 13.4 |
| 18 | Rush | 13.4 | 0.0 | 0.0 | 13.4 |
| 19 | Linkin Park | 1.7 | 4.7 | 6.3 | 13.1 |
| 20 | The Who | 12.6 | 0.0 | 0.0 | 12.6 |
| 21 | Red Hot Chili Peppers | 6.1 | 3.4 | 2.7 | 12.1 |
| 22 | Brian “Baby” Williams | 0.2 | 2.7 | 0.9 | 11.8 |
| 23 | Nsync | 7.7 | 0.5 | 0.9 | 9.4 |
| 24 | Barry Manilow | 8.0 | 1.2 | 0.0 | 9.2 |
| 25 | Britney Spears | 5.5 | 1.8 | 1.0 | 9.1 |
| 26 | Alan Jackson | 4.6 | 3.0 | 1.4 | 9.0 |
| 27 | Rod Stewart | 6.6 | 1.4 | 0.8 | 8.8 |
| 28 | Andrea Bocelli | 8.1 | 0.2 | 0.4 | 8.7 |
| 29 | Brooks and Dunn | 6.7 | 0.4 | 1.4 | 8.1 |
| 30 | Enrique Iglesias | 4.4 | 1.5 | 1.7 | 7.6 |
| 31 | Tom Petty | 6.6 | 0.2 | 0.7 | 7.5 |
| 32 | Tool | 7.3 | 0.0 | 0.0 | 7.4 |
| 33 | Kid Rock | 3.4 | 0.8 | 1.3 | 7.0 |
| 34 | Kenny Chesney | 5.8 | 1.1 | 0.1 | 7.0 |
| 35 | Santana | 6.0 | 0.0 | 0.7 | 6.9 |
| | Average | 12.7 | 1.7 | 1.3 | 17.4 |

Note: Figures are estimates of pretax gross income in 2002.

Source: Connolly and Krueger (2006).

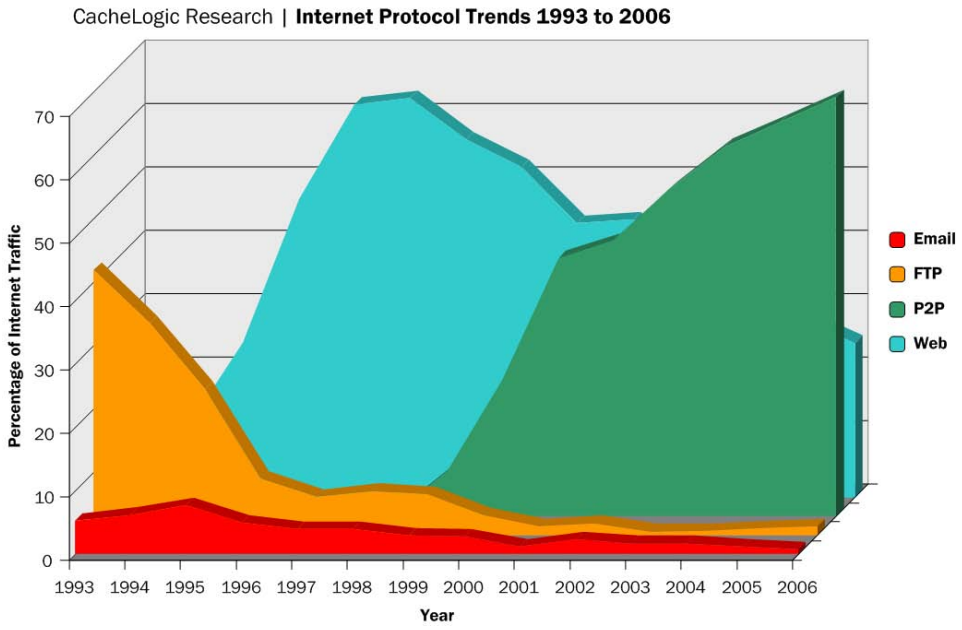
FIGURE 1
TRENDS IN U.S. FILE-SHARING ACTIVITY, 2003-2009




Notes: Bulk traffic is a TCP flow that transferred more than 10MB of data. No date is available for the following weeks: 2/3/03, 7/28/03, 2/23/04, 12/20/04-5/2/05, 7/11/05, 2/27/06-3/27/06, 4/17/06, 5/8/06-10/9/06, 2/19/07-3/5/07, 6/18/07, and 11/19/07.

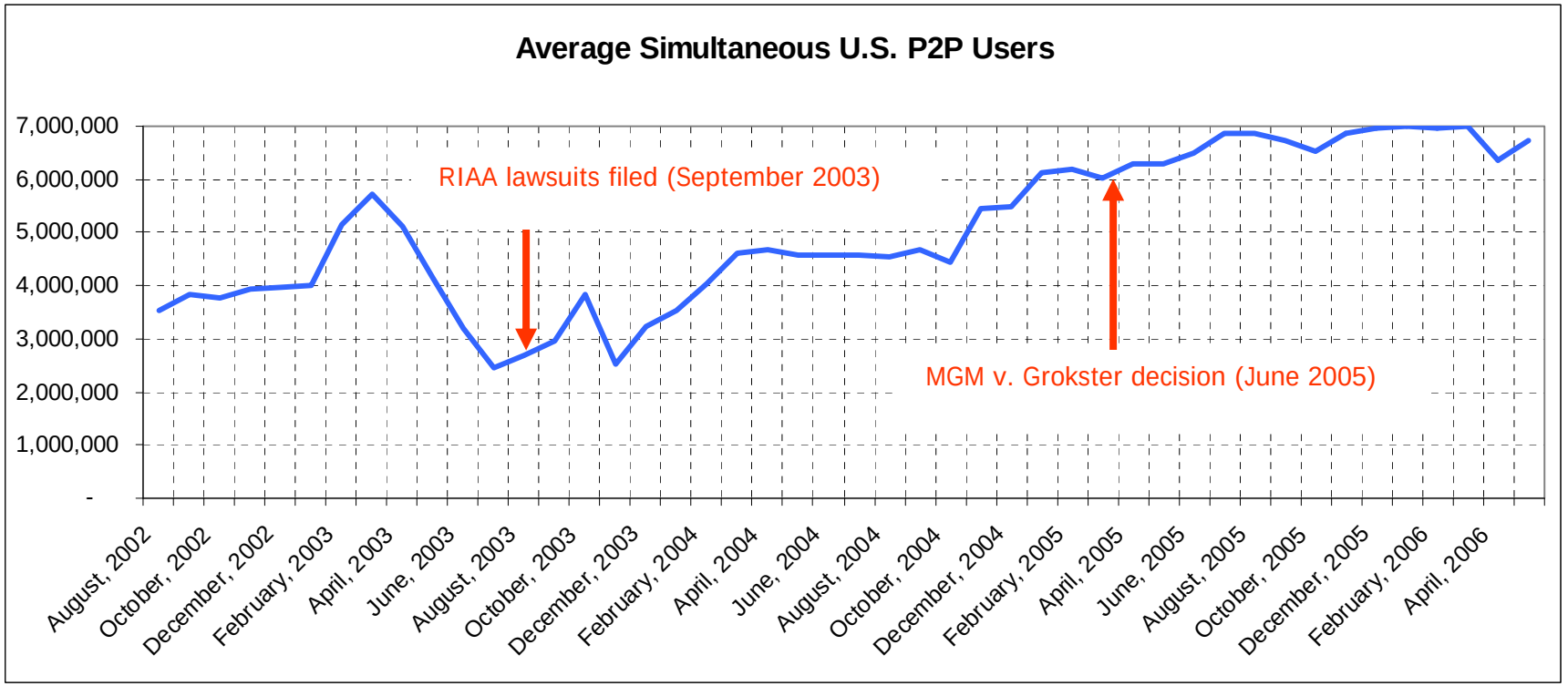
Source: Data from Internet2 Netflow Statistics (2009).

FIGURE 2
GLOBAL FILE SHARING, 1999-2006



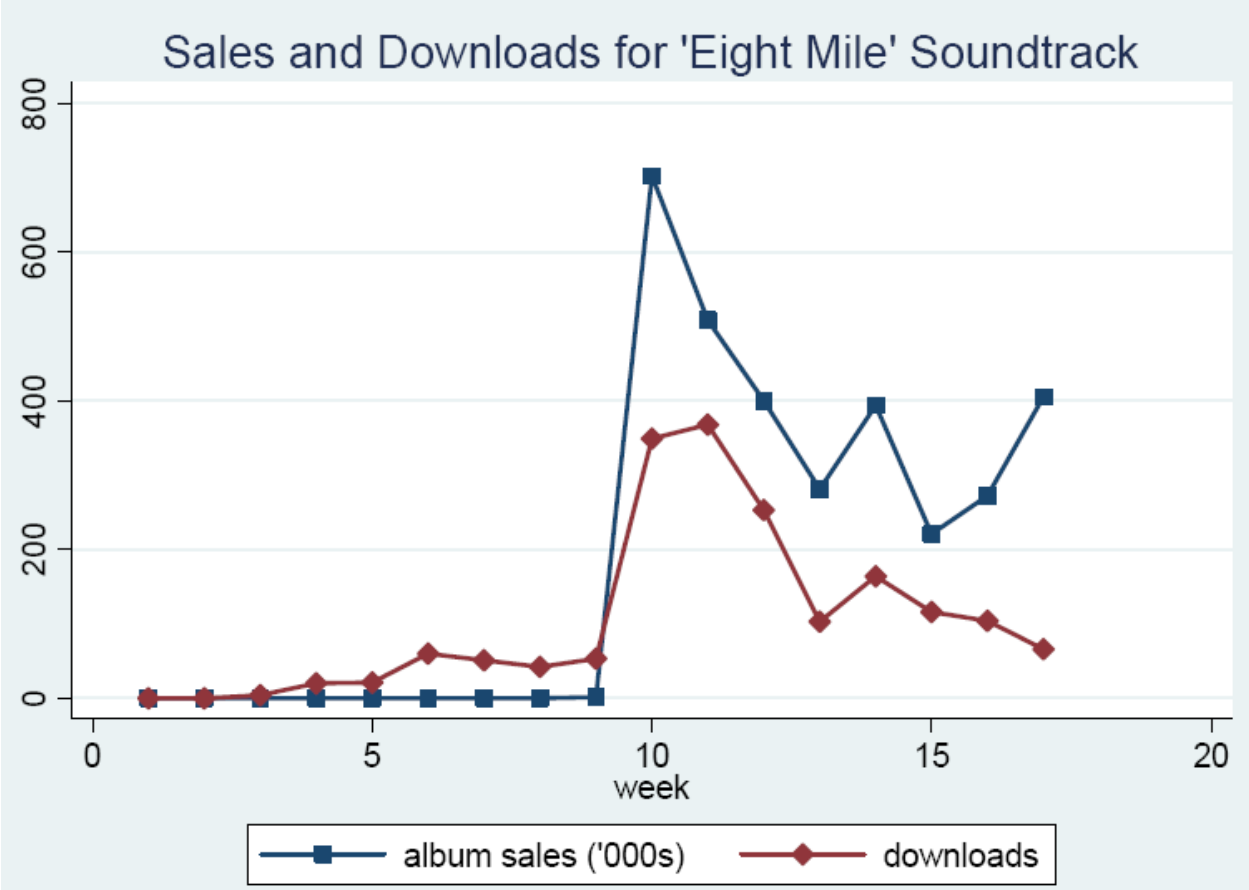

Source: Ferguson (2006)

FIGURE 3
TRENDS IN THE NUMBER OF U.S. FILE-SHARING USERS



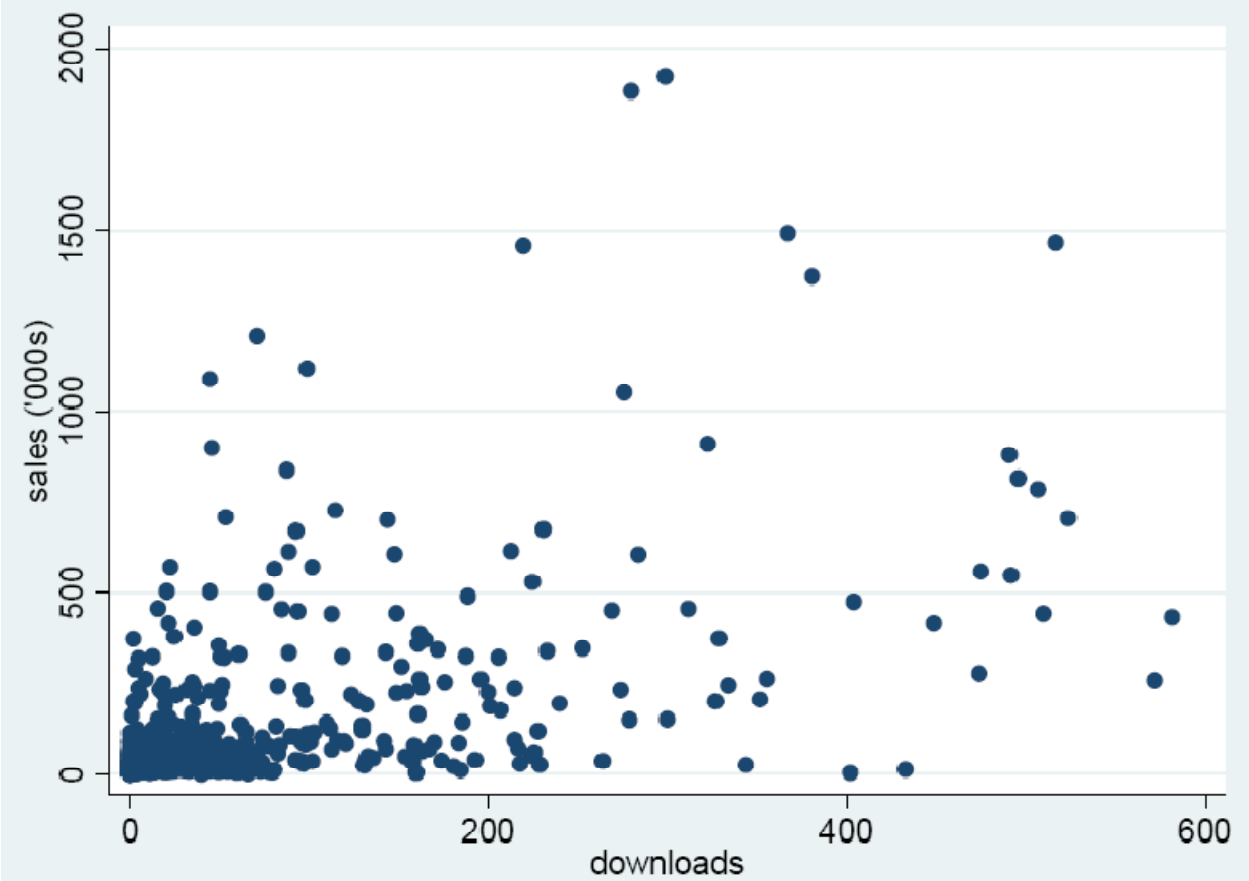

Source: BigChampagne.com

FIGURE 4
INDUSTRY MARKETING AND FILE-SHARING



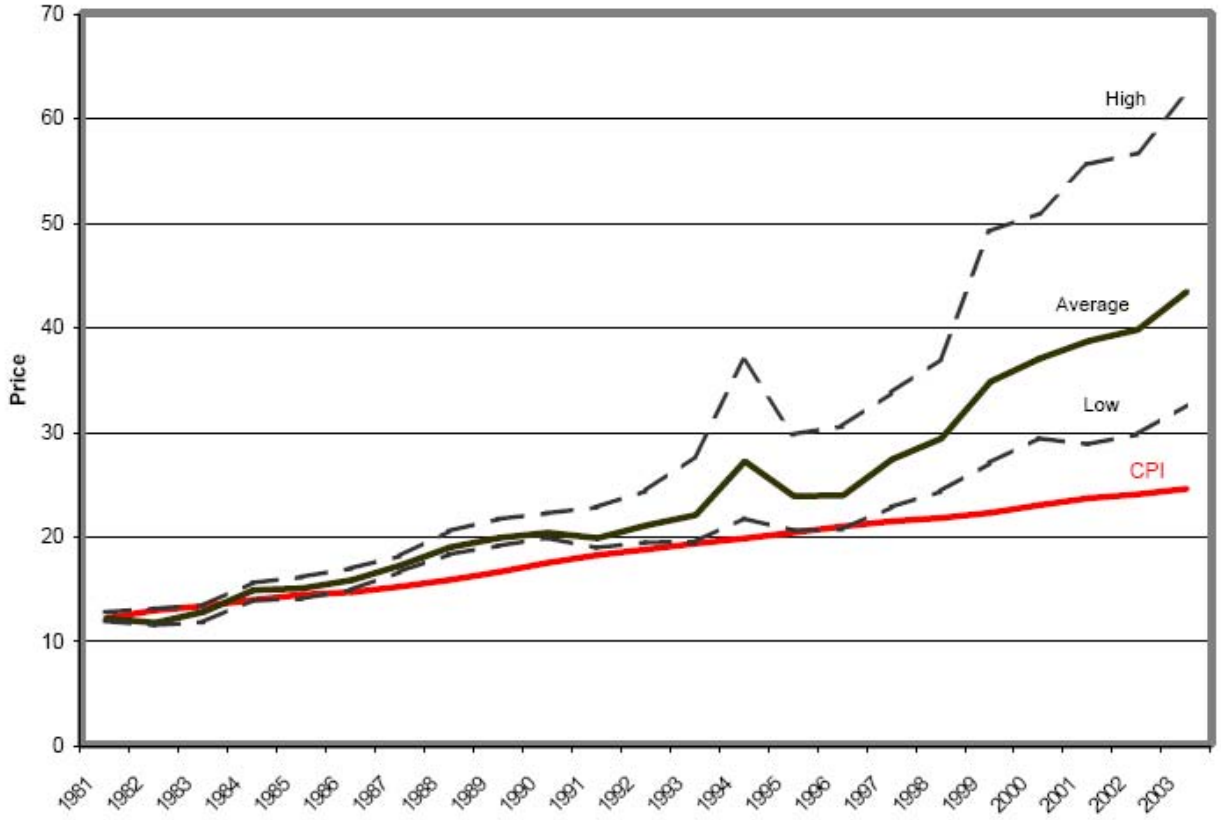

Data from Oberholzer-Gee and Strumpf (2007)

FIGURE 5
ENDOGENEITY OF FILE SHARING



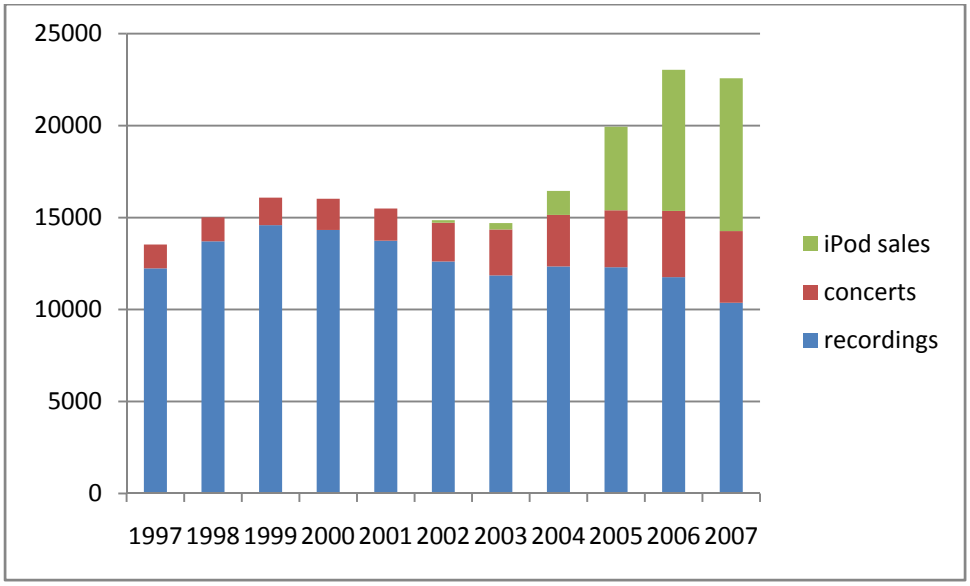

Data from Oberholzer-Gee and Strumpf (2007)

FIGURE 6
CONCERT PRICES 1981-2004




Source: Krueger, 2005

FIGURE 7
U.S. MUSIC INDUSTRY SALES TRENDS




Sources: Recording Industry Association of America, "2007 Year-End Shipment Statistics" (www.riaa.com), Pollstar (www.pollstar.com), Apple, Inc. Annual Reports (www.apple.com), accessed 18 March 2008.

# ATTACHMENT B

# THE EFFECT OF FILE SHARING ON RECORD SALES: AN EMPIRICAL ANALYSIS

FELIX OBERHOLZER-GEE AND KOLEMAN STRUMPF

# The Effect of File Sharing on Record Sales
# An Empirical Analysis*

Felix Oberholzer-Gee
Harvard University

Koleman Strumpf
University of Kansas

* We would like to thank Bharat Anand, Gary Becker, Bob Frank, Shane Greenstein, Austan Goolsbee, Alan Krueger, Steven Levitt, Tom Mroz, Alan Sorensen, Joel Waldfogel, Steven Wildman, Pai-Ling Yin, participants at numerous seminars, and two anonymous referees for helpful comments. This project would not have been possible without the assistance of several individuals and organizations. MixMasterFlame and the FlameNap network shared P2P data with us, and BigChampagne LLC, the CMJ Network, Nathaniel Leibowitz, and Nevil Brownlee generously provided auxiliary data. We thank Keith Ross and David Weekly for assistance in understanding the KaZaA, OpenNap, and WinMX search protocols and database indices. Sarah Woolverton and Christina Hsiung Chen provided superb research assistance. The financial support of the George F. Baker Foundation (Oberholzer-Gee) and the Kenan Faculty Fund (Strumpf) is gratefully acknowledged. We appreciated the aural support from Massive Attack, Sigur Ros and The Mountain Goats.

Abstract
For industries ranging from software to pharmaceuticals and entertainment, there is an intense debate about the appropriate level of protection for intellectual property. The Internet provides a natural crucible to assess the implications of reduced protection because it drastically lowers the cost of copying information. In this paper, we analyze whether file sharing has reduced the legal sales of music. While this question is receiving considerable attention in academia, industry and in Congress, we are the first to study the phenomenon employing data on actual downloads of music files. We match an extensive sample of downloads to U.S. sales data for a large number of albums. To establish causality, we instrument for downloads using data on international school holidays. Downloads have an effect on sales which is statistically indistinguishable from zero. Our estimates are inconsistent with claims that file sharing is the primary reason for the decline in music sales during our study period.

## I. Introduction

File sharing is now one of the most common online activities. U.S. households swap more than 300 million files each month, a figure that has grown by over 50% in the last two years (Karagiannis, Broido, Brownlee, claffy and Faloutsos 2004; Billboard 2006). Sharing files is largely non-rivalrous because the original owner retains his copy of a downloaded file. The low cost of sharing and significant network externalities are key reasons for the dramatic growth in file-sharing. While few participated prior to 1999, the founding year of Napster, in 2006 there were about ten million simultaneous users on the major peer-to-peer (P2P) networks (BigChampagne 2006). Because physical distance is largely irrelevant in file sharing, individuals from virtually every country in the world participate.

There is great interest in understanding the economic effects of file sharing, in part because the music industry was quick to blame the phenomenon for the recent decline in sales. Between 2000 and 2005, the number of CDs shipped in the United States fell by 25% to 705 million units (RIAA 2006). Claiming that file sharing was the culprit, the recording industry started suing thousands of individuals who share files. The industry also asked the Supreme Court to rule on the legality of file-sharing services, a question which critically hinges on the "market harm" caused by the new technology. Congress is currently considering a number of measures designed to counter the perceived threat of file sharing.

While concerns about P2P are widespread, the theoretical effect of file sharing on record sales and industry profits is ambiguous (Bakos, Brynjolfsson and Lichtman 1999; Takeyama 1997; Varian 2000). Participants could substitute downloads for legal purchases, thus reducing sales. The inferior sound quality of downloads and the lack of features such as liner notes or cover art

perhaps limit such substitution. Alternatively, file sharing allows users to learn about music they would not otherwise be exposed to. In the file sharing community, it is common practice to browse the files of others and discuss music in file server chat rooms. This learning may promote new sales. Other mechanisms proposed in the theoretical literature have unclear effects on sales. Individuals can use file sharing to sample music, which will increase or decrease sales depending on whether users like what they hear (Shapiro and Varian 1999). The availability of file sharing could also change the willingness to pay for music – it could either decrease it due to the ever present option of downloading, or it could increase it through network effects and the greater ease of sharing (Takeyama 1994). Finally, it is possible there is little effect on sales. File sharing lowers the price of music, which draws in low-valuation individuals who would otherwise not have purchased albums. Rob and Waldfogel (2006) find in a recent survey that college students value albums they purchased in the store at $15.91. In contrast, respondents' willingness to pay for albums they downloaded was only $10.66, a value below the average purchase price of a CD.

With no clear theoretical prediction, the effect of file sharing on sales is an empirical question.[1] Most of what we know about the effects of file sharing is based on surveys. The evidence is mixed. File sharers generally acknowledge both sales displacement and learning effects, and it is unclear if either effect dominates. Rather than relying on surveys, this study is the first to use observations of actual file-sharing behavior of a large population to assess the impact of downloads on sales. Our dataset includes 0.01% of the world's downloads (1.75 million file

[1]The entertainment industry's opposition to file sharing is not a priori evidence that file sharing imposes economic damages. The industry has often blocked new technologies which later become sources of profit. For example, Motion Picture Association of America President Jack Valenti argued that "the VCR is to the American film producer as the Boston strangler is to the woman home alone" (Congressional Hearings on Home Recording, 12 April 1982). By 2004, 72% of domestic industry revenues came from VHS and DVD rentals or sales (DEG 2005; MPAA 2005). Other examples include the record industry's initial opposition to radio in the 1920s and 1930s and to home taping in the 1980s.

transfers) from the last third of 2002, a period of rapid growth in file sharing. We match audio downloads of users in the United States to a representative set of commercially relevant albums for which we have concurrent weekly sales, resulting in a database of over ten thousand album-weeks. This allows us to directly study the relationship between downloads and sales. To establish causality, we instrument for downloads using international school holidays, a supply shock that is plausibly exogenous to sales. Our instruments are relevant since they have a large impact on file transfer time, which in turn is a key determinant of the number of downloads.

We find that file sharing has only had a limited effect on record sales. After instrumenting for downloads, the estimated effect of file sharing on sales is not statistically distinguishable from zero. The economic effect of the point estimates is also small. When considering the policy implications of these results, it is important to take into account the precision of our estimates. Based on all specifications presented in this paper, even our least precise results, we can reject the hypothesis that file sharing cost the industry more than 24.1 million albums annually (3% of sales and less than one third of the observed decline in 2002). Models that consider the dynamics of file sharing allow us to make more precise statements. For example, if we account for the growth in file sharing during our study period we can reject a null that P2P displaced more than 6.6 million in CD sales or less than 10% of the 2002 decline. We arrive at similar conclusions if we allow the effect of international school holidays to vary by album. Our results continue to hold after permitting downloads to influence sales with a lag, omitting data from the holiday shopping season, and restricting our sample to popular titles. In total the estimates indicate that the sales decline over 2000-2002 was not primarily due to file sharing. While downloads occur on a vast scale, most users are likely individuals who in the absence of file sharing would not have bought the music they downloaded.

Our conclusion is supported by other data and methods of analysis. For instance, in the most recent Consumer Expenditure Survey (2004) for the U.S., households without a computer, who seem unlikely to engage in file sharing, report that they reduced their spending on CDs by 43% since 1999. Quasi-experimental evidence on the long-term effect of P2P on music sales also lead to similar results. For example, we document that the share of sales during the summer months when fewer students have access to high-speed campus Internet connections did not change as a result of P2P. Similarly, sales did not decline more precipitously in the Eastern Time Zone of the United States where P2P users can more conveniently download files provided Europeans. Using several years of data, we also show that the number of P2P users is not correlated with album sales. Finally we document that the recording industry often experiences sales reductions, including a recent episode with a sharper reduction than the current period. These experiments are an important complement to our micro-data results. While the main estimates focus on high-frequency variation over several months, the experiments focus on long-term trends using data spanning several years.

Our results have broader implications beyond the specific case of file sharing. A longstanding question in economics concerns the level of protection for intellectual property that is necessary to ensure innovation (Posner 2005). Economic research on the role of patents and copyrights likely began with the critique in Plant (1934) and continues today in the debate between Boldrin and Levine (2002) and Klein, Lerner and Murphy (2002). We provide specific evidence on the impact of weaker property rights for the case of a single industry, recorded music. The file-sharing technology available in 2002 had markedly lowered the protection that copyrighted music recordings enjoyed, so it is interesting to analyze to what extent this reduced protection adversely affected sales. For our study period, we do not detect a significant impact. The paper

also contributes to a growing literature which studies the interactions between the Internet and brick and mortar economies (Goolsbee 2000; Gentzkow forthcoming).

The outline of the paper is as follows. The next section provides an overview of the empirical literature. Section III describes the mechanics of file sharing, and we discuss our data in Section IV. Next we describe the econometric approach. Section VI presents the results, and the last section discusses the implications of this study.

## II. The Literature

Empirical research on file sharing and record sales has been limited and inconclusive, primarily, we believe, due to shortcomings with the data. Most of what we know about the effect of file sharing on sales is based on surveys. There are numerous industry studies which arrive at a diverse range of conclusions. For instance, Forrester Research (2002) and Jupiter Media Metrix (2002) find neutral or positive effects, while the International Federation of the Phonographic Industry (2002), Edison Media Research (2003) and Forrester Research (2004) document a sales displacement. A general difficulty with these studies is that they compare the purchases of individuals who download files with the purchases of those who do not. While downloaders may in fact buy fewer records, this could simply reflect a selection effect. File sharing is attractive to those who are time-rich but cash-poor, and these individuals would purchase fewer CDs even in the absence of P2P networks.

A handful of academic studies rely on micro data to address the issue of unobserved heterogeneity among file sharers.[2] Rob and Waldfogel (2006) study the survey responses of a convenience sample of U.S. college students. For hit albums which sold more than 2 million

[2] The Journal of Law and Economics published additional papers in a symposium on file sharing in 2006. Oberholzer-Gee and Strumpf (2005) discusses these studies and additional work.

copies since 1999, they find no relationship between downloading and sales. Expanding the set of albums to include all music the students acquired in 2003, downloading five albums displaces the sale of one CD. These results could mean that piracy does not affect hit albums but hurts smaller artists, or it is also possible that file sharing had less of an effect on sales in earlier years. After instrumenting for downloads with the school the students attend – everyone at Penn has broadband access while this is not true for the other schools – the resulting estimates are too imprecise to draw any firm conclusions. Zentner (2006) employs European survey data to study the relation between file sharing and sales. Using measures of Internet sophistication and access to broadband as instruments, Zentner finds some displacement. Unfortunately, neither the Rob and Waldfogel study nor Zentner's work allow inferences about the total impact of file sharing on record sales because neither paper studies a representative sample of file sharers. Zentner also lacks information about the number of downloads and CD purchases.

Our approach differs from the current literature in that we directly observe file sharing. Our results are based on a large and representative sample of downloads, and individuals are generally unaware that their actions are being recorded.

## III. File sharing Networks

File sharing relies on computers forming networks which allow the transfer of data. Each computer may agree to share some files and has the ability to search for and download files from other computers in the network. Our data come from the OpenNap network, an open-source descendant of Napster. OpenNap is an example of a centralized P2P network in which users log on to a central server that tracks all search requests and file downloads. During our study period in the fall of 2002, P2P networks were already quite large. FastTrack (which includes the

popular KaZaA service (see Liang, Kumar and Ross 2004) had grown to 3.5 million simultaneous users by December 2002. The second largest network was WinMX, which had about 1.5 million simultaneous users in 2002. Even the smaller networks were fairly large. OpenNap, the choice of about one percent of all P2P users, had at least 25,000 simultaneous users sharing over 10 million files. Napster no longer operated in the fall of 2002.

## IV. Data

We use two main data sources for this study. Logs for two OpenNap servers allow us to observe what files users download. Weekly album-level sales data come from Nielsen SoundScan (2005). SoundScan tracks music purchases at over 14,000 retail, mass merchant and online stores in the United States. Nielsen SoundScan data are the source for the well-known Billboard music charts. To develop our instruments, we rely on a large number of additional data sources which we discuss in the next section.

### *File Sharing Data*

Our data were collected from two OpenNap servers, which operated continuously for seventeen weeks from 8 September to 31 December 2002. The information on file transfers is collected as part of the log files which the servers generate, and most users are unaware their actions are being observed and recorded. An excerpt of a typical log file is:

```
[2:53:35 PM]: User evnormski "(XNap 2.2-pre3, 80.225.XX.XX)" logged in
[2:55:31 PM]: Search: evnormski "(XNap 2.2-pre3)": FILENAME CONTAINS "kid rock devil"
   MAX_RESULTS 200 BITRATE "EQUAL TO" "192" SIZE "EQUAL TO" "4600602" "(3 results)"
[3:02:15 PM]: Transfer: "C:\Program Files\KaZaA\My Shared Folder\Kid Rock –Devil
   Without A Cause.mp3" (evnormski from bobo-joe)
```

The last two lines in the log file show user "evnormski" downloading the song "Devil Without a Cause" by Kid Rock from user "bobo-joe". Information on downloads are the building blocks of our analysis. We focus on downloads because these are the files users actually obtain and they can potentially displace sales. Over the sample period we observe 1.75 million file downloads, or about 0.01% of all downloads in the world. We restrict the analysis to audio files by users in the U.S. The server logs include the I.P. address for each client which we use to identify our users' home country.

An important question is whether our sample is representative of data on all P2P networks.[3] While we are unaware of any database spanning the universe of music downloads, we were able to compare the data from our servers with a sample of more than 25,000 downloads from FastTrack/KaZaA, the leading network at the time. We find that the availability of titles is highly correlated on the two networks. Using a standard homogeneity test based on 1,789 unique songs, we cannot reject a null that the two download samples are drawn from the same population (Pearson $\chi^2$ statistic is 1824.1). The resemblance of files is not surprising. Individuals in our data are similar to those on the most popular networks because the user experience is quite similar and many individuals employ software which allows them to simultaneously participate on several networks. For example, roughly one third of OpenNap participants uses the WinMX software, which allows them to simultaneously access the two largest networks during our study period. We also find that users on these larger networks and those on our servers have access to a comparable number of files and that network size has little effect on the *distribution* of downloads. Based on these tests, we conclude that our sample is representative of the file transfers on the major P2P networks during our study period.

[3] A more comprehensive discussion of this point is in Appendix A of Oberholzer-Gee and Strumpf (2005).

*Sales Data and Album Sample*

In this study, we focus on a *sample* of albums sold in U.S. stores in the second half of 2002. The sample is representative of all commercially relevant albums, allowing us to draw meaningful inferences about P2P's impact on overall music sales.[4] The sample is drawn from a *population* of albums on 11 charts produced by Nielsen SoundScan (2005): Alternative Albums (a chart with 50 positions), Hard Music Top Overall (100), Jazz Current (100), Latin Overall (50), R&B Current Albums (200), Rap Current Albums (100), Top Country Albums (75), Top Soundtracks (100), Top Current (200), New Artists (150), and Catalogue Albums (200). The charts are published on a weekly basis, and we include an album in the population if it appears on any chart in any week during the second half of 2002. The original population is extensive (2,282 albums) and includes many poorer-selling albums. For instance, our data include two albums which sold fewer than 100 copies during our study period, and the 25$^{th}$ percentile of sales in our data is only 12,493 copies.[5] While we study the commercially most relevant music, it would be incorrect to think of our population as a set of superstar albums. From this population, we draw a genre-based, stratified random sample of 680 releases. To reflect the popularity of different music styles, we set the sample share of a genre equal to its fraction of CD sales in 2002.[6] Within each genre, we randomly select individual titles.

The average album in the resulting sample sold 143,096 copies during our study period. Table 1 reports sales statistics for the full sample and for individual categories. Across all categories,

[4]The genre charts we sample from made up 81.8% of all CD sales in the United States in the last third of 2002. This is virtually identical to the 2002 share of 83.6% for the Big Five record companies, and 97% of the albums on the annual version of these charts were released on RIAA-associated labels.

[5]A typical measure of album success is gold certification which occurs at sales of half a million copies.

[6]Albums can appear on more than one chart because some charts (e.g., New Artists, Top Current) comprise many musical styles. For sampling purposes, we grouped all albums by style; a Rap album on the Top Current list is grouped with all other Rap albums during the sampling process. In the descriptive statistics, we classify albums by their original charts.

44% of population sales are represented in the sample. A two-sample Kolmogorov-Smirnov test comparing the distribution of sales on the original charts and in our sample is unable to reject the null that sample sales are representative of the population of all albums ($p$=0.991). We also reject this null comparing each of our 11 original charts with the sample sales for that particular chart ($p$>0.539 for all 11 charts.)

In order to compare sales and downloads, we match the 260,889 songs which U.S. users successfully transferred during our study period to the 10,271 songs on the 680 albums in our sample. The matching procedure is hierarchical in that we first parse each transfer line, identifying text strings that could be artist names. These text strings are then compared to the artist names in our set of albums. The list of artists contains the name on the cover and up to two other performing artists or producers that are associated with a particular song. For example, the song "Dog" on the B2K album "Pandemonium" is performed by Jhene featuring the rapping of Lil Fizz. For "Dog," B2K, Jhene and Lil Fizz are recognized as artists. Once an artist is identified, the program then matches strings of text to the set of songs associated with that particular artist. Using this algorithm, we match 47,709 downloads in the server log files to our list of songs, a matching rate of about 18%.

There are two reasons why this rate is less than 100%. First, a download may be for a song that is not in our sample. These transfers are not of any concern, they simply reflect the fact that we are working with a sample. A second reason for a match rate of less than 100% could be that our matching algorithm fails to recognize songs. To investigate this possibility, we hand-checked a file with 2,000 randomly chosen unmatched transfers, comparing these downloads against our sample. Only five of the unmatched songs were in our sample. As a result, we believe that the 18% match rate mostly reflects transfers of songs that are not in our sample.

*Descriptive Statistics*

As this is one of the few data sets that allow us to directly observe P2P users, we describe our data in some detail. A first stylized fact is that file sharing is truly global in nature. While over ninety percent of users are in developed countries, a total of 150 countries are represented in the data. U.S. users make up 31% of the sample. Table 2 shows the top countries for users and downloads. As the data indicate, there is only a loose correlation between user share and other country covariates such as Internet use or the software piracy rate. Column 3 in Table 2 confirms that interactions among file sharers transcend geography and language. U.S. users download only 45.1% of their files from other U.S. users, with the remainder coming from a diverse range of countries including Germany (16.5%), Canada (6.9%) and Italy (6.1%).

While file sharing activities are dispersed geographically, only a limited number of songs are transferred with any frequency. Table 3 shows the average song is downloaded 4.6 times over the study period, but the median number of downloads is zero.[7] Although our sample is representative of all commercially relevant music in the second half of 2002, it is striking to see that more than 60% of the songs in our sample are never downloaded. Aggregated up to the album level, users made 70 downloads from the average album in our sample. The most popular album among file sharers (and the second-best seller) has 1799 downloads, while the median number of downloads per album is 16, the 75th percentile is 63, the 90th percentile is 195, and the 95th percentile is 328. Both downloads and sales closely follow a power-law (pareto) distribution.

File sharing is limited to a select number of songs and most of these songs come from just a few charts. Table 3 shows that songs on the Top Current chart ("Billboard 200") are most frequently

[7]The 75th percentile of downloads per song is 2, the 90th percentile is 11, and the 95th percentile is 22.

downloaded. Downloads from this chart alone make up 48% of all file transfers. Another 25% come from the "Alternative" category. The remaining 9 charts are not particularly popular among file sharers. In view of the low cost of sharing and sampling music on P2P, one could expect users to seek out a great variety of songs representing many musical styles. But this is not the case. P2P downloads closely resemble the play lists of Top 40 radio stations. As a result, it is not surprising that songs from higher-selling albums are downloaded more frequently (Table 4). In the top quartile of sales, albums average 200 downloads. In the bottom category, the mean number of downloads is only 11. This suggests that common factors drive downloads and sales, which is a key concern for the development of our empirical strategy.

## V. Empirical Strategy

*Econometrics*

Our goal is to measure the effect of file sharing on sales. We observe sales and downloads at the album-week level for seventeen weeks. These panel data allow us to estimate a model with album fixed effects,

$$S_{it} = X_{it}\beta + \gamma D_{it} + \omega_s t^s + \nu_i + \mu_{it}. \qquad (1)$$

$i$ indicates the album, $t$ denotes time in weeks, $S_{it}$ is observed sales, $X_{it}$ is a vector of time-varying album characteristics that includes a measure of the title's popularity in the U.S., $D_{it}$ is the number of downloads for all songs on an album, and $\omega_s$ controls for time trends (a flexible polynomial or week fixed effects). The key concern in our empirical work is that the number of downloads is likely to be correlated with unobserved album-level heterogeneity. As the

descriptive statistics suggest, the popularity of an album is likely to drive both file sharing and sales, implying the parameter of interest $\gamma$ will be estimated with a positive bias. The album fixed effects $v_i$ control for some aspects of popularity, but only imperfectly so because the popularity of many releases in our sample changes quite dramatically during the study period.

We address this issue by instrumenting for $D_{it}$ in a 2SLS model. Valid instruments $Z_{it}$ predict file sharing but are uncorrelated with the second-stage error $\mu_{it}$. As in the differentiated products literature, where the problem is correlation between prices and unobserved product quality, we use cost shifters to break the link between unobserved popularity, downloads and sales. An advantage of our instruments, which we discuss below, is that they do not rely on the common but potentially problematic assumption that product characteristics are exogenous (Nevo 2001).[8]

*Instruments*

Our most important instrument is the number of German secondary school kids who are on vacation in a given week. German users provide about one out of every six U.S. downloads, making Germany the most important foreign supplier of songs.[9] German school vacations produce an increase in the supply of files and make it easier for U.S. users to download music.[10] During holidays German teens can spend more time trading music online, since they do most of their file sharing at home (Niesyto 2002). School vacations also allow the German kids to stay up later, which means they can engage in file sharing during the peak U.S. trading hours (early evening, EST). Supporting this intuition, we find that the number of German kids on vacation is

[8]Appendix B of Oberholzer-Gee and Strumpf (2005) presents a formal model of purchase and download behavior which is the foundation for our econometric approach. In particular it shows why we can use linear demand equations rather than the more complicated transformations which are typical in this literature (Berry 1994; Bresnahan, Stern and Trajtenberg 1997).

[9]The important role of German file sharing users is documented in the authoritative BigChampagne database (OECD 2004). Oberholzer-Gee and Strumpf (2005) provides intuition on why this connection is so strong.

[10]Appendix C of Oberholzer-Gee and Strumpf (2005) shows German users are always net suppliers to file sharing networks, and this effect is accentuated during weeks when many kids are on vacation.

a significant predictor of the number of files uploaded from Germany to the United States ($p$=0.011). The effect is particularly large for music genres that are popular in Germany.

For German vacations to be a valid instrument, they must not be directly related to U.S. music demand. This seems likely because the vacation variable varies over time for reasons that are specific to Germany. The sixteen German Bundesländer (states) start their academic year at different points in time to smooth the demand for the German tourism industry and avoid traffic jams (Kultusministerkonferenz 2002). For example, Bavarian students were still on summer vacation during the first week of our study period while Rheinland-Pfälzer kids were already back in school (see Figure 1). A second difference to a typical U.S. vacation schedule is that many, but not all Bundesländer grant their students one or two weeks of fall vacation. In Rheinland-Pfalz, this happened in weeks 4 and 5. Bavaria, in contrast, did not schedule a longer fall recess. These länder-specific holidays move from year to year. A Bundesland with early summer vacations in one year is given a later slot in the following year (Agentur Lindner 2004). As we explain in greater detail below, there are additional reasons to believe this variable is exogenous. If file sharing were eliminated tomorrow, German school holidays would have no relation to U.S. record sales.

We create three additional instruments by interacting the German-kids-on-vacation variable with album-specific characteristics. These instruments are particularly useful because they vary across both time and albums and provide identification even if a full set of week and album fixed effects is included.

*German-kids-on-vacation × band is on tour in Germany:* Tours spur local interest and sales of an album, and they are likely to create a positive supply shock of downloadable files. This instrument is not directly related to U.S. sales because the promotional effect of tours will not

spill across the Atlantic and because the timing of fall and winter concerts in Germany typically reflects idiosyncratic features like venue availability and weather. We expect the effect of German vacations to be even larger if an artist happens to be on tour in Germany that week.

*German-kids-on-vacation × indicator for misspellings in song titles:* To download a song, a user's search query must match a shared file. At the time of our study, file sharing programs were rather rigid in determining matches.[11] Unless both the searcher and sharer agree on the naming convention, no match will occur. This two-sided search problem suggests that songs with unconventionally spelled titles may be more difficult to find. We use MS Word's spell checker to determine if an album has any song titles with an unconventional spelling. We expect misspellings to reduce the size of the positive supply shock coming from German vacations.

*German-kids-on-vacation × rank of album on German charts:* Songs from popular albums in Germany are easier to download because the supply of these files is larger. Our measure for German popularity is the rank of the album on the weekly German Top 100 chart (Musikmarkt 2002). Obviously, there is a concern that these chart positions might also measure U.S. popularity. However, the instrument is included along with album fixed effects, so it is the timing of the chart rankings in Germany that identifies downloads. There are important differences in the dynamics of song popularity in the two countries due to taste differences and differences in album release dates.

For all our instruments, we provide additional evidence for their exogeneity in the following sections. Summary statistics for the instruments are in Table 5. Each measure exhibits noticeable variation.

[11]For example, "lose yourself," the name of a popular song, would typically return over a thousand results, but mistyping even one character (such as "lose yourse;f") or omitting part of a word ("lose yours") returned zero results.

*Mechanisms Underlying the Main Instruments*

Our analysis presumes that each instrument influences download costs, and that these costs impact the number of file transfers. We test this idea by analyzing more detailed server log files which allow us to calculate the download time and success rate of download attempts. We construct five measures of download costs: the time between a download request and the successful initiation of the download ($C_1$), the time between a search request and a download request ($C_2$), the time between the initiation of the download and its successful completion ($C_3$), the ratio of search requests to the number of successful downloads ($C_4$), and the percentage of failed or canceled download requests ($C_5$). Each $C_i$ term captures aspects of delay or frustration which a U.S. downloader might experience. The measures are aggregated up to the album-week. For example, $C_1$ is the average time until download initiation among all observed requests for that album in a particular week.

Mean $C_i$ values are presented in the last row of Table 6. The first three columns show that the typical file takes twenty minutes to download, starting from the initial search until the transfer is complete.[12] There are also long delays for top-selling albums, suggesting there is an ubiquitous scarcity of supply. While slow download speeds are the norm in our data, the estimates in Table 6 show that searching and downloading audio files in the U.S. is considerably easier when a larger number of German school children are on vacation. This reduction is even larger when the artist is on tour and when the album is highly ranked on the German charts.[13] The misspellings interaction significantly increases the time between a search and a download request as well as the number of unfulfilled downloads ($C_2$, $C_4$, $C_5$), but it has little effect on the time it

[12]Gummadi, Dunn, Saroiu, Gribble, Levy and Zahorjan (2003) independently document these long download times. This likely reflects the fact that only a third of the U.S. users in our data had a broadband connection.

[13]Note that the German tour and singles chart variable parameters are identified using only within album variation since fixed effects are included. This mitigates concerns that album popularity in the U.S. is driving the parameter estimates.

takes to transfer a file ($C_1$, $C_3$). This is consistent with the argument that misspellings create confusion, though they do not slow down the file transfer itself. The estimated effects on download times are economically significant. For example, a one standard deviation increase in the German vacation variable implies a 1.25 minute reduction in the time for a download to begin ($C_1$), which is an eighth of the typical delay..

These results are meaningful only if the cost of downloading influences the number of file transfers. This is not obviously true because P2P users can engage in other activities while files are being downloaded, which could mean they are insensitive to the time cost of file sharing. To check if the variation in download time that is due to our instruments has a significant impact on the number of transfers, we estimate the system

$$\begin{aligned} C_{it} &= Z_{it}\delta + v_i + \mu_{it} \\ D_{it} &= C_{it} + v_i + \varepsilon_{it} \end{aligned}, \qquad (2)$$

where $Z_{it}$ is the full list of instruments and $C_{it}$ denotes total download time ($C_1$+$C_2$+$C_3$). The last two columns of Table 6 shows that P2P users are fairly sensitive to the time cost of file sharing: a one standard deviation increase in download time reduces downloads by almost half of their mean. We find similar effects when we separately estimate equation (2) for each of the five $C_i$ terms. These estimates confirm our initial claims. German vacations influence the cost of downloading, and this effect has an important impact on the number of downloads in the U.S.[14]

*Specific Concerns with Individual Instruments*[15]

[14] A different approach to show that German vacations influence downloading activity is to look at international data. We find that school holidays have an important effect only in countries whose time zones are complementary to Germany's. Appendix C of Oberholzer-Gee and Strumpf (2005) presents this point in detail.

[15] A general concern is that the instruments are based on high frequency variation in download costs. Unfavorable conditions might lead users to simply defer downloads to a later time, in which case our second stage estimates will be attenuated to zero. Oberholzer-Gee and Strumpf (2005) shows this concern is not warranted, since users are impatient and quickly lose interest in an album.

*German-kids-on-vacation:* A potential difficulty with the vacation variable is that it might be correlated with time-varying album popularity in the U.S. We perform a number of tests to see if this is the case. First, we check if German vacations happen to coincide with official U.S. holidays. We find that there is little overlap.[16] A second possibility is that German school vacations proxy for American vacations which are likely to have a direct impact on music sales. As there is no centralized data on holidays for all 14,000 U.S. school districts, we collect information on the number college students who are out of school during our study period. The sample includes all schools in the top two tiers of U.S. News and World Report's 2002 ranking. Information on school breaks is available for 157 schools, leaving us with data for 2.17 million students, almost a quarter of all U.S. college students. Figure 1 compares the vacation patterns in Germany and the U.S. There are marked differences. When some German kids are off in early fall, U.S. students are mostly in school. During the Thanksgiving break in the U.S., German kids are in school. Both populations are off during the Christmas break, although the break starts earlier for U.S. students. To test more formally if the number of German kids on vacation proxies for the number of U.S. kids, we include the latter in the first stage of equation (1). We find no evidence that the measured effect of German vacations on American music downloads is mediated by U.S. vacations.[17]

In a final test, we check more directly if the German vacation variable is in fact uncorrelated with U.S. demand for music albums. We do this by interacting the instrument with an album's rank on the U.S. MTV charts.[18] MTV rankings have the advantage that videos are often shown prior

[16] Estimates over our 17 week observation period yield: $\text{US Holidays}_t = 1.148\ (1.61) - 0.182\ (0.16) \times \text{German Kids}$, where US Holidays is the number of official American holidays (such as Columbus Day or Thanksgiving) in week t and German Kids is the German holiday instrument.

[17] Controlling for the entire set of instruments, the estimated effect of German vacations on downloads changes from 0.667 (0.054) without the U.S. students-on-break variable to 0.643 (0.057) with this variable.

[18] We thank one of our referees for this suggestion. We also used the Billboard Airplay ranking to explore these effects, with similar results.

to the release of a CD, at a time when songs from a forthcoming album first appear on file-sharing networks. This interaction is included in both stages of equation (1).

$$\begin{aligned} D_{it} &= X_{it}\beta + Z_{it}\delta + \varphi_1 Gkids_t \times MTV_{it} + \omega_{1s}t^s + \nu_i + \varepsilon_{it} \\ S_{it} &= X_{it}\beta + \gamma\,\hat{D}_{it} + \varphi_2 Gkids_t \times MTV_{it} + \omega_{2s}t^s + \nu_i + \mu_{it} \end{aligned}, \qquad (3)$$

where $Z_{it}$ is our full set of instruments. As required under our assumptions, $\varphi_1$ is positive: German vacations have a larger effect for files that are more popular in the U.S. In the second stage, however, $\varphi_2$ is economically small and statistically insignificant. When an album becomes more popular in the U.S., this boost in popularity is not directly related to German vacations, supporting our claim that the holiday shocks are exogenous.

A second concern is that Germans supply only a narrow slice of music that is of interest to U.S. file sharers. If those who like the type of music that Germans make available substitute downloads for purchases in an atypical fashion, we measure a local average treatment effect, not a true population effect (Imbens and Angrist 1994). Fortunately, there is substantial overlap between American and German musical tastes. Of the albums that entered our sample via the Billboard 200, 62.65% are also on the top 100 German charts. More generally, we study Amazon rankings to compare sales ranks in the two countries (Goolsbee and Chevalier 2003). With the exception of Latin and Country music, Wilcoxon matched-pairs signed-ranks tests cannot reject the null of equal distributions for the eleven genres in our sample. In the robustness section of the paper, we test if the undersupply of Latin and Country music affects our estimates. We show that this is not the case, suggesting the measured effect of downloads on sales is likely to be a good estimate of the average population effect.

*German-kids-on-vacation × indicator for misspellings in song titles:* Because misspellings appear to be more likely in some genres than in others, one might argue that this indicator is

likely to proxy for album popularity. In our application, this concern is not valid for two reasons. First, as an empirical matter, we find that misspellings are not correlated with sales, even in models without album or genre fixed effects.[19] Second, all our specifications presented in the results section include album fixed effects which control for an album's time-invariant popularity.

A second difficulty with the misspelling instrument could be that misspellings cause our song matching algorithm to fail. This would result in a negative relationship between misspellings and measured downloads, even if misspellings had no effect on actual downloads. More importantly, the second-stage estimates would be attenuated towards zero, since the variation in fitted downloads would be largely due to noise. Several pieces of evidence suggest this is not true. First, the estimates in the last sub-section show that misspellings do in fact have real effects on transfer times and user behavior. Second, we can check for misspellings in unmatched downloads. If the criticism is correct, there should be more misspellings in the unmatched than in the matched sample. This is not the case.[20]

*German-kids-on-vacation × rank of album on German charts:* The idea underlying this instrument is that vacation periods in Germany will boost downloads in the U.S. more when many German users make a particular file available. Because the instrument is included along with album fixed effects, it is the timing of the chart rankings in Germany that identify downloads. However, if U.S. popularity shocks happen to coincide with high German chart positions, we would measure the effect of downloads on sales with a positive bias. We can test for this spurious correlation in two ways. First, assuming that the German vacation variable is a

[19] The effect of misspellings on sales is statistically insignificant and economically small. A one-standard-deviation increase in misspellings raises sales by a mere 11,000 copies (less than ten percent of the mean) during our entire study period.

[20] The rates are 0.041 (N=35614) and 0.038 (N=7163), in the unmatched and matched samples respectively. The Pearson $\chi 2$ statistic is 1.402.

valid instrument, we can perform overidentification tests for this and the other interactions that we use as instruments. These tests, reported in the results section of the paper, provide no indication that any of our instruments are invalid. A second and more direct test is to see whether shocks in U.S. demand are correlated with German popularity.[21] Under our hypotheses, U.S. demand shocks must not get magnified when albums become more popular in Germany. For example, we expect U.S. vacations to increase P2P activity, but this increase must not vary with German popularity. The model is,

$$D_{it} = Z_{it}\delta + \varphi_1 Ukids_t + \varphi_2 Ukids_t \times Gcharts_{it} + \varphi_3 Ukids_t \times MTV_{it} + \varphi_4 Gkids_t \times MTV_{it} + \omega_s t^s + \nu_i + \varepsilon_{it} \quad (4)$$

$Ukids_t$ denotes the number of U.S. college students on break (our measure of U.S. demand shocks), $Gcharts_{it}$ is a title's rank on the German charts, and $MTV_{it}$ is the position on the MTV chart (our measure of U.S. popularity). The effect of interest in this specification, $\varphi_2$, shows whether a shock in demand in the U.S. is mediated by German popularity. This is not the case: $\varphi_2$ is -0.0008 with a standard error of 0.0134, and this effect is only one tenth of the size of the German kids × German chart interaction in our later specifications. The data show that relative popularity in Germany interacts with German but not with U.S. vacations.

## VI. Results

Before turning to the estimates, it is instructive to graph some of the data.. Figure 2 shows the weekly time series of sales and purchases for one of the most popular albums in our sample. This "Superstar" album was largely ignored in file sharing networks until it became available for sale in week ten of our sample. This suggests it is the publicity associated with an official

[21] We thank one of our referees for this suggestion.

release which drives downloads as well as sales. Notice also the rapid but non-monotone decay in sales and downloads, which highlights the importance of using high-frequency data.

*Panel Analysis*

In Table 7 we report results for equation (1). The unit of observation is the album-week. The models include a control in both stages for time-varying U.S. popularity, the album's position on the American MTV charts, and a polynomial time trend of degree six. As expected, a simple OLS specification yields a large positive effect of 1.093 with a standard error of 0.023. A model which adds album fixed effects is given in column (1). While we continue to find a positive effect of downloads on sales, the relationship is now much weaker. The remaining estimates in Table 7 instrument for downloads. We begin by using the number of German kids on school vacation (column II). The first-stage estimates imply that a one standard deviation increase in the number of children on vacation boosts weekly album downloads by slightly more than one half of their mean,, an effect that is statistically significant and economically meaningful. Once we instrument for downloads, the estimated effect of file sharing on sales is small and statistically indistinguishable from zero.

We next consider specifications in which we add the band-on-tour-in-Germany interaction and the remaining time-varying instruments (columns III and IV). The tour and the German-chart interactions are of particular interest since they vary across albums as well as over time and provide an additional source of identification. The instruments have the expected first-stage signs. Tours and better chart positions magnify the effect of German students on vacation. The reverse is true for misspellings, which make it more difficult to search for files. Sargan

overidentification tests are reported at the bottom of the table. In these richer models downloads continue to have economically small and statistically insignificant effects on sales.

To help improve the precision of our second-stage estimates, in column (5), we allow the effect of the German vacation instrument to vary by album. The logic for including these interactions follows from the same arguments used for the other instruments. When German kids spend more time on P2P networks, the resulting supply shock will vary across albums because the students supply the files that happen to be popular in Germany at the time of the shock. As before, we face a potential problem with using this type of variation: If it so happens that the exogenous German shock is spuriously correlated with album-specific surges in popularity in the U.S., our estimates would be biased. The specification in column (5) addresses this issue in four ways. As before, we include album fixed effects to make sure it is the timing of the supply shocks that identify downloads. Second, we introduce album-specific U.S. popularity effects at both stages of the model by interacting the MTV variable with the album fixed effects. The model thus controls for changes in the U.S. popularity of a release. Third, relying on the assumption that the number of German kids on vacation is a valid instrument, we conduct overidentification tests in a specification that includes only two instruments: the vacation variable and one of the vacation × album-fixed-effect interactions. There are 680 such tests. To err on the side of caution, we exclude from the final specification all interactions whose overidentification tests cannot reject the null at a significance level of greater than 0.20. There are 21 such interactions. Fourth, we estimate a variant of equation (3), now with German kids × album fixed effect × U.S. MTV interactions. In the sales equation, these interactions are individually and collectively not different from zero.

Column (5) in Table 7 reports results with the album interactions. Our instruments retain their statistical significance.[22] The mean of the coefficients on the vacation-album-fixed-effect interactions is -1.143, leaving the average effect of vacations on downloads almost unchanged from the earlier specifications. Grouping the album interactions by genre, we find that vacations increase downloads the most for music types that are popular in Germany: the mean of the vacation-album-fixed-effect coefficients is -0.71 for International albums and -0.91 for Rock. In contrast, the effect of vacations is much smaller, but still positive, for genres that are less popular in Germany (the mean interactions are -1.52 for Latin music, -1.54 for Country, and -1.57 for Holiday music.) At the second stage, the estimated effect of downloads on sales is virtually unchanged in this specification, but the standard error drops considerably.

To see if our results are driven by our modeling choice for the time trend in downloads and sales, we replace the polynomial time trend with week fixed effects in columns (6) and (7) of Table 7. In these specifications, we lose the German-kids-on-vacation instrument because it does not vary across releases. The results remain similar, with more precise second-stage estimates when we allow the effect of vacations to vary by release (column VII).

Table 7 suggests file sharing had a surprisingly small effect on sales that is statistically indistinguishable from zero. The instrumented point estimates fall within a very narrow range and suggest that file sharing did not heavily impact the music industry as a whole. If file sharing were to be eliminated, the most negative estimate (column VI) implies industry sales for all of 2002 would increase by 6.5 million albums. Using the most positive estimate (column VII), industry sales would fall by 8.9 million copies.[23] In 2002, the industry sold 803 million CDs.

[22] The vacations × misspellings interaction is collinear with the vacations × album fixed effects and cannot be included in this specification.

[23] The impact is the difference between predicted sales and the fitted value when downloads are set at zero. Using equation (1), the summed impact for our album sample and for our 17 week observation period is $\sum_t \sum_i S_{it}(D_{it}) - S_{it}(0)$

The robustness of these results extends to specifications not reported in Table 7. For example, we arrive at the same conclusions if we omit the misspelling or the German rank instrument.

*Dynamic Analysis*

The models in Table 7 only allow for a contemporaneous effect of downloads on sales, but it is quite possible that downloads influence sales at a later point in time. For example, users might sample music which they consider buying in the future. In Table 8, we address this issue by studying the effect of several weeks of downloads on sales and by estimating Generalized Methods of Moments (GMM) models.

A difficulty with the first approach is that downloads are highly correlated across time, which prevents us from including downloads in past weeks as individual covariates. Instead, we study the effect of a weighted sum of current and past downloads on current sales. Downloads are instrumented using the core set of instruments (specification IV in Table 7) or the extended set (specification V). Our formal measure is the weighted stock of current and previous weekly downloads, $D_t^{Stock} = \sum_{s\geq 0}\delta_s \times D_{t-s}$.[24] In these models, we continue to find small and statistically insignificant effects for the weighted sum of three weeks of downloads, both in specifications with a polynomial time trend (Table 8, I&II) and with week fixed effects (III&IV). As in the panel results, standard errors drop significantly with the extended set of instruments (II&IV). We also constructed stock variables for the sum of downloads during the past four and six weeks and found no evidence of a sales crowd-out in these models.

---

$= \gamma \times \sum_t \sum_i D_{it}$. We multiply this number by a scaling factor to get the annual impact for the entire music industry, $\gamma \times 240$m (this calculation is described in more detail below Table 11).

[24] The weights $\delta_s$ are chosen in a grid search that minimizes the unexplained fraction of the variance in our sales equation subject to $\delta_s \geq \delta_{s+1}$. The optimal weights $(\delta_0, \ldots, \delta_T)$ are (1,0.1,0.1). It is interesting that the weights which best fit our data give much importance to downloads in the current week, while downloads further back in the past do not heavily influence sales. Oberholzer-Gee and Strumpf (2005) presents additional results showing that file sharers are impatient. These findings are consistent with those of Einav (2004) for movie consumption.

Models (5) and (6) in Table 8 use the GMM estimator developed by Arellano and Bond (1991). The GMM models are more general than the previous specifications in the sense that we do not need to make any assumptions about the appropriate lag structure. The lag of sales that is included on the right-hand side accounts for any effect that past downloads might have had on current sales. The model is estimated in first differences. We instrument for past sales using suitable lags of their own levels and our core set of first-differenced instruments.[25] Arellano-Bond tests for autocorrelation are applied to the first-difference equation residuals. Second-order autocorrelation would indicate that some lags of the dependent variable which are used as instruments are endogenous, but the tests reveal no such problem. The results of these models, with a polynomial time trend as in (5) or with week fixed effects as in (6), are similar to our previous findings. The estimates are fairly precise, making these GMM models an alternative to using our extended set of instruments.

*"Drop-out" Hypothesis*

A possible explanation for our inability to find a statistically significant relationship between file sharing and sales is that file sharers and consumers who purchase music are in fact two separate groups. According to this hypothesis, growth in file sharing does displace sales but we cannot identify this effect because our data do not reflect the increasing number of file sharers.

There are three responses to this conjecture. First, it is inconsistent with what we know about consumer behavior. The premise underlying the "drop-out" hypothesis is that file sharers no

[25]The formal model is,

$$S_{it} = \alpha S_{i,t-1} + X_{it}\beta + \gamma D_{it} + \omega_s t_s + \nu_i + \mu_{it}.$$

The lagged sales term soaks up any delayed effect of downloads, regardless of how far in the past they occurred (taking a Koyck transformation yields a specification with infinite lags of downloads on the right hand side). Estimating in first differences purges the album fixed effects. We instrument for the first-differenced $S_{i,t-1}$ which are now endogenous.

longer buy CDs. However, every survey we are aware of, including the industry studies listed in the literature section, indicates that downloaders, even heavy ones, continue to purchase legal CDs. We corroborated these findings with our own survey of individuals who were engaged in file sharing (Oberholzer-Gee and Strumpf 2005). Ninety percent reported that they recently purchased a CD, a value reaching one hundred percent among the most active downloaders.

Secondly, we can test the "drop-out" hypothesis directly by controlling for the increasing number of users. An implication of the hypothesis is that our download sampling rate declines over time because the servers for which we have data handle a limited number of users. Growth in file sharing, however, is managed by additional server capacity which we do not observe. If we accounted for this growth, the hypothesis suggests, we would find a displacement effect because the "drop-outs" are replacing purchases with transfers. We address this issue by scaling up the number of downloads in our sample to reflect the growth in file sharing. We use the number of FastTrack/KaZaA users as a proxy for the rate of growth.[26] Because the number of users increased by over a third over our observation period, we should be able to detect a drop-out effect if it exists. Table 9 reports these estimates for three panel models, three models using a stock of previous downloads, and for two GMM models. In all these specifications, downloads still do not have a significant effect on sales. A third approach to testing the drop-out hypothesis is to compare the long-run sales growth of individual genres of music. We return to this point in Section VII.

[26] We use 22 data points on the number of KaZaA users in the period from 9/9/2002 to 2/4/2003 to fit a fractional polynomial trend in the number of users. The model explains 85% of the variation.

*Robustness Tests*

To further corroborate our results, we perform a large number of robustness checks, some of which we report in Table 10.[27] The tests fall in three broad categories: models for subsets of our sample, alternative econometric specifications, and models that allow the effect of file sharing on sales to vary by popularity. We first investigate the importance of the holiday season when many consumers purchase CDs as gifts. It is possible that downloads are less substitutable for sales during this period due to the reluctance to give downloaded music as a present. Note that this is also an argument against the idea that file sharing is the main cause of the sales decline, since purchases are heavily concentrated in the holiday season. Still, it is straightforward to test for this effect. In Table 10, we exclude the December data from our sample. We report these results for specifications IV, VI and VII of Table 7. Even without the December data, there is no statistically significant effect of file sharing on sales. In a second test, we omit albums that are not downloaded during our study period. These less popular releases might have little sales even in the absence of file sharing, making the effect of P2P on sales miniscule by definition. Omitting these albums, however, does not change our conclusions. The same holds if we restrict our sample to better-selling albums.

We next test if the undersupply of Latin and Country music influences our estimates. Recall from Section V.D. that this would cause a problem only if the substitutability of downloads and album purchases varies across music genres. The last specification in the first panel of Table 10 re-estimates our models without Latin or Country releases. As expected, this increases the effect of vacations on downloads, from a coefficient estimate of 0.667 in model IV of Table 7 to 0.744 in this model. However, the measured effect of downloads on sales remains similar, a finding

[27]We thank our referees for suggesting several of these points. Many additional robustness tests can be found in Oberholzer-Gee and Strumpf (2005). This working paper also presents pooled specifications utilizing only cross-album variation, and these estimates also show file sharing has little impact on sales.

that is consistent with the idea that the substitutability of downloads and purchases is roughly similar across genres.

In the second panel in Table 10, we explore two alternative specifications. To reduce the importance of outlier albums with a large number of sales, we use log(sales) as the dependent variable. The impact on sales continues to be insignificant in all three specifications. In the next model, we first-difference both sales and downloads and express them as percentage changes. An advantage of this model is that it nicely captures album-specific trends in popularity. Unfortunately, this advantage comes at the cost of a reduced number of observations due to the first-differencing and the weeks with zero downloads or sales. Using our core set of instruments, we now find a positive and statistically significant but economically small effect of downloads on sales. However, the estimated coefficient drops considerably and is insignificant when we introduce week fixed effects.

The previous models constrained the effect of downloads on sales to be identical for all releases. In the bottom panel of Table 10, we relax this assumption. We first explore the idea that the effect varies by artist popularity. We do this by interacting the download variable with two measures of popularity: an artist's last and his best-ever Billboard ranking. The rankings themselves are subsumed in the album fixed effects, but the interaction term varies by week. To make it easier to interpret the results, Billboard ranks are coded as [201 − actual rank] so that larger numbers indicate greater popularity.[28] We estimate these models using specification IV in Table 7. There is no indication that more popular artists are affected differentially. Neither the interaction terms nor the joint effect of the main and interaction terms are statistically significant.

[28]More precisely, the term is a three-way interaction: [downloads × indicator that the artist had a Billboard ranking × (201−Billboard rank)].

From a welfare point of view, it is particularly interesting to study variations in the effect of file sharing across younger and older artists because such differences might influence their decision to start and continue a career in music. Interacting downloads with the number of albums an artist produced, we find no significant differences across more or less experienced performers. Finally, we investigate whether the effect of downloads on sales varies with the number of popular songs on an album. As documented earlier, most file sharers obtain just a few songs from an album. One might suspect that P2P is a fairly good substitute for albums with only one or two popular songs. We calculate a Herfindahl index for each album-week as a measure of concentration of downloads. The index is included in both the first and the second stage. There is no evidence that albums with more concentrated downloads suffer disproportionately from file sharing.

## VII. Quasi-experimental Evidence

Our data also allow us to study the impact of P2P on sales in a quasi-experimental context. In particular we can examine how album sales respond to exogenous variation in file sharing intensity due to seasonality, geography, music genre, or secular growth. One of the advantages of this approach is that we can utilize several years of data, which allows us to investigate the long-term impact of file sharing. In all cases we continue to use sales data from Nielsen SoundScan (2005).

The first experiment involves variation over time. The number of file sharing users in the U.S. drops twelve percent over the summer (estimated from BigChampagne 2006) because college students are away from their high-speed campus Internet connections. If downloads crowd out sales, we should observe that the share of albums sold in the summer increases following the

advent of file-sharing. We consider a differences-in-differences approach and compare the share of summer sales in the period prior to file sharing (the control group) with sales following the introduction of file sharing (the treatment group). We calculate the share of album sales occurring in the May to September period using weekly SoundScan data. We find that the introduction of widespread file-sharing has had virtually no impact on summer sales. In the four years (1995-1998) preceding the introduction of Napster, the average share of summer sales was 37.0% with a range of 36.4-37.8%. During the more recent period of extensive file-sharing (1999-2005), the average share of summer sales was 37.2% with a range of 35.9-37.8%.

A second experiment considers spatial variation. Recall that U.S. users download over a third of their music files from Western European countries such as Germany and Italy. Due to time zone differences, such transfers are easier for East rather than West Coast users. This is because the peak file-sharing period (7pm to 3am) overlaps between Western Europe and the East Coast, which have a six hour time difference, but not between Europe and the West Coast, which have a nine hour difference. So East Coast users can draw on a larger base of files from international users than West Coast users. Consistent with these differences, we find that there is more file sharing on the East Coast than on the West Coast.[29] If file sharing had a large negative effect on record sales, then sales during the file sharing era should decrease more on the East Coast than on the West Coast. For the period 1998-2002, we obtained total album sales for the one hundred one largest "Designated Market Areas" from SoundScan. Despite the differences in the availability of files, sales have not noticeably varied across the country. In 1998, the last year in the pre-P2P period, the share of album sales in the Eastern Time Zone was 43.9%. This share has hardly moved since then. In 1999-2002, the mean was 43.5% and the range was 42.7-44.0%.

[29] Unfortunately, IP addresses can only be matched imperfectly to locations, so this finding is merely suggestive.

This is consistent with some common national factors, rather than file-sharing, driving sales trends.

A third experiment, which also provides a test of the "drop-out" hypothesis, is to see whether download intensity influences long-run sales growth after explicitly controlling for trends in music format popularity. The model for the period 1999-2005 is,

$$\text{Sales Growth}_g = \alpha + \gamma \times \text{Downloads}_g + \lambda \times \text{Listenership}_g + e_g \qquad (5)$$

where g indicates genre, $\text{Sales Growth}_g$ is the percentage growth in sales over 1999-2005, $\text{Downloads}_g$ are measures of genre-specific download intensity from our data, and $\text{Listenership}_g$ is the genre-specific radio listenership growth rate (Arbitron 2006) which controls for trends in popularity. Since downloading is relatively concentrated across genres (Table 3), the "drop-out" hypothesis predicts a greater sales reduction for genres which are popular on file sharing networks. The estimated $\gamma$ is not statistically significant using either download levels or downloads relative to purchases. For example, using mean downloads per album and controlling for genre sales levels, the estimated $\gamma$ is 0.05 with a standard error of 0.52 (the mean for downloads is 61.2, and for sales growth it is -5.8).

Finally, we consider whether growth in file sharing can be linked to changes in total album sales. The key question is whether periods of particularly rapid growth in the user-base are linked to sharper sales reductions. A simple test is to consider annual sales since the advent of widespread file sharing in 1999. According to SoundScan, album sales increased in three of the seven years over this period, in contrast to movie ticket sales which rose in only two years. It is worth stressing that extended sales slumps are common in the music business, even prior to file sharing. While real revenues have fallen 28% over 1999-2005, real revenue fell 35% during the

collapse of disco music in 1978-1983. Real sales also dropped 6% over 1994-1997.[30] More direct evidence comes from regressing total album sales, including paid digital downloads, on the average number of simultaneous file sharing users in the U.S. (BigChampagne 2006),

$$Sales_t = \gamma \times Users_t + \nu_m + \mu_t \qquad (6)$$

where t indicates a month, and $\nu_m$ are monthly fixed effects which account for seasonality. Using monthly data from August 2002-May 2006 (N=46) and defining Sales and Users in millions (with respective sample means of 56.0m and 5.0m), the estimated $\gamma$=-0.427 with a robust standard error of 0.33. There is little evidence that growth in the number of users has had a statistically or economically significant effect on sales.[31] The estimates remain insignificant if equation (6) is estimated in first differences.

The results of these quasi experiments are consistent with our earlier findings. Looking at variation in downloading intensity that is due to geography, seasonality, the genre of music, or secular growth, we find no evidence that the advent of P2P technology is the primary cause of the recent slump in music sales.

## VIII. Conclusions

Using detailed records of transfers of digital music files, we find that file sharing has had no statistically significant effect on purchases of the average album in our sample. Even our most negative point estimate (Table 7, model VI), implies that a one standard deviation increase in file-sharing reduces an album's weekly sales by a mere 368 copies, an effect that is too small to be statistically distinguishable from zero. Because our sample was constructed to be representative of the population of commercially relevant albums, we can use our estimates to

[30]These are calculated from nominal RIAA revenues listed in Lesk (2003) and RIAA (1998; 2006).
[31]If file sharing were eliminated, the point estimates imply monthly sales would only increase by 2.1m.

test hypotheses about the impact of P2P on the entire industry. Using ninety-five percent confidence bands, these tests are presented in Table 11. Taking into account all our (instrumented) estimates including the least precise results in Tables 7-9, we can reject a null that P2P caused a sales decline greater than 24.1 million albums. For reference, the music industry sold 803m CDs in 2002, which was a loss of 80m from the previous year (RIAA 2004). Our estimates become more precise if we relax the assumption that file sharing only impacts contemporaneous sales and if we allow for growth in the number of file sharers. For example, the scaled GMM models in Table 9 reject a null of losses greater than 6.6 million. Relying on our five most precise estimates, we conclude that the impact could not have been larger than 6.0 million albums. While file sharers downloaded billions of files in 2002, the consequences for the industry amounted to no more than 0.7% of sales

If file sharing is not the culprit, what other factors can explain the decline in music sales? Several plausible candidates exist. A first reason is the change in how music is distributed. Between 1999 and 2003, more than 14% of music sales shifted from record stores to more efficient discount retailers such as Wal-Mart, possibly reducing inventories. As a result, album shipments, which are often cited to document the decline in the legal demand of music, fell much more than actual sales.[32] A second factor is the ending of a period of atypically high sales, when consumers replaced older music formats with CDs. Perhaps more important than these developments is the growing competition from other forms of entertainment. A shift in entertainment spending towards recorded movies alone can largely explain the reduction in sales. The sales of DVDs and VHS tapes increased by over $5 billion between 1999 and 2003. This figure more than offsets the $2.6 billion reduction in album sales since 1999. Consumers also

[32] In the 1999 to 2003 period, the number of shipped albums fell by 301 million but the number of albums that were sold declined by only 99 million.

spent more on video games, where spending increased by 40%, or $3 billion, between 1999 and 2003, and on cell phones. Teen cell phone use alone tripled between 1999 and 2003.

An interesting question is whether our results continue to hold in more recent years. Since the time of our study, P2P technology has become more efficient, broadband access is much more widespread, and the number of file sharers has doubled. While a full analysis is outside the scope of this paper, there are several trends that are inconsistent with the view that P2P now displaces sales on a large scale. First, our natural experiments, for which we have data up to 2005, give no indication that file sharing has caused a sales decline in more recent years. Second, music sales have been flat or even rising in major markets with a quickly growing file-sharing population. For example, in 2005 retail music sales rose in four of the five largest national markets. Third, in the United States the entire drop in 2005 album sales is due to losses at a single firm, the recently merged Sony-BMG, which has experienced severe post-merger integration difficulties. If file sharing were responsible for the observed sales decline in the U.S., we would not expect this activity to only affect the products of a single firm.

The advent of the new P2P technologies can be considered in a broader context. A key question is how social welfare changes with weaker property rights for information goods. To make such a calculation, we would need to know how the production of music responds to the presence of file sharing. Based on our results, we do not believe file sharing had a significant effect on the supply of recorded music. For artists who produce commercially relevant products, the effects documented in this study are simply too small to change the number or quality of recordings that they release. And for new bands that are about to launch their career, the probability of success is so low as to make the expected income from producing music virtually zero, so file sharing will not change the relevant incentives. If we are correct in arguing that downloading has had

little effect on the incentives to produce music, we agree with Rob and Waldfogel (2006) who find that file sharing likely increased aggregate welfare. The limited shifts from sales to downloads are simply transfers between firms and consumers. But the sheer magnitude of P2P activity, the billions of songs downloaded each year, suggests the added social welfare from file sharing is likely to be high.

**References**


Agentur Lindner 2004. http://www.agentur-lindner.de/special/schulferien/index.html.

Arbitron 2006. *Format Trends Report.* http://wargod.arbitron.com/scripts/ndb/fmttrends2.asp.

Arellano, Manuel and Stephen Bond 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *The Review of Economic Studies* 58 (2): 277-97.

Bakos, Yannis, Erik Brynjolfsson and Douglas Lichtman 1999. "Shared Information Goods." *Journal of Law and Economics* 42: 117-156.

Berry, Steven 1994. "Estimating Discrete-Choice Models of Product Differentiation." *Rand Journal of Economics* 25: 242-262.

BigChampagne 2006. "Average Simultaneous U.S. Users: August 2002-May 2006." personal correspondence.

Billboard 2006. Fuzzy Math. 1 July: 24-26.

Blundell, Richard and Stephen Bond 1998. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics* 87 (1): 115-43.

Boldrin, Michele and David Levine 2002. "The Case Against Intellectual Property." *American Economic Review: Papers and Proceedings* 92: 209-212.

Bresnahan, Timothy, Scott Stern, and Manuel Trajtenberg 1997. "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the late 1980s." *Rand Journal of Economics* 28: S17-S44.

Business Software Alliance 2003. *Eighth Annual BSA Global Software Piracy Study.* http://www.bsa.org/globalstudy2003/index.cfm.

Central Intelligence Agency 2002. The World Factbook. https://www.cia.gov/cia/publications/factbook/index.html.

Central Intelligence Agency 2003. The World Factbook. https://www.cia.gov/cia/publications/factbook/index.html

Consumer Expenditure Survey 2004. http://www.bls.gov/cex/.

DEG 2005. "Industry Boosted by $21.2 Billion in Annual DVD Sales and Rentals." The Digital Entertainment Group. http://www.dvdinformation.com.

Edison Media Research 2003. *The National Record Buyers Study III.* Sponsored by Radio & Records. http://www.edisonresearch.com.

Einav, Liran (forthcoming). "Seasonality in the U.S. Motion Picture Industry". *Rand Journal of Economics*.

Forrester 2002. "Downloads Save the Music Business." http://www.forrester.com.

Forrester 2004. "US Antipiracy Bill Won't Stop File Sharing." http://www.forrester.com.

Gentzkow, Matthew forthcoming. "Valuing New Goods in a Model with Complementarity: Online Newspapers." *American Economic Review.*

Goolsbee, Austan 2000. "In a World Without Borders: The Impact of Taxes on Internet Commerce." *Quarterly Journal of Economics* 115: 561-576.

Goolsbee, Austan and Judy Chevalier 2003. "Price Competition Online: Amazon Versus Barnes And Noble," *Quantitative Marketing and Economics* 1 (2), June 2003, 203-222.

Gummadi, Krishna, Richard Dunn, Stefan Saroiu, Steven Gribble, Henry Levy, and John Zahorjan 2003. "Measurement, Modeling, and Analysis of a Peer-to-Peer File-Sharing Workload." Proceedings of the 19th ACM Symposium on Operating Systems Principles (SOSP-19).

International Federation of the Phonographic Industry 2002. *Recording Industry in Numbers 2001*. International Federation of Phonographic Industry.

Imbens, Guido and Joshua Angrist 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62: 467-475.

Internet2 Netflow Statistics (2004). *Internet2 NetFlow: Weekly Reports.* `http://netflow.Internet2.edu/weekly/.`

Jupiter Media Metrix (2002). "File Sharing: To Preserve Market Value Look Beyond Easy Scapegoats." `http://www.jupiterresearch.com.`

Karagiannis, Thomas, Andre Broido, Nevil Brownlee, kc claffy, and Michalis Faloutsos (2004). "Is P2P dying or just hiding?" Presented at Globecom 2004 in November-December 2004. `http://www.caida.org/outreach/papers/2004/p2p-dying/.`

Klein, Benjamin, Andres Lerner, and Kevin Murphy (2002). "The Economics of Copyright 'Fair Use' in a Networked World." *American Economic Review: Papers and Proceedings.* 92: 205-208.

Kultusministerkonferenz, Statistische Veröffentlichungen (2002). Nummer 162 vom August.

Lesk, Michael 2003. "Chicken Little and the Recorded Music Crisis." IEEE Security & Privacy (September): 73-75.

Liang, Jian, Rakesh Kumar, and Keith Ross 2004. "Understanding KaZaA." Manuscript, Polytechnic University.

MPAA 2005. "U.S. Entertainment Industry: 2004 MPA Market Statistics." Motion Picture Association: Worldwide Market Research.

Musikmarkt 2002. "Deutschland Single-Charts." `http://musikmarkt.lw-t1.thuecom-medien.de/content/charts/history.php3?jahr=2002.`

Nevo, Aviv 2001. "Measuring Market Power in the Ready-to-Eat Cereal Industry." *Econometrica* 69: 307-342.

Nielsen SoundScan (2005). http://home.soundscan.com/about.html.

Niesyto, Horst 2002. Digitale Spaltung - digitale Chancen: Medienbildung mit Jugendlichen aus benachteiligenden Verhältnissen. Mimeo, Pädagogische Hochschule Ludwigsburg.

Oberholzer-Gee, Felix and Koleman Strumpf 2005. "The Effect of File Sharing on Sales: An Empirical Analysis." Manuscript, Harvard Business School and the University of North Carolina at Chapel Hill.

OECD 2004. *OECD Information Technology Outlook 2004.* Paris: Organisation for Economic Co-operation and Development.

Plant, Arnold 1934. "The Economic Aspects of Copyright in Books." *Economica* 1: 167-195.

Posner, Richard 2005. "Intellectual Property: The Law and Economics Approach." *Journal of Economic Perspectives* 19 (2): 57-73.

RIAA 1998. *RIAA 1996 Statistical Overview.* Archived copy from the Internet archive, http://web.archive.org/web/19980124173720/www.riaa.com/market/releases/statover.htm.

RIAA 2004. *RIAA Market Data: The Cost of a CD.* Archived copy from the Internet archive, http://web.archive.org/web/20030416004543/http://www.riaa.com/MD-U.S.-7.cfm.

RIAA 2006. *The Recording Industry Association of America's 2005 Yearend Statistics.* http://www.riaa.com.

Rob, Rafael and Joel Waldfogel 2006. "Piracy on the High C's: Music Downloading, Sales Displacement, and Social Welfare in a Sample of College Students." *Journal of Law and Economics* 49(1): 29-62.

Shapiro, Carl and Hal Varian 1999. *Information Rules: A Strategic Guide to the Network Economy*. Boston: Harvard Business School Press.

Takeyama, Lisa 1994. "The Welfare Implications of Unauthorized Reproduction of Intellectual Property in the Presence of Demand Network Externalities." *The Journal of Industrial Economics* 42: 155-166.

Takeyama, Lisa 1997. "The Intertemporal Consequences of Unauthorized Reproduction of Intellectual Property." *Journal of Law & Economics* 40: 511-22.

Varian, Hal 2000. "Buying, Sharing and Renting Information Goods." *The Journal of Industrial Economics* 48: 473-488.

Windmeijer, Frank 2000. "A finite sample correction for the variance of linear two-step GMM estimators." Institute for Fiscal Studies, IFS Working Papers: W00/19.

Zentner, Alejandro 2006. "Measuring the Effect of Music Downloads on Music Purchases." *Journal of Law & Economics* 49(1): 63-90.

TABLE 1

SAMPLE SALES BY CATEGORY

| | Observations | Mean sales | Std dev | Min | Max |
|---|---|---|---|---|---|
| Full sample | 680 | 143,096 | 344,476 | 74 | 3,430,264 |
| Catalogue | 50 | 46,833 | 40,031 | 219 | 223,085 |
| Current Alternative | 117 | 118,599 | 130,257 | 9,210 | 785,747 |
| Hard Music Top Overall | 19 | 28,304 | 22,103 | 2,945 | 86,416 |
| Jazz Current | 21 | 21,940 | 62,522 | 86 | 290,026 |
| Latin | 21 | 27,590 | 35,840 | 3,143 | 153,209 |
| New artists | 50 | 15,816 | 13,635 | 319 | 61,673 |
| R&B | 144 | 46,512 | 67,050 | 2,151 | 457,338 |
| Rap | 76 | 39,307 | 61,278 | 1,069 | 324,426 |
| Top Current ("Billboard 200") | 83 | 744,022 | 710,054 | 4,092 | 3,430,264 |
| Top Current Country | 66 | 87,839 | 130,096 | 74 | 669,575 |
| Top Soundtrack | 33 | 44,920 | 79,264 | 1,788 | 318,538 |

NOTE.— These figures only include sales over our seventeen week observation period. Most of the top-selling albums are classified as "Current" for the purposes of this table

TABLE 2
THE GEOGRAPHY OF FILE SHARING
(numbers in %)

| Country | Share of users | Share of downloads | Users in U.S. download from (%) | Users in U.S. upload to (%) | Share World Population | Share World GDP | Share World Internet Users | Software Piracy Rate |
|---|---|---|---|---|---|---|---|---|
| United States | 30.9 | 35.7 | 45.1 | 49.0 | 4.6 | 21.2 | 27.4 | 23 |
| Germany | 13.5 | 14.1 | 16.5 | 8.9 | 1.3 | 4.5 | 5.3 | 32 |
| Italy | 11.1 | 9.9 | 6.1 | 5.7 | 0.9 | 2.9 | 3.2 | 47 |
| Japan | 8.4 | 2.8 | 2.5 | 1.8 | 2.0 | 7.2 | 9.3 | 35 |
| France | 6.9 | 6.9 | 3.8 | 4.7 | 1.0 | 3.1 | 2.8 | 43 |
| Canada | 5.4 | 6.1 | 6.9 | 7.9 | 0.5 | 1.9 | 2.8 | 39 |
| United Kingdom | 4.1 | 4.0 | 4.2 | 4.2 | 1.0 | 3.1 | 5.7 | 26 |
| Spain | 2.5 | 2.6 | 1.8 | 2.0 | 0.6 | 1.7 | 1.3 | 47 |
| Netherlands | 2.1 | 2.1 | 1.9 | 1.6 | 0.3 | 0.9 | 1.6 | 36 |
| Australia | 1.6 | 1.9 | 0.8 | 2.2 | 0.3 | 1.1 | 1.8 | 32 |
| Sweden | 1.5 | 1.7 | 1.8 | 1.5 | 0.1 | 0.5 | 1.0 | 29 |
| Switzerland | 1.4 | 1.5 | 0.9 | 1.0 | 0.1 | 0.5 | 0.6 | 32 |
| Brazil | 1.3 | 1.4 | 1.2 | 1.3 | 2.9 | 2.7 | 2.3 | 55 |
| Belgium | 0.9 | 1.2 | 0.5 | 1.0 | 0.2 | 0.6 | 0.6 | 31 |
| Austria | 0.8 | 0.6 | 0.6 | 0.4 | 0.1 | 0.5 | 0.6 | 30 |
| Poland | 0.5 | 0.7 | 0.7 | 0.5 | 0.6 | 0.8 | 1.1 | 54 |

NOTE.— Shares of users and downloads is from the file sharing dataset described in the text. All other statistics are from the Central Intelligence Agency (2002, 2003), except the software piracy rates which are from the Business Software Alliance (2003). All values are world shares, except the piracy rates are the fractions of business application software installed without a license in the country. All non-file sharing data are for 2002 except population which is for 2003.

TABLE 3
DOWNLOADS BY GENRE

| | # songs (# albums) in sample | Mean # of downloads | Std dev | Min | Max |
|---|---|---|---|---|---|
| | | | Song level | | |
| All genres | 10271 | 4.645 | 21.462 | 0 | 1258 |
| Catalogue | 714 | 4.361 | 10.370 | 0 | 152 |
| Alternative | 1707 | 7.021 | 18.153 | 0 | 312 |
| Hard | 270 | 4.830 | 8.684 | 0 | 52 |
| Jazz | 261 | 0.333 | 0.920 | 0 | 7 |
| Latin | 309 | 0.550 | 2.927 | 0 | 28 |
| New artists | 711 | 0.609 | 7.039 | 0 | 184 |
| R&B | 2249 | 1.635 | 7.680 | 0 | 159 |
| Rap | 1227 | 0.920 | 4.887 | 0 | 82 |
| Current | 1342 | 17.182 | 51.286 | 0 | 1258 |
| Country | 913 | 1.974 | 6.382 | 0 | 128 |
| Soundtrack | 568 | 1.673 | 5.301 | 0 | 61 |
| | | | Album level | | |
| All genres | 680 | 70.162 | 158.628 | 0 | 1799 |
| Catalogue | 50 | 62.280 | 103.114 | 0 | 680 |
| Alternative | 117 | 102.436 | 122.794 | 0 | 674 |
| Hard | 19 | 68.632 | 82.899 | 0 | 264 |
| Jazz | 21 | 4.143 | 4.542 | 0 | 13 |
| Latin | 21 | 8.095 | 26.344 | 0 | 121 |
| New artists | 50 | 8.660 | 33.097 | 0 | 229 |
| R&B | 144 | 25.542 | 56.494 | 0 | 433 |
| Rap | 76 | 14.855 | 24.487 | 0 | 119 |
| Current | 83 | 277.807 | 333.935 | 2 | 1799 |
| Country | 66 | 27.303 | 51.649 | 0 | 344 |
| Soundtrack | 33 | 28.788 | 36.611 | 0 | 185 |

TABLE 4

DOWNLOADS BY SALES – ALBUM LEVEL

| | Obs | Mean # of downloads | Std dev | Min | Max | Mann-Whitney |
|---|---|---|---|---|---|---|
| 1st quartile: mean 7,235 copies [up to 12,493 copies] | 170 | 11.358 | 38.472 | 0 | 402 | - 14.067** |
| 2nd quartile: mean 21,022 copies [up to 31,115 copies] | 170 | 20.929 | 52.082 | 0 | 433 | -12.431** |
| 3rd quartile: mean 57.940 copies [up to 100,962 copies] | 170 | 48.088 | 55.223 | 0 | 264 | -8.187** |
| 4th quartile: mean 486,184 copies [max 3,430,264 copies] | 170 | 200.270 | 265.369 | 0 | 1799 | |

NOTE.— Mann Whitney test statistics are for the null that the 4th quartile with the highest sales comes from the same population as the other sales quartiles.

** significant at the 1% level

TABLE 5

SUMMARY STATISTICS

| | Observations | mean (std dev) | min | max |
|---|---|---|---|---|
| Sales (1,000s) | 10093 | 9.580 | 0 | 874.137 |
| | | (34.361) | | |
| Downloads | 10093 | 4.360 | 0 | 368 |
| | | (13.644) | | |
| German kids on Vacation (million) | 10093 | 9.855 | 0 | 12.491 |
| | | (3.576) | | |
| Band on tour in Germany | 10093 | 0.003 | 0 | 1 |
| | | (0.053) | | |
| Misspelling indicator | 10093 | 0.062 | 0 | 1 |
| | | (0.187) | | |
| Rank of single on German charts (calculated as 101 minus rank) | 10093 | 1.576 | 0 | 100 |
| | | (10.268) | | |
| Rank of single on MTV charts (calculated as 101 minus rank) | 10093 | 2.158 | 0 | 100 |
| | | (13.568) | | |
| Billboard rank previous album (calculated as 201 minus rank) | 10093 | 61.136 | 0 | 200 |
| | | (82.314) | | |
| Best Billboard rank ever (calculated as 201 minus rank) | 10093 | 83.548 | 3 | 200 |
| | | (89.994) | | |
| # previous releases | 10093 | 6.718 | 0 | 194 |
| | | (15.574) | | |
| HHI downloads | 10093 | 2.460 | 0 | 10000 |
| | | (3.672) | | |

TABLE 6

DOWNLOAD TIMES: RELATION TO INSTRUMENTS AND IMPACT ON NUMBER OF TRANSFERS

| | (1) | (2) | (3) | (4) | (5) | (6) | |
|---|---|---|---|---|---|---|---|
| | | | | | | Impact of download time on download quantity | |
| | Time: Download Request to Initiation (sec) $C_1$ | Time: Search Request to Download Request (sec) $C_2$ | Time: Initiation Download to Completion (sec) $C_3$ | Ratio: # Search Requests to # Downloads $C_4$ | Percentage: Download Requests which are not completed $C_5$ | Download Time ($1^{st}$ stage) $C_1+C_2+C_3$ | Downloads ($2^{nd}$ stage) $D_{it}$ |
| German kids on Vacation (million) | -32.005 | -4.336 | -26.031 | -0.453 | -2.351 | -62.420 | |
| | (5.51)** | (0.29)** | (2.69)** | (0.05)** | (0.10)** | (5.24)** | |
| German kids × Band on tour | -49.914 | -3.966 | -35.015 | -0.480 | -2.927 | -89.010 | |
| | (20.31)* | (1.73)* | (13.35)** | (0.22)* | (0.51)** | (17.83)** | |
| German kids × Misspellings | 22.494 | 6.157 | 8.609 | 0.672 | 1.963 | 7.302 | |
| | (33.66) | (2.182)** | (17.76) | (0.25)** | (0.58)** | (40.59) | |
| German kids × rank German charts | -0.347 | -0.034 | -0.471 | -0.005 | -0.024 | -0.849 | |
| | (0.18)* | (0.02) | (0.16)* | (0.00)* | (0.01)* | (0.22)** | |
| Download time | | | | | | | -0.006 |
| | | | | | | | (0.00)** |
| Album Fixed Effects? | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1662 | 1952 | 1332 | 2164 | 1952 | 1332 | 1332 |
| Mean for Dependent Variable | 609.08 | 91.02 | 796.20 | 12.21 | 62.96 | 1491.18 | 7.25 |

NOTE.— Albums or album-weeks are omitted when the dependent variable is undefined (e.g. for $C_1$ when there are no successful album download initiations). Robust standard errors are in parentheses. These estimates are based on data from weeks 3-6 of our observation period (the data come from more detailed log files which are only available during these weeks).

* significant at the 5% level

** significant at the 1% level

TABLE 7

PANEL ANALYSIS - DOWNLOADS AND ALBUM SALES

| | (1) | (2) | | (3) | | (4) | | (5) | | (6) | | (7) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Sales | 1st stage down-loads | 2nd stage Sales | 1st stage down-loads | 2nd stage sales | 1st stage down-loads | 2nd stage Sales | 1st stage down-loads | 2nd stage sales | 1st stage down-loads | 2nd stage Sales | 1st stage down-loads | 2nd stage sales |
| # downloads | 0.277 | | 0.003 | | 0.024 | | -0.010 | | 0.005 | | -0.027 | | 0.037 |
| | (0.025)** | | (0.194) | | (0.189) | | (0.158) | | (0.062) | | (0.270) | | (0.065) |
| German kids on vacation | | 0.671 | | 0.670 | | 0.667 | | 1.818 | | | | | |
| | | (0.054)** | | (0.054)** | | (0.054)** | | (0.125)** | | | | | |
| German kids × band on tour | | | | 0.469 | | 0.474 | | 0.470 | | 0.464 | | 0.451 | |
| | | | | (0.168)** | | (0.167)** | | (0.161)** | | (0.167)** | | (0.161)** | |
| German kids × Misspellings | | | | | | -0.288 | | | | -0.290 | | | |
| | | | | | | (0.124)* | | | | (0.124)* | | | |
| German kids × Germ charts | | | | | | 0.012 | | 0.007 | | 0.012 | | 0.007 | |
| | | | | | | (0.001)** | | (0.002)** | | (0.001)** | | (0.002)** | |
| U.S. MTV rank | 0.079 | 0.036 | 0.089 | 0.037 | 0.088 | 0.035 | 0.089 | 0.058 | -0.194 | 0.036 | 0.092 | -0.042 | -0.183 |
| | (0.020)** | (0.008)** | (0.021)** | (0.008)** | (0.021)** | (0.008)** | (0.021)** | (0.103) | (0.256) | (0.008)** | (0.022)** | (0.102) | (0.255) |
| German kids × album FE | No | No | No | No | No | No | No | Yes | No | No | No | Yes | No |
| MTV × album FE | No | No | No | No | No | No | No | Yes | Yes | No | No | Yes | Yes |
| Polynomial time trend | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | No | No | No | No |
| Week FE | No | No | No | No | No | No | No | No | No | Yes | Yes | Yes | Yes |
| Album FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 10093 | 10093 | 10093 | 10093 | 10093 | 10093 | 10093 | 10093 | 10093 | 10093 | 10093 | 10093 | 10093 |
| Prob $\chi^2$>0 on excluded instruments | | 0.0000 | | 0.0000 | | 0.0000 | | 0.0000 | | 0.0000 | | 0.0000 | |
| Sargan test (p-value) | | | | | 0.73 | | 0.70 | | 0.98 | | 0.50 | | 0.97 |
| *R*-squared | 0.75 | 0.74 | 0.76 | 0.74 | 0.76 | 0.73 | 0.76 | 0.74 | 0.79 | 0.82 | 0.77 | 0.85 | 0.79 |

NOTE.— The unit of analysis is the album-week. Dependent variables are the number downloads at the 1st stage (summing all songs on an album) and album sales (1,000s).

Robust standard errors are in parentheses. Since all models include album fixed effects, the reported *R*-squared is the sum of the explained within-variance and the fraction of the variance that is due to the fixed effects. Album-weeks prior to the release date are excluded from the sample.

* significant at the 5% level

** significant at the 1% level

TABLE 8

DYNAMIC PANEL ANALYSIS - DOWNLOADS AND LAGGED ALBUM SALES

| | (1) 2nd stage Sales | (2) 2nd stage sales | (3) 2nd stage Sales | (4) 2nd stage sales | (5) GMM Δ sales | (6) GMM Δ sales |
|---|---|---|---|---|---|---|
| Weighted ∑ of three weeks of downloads (instrumented) | 0.097 (0.115) | 0.048 (0.039) | 0.022 (0.170) | 0.045 (0.041) | | |
| Δ downloads | | | | | 0.029 (0.074) | 0.047 (0.078) |
| U.S. MTV rank | 0.092 (0.015)** | -0.016 (0.169) | 0.097 (0.016)** | -0.022 (0.168) | 0.085 (0.091) | 0.041 (0.080) |
| lagged sales | | | | | 0.166 (0.100)) | 0.261 (0.117)* |
| German kids × album FE in 1st stage | No | Yes | No | Yes | No | No |
| MTV × album FE | No | Yes | No | Yes | No | No |
| Polynomial time trend? | Yes | Yes | No | No | Yes | No |
| Week Fixed Effects? | No | No | Yes | Yes | No | Yes |
| Album Fixed Effects? | Yes | Yes | Yes | Yes | No | No |
| 1st-stage specification is as in Table 7, model | 4 | 5 | 6 | 7 | | |
| Observations | 8739 | 8739 | 8739 | 8739 | 8739 | 8739 |
| Arellano-Bond test for AR(1) in first differences: Pr > z | | | | | 0.302 | 0.204 |
| Arellano-Bond test for AR(2) in first differences: Pr > z | | | | | 0.638 | 0.522 |
| *R*-squared | 0.92 | 0.96 | 0.92 | 0.97 | | |

NOTE.— The dependent variable is album sales (1,000s). The number of downloads is instrumented using the Table 7 specification listed in the fifth row from the bottom. The weighted sum of three weeks of downloads includes the current week. The weights are chosen in a grid search which minimizes the unexplained fraction of the variance in our models. Models (5) and (6) use the Generalized Method of Moments estimator developed by Arellano and Bond (1991). In this model, the typical standard error estimator tends to be downwards biased (Blundell and Bond 1998). Standard errors are corrected using the two-step covariance matrix derived by Windmeijer (2000). Arellano-Bond tests for autocorrelation are applied to the first-difference equation residuals. Second-order autocorrelation would indicate that some lags of the dependent variable which are used as instruments are endogenous. The tests reveal no such problem. Album-weeks prior to the release date are excluded from the sample.

* significant at the 5% level

** significant at the 1% level

TABLE 9

ROBUSTNESS CHECK WITH SCALED DOWNLOADS – TESTING THE "DROP-OUT" HYPOTHESIS

| | (1) | | (2) | | (3) | | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1st stage downloads | 2nd stage Sales | 1st stage downloads | 2nd stage sales | 1st stage downloads | 2nd stage sales | 2nd stage Sales | 2nd stage Sales | 2nd stage Sales | GMM Δ sales | GMM Δ sales |
| Scaled downloads | | -0.009 | | 0.022 | | 0.029 | | | | | |
| | | (0.126) | | (0.046) | | (0.049) | | | | | |
| Weighted ∑ of three Weeks downloads | | | | | | | 0.078 | 0.038 | 0.037 | | |
| | | | | | | | (0.093) | (0.030) | (0.031) | | |
| Δ downloads | | | | | | | | | | 0.072 | 0.123 |
| | | | | | | | | | | (0.053) | (0.072) |
| German kids on Vacation (million) | 0.856 | | 2.608 | | | | | | | | |
| | (0.073)** | | (0.171)** | | | | | | | | |
| German kids × Band on tour | 0.602 | | 0.600 | | 0.585 | | | | | | |
| | (0.225)** | | (0.216)** | | (0.216)** | | | | | | |
| German kids × Misspellings | -0.377 | | | | | | | | | | |
| | (0.167)* | | | | | | | | | | |
| German kids × rank German charts | 0.014 | | 0.008 | | 0.008 | | | | | | |
| | (0.002)** | | (0.002)** | | (0.002)** | | | | | | |
| U.S. MTV rank | 0.036 | 0.089 | -0.084 | -0.198 | -0.059 | -0.182 | 0.093 | 0.139 | -0.023 | 0.085 | 0.044 |
| | (0.011)** | (0.020)** | (0.137) | (0.255) | (0.137) | (0.255) | (0.015)** | (0.158) | (0.168) | (0.097) | (0.077) |
| Lagged sales | | | | | | | | | | 0.166 | 0.261 |
| | | | | | | | | | | (0.101) | (0.118)* |
| German kids × album FE in 1st stage | No | No | Yes | Yes | Yes | Yes | No | Yes | Yes | No | No |
| MTV × album FE | No | No | Yes | Yes | Yes | Yes | No | Yes | Yes | No | No |
| Polynomial time trend | Yes | Yes | Yes | Yes | No | No | Yes | Yes | No | Yes | No |
| Week Fixed Effects? | No | No | No | No | Yes | Yes | No | No | Yes | No | Yes |
| Album Fixed Effects? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | No | No |
| Specification as in Table (model) | 7 (4) | 7 (4) | 7 (5) | 7 (5) | 7 (7) | 7 (7) | 8 (1) | 8 (2) | 8 (4) | 8 (5) | 8 (6) |
| Observations | 10093 | 10093 | 10093 | 10093 | 10093 | 10093 | 8739 | 8739 | 8739 | 8739 | |
| *R*-squared | 0.74 | 0.76 | 0.85 | 0.79 | 0.87 | 0.79 | 0.82 | 0.86 | 0.87 | | |
| AB test for AR(1) | | | | | | | | | | 0.305 | 0.201 |
| AB test for AR(2) | | | | | | | | | | 0.643 | 0.531 |

NOTE.— Dependent variables are album sales (1,000s) and scaled downloads at the 1st stage. Downloads are scaled to reflect the growth of KaZaA users over the sample period. For the fixed-effects models, the reported *R*-squared is the sum of the explained within-variance and the fraction of the variance that is due to the fixed effects. Album-weeks prior to the release date are excluded from the sample.

* significant at the 5% level

** significant at the 1% level

TABLE 10

ROBUSTNESS CHECKS

| Table 7 (4) Coefficient downloads (std. error) | Table 7 (6) Coefficient downloads (std. error) | Table 7 (7) Coefficient downloads (std. error) | *N* | Specification |
|---|---|---|---|---|
| -0.010<br>(0.158) | 0.005<br>(0.062) | 0.037<br>(0.065) | 10093 | Benchmark specifications, models (4), (6) and (7) in Table 7 |
| | Changes in Sample | | | |
| 0.064<br>(0.376) | -0.001<br>(0.108) | -0.013<br>(0.112) | 7399 | Without holiday sales |
| 0.018<br>(0.166) | 0.034<br>(0.071) | 0.079<br>(0.075) | 7890 | Without albums that are not downloaded |
| 0.051<br>(0.184) | 0.083<br>(0.090) | 0.161<br>(0.097) | 5033 | Albums that sell more than 151,284 copies ($50^{th}$ percentile) during the sample period |
| 0.037<br>(0.135) | 0.062<br>(0.055) | 0.092<br>(0.058) | 8567 | Without Latin and Country albums |
| | Changes in Model Specification | | | |
| -0.006<br>(0.007) | 0.001<br>(0.003) | 0.004<br>(0.003) | 10093 | Dependent variable is log of sales |
| 0.083<br>(0.029)** | 0.019<br>(0.026) | 0.005<br>(0.022) | 3232 | Sales and downloads are expressed as percentage changes |
| | Does the estimated effect vary by popularity? | | | |
| Main effect downloads | Interaction | $H_0$ sum = 0 (Prob > F) | | Downloads (instrumented) are interacted with… |
| -0.095<br>(0.185) | 0.001<br>(0.001) | 0.6119 | 10093 | Billboard rank of artist's prior album |
| -0.130<br>(0.192) | 0.001<br>(0.001) | 0.5015 | 10093 | Best Billboard rank for artist during career |
| 0.002<br>(0.181) | 0.002<br>(0.007) | 0.9822 | 10093 | Number of previous albums |
| -0.128<br>(0.175) | 0.039<br>(0.026) | 0.5917 | 10093 | Herfindahl index measuring concentration of downloads |

NOTE.— Dependent variables are album sales (1,000s) and # downloads at the $1^{st}$ stage. Robust standard errors are in parentheses. For the popularity results in the lower panel, the specification is model (5) in Table 7. Album-weeks prior to the release date are excluded from the sample.

* significant at the 5% level

** significant at the 1% level

TABLE 11
HYPOTHESES TESTS

| Class of Models | Lower bound of 95% confidence interval Can reject hypothesis that the impact of file sharing is larger than (in million albums) |
|---|---|
| All models (Tables 7 through 9) | -24.1 |
| Models with German vacation × Album FE interactions | -12.7 |
| Models with scaled downloads (Table 9) | -12.4 |
| GMM models with scaled downloads (Table 9) | -6.6 |
| 5 models with smallest standard errors | -6.0 |

NOTE.— These values represent the overall, industry-wide impact of file sharing for 2002 as implied by the various specifications. The lower bound is the minimum of the 95% confidence interval around the mean impact. Details of this calculation are listed below. The second column of each row reports the median lower bound for that class of models.

The lower bound is calculated as $\sum_t\sum_i (D_{it}\times 5.04\times 1000)\times(\gamma - 2\times se(\gamma)) = 240m\times(\gamma - 2\times se(\gamma))$, where $\gamma$ is the point estimate from equation (1). The factor 5.04 scales the results from our sample to all releases and the entire year 2002. It is calculated as: Aggregate impact = (Effect of file sharing on sample sales over observation period) × (population sales/sample sales) × (file sharing activity over year/file sharing activity in observation period). From our sales data, the ratio (population sales/sample sales) is 2.27. The second ratio is (File sharing activity over year/file sharing activity in observation period) = 2.22, which is calculated from weekly file sharing traffic rates over the 2002 calendar year on the Internet2 backbone (Internet2 Netflow Statistics 2004) and the monthly average number of U.S. file sharing users (BigChampagne 2006). Note that the second conversion factor is close to a naïve correction based simply on time, (52 weeks in year/17 weeks in observation period) = 3.06.

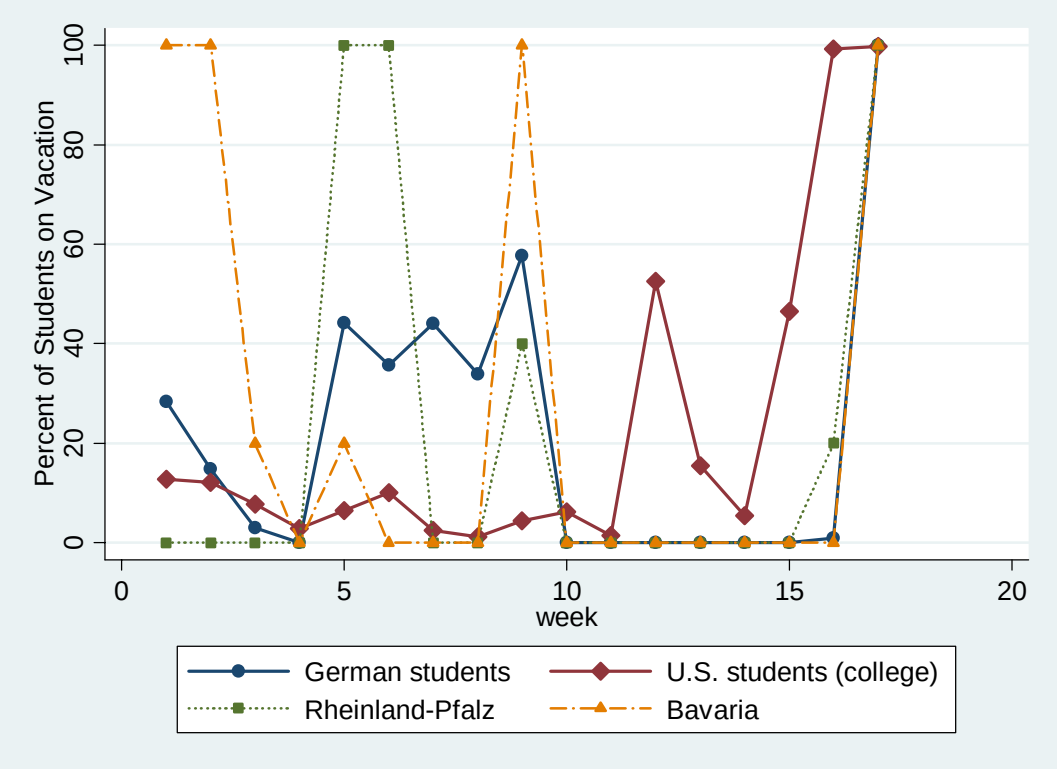



Fig. 1.— Timing of German and U.S. School Vacations

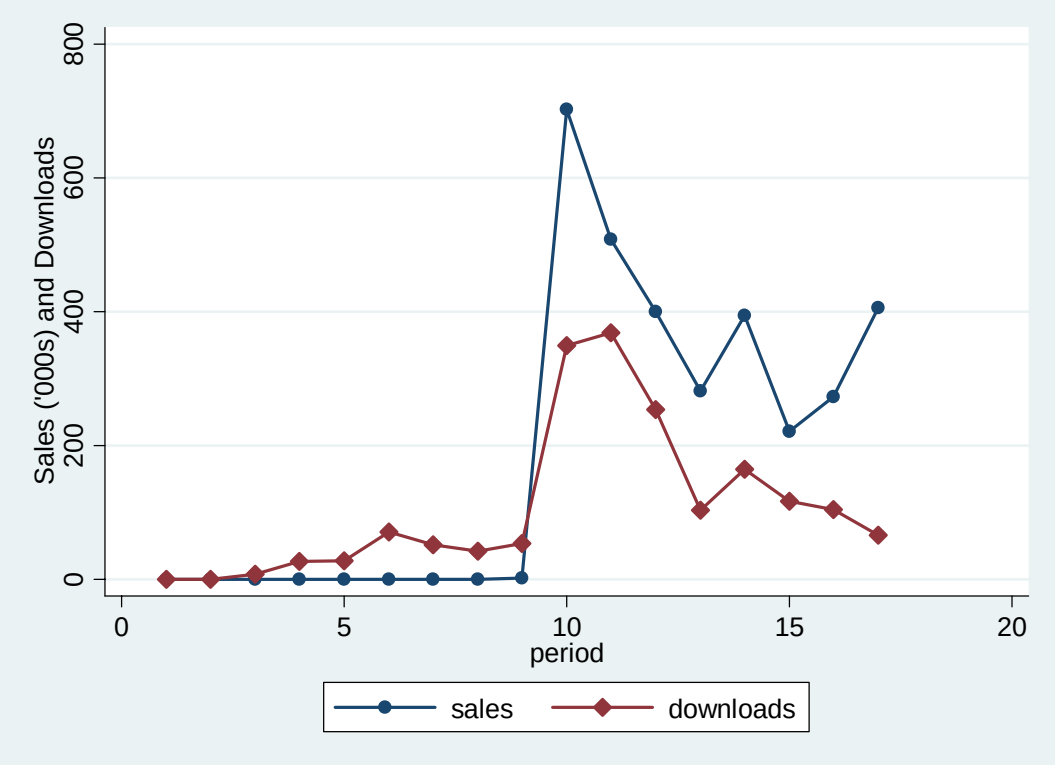



Fig. 2.– Dynamics of Downloads and Albums Purchases for a Popular Album
(by week, sales in thousands)